Research Article

Hao Wang and Yu Cheng*

Piecewise Cause-Specific Association Analyses of Multivariate Untied or Tied Competing Risks Data

Abstract: In this paper we extend the bivariate hazard ratio to multivariate competing risks data and show that it is equivalent to the cause-specific cross hazard ratio. Two approaches are proposed to estimate these two equivalent association measures. One extends the plug-in estimator, and the other adapts the pseudo-likelihood estimator for bivariate survival data to multivariate competing risks data. The asymptotic properties of the extended estimators are established by using empirical processes techniques. The extended plug-in and pseudo-likelihood estimators have comparable performance with the existing U-statistic when the data have no tied events. However, in many applications, there are tied events in which all the three estimators are found to produce biased results. To our best knowledge, we are not aware of any association analysis for multivariate competing risks data that has considered tied events. Hence we propose a modified U-statistic to specifically handle tied observations. The modified U-statistic clearly outperforms the other estimators when there are rounding errors. All methods are applied to the Cache County Study to examine mother–child and sibship associations in dementia among this aging population, where the event times are rounded to the nearest integers. The modified U performs consistently with our simulation results and provides more reliable results in the presence of tied events.

Keywords: cause-specific hazard function, cross hazard ratio, pseudo-likelihood estimator, tied events, U-statistic

DOI 10.1515/ijb-2013-0023

1 Introduction

Multivariate competing risks data occur often in genetic familial studies and demography as there are usually more than one risk involved. As a motivating example, we consider a study conducted in Cache County, Utah on dementia in an aging population. One goal of this study is to measure the association in dementia onset among family members. For each family member, the researchers recorded his or her onset age of dementia, age at death, or age at termination of the study, whichever occurred first, along with the corresponding cause indicator. This falls into the paradigm of multivariate competing risks, where there is within-subject dependence between the risk of interest and competing risks, as well as between-subject dependence among family members induced by common genetic or environmental factors.

Many researchers have studied association analysis for bivariate survival data with independent right censoring. An overview of their work was given by Hougaard [1]. The previous work, however, cannot be applied directly to multivariate competing risks data because the event of interest may not be observed due to the occurrence of a competing event. One limitation of competing risks data is that the within-subject dependence between the target event and competing events is nonparametrically untestable, and their marginal distributions are nonparametrically nonidentifiable [2, 3]. The nonidentifiability issue raises difficulties in analyzing the association in dementia in the presence of possibly dependent censoring by death, which occurred frequently among this aging population.

*Corresponding author: Yu Cheng, Department of Statistics and Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA 15260, USA, E-mail: yucheng@pitt.edu

Hao Wang, Amgen, Thousand Oaks, CA, USA, E-mail: hao@amgen.com

Some of the association analyses for bivariate survival data with independent censoring have been recently adapted to bivariate competing risks data Bandeen-Roche and Liang [4] modified Oakes's [5] cross hazard ratio to the bivariate competing risks setting. They defined a cause-specific association measure $\theta_{CS}(s, t; k, l)$ which gives the factor by which a child's risk of having cause k event at time s compares if his/her mother has experienced the cause l event by time t and has not yet experienced any event by time t . To estimate $\theta_{CS}(s, t; k, l)$, Bandeen-Roche and Liang [4] proposed a U-type estimator based on a local version of Kendall's τ with inference using standard theory for U-statistics [6]. Cheng and Fine [7] defined a ratio $\zeta(s, t; k, l)$ of the bivariate hazard with cause-specific events from both subjects to the products of the bivariate hazards with a single event from each subject as an association measure. They established the equivalence of $\zeta(s, t; k, l)$ to the association measure $\theta_{CS}(s, t; k, l)$ and developed a plug-in estimator with an explicit variance form.

More recently, Shih and Albert [8] examined familial association in breast cancer based on a cross hazard ratio for paired times of first event and an odds ratio for two competing causes (cancer and non-cancer deaths) and estimated the two association measures by adopting the two-stage estimation method [9]. They noted that the cause-specific hazard ratio $\theta_{CS}(s, t; k, l)$ may be more clinically relevant than their cross hazard ratio for the first event times of paired members. Hence, we focus on the cause-specific association measures θ and ζ which have straightforward interpretations as the two are essentially a local correlation between indicator functions.

Most of the previous works related to θ_{CS} or ζ focused on bivariate competing risks data with few exceptions. Cheng et al. [10] adapted the cause-specific cross hazard ratio θ_{CS} for bivariate data to sibship data and multiple mother-child pairs and greatly broadened its scope of application. Gorfine and Hsu [11] proposed a flexible frailty model for multivariate competing risks data based on cause-specific hazard functions, which includes the frailty model considered in Bandeen-Roche and Liang [4] and the positive-stable model in Katki et al. [12] as special cases. However, the modeling and estimation approach in Gorfine and Hsu [11] requires the specification of a proper frailty model and the use of the EM algorithm with numerical integration. The focus of this paper is on nonparametric estimation of the cause-specific association measures for multivariate competing risks data within some pre-specified time regions, in line with the work by Cheng et al. [10].

Therefore, we first extend the association measure ζ from bivariate competing risks data to clustered data and develop a plug-in estimator as is done in the bivariate case. Next, we adapt a pseudo-likelihood estimator for bivariate survival data [13, 14] to multivariate competing risks data. The asymptotic properties of the two proposed estimators are established by using empirical processes techniques. The finite performances of the two proposed methods are compared with those of the existing U-estimator [10] for clustered competing risks data. Interestingly, we have found that all the three estimators produce biased results when the data have tied events. Rounding errors are common in many applications as the event times are often rounded to their nearest integers. Somewhat surprisingly, we are not aware of any work that addresses how to handle tied events in evaluating concordant status between two pairs of event times of interest. Therefore, we develop a modified U-estimator that explicitly takes into account the tied events and show that it outperforms the other three estimators for tied data. To our best knowledge, it is the first attempt to address the issue of rounding errors in nonparametric association analyses of bivariate and multivariate competing risks data.

The rest of the paper is organized as follows. We extend the plug-in estimator to multivariate competing risks data in Section 2.2, briefly describe the notation and the method for the existing U-estimator [10] in Section 2.3, and propose the pseudo-likelihood estimator in Section 2.4. In Section 2.5 tied event data are considered, and a modified U-statistic is proposed to specifically handle rounding errors. In Section 2.6 we will discuss some formal tests on whether the associations are constant over time. The extended plug-in estimators and pseudo-maximum-likelihood estimators are compared with the U-statistics for untied data via simulation studies in Sections 3.1–3.3. In Section 3.4 we investigate the impact of rounding errors on the performance of the three estimators in comparison to the modified U-estimator. All methods are applied to the Cache County Study in Section 4, and some remarks are given in Section 5.

2 Methods

2.1 Multivariate competing risks data

In this paper, we consider clustered competing risks data. Suppose that in each cluster there are $M + 1$ failure times $(T_0, T_1, \dots, T_M)$ and their corresponding cause indicators $(\varepsilon_0, \varepsilon_1, \dots, \varepsilon_M)$ from a mother and her M children, where M is random and is assumed to bear no information on the association. We only consider two causes here $\varepsilon_j = 1, 2, j = 0, \dots, M$, as we are primarily interested in association in cause 1 events and can always group multiple competing events together. In addition, there may also be independent and non-informative censoring times $(C_0, C_1, \dots, C_M)$. Note that the data were collected retrospectively, hence it is reasonable to assume censoring times are independent of event times. This is in contrast with a perspective study where there may be some dependence among censoring times from mothers and children due to their inherent age differences. Hence one observes $Y_j = \min(T_j, C_j)$, and $\eta_j = \varepsilon_j I\{T_j \leq C_j\}$, for $j = 0, 1, \dots, M$, where I is the indicator function. This notation incorporates two types of multivariate competing risks data. One is the sibship data, where pairwise exchangeability can be assumed among siblings, i.e. all siblings have the same marginals and the pairwise joint densities are symmetric. The other consists of multiple mother and child pairs where we do not impose the exchangeability assumption between a mother and her child, but assume exchangeability among all children. The observed data containing n i.i.d. clusters are denoted as $\{(Y_{i0}, \dots, Y_{iM}, \eta_{i0}, \dots, \eta_{iM}, m_i), i = 1, \dots, n\}$. The following two association measures are considered for each of the two types of multivariate competing risks data.

2.2 Bivariate cause-specific hazard ratio for clustered data and its plug-in estimator

Cheng and Fine [7] proposed a cause-specific association measure based on bivariate cause-specific hazard functions and demonstrated its equivalence to the conditional cross hazard ratio [4]. We now extend their association measure to the clustered competing risks data. To simplify the notation we focus on association analysis of cause 1 events, though the methods proposed here are applicable to other causes. Analogous to Cheng and Fine [7], for any child–mother pair $(T_j, T_0), 1 \leq j \leq M$, we define a bivariate cause-specific hazard function for double cause 1 events from both subjects under the exchangeability assumption for all children

$$\lambda_{11}(s, t) = \lim_{(\Delta s, \Delta t) \downarrow 0} P(T_j \in [s, s + \Delta s], \varepsilon_j = 1, T_0 \in [t, t + \Delta t], \varepsilon_0 = 1 | T_j \geq s, T_0 \geq t) / (\Delta s \Delta t),$$

and define bivariate hazard functions of a single cause 1 event from the child or mother with the other member being at risk as the following:

$$\lambda_{10}(s, t) = \lim_{\Delta s \downarrow 0} P(T_j \in [s, s + \Delta s], \varepsilon_j = 1 | T_j \geq s, T_0 \geq t) / \Delta s, \text{ and}$$

$$\lambda_{01}(s, t) = \lim_{\Delta t \downarrow 0} P(T_0 \in [t, t + \Delta t], \varepsilon_0 = 1 | T_j \geq s, T_0 \geq t) / \Delta t.$$

Then the association measure for cause 1 events based on these bivariate hazard functions is

$$\zeta_{CM}(s, t; 1, 1) = \frac{\lambda_{11}(s, t)}{\lambda_{10}(s, t)\lambda_{01}(s, t)}. \quad (1)$$

It quantifies the elevated risk that both a child and his/her mother failed at times (s, t) , as compared to the product of the risks that only one of them failed at s or t , conditionally on both the child and the mother being at risk at (s, t) .

A simple plug-in estimator of $\zeta_{CM}(s, t; 1, 1)$ can be constructed analogously to Cheng and Fine [7] based on all child–mother pairs. We first need to extend the event processes to the more complicated family

structure with multiple children in a family. Let $N_{11}^*(s, t) = \sum_j^M I(Y_j \leq s, \eta_j = 1, Y_0 \leq t, \eta_0 = 1)$, $N_{10}^*(s, t) = \sum_j^M I(Y_j \leq s, \eta_j = 1, Y_0 \geq t)$, $N_{01}^*(s, t) = \sum_j^M I(Y_j \geq s, Y_0 \leq t, \eta_0 = 1)$, $H^*(s, t) = \sum_j^M I(Y_j \geq s, Y_0 \geq t)$, where M is the number of children in a family which may vary from family to family. Cheng et al. [10] divided the time scale as follows: $0 = \tau_{1,0} < \tau_{1,1} < \tau_{1,2} < \dots < \tau_{1,n_1} = \tau_1$, $0 = \tau_{2,0} < \tau_{2,1} < \tau_{2,2} < \dots < \tau_{2,n_2} = \tau_2$, where τ_1 and τ_2 are the largest observed event times. These grids are predetermined based on practical interests as well as the feasibility of analyzing the data (i.e. the number of at risk or the number of the events should not be too small in any time grid). Assuming constant association measure $\zeta_{CM}(\Omega_{qr})$ over the region $\Omega_{qr} \equiv (\tau_{1,q}, \tau_{1,q+1}) \times (\tau_{2,r}, \tau_{2,r+1})$ with $0 \leq q \leq n_1 - 1, 0 \leq r \leq n_2 - 1$, we have

$$\int_{(s,t) \in \Omega_{qr}} w(s, t) \{EN_{11}^*(ds, dt)EH^*(s, t) - \zeta_{CM}(\Omega_{qr})EN_{10}^*(ds, t)EN_{01}^*(s, dt)\} = 0,$$

where E is the expectation operator and $w(s, t)$ is some bounded deterministic function. The above equation is derived by recasting the hazard functions in eq. (1) in terms of the expectations of N_{11}^* , N_{10}^* , N_{01}^* , and H^* , similar to the one given in Cheng and Fine [7]. That is, $EN_{11}^*(s, t) = E(M)P(Y_1 \leq s, \eta_1 = 1, Y_0 \leq t, \eta_0 = 1)$ and the expectations of the other three stochastic processes can also be expressed as the products of $E(M)$ and the probabilities of a single event or at risk. Therefore, $\zeta_{CM}(\Omega_{qr}) = \frac{EN_{11}^*(ds, dt)EH^*(s, t)}{EN_{10}^*(ds, t)EN_{01}^*(s, dt)}$ and the above equation follows automatically.

Define the empirical process $\mathbb{P}_n N_{11}^*(s, t) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{m_i} I(Y_{ij} \leq s, Y_{i0} \leq t, \eta_{ij} = 1, \eta_{i0} = 1)$, and similarly for $\mathbb{P}_n H^*(s, t)$, $\mathbb{P}_n N_{10}^*(s, t)$, and $\mathbb{P}_n N_{01}^*(s, t)$. Since the number of children M is random, the above formula naturally handles different cluster sizes. Plugging the empirical processes into the above equation yields a plug-in estimator

$$\hat{\zeta}_{CM}(\Omega_{qr}) = \frac{\int_{(s,t) \in \Omega_{qr}} \hat{w}(s, t) \mathbb{P}_n N_{11}^*(ds, dt) \mathbb{P}_n H^*(s, t)}{\int_{(s,t) \in \Omega_{qr}} \hat{w}(s, t) \mathbb{P}_n N_{10}^*(ds, t) \mathbb{P}_n N_{01}^*(s, dt)}, \quad (2)$$

where $\hat{w}$ is a weight function satisfying $\int_{(s,t) \in \Omega_{qr}} \hat{w}(s, t) ds dt = 1$, $\|\hat{w} - w\|_\infty \rightarrow 0$, and $\sqrt{n}(\hat{w} - w) = O_P(1)$. The selection of the weight function $\hat{w}$ deserves some attention. If the association is constant within the region Ω_{qr} , different weight functions usually would yield comparable estimates of ζ_{CM} . In this case, the uniform weight $\hat{w}(s, t) = \frac{1}{(\tau_{1,q+1} - \tau_{1,q})(\tau_{2,r+1} - \tau_{2,r})}$ is a convenient choice. However, if the association is time-varying within Ω_{qr} , the weight will play an important role in summarizing the time-varying association. When $\hat{w}(s, t) = \frac{1}{\mathbb{P}_n N_{10}^*(ds, t) \mathbb{P}_n N_{01}^*(s, dt)} / \int_{\Omega_{qr}} \frac{1}{\mathbb{P}_n N_{10}^*(ds, t) \mathbb{P}_n N_{01}^*(s, dt)}$, we are averaging the point-wise association across Ω_{qr} with equal weight. One may consider other weighting schemes such as putting more weights on early time points at which the number of at risk is large.

The consistency and asymptotic normality of $\hat{\zeta}_{CM}(\Omega_{qr})$ can be established using empirical processes techniques [15, 16]. Given $EM^2 < \infty$ we can show that all the above indicator processes are P-Donsker. Since $\hat{\zeta}_{CM}$ is a compact differential map of empirical processes indexed by Donsker classes, its consistency and asymptotic normality follow. In addition, there is an asymptotically linear representation $n^{1/2}(\hat{\zeta}_{CM} - \zeta_{CM}) = n^{-1/2} \sum_i I_i + o_P(1)$, where I_i are influence functions, upon which the variance of $\hat{\zeta}_{CM}$ can be estimated explicitly. Detailed proofs are given in the section “Asymptotic properties of $\hat{\zeta}_{CM}$ ” in Appendix.

The association measure $\zeta(s, t; 1, 1)$ for bivariate competing risks data can be similarly adapted to sibship data denoted as $\zeta_{CC}(s, t; 1, 1)$. To estimate ζ_{CC} , we define the empirical processes $\mathbb{P}_n N_{11}^\#(s, t) = \frac{1}{n} \sum_{i=1}^n \sum_{1 \leq j' \leq m_i} I(Y_{ij} \leq s, \eta_{ij} = 1, Y_{ij'} \leq t, \eta_{ij'} = 1)$ and similarly $\mathbb{P}_n H^\#(s, t)$, $\mathbb{P}_n N_{10}^\#(s, t)$, and $\mathbb{P}_n N_{01}^\#(s, t)$ based on all within-cluster pairs. The asymptotic property of the plug-in estimator for ζ_{CC} can be established using similar arguments to those in the section “Asymptotic properties of $\hat{\zeta}_{CM}$ ” in Appendix except assuming $EM^4 < \infty$ as a sufficient condition, which is reasonable as the family size M is random but bounded.

2.3 Cause-specific cross hazard ratio for clustered data

For mother and all children data, under the exchangeability assumption for the children, the child–mother cause-specific density was defined as

$$g(s, t; k, l) = \lim_{(\Delta s, \Delta t) \downarrow 0} P(s \leq T_j \leq s + \Delta s, t \leq T_0 \leq t + \Delta t, \varepsilon_j = k, \varepsilon_0 = l) / (\Delta s \Delta t),$$

for $k, l = 1, 2$ and $j = 1, \dots, M$. Assuming g is absolutely continuous, a child–mother pair (T_j, T_0) has a joint overall survival function

$$G(s, t) = \int_s^\infty \int_t^\infty \sum_{k=1}^2 \sum_{l=1}^2 g(u, v; k, l) du dv.$$

Cheng et al. [10] defined a child–mother cause-specific association measure denoted by $\theta_{CM}(s, t; k, l)$ as follows:

$$\theta_{CM}(s, t; k, l) = \frac{G(s, t)g(s, t; k, l)}{\left\{ \int_s^\infty \sum_{m=1}^2 g(u, t; m, l) du \right\} \left\{ \int_t^\infty \sum_{h=1}^2 g(s, v; k, h) dv \right\}}, \quad (3)$$

which shares the same interpretation as the conditional cross hazard ratio proposed by Bandeen-Roche and Liang [4]. That is, the risk of a child experiencing a cause k event by time s would be accelerated for $\theta_{CM} > 1$ or decelerated for $\theta_{CM} < 1$ had his mother developed a cause l event by time t as compared to the case that his mother had no event by time t . This extended conditional cross hazard ratio was estimated in a piecewise manner by a U-statistic [10]. We now briefly describe the notation and methods given in Cheng et al. [10]. Similar notation will be used when we develop our modified U-statistic in Section 2.5 that specifically deals with tied events.

Let $Y_{i,a}, Y_{i,b}, 1 \leq a < b \leq m_i$ be the failure times of two distinct members of the i th cluster, and $Y_{j,c}, Y_{j,d}, 1 \leq c < d \leq m_j$ be the failure times from the j th cluster. $\theta_{CM}(s, t; k, l)$ was assumed to be a constant over each grid Ω_{qr} , $0 \leq q \leq n_1 - 1, 0 \leq r \leq n_2 - 1$. Cheng et al. [10] defined $Y_{(iajc)} = \min(Y_{ia}, Y_{jc})$, $\eta_{(iajc)} = \eta_{ia}I(Y_{ia} < Y_{jc}) + \eta_{jc}I(Y_{ia} > Y_{jc})$ and similarly $Y_{(ibjd)}$ and $\eta_{(ibjd)}$ and defined the concordant indicator

$$\phi_{ij,acbd}^{qr} = I\left\{ (Y_{ia} - Y_{jc})(Y_{ib} - Y_{jd}) > 0, (Y_{(iajc)}, Y_{(ibjd)}) \in \Omega_{qr}, (\eta_{(iajc)}, \eta_{(ibjd)}) = (k, l) \right\},$$

and the discordant indicator

$$\psi_{ij,acbd}^{qr} = I\left\{ (Y_{ia} - Y_{jc})(Y_{ib} - Y_{jd}) < 0, (Y_{(iajc)}, Y_{(ibjd)}) \in \Omega_{qr}, (\eta_{(iajc)}, \eta_{(ibjd)}) = (k, l) \right\}.$$

The extended child–mother association measure was estimated by

$$\hat{\theta}_{CM}^U(\Omega_{qr}; k, l) = \frac{\binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \sum_{1 \leq a \leq m_i} \sum_{1 \leq c \leq m_j} \phi_{ij,ac00}^{qr}}{\binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \sum_{1 \leq a \leq m_i} \sum_{1 \leq c \leq m_j} \psi_{ij,ac00}^{qr}} \quad (4)$$

which was the ratio of the number of concordant pairs to that of discordant pairs, among those pairs where concordance status was determinable, and the times and causes of failure matched those in θ_{CM} .

For any $1 \leq j < j' \leq M$ in the sibship data, Cheng et al. [10] defined the bivariate cause-specific densities $f_{jj'}(s, t; k, l) = \lim_{(\Delta s, \Delta t) \downarrow 0} P(s \leq T_j \leq s + \Delta s, \varepsilon_j = k, t \leq T_{j'} \leq t + \Delta t, \varepsilon_{j'} = l) / (\Delta s \Delta t)$, for $k, l = 1, 2$ and the overall survival function $S_{jj'}(s, t) = \int_s^\infty \int_t^\infty \sum_{k=1}^2 \sum_{l=1}^2 f_{jj'}(u, v; k, l) du dv$ assuming that $f_{jj'}$ is absolutely continuous. The subscripts j and j' were dropped under the exchangeability assumption among siblings, such that

$f(s, t; k, l) = f(t, s; l, k)$, for any pair of causes k, l at any time point (s, t) . The sibship cause-specific association measure assuming exchangeability was defined as

$$\theta_{CC}(s, t; k, l) = \frac{f(s, t; k, l)}{\int_s^\infty \sum_{m=1}^2 f(u, t; m, l) du} \frac{S(s, t)}{\int_t^\infty \sum_{h=1}^2 f(s, v; k, h) dv}. \quad (5)$$

The estimation of the sibship association θ_{CC} is more complicated than that of the child–mother association θ_{CM} , as under pairwise exchangeability it is meaningful to define concordance and discordance indicators based on two different pairings (Y_{ia}, Y_{jc}) with (Y_{ib}, Y_{jd}) as well as (Y_{ia}, Y_{jd}) with (Y_{ib}, Y_{jc}) [17]. Hence θ_{CC} is estimated by

$$\hat{\theta}_{CC}^U(\Omega_{qr}; k, l) = \frac{\binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \sum_{1 \leq a < b \leq m_i} \sum_{1 \leq c < d \leq m_j} (\phi_{ij,acbd}^{qr} + \phi_{ij,adbc}^{qr})}{\binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \sum_{1 \leq a < b \leq m_i} \sum_{1 \leq c < d \leq m_j} (\psi_{ij,acbd}^{qr} + \psi_{ij,adbc}^{qr})}. \quad (6)$$

The estimators (4) and (6) are U-statistics whose asymptotic properties can be obtained with U-statistic theory [10].

2.4 Maximum-pseudo (composite)-likelihood estimator

One can show that ζ_{CC} is equivalent to θ_{CC} defined in eq. (5) and that ζ_{CM} is equivalent to θ_{CM} defined in eq. (3), following the same argument as shown in Cheng and Fine [7]. Even though the data are multivariate, the association measures are still defined for paired event times, and the argument in Cheng and Fine [7] can be applied directly. Hence the two sets of estimators, the U-statistics and the plug-in estimators, are estimating the same quantities. We now discuss another estimating procedure for θ s adapted from the Clayton's [13] pseudo-likelihood function for bivariate survival data. The likelihood was also called a composite likelihood as discussed in Lindsay [18]. Ning and Bandeen-Roche [19] recently adapted the pseudo-likelihood to bivariate competing risks data under a regression setting. We now further extend the likelihood to multivariate competing risks data using a stratified method.

We first divide mother–multiple children data into several strata, each containing pairs of mothers and their d th children. Here d ranges from 1 to $m_{\max}$, where $m_{\max} = \max_{i=1}^n m_i$ is the maximum number of children among all families. Without loss of generality, we consider the first stratum of mothers and the corresponding eldest children. Define the cause-specific set $A(s, t; k, l)$ of size $a(s, t; k, l)$ pairs satisfying (1) both members are not censored by time (s, t) , and $a(s, t; k, l) \geq 1$; (2) there is at least one type l event that occurred at time t from one of the mothers and no event just prior to time s from her eldest child; and (3) there is at least one type k event that happened at time s from one of the eldest children and no event just prior to time t from the corresponding mother. Let $\Delta(s, t; k, l) = 1$ if the types k and l events were from the same pair and 0 otherwise. It can be shown that $P\{\Delta(s, t; k, l) = 1 | \mathcal{A}(s, t; k, l)\} = \frac{\theta_{CS}(s, t; k, l)}{a(s, t; k, l) - 1 + \theta_{CS}(s, t; k, l)}$, and $P\{\Delta(s, t; k, l) = 0 | \mathcal{A}(s, t; k, l)\} = \frac{a(s, t; k, l) - 1}{a(s, t; k, l) - 1 + \theta_{CS}(s, t; k, l)}$.

We now construct the pseudo-likelihood function based on the observed mother–multiple children data $\{(Y_{i0}, \dots, Y_{i, m_i}, \eta_{i0}, \dots, \eta_{i, m_i}, m_i), i = 1, \dots, n\}$ from the Cache County Study. To incorporate the multiple children structure, we adopt a stratified method. Let D_d denote the set containing all the families that have a d th child, where $d = 1, \dots, m_{\max}$. Then mother–child pairs belonging to D_d form a bivariate competing risks data set $\{(Y_{id}, \eta_{id}, Y_{i0}, \eta_{i0}), i \in D_d\}$ for which the pseudo-likelihood can be constructed. As discussed before, we assume a constant $\theta_{CM}(\Omega_{qr})$ over the region Ω_{qr} . To find all (s, t) pairs in Ω_{qr} such that $a(s, t; k, l) \geq 1$ and $\Delta(s, t; k, l) = 1$, we count each child–mother pair in D_d as a concordant pair, if two event times from a pair are contained in Ω_{qr} and their causes are (k, l) , denoted by the indicator function

$\phi_{i,d0}^{qr} = I\{(Y_{id}, Y_{i0}) \in \Omega_{qr}, (\eta_{id}, \eta_{i0}) = (k, l), \}$, $i \in D_d$. The contribution of such pair to the log-likelihood function is $\phi_{i,d0}^{qr} \log \left(\frac{\theta_{CM}}{a(Y_{id}, Y_{i0}; k, l) - 1 + \theta_{CM}} \right)$. Next, to find (s, t) such that $\Delta(s, t; k, l) = 0$, i.e. the events k and l are not from the same pair, we check any two different pairs $(Y_{id}, \eta_{id}, Y_{i0}, \eta_{i0})$ and $(Y_{jd}, \eta_{jd}, Y_{j0}, \eta_{j0})$ in D_d . If the times of the two pairs are discordant, s will be the minimum of children's failure times $Y_{(idj)}$ and t will be the minimum of mothers' failure times $Y_{(ioj)}$. The pair of the minimum failure times $(Y_{(idj)}, Y_{(ioj)})$ needs to be in Ω_{qr} and the cause indicators of the minimums $(\eta_{(idj)}, \eta_{(ioj)})$ need to match (k, l) . We call two such pairs "discordant pairs" and define the discordant indicator $\psi_{ij,dd00}^{qr} = I\{(Y_{id} - Y_{jd})(Y_{i0} - Y_{j0}) < 0, (Y_{(idj)}, Y_{(ioj)}) \in \Omega_{qr}, (\eta_{(idj)}, \eta_{(ioj)}) = (k, l)\}$, where $i < j$ and $i, j \in D_d$. Their contribution to the log-likelihood function is $\psi_{ij,dd00}^{qr} \log \left(\frac{a(Y_{(idj)}, Y_{(ioj)}; k, l) - 1}{a(Y_{(idj)}, Y_{(ioj)}; k, l) - 1 + \theta_{CM}} \right)$. Combining all the bivariate competing risks data from different strata D_d together, $d = 1, \dots, m_{\max}$, we have the following pseudo-log-likelihood function:

$$\begin{aligned} L_n(\theta_{CM}(\Omega_{qr})) &= \sum_{d=1}^{m_{\max}} \sum_{i \in D_d} \phi_{i,d0}^{qr} \log \left\{ \frac{\theta_{CM}}{a(Y_{id}, Y_{i0}; k, l) - 1 + \theta_{CM}} \right\} \\ &\quad + \sum_{d=1}^{m_{\max}} \sum_{i < j, i, j \in D_d} \psi_{ij,dd00}^{qr} \log \left\{ \frac{a(Y_{(idj)}, Y_{(ioj)}; k, l) - 1}{a(Y_{(idj)}, Y_{(ioj)}; k, l) - 1 + \theta_{CM}} \right\}. \end{aligned} \quad (7)$$

The maximum-pseudo-likelihood estimator $\tilde{\theta}_{CM}^L$ can be obtained by using some standard minimization procedure, e.g. function "nlminb" in R, for the negative log likelihood. As pointed out by Oakes [14], the variance of $\tilde{\theta}_{CM}^L$ cannot be obtained from the second derivative of the pseudo-log-likelihood function as $\psi_{ij,dd00}^{qr}$ may be dependent when $\theta_{CM}(\Omega_{qr}) \neq 1$. Instead, we use some results for Z-estimators in empirical processes [15] to derive the asymptotic property of $\tilde{\theta}_{CM}^L$ based on the estimation equation $\Psi_n(\theta_{CM}) = \frac{\partial \theta_{CM}}{\partial n} \dot{L}_n(\theta_{CM}) = 0$, where the superscript dot denotes the derivative. The inference will be based on the nonparametric bootstrap method.

The above framework can be further extended to clustered sibship data for which exchangeability can be assumed among all siblings. The adaptation of the maximum-pseudo-likelihood estimation to the sibship data is more complicated as there are two meaningful ways to form a pair, (Y_{ia}, Y_{ib}) and (Y_{ib}, Y_{ia}) , for any two children a, b from the i th family, under the exchangeability assumption. In addition, there may be too many different combinations of a and b from multiple children families. To simplify the problem, we continue to use our stratified method and denote the set D_{ab} , $1 \leq a < b \leq m_{\max}$, which consists of all families having a th and b th children. To adapt the exchangeability assumption, we consider the bivariate competing risks set consisting of all child-child pairs $\{(Y_{ia}, \eta_{ia}, Y_{ib}, \eta_{ib}), i \in D_{ab}\}$ together with its switched set containing all pairs $\{(Y_{ib}, \eta_{ib}, Y_{ia}, \eta_{ia}), i \in D_{ab}\}$. The cause-specific set $\mathcal{A}(s, t, k, l)$ is naturally extended to this bivariate set and its switched set, with the size of $a(s, t, k, l)$. As before let $\Delta(s, t; k, l) = 1$ if cause k event at time s and cause l event at time t were from the same pair and $\Delta(s, t; k, l) = 0$ otherwise. We can still show that $P\{\Delta(s, t; k, l) = 1 | \mathcal{A}(s, t; k, l)\} = 1 - P\{\Delta(s, t; k, l) = 0 | \mathcal{A}\} = \frac{\theta_{CC}(s, t; k, l)}{a(s, t; k, l) - 1 + \theta_{CC}(s, t; k, l)}$.

We now construct the pseudo-log-likelihood function based on the stratum D_{ab} . As before we assume constant sibship association θ_{CC} within time region Ω_{qr} . To find all (s, t) pairs in $\mathcal{A}(s, t; k, l) \cap \Omega_{qr}$ such that $\Delta(s, t; k, l) = 1$, we count each concordant sibship pair in the bivariate set and its reversed set and define concordant indicators $\phi_{i,ab}^{qr} = I\{(Y_{ia}, Y_{ib}) \in \Omega_{qr}, (\eta_{ia}, \eta_{ib}) = (k, l)\}$ and $\phi_{i,ba}^{qr} = I\{(Y_{ib}, Y_{ia}) \in \Omega_{qr}, (\eta_{ib}, \eta_{ia}) = (k, l)\}$, $i \in D_{ab}$. It is more complicated to find all time points $(s, t) \in \Omega_{qr}$ such that $\Delta(s, t; k, l) = 0$, because for any two pairs, $i < j$, from D_{ab} , we may have four different pairings: (Y_{ia}, Y_{ib}) with (Y_{ja}, Y_{jb}) , (Y_{ia}, Y_{ib}) with (Y_{jb}, Y_{ja}) , (Y_{ib}, Y_{ia}) with (Y_{ja}, Y_{jb}) , and (Y_{ib}, Y_{ia}) with (Y_{jb}, Y_{ja}) . For the first pairing we define the discordant indicator $\psi_{ij,aabb}^{qr} = I\{(Y_{ia} - Y_{ja})(Y_{ib} - Y_{jb}) < 0, (Y_{(iaja)}, Y_{(ibjb)}) \in \Omega_{qr}, (\eta_{(iaja)}, \eta_{(ibjb)}) = (k, l)\}$. Similarly, we

define discordant indicators $\psi_{ij,abba}^{qr}$, $\psi_{ij,baab}^{qr}$, and $\psi_{ij,bbaa}^{qr}$. Combining all the strata D_{ab} together we have the following pseudo-log-likelihood function:

$$\begin{aligned}
 L_n(\theta_{CC}) = & \sum_{1 \leq a < b \leq m_{\max}} \sum_{i \in D_{ab}} \left[\phi_{i,ab}^{qr} \log \left\{ \frac{\theta_{CC}}{a(Y_{ia}, Y_{ib}; k, l) - 1 + \theta_{CC}} \right\} \right. \\
 & \left. + \phi_{i,ba}^{qr} \log \left\{ \frac{\theta_{CC}}{a(Y_{ib}, Y_{ia}; k, l) - 1 + \theta_{CC}} \right\} \right] \\
 & + \sum_{1 \leq a < b \leq m_{\max}} \sum_{i < j \text{ and } i, j \in D_{ab}} \left[\psi_{ij,aabb}^{qr} \log \left\{ \frac{a(Y_{iaja}, Y_{ibjb}); k, l) - 1}{a(Y_{iaja}, Y_{ibjb}); k, l) - 1 + \theta_{CC}} \right\} \right. \\
 & + \psi_{ij,abba}^{qr} \log \left\{ \frac{a(Y_{iajb}, Y_{ibja}); k, l) - 1}{a(Y_{iajb}, Y_{ibja}); k, l) - 1 + \theta_{CC}} \right\} \\
 & + \psi_{ij,baab}^{qr} \log \left\{ \frac{a(Y_{ibja}, Y_{iajb}); k, l) - 1}{a(Y_{ibja}, Y_{iajb}); k, l) - 1 + \theta_{CC}} \right\} \\
 & \left. + \psi_{ij,bbaa}^{qr} \log \left\{ \frac{a(Y_{ibjb}, Y_{iaja}); k, l) - 1}{a(Y_{ibjb}, Y_{iaja}); k, l) - 1 + \theta_{CC}} \right\} \right]. \quad (8)
 \end{aligned}$$

The maximum-likelihood estimator $\tilde{\theta}_{CC}^L$ for the constant association measure θ_{CC} can be similarly obtained by using the function “nlminb” in R. The consistency and asymptotic properties of $\tilde{\theta}_{CC}^L$ can be proved analogously as for $\tilde{\theta}_{CM}^L$ in the section “Asymptotic properties of $\tilde{\theta}_{CM}^L$ ” in Appendix.

For each extended association measures $\theta_{CC}(\zeta_{CC})$ and $\theta_{CM}(\zeta_{CM})$, we have discussed three sets of estimators. The plug-in estimators $\hat{\zeta}_{CC}$ and $\hat{\zeta}_{CM}$ and the U-type estimators discussed in Cheng et al. [10] ($\hat{\theta}_{CM}^U, \hat{\theta}_{CC}^U$) are computationally straightforward to obtain. The extended maximum-pseudo-likelihood estimators proposed in this paper ($\tilde{\theta}_{CM}^L, \tilde{\theta}_{CC}^L$) do not have explicit variance estimators and hence are more computationally intensive. Asymptotic normality can be established for all the three estimators. Their relative efficiencies are examined by numerical studies in Section 3.

2.5 Modified U-statistic for tied data

In practice, the data are often rounded to the nearest integers, resulting in tied events. The adaptation of the plug-in estimator to the tied data is straightforward. For the U-statistic, we followed Cheng et al. [10] and excluded those pairs where one or both event times were tied. For the pseudo-likelihood function (7), the tied observations do not affect concordant pairs much, since a mother–child pair with double events is still counted as a concordant pair after rounding. However, rounding tends to underestimate the number of discordant pairs. Suppose we have two discordant pairs $(Y_{i0}, \eta_{i0}, Y_{i1}, \eta_{i1})$ and $(Y_{j0}, \eta_{j0}, Y_{j1}, \eta_{j1})$ from two distinct families i and j , where $Y_{i0} \leq Y_{j0}, \eta_{i0} = k$ and $Y_{i1} \geq Y_{j1}, \eta_{j1} = l$. When ties exist, e.g. $Y_{i1} = Y_{j1}$, and if $\eta_{i1} = l$, then the pairing of these two pairs would not be considered as discordant any longer.

We will show through simulation studies in Section 3.4 that the plug-in estimator tends to underestimate while the other two tend to overestimate the true positive association with the presence of tied data. For the pseudo-likelihood estimator, the ties decrease the number of discordant pairs, and hence increase the estimated strength of association. For the U-statistic, rounding causes the removal of roughly similar amounts of concordant and discordant pairs. When the association is positive, i.e. the numerator is larger than the denominator, subtracting the same positive value from both the numerator and the denominator makes the ratio larger. Therefore, the U-statistic overestimates the positive association when ties are present. The similar arguments can be used to show that the U-statistic results in an estimate that is

smaller than the true value when the association is negative ($0 < \theta(\zeta) < 1$) and ties are present. On the other hand, the pseudo-likelihood estimator still overestimates θ , since the number of discordant pairs is getting smaller. However, it is not clear theoretically how rounding errors would affect the plug-in estimator. Our simulation study suggests that the plug-in estimator tends to bias toward independence for both positive and negative associations.

To address potential loss of accuracy in the three estimators for tied data, we propose a modified U-statistic which essentially adds back those removed concordant and discordant pairs due to tied events. When two mother–child pairs have tied observations – for example, when both mothers had the same event time – we consider their corresponding cause indicators. If mother 1 died and mother 2 had dementia at the same age, we assume that mother 1 would have dementia later than mother 2, had mother 1 not died. If child event times were not tied, we could determine concordant status. For those tied pairs that can be equally concordant or discordant, we add half a pair to both concordant and discordant pairs. It becomes more complicated for sibship data under the exchangeability assumption. The section “Formula for corrected U” in Appendix provides details on how to add back all missing concordant and discordant pairs for sibship data.

The idea may be used analogously for the pseudo-likelihood estimator. In the section “Asymptotic properties of $\hat{\theta}_{CM}^L$ ” in Appendix, we show that the pseudo-likelihood estimator is a Z-estimator based on the estimating equation which is a U-statistic of order 2 with the kernel function $h(X_i, X_j; \theta)$ given in eq. (11). One may let both $\phi_{ij,dd00}^{qr}$ and $\psi_{ij,dd00}^{qr}$ in the numerator of eq. (11) be 0.5 for tied events following the same scheme as in the section “Formula for corrected U” in Appendix. However, it is not clear how to fix the plug-in estimator directly when there are rounding errors. Some random perturbation method may be used by adding simulated small error terms to tied observations. However, multiple imputation is needed which is beyond the scope of this paper.

2.6 Grid selection

The selection of grids plays an important role in capturing the local association structure. If we use too few grid points, the resulting estimates may be too smooth, missing crucial local associations. On the other hand, if there are too many grid points, the number of events would be too few and there would be too much variability in the estimators. In practice, if there is no particular scientific definition of time regions in place, we recommend that we start with some crude divisions and then further refine the grid points. We will carry out formal hypothesis tests on two neighboring associations being the same, based on which we decide whether we should keep the finer grid points. That is, we would like to test the following hypotheses:

$$H_0 : \zeta_{CM}(\Omega_{qr}) = \zeta_{CM}(\Omega_{q+1,r}), q = 0, 1, \dots, n_1 - 2, \text{ and}$$

$$H_0 : \zeta_{CM}(\Omega_{qr}) = \zeta_{CM}(\Omega_{q,r+1}), r = 0, 1, \dots, n_2 - 2,$$

where $\Omega_{qr} \equiv (\tau_{1,q}, \tau_{1,q+1}) \times (\tau_{2,r}, \tau_{2,r+1})$ with $0 \leq q \leq n_1 - 1, 0 \leq r \leq n_2 - 1$. For any pair of two adjacent regions, e.g. Ω_{qr} and $\Omega_{q+1,r}$, we construct a Wald type of statistic based on the difference in the two local associations. That is,

$$\chi_W^2 = \frac{\left\{ \hat{\zeta}_{CM}(\Omega_{qr}) - \hat{\zeta}_{CM}(\Omega_{q+1,r}) \right\}^2}{\text{var}\left\{ \hat{\zeta}_{CM}(\Omega_{qr}) \right\} + \text{var}\left\{ \hat{\zeta}_{CM}(\Omega_{q+1,r}) \right\}},$$

where $\text{var}\{\hat{\zeta}_{CM}(\Omega_{qr})\}$ and $\text{var}\{\hat{\zeta}_{CM}(\Omega_{q+1,r})\}$ are the variance estimators of $\hat{\zeta}_{CM}(\Omega_{qr})$ and $\hat{\zeta}_{CM}(\Omega_{q+1,r})$ computed based on the influence functions given in eq. (10). When the number of events in each grid is not too few, we can compare this test statistic with its asymptotic chi-squared distribution with one degree of

freedom. For the two neighboring regions Ω_{qr} and $\Omega_{q,r+1}$, similar tests can be constructed. Due to the equivalence of ζ and θ , the above hypotheses can be equally expressed in terms of θ and be tested based on the U-estimator or the pseudo-likelihood estimator, though for the latter estimator the bootstrap method is used to obtain the variance. The same idea can be applied analogously to the sibship association analysis.

3 Simulation studies

3.1 Constant association

We conducted simulation studies to evaluate small-sample performance of the plug-in estimator $\hat{\zeta}$, the maximum-pseudo-likelihood estimator $\hat{\theta}^L$, and the U-statistic $\hat{\theta}^U$. We first simulated mother–multiple children data by extending Bandeen-Roche and Liang’s [4] frailty model. They modeled cause-specific between-subject dependency by an overall frailty A and proportions of cause allocations B which was between 0 and 1 and independent of A . As the Archimedean copula is associative, we naturally extended their frailty model for the bivariate competing risks to multivariate settings in our simulations. Three different strengths of association were considered. For each scenario, we generated 500 replicates of 500 families with varying family sizes. More details of our simulation are given below.

For each family, we first drew the familywise propensity A from a Gamma (1, 1). The number of children for a family M was determined from a multinomial distribution ranging from 1 to 5 with probabilities 0.09, 0.18, 0.24, 0.24, and 0.25 which were estimated from the Cache County Data. Next we generated $M + 1$ failure times from $\text{Exp}(A)$ for the mother and her M children. Since the Cache County Study data were collected retrospectively, some previous investigation [20] showed that the marginal cumulative incidence function (CIF) for a mother is similar to that for her eldest child. For the sake of simplicity, we used the same distribution for mother and children in this simulation. Cause indicators were generated from a Bernoulli distribution with probability B for each subject, where B was drawn from a Beta($R_1, 1 - R_1$). Three values of R_1 , 0.2, 0.5, and 0.8, were considered with corresponding theoretical values of $\theta_{CM}(s, t; 1, 1)$ being 6.0, 3.0, and 2.25 for all s and t . Finally, independent censoring times were generated from Uniform(5, 10) for each subject in this family, imposing about 12% censoring.

We first adopted the same frailty model as those used in Bandeen-Roche and Liang [4] and Cheng and Fine [7]. The model assumes the same strength of association between causes and results in constant association over time. Hence the estimation is carried out over the entire observation region with uniform weight. We applied the three methods to each of the 500 simulated data sets. The mean of estimates, mean of standard error (for the plug-in and U-estimators), or bootstrap standard error (for the pseudo-likelihood estimator) using family-based resampling with replacement with 500 replications, empirical standard error, and coverage rate of 95% confidence intervals are summarized in the upper left panel of Table 1. With moderate sample size, light censoring, and 50–80% of cause 1 events, all the three estimators perform well. The estimates are very close to the true values. The model-based or bootstrap standard errors agree well with the empirical standard errors. The coverage rates are close to the nominal level 0.95. There is some bias in the pseudo-likelihood estimate and U-statistic when $R_1 = 0.2$, since there were fewer cause 1 events. Similarly, we simulated 500 sibship data sets using the same simulation strategy. Hence the child–child association θ_{CC} takes on the same values as the mother–child association θ_{CM} . Each data set contained 250 families. The results are summarized in upper right panel of Table 1. Since there are 250 families instead of 500 in each sibship data, there is some discrepancy between the estimates and the true value of association when the number of cause 1 events is small ($R_1 = 0.2$).

Next, we evaluated the performances of the three estimators when the number of clusters was smaller and censoring was heavier. We reduced the cluster sizes for mother–children data and sibship data to 200 or 100 families and generated censoring times from an exponential function with mean 4 yielding ~35% censoring. The results are summarized in the middle and bottom panels of Table 1. The bias is more

Table 1 Simulation results for mother–multiple children data (CM) and sibship data (CC)

R_1	$\theta(\zeta)_{CM}$ or $\theta(\zeta)_{CC}$	STAT	$\hat{\zeta}_{CM}$	$\hat{\theta}_{CM}^L$	$\hat{\theta}_{CM}^U$	$\hat{\zeta}_{CC}$	$\hat{\theta}_{CC}^L$	$\hat{\theta}_{CM}^U$
			Mother–Children			Sibship		
			$n = 500, CP = 12\%$			$n = 250, CP = 12\%$		
0.2	6	EST	6.03	6.16	6.15	5.98	6.23	6.22
		MSE	0.82	0.84	0.84	1.11	1.15	1.18
		ESE	0.83	0.79	0.86	1.21	1.23	1.31
		COV	0.93	0.96	0.95	0.90	0.93	0.92
0.5	3	EST	2.99	3.01	3.01	3.00	3.04	3.05
		MSE	0.24	0.23	0.24	0.32	0.33	0.33
		ESE	0.23	0.23	0.23	0.32	0.33	0.33
		COV	0.95	0.95	0.96	0.95	0.94	0.96
0.8	2.25	EST	2.24	2.26	2.25	2.24	2.26	2.26
		MSE	0.13	0.13	0.13	0.18	0.18	0.18
		ESE	0.13	0.13	0.13	0.16	0.16	0.17
		COV	0.94	0.96	0.94	0.97	0.96	0.97
			$n = 200, CP = 35\%$			$n = 200, CP = 35\%$		
0.2	6	EST	6.11	6.49	6.45	6.08	6.41	6.42
		MSE	1.60	2.03	1.77	1.51	1.59	1.64
		ESE	1.73	1.80	1.94	1.64	1.66	1.81
		COV	0.94	0.98	0.95	0.91	0.95	0.94
0.5	3	EST	3.02	3.08	3.08	3.00	3.07	3.06
		MSE	0.46	0.47	0.48	0.43	0.41	0.45
		ESE	0.47	0.44	0.49	0.44	0.41	0.45
		COV	0.95	0.97	0.95	0.94	0.96	0.95
0.8	2.25	EST	2.22	2.26	2.25	2.23	2.25	2.26
		MSE	0.25	0.24	0.26	0.23	0.22	0.24
		ESE	0.24	0.23	0.25	0.23	0.22	0.24
		COV	0.94	0.95	0.94	0.95	0.94	0.95
			$n = 100, CP = 35\%$			$n = 100, CP = 35\%$		
0.2	6	EST	6.30	7.15	7.16	6.05	6.79	6.81
		MSE	2.25	4.44	2.83	2.03	2.66	2.40
		ESE	2.58	3.05	3.39	2.38	2.73	2.93
		COV	0.89	0.98	0.92	0.87	0.93	0.90
0.5	3	EST	3.02	3.17	3.15	2.97	3.10	3.10
		MSE	0.64	0.75	0.70	0.60	0.60	0.64
		ESE	0.63	0.66	0.70	0.64	0.64	0.69
		COV	0.93	0.97	0.94	0.90	0.93	0.92
0.8	2.25	EST	2.25	2.32	2.30	2.23	2.28	2.28
		MSE	0.35	0.37	0.37	0.33	0.32	0.34
		ESE	0.35	0.34	0.37	0.33	0.33	0.35
		COV	0.94	0.96	0.95	0.93	0.92	0.95

Note: Mean of estimates (EST), mean of standard error/bootstrap standard error (MSE), empirical standard error (ESE), coverage rate of 95% confidence interval (COV), number of clusters (n), and censoring proportion (CP).

noticeable in the estimates as compared to the true values. Both the empirical standard error and the model-based or bootstrap standard error increase, though they in general agree with each other. The coverage rates are still pretty close to the nominal level 0.95, except when the number of cause 1 event is very small ($n = 100$ and $R_1 = 0.2$).

3.2 Time-varying association

In practice, familial association may not be constant over each time grid. In this case, θ or ζ can be viewed as a weighted average of time-varying association across all time points within a specific region. It would be interesting to examine how the three estimators perform when the underlying association is actually time-varying and how the selection of grid may affect the results. Hence we simulated mother–multiple children data with time-varying association between the cause 1 events by adopting a gamma frailty model for the bivariate CIF as discussed in Cheng and Fine [21]. More specifically, the bivariate CIF of a child–mother pair $F_{11}^{CM}(s, t) = P(T_j \leq s, \varepsilon_j = 1, T_0 \leq t, \varepsilon_0 = 1), j = 1, \dots, M$ was connected with the marginal CIF for the child $F_1^C(s) = P(T_j \leq s, \varepsilon_j = 1)$ and that for the mother $F_1^M(t) = P(T_0 \leq t, \varepsilon_0 = 1)$ through a gamma frailty model. That is,

$$F_{11}^{CM}(s, t) = \left\{ F_1^C(s)^{1-\alpha} + F_1^M(t)^{1-\alpha} - 1 \right\}^{\frac{1}{1-\alpha}},$$

where we chose $\alpha = 1.14$, the same value as estimated from the Cache County data [21]. For the marginals, we assumed equal distribution and adopted a modified three-parameter logistic model [22] as suggested by the Cache County data which were collected retrospectively. That is,

$$F_1^C(s) = F_1^M(s) = \frac{p \exp\{b(s - c)\} - p \exp(-bc)}{1 + \exp\{b(s - c)\}},$$

with $p = 0.5$, $b = 1$, and $c = 2$. The competing events were simulated from a conditional three-parametric logistic model, where $p = 1$, $b = 2$, and $c = 4$; see Cheng and Fine [21, Section 4.1] for details. The independent censoring time was generated from Uniform (4,5) resulting in about 20% censored observations.

We generated 500 clusters of mother and children data, each containing 500 or 100 families, and obtained the mean estimates, the model-based or bootstrap standard errors, the empirical standard errors, and the coverage rates where the true values were computed by averaging the estimates from 200 replications of 2,000 mother–child pairs. We examined the performance of the three estimators over a four-region grid, as well as over the entire observed time region, with uniform weight. The results are summarized in Table 2. The true values for the plug-in estimator and the U-estimator are identical and slightly different from those for the pseudo-likelihood estimator, since the latter weights the time-varying associations at various time points differently. The true values change across time regions. When we employ the entire region, the true values fall somewhere between the smallest and largest associations among all subregions. In general, all the three estimators still have good practical performance, except for the pseudo-likelihood estimator when the cluster size is $n = 100$. The coverage rates for the average association across the entire region are slightly lower than those over smaller regions. It may be explained that the estimators over the entire region have smaller variability, and some slight departure from the true value would lead to lower coverage. When the association is time-varying, the selection of grids plays a more important role in summarizing the strength of association than in the situation of constant association. As suggested by the results from $n = 100$, both the bias and the standard errors of the estimators increase when the sample size is much smaller than 500. To reduce the variability, we may reduce the number of grids, which on the other hand will restrict our ability of examining time-varying association.

Table 2 Simulation results for mother–multiple children data with time-varying association

$(s, t) \in \Omega_{gr}$		$s \in [0, 2.5]$	$s \in [0, 2.5]$	$s \in [2.5, 5]$	$s \in [2.5, 5]$	$s \in [0, 5]$
		$t \in [0, 2.5]$	$t \in [2.5, 5]$	$t \in [0, 2.5]$	$t \in [2.5, 5]$	$t \in [0, 5]$
$n = 500$						
$\hat{\zeta}_{CM}$	TRUE	1.40	1.13	1.13	1.06	1.26
	EST	1.40	1.13	1.13	1.07	1.27
	MSE	0.13	0.14	0.15	0.17	0.08
	ESE	0.13	0.14	0.16	0.18	0.08
	COV	0.93	0.93	0.94	0.95	0.90
$n = 100$						
$\hat{\zeta}_{CM}$	TRUE	1.40	1.13	1.13	1.06	1.26
	EST	1.41	1.17	1.16	1.10	1.27
	MSE	0.28	0.30	0.32	0.37	0.18
	ESE	0.30	0.34	0.36	0.42	0.18
	COV	0.94	0.92	0.89	0.90	0.94
$n = 500$						
$\hat{\theta}_{CM}^U$	TRUE	1.40	1.13	1.13	1.06	1.26
	EST	1.40	1.13	1.13	1.07	1.27
	MSE	0.13	0.13	0.15	0.17	0.08
	ESE	0.13	0.14	0.15	0.18	0.08
	COV	0.93	0.93	0.94	0.94	0.91
$n = 100$						
$\hat{\theta}_{CM}^U$	TRUE	1.40	1.13	1.13	1.06	1.26
	EST	1.41	1.17	1.16	1.11	1.28
	MSE	0.29	0.31	0.34	0.40	0.19
	ESE	0.31	0.35	0.38	0.45	0.16
	COV	0.94	0.92	0.89	0.89	0.95
$n = 500$						
$\hat{\theta}_{CM}^L$	TRUE	1.36	1.12	1.12	1.05	1.19
	EST	1.36	1.12	1.12	1.06	1.21
	BSE	0.12	0.13	0.14	0.16	0.07
	ESE	0.12	0.13	0.15	0.16	0.07
	COV	0.92	0.93	0.94	0.94	0.88
$n = 100$						
$\hat{\theta}_{CM}^L$	TRUE	1.36	1.12	1.12	1.05	1.19
	EST	1.38	1.17	1.15	1.12	1.21
	BSE	0.13	0.13	0.15	0.20	0.07
	ESE	0.30	0.33	0.35	0.45	0.16
	COV	0.62	0.59	0.61	0.61	0.60

Note: Mean of estimates (EST), mean of standard error/bootstrap standard error (MSE/BSE), empirical standard error (ESE), and coverage rate of 95% confidence interval (COV).

3.3 Tied observations

Real data are often rounded as in the Cache County Study, resulting in tied observations. To evaluate how the three estimators perform when there are tied events, we rounded simulated mother–children data and sibship data in Section 3.1 to one decimal place containing 100 time points from 0 to 9.9. This is roughly comparable to 53 age points from 55 to 107 in the Cache County Study. The results are summarized in Table 3. The plug-in estimator $\hat{\zeta}$ tends to underestimate θ_{CM} while the maximum-pseudo-likelihood

Table 3 Simulation results for mother–multiple children data (CM) and sibship data (CC) with tied events

R_1	$\theta(\zeta)_{CM}$ or $\theta(\zeta)_{CC}$	STAT	$\hat{\zeta}_{CM}$	$\hat{\theta}_{CM}^L$	$\hat{\theta}_{CM}^U$	$\hat{\theta}_{CM}^{MU}$	$\hat{\zeta}_{CC}$	$\hat{\theta}_{CC}^L$	$\hat{\theta}_{CC}^U$	$\hat{\theta}_{CC}^{MU}$
			Mother–Children				Sibship			
			$n = 500, CP = 12\%$				$n = 250, CP = 12\%$			
0.2	6	EST	5.74	6.47	6.45	5.98	5.66	6.54	6.50	6.12
		MSE	0.75	0.91	0.93	0.81	1.01	1.23	1.28	1.14
		ESE	0.76	0.89	0.94	0.81	0.99	1.23	1.29	1.14
		COV	0.90	0.95	0.96	0.96	0.88	0.97	0.96	0.95
0.5	3	EST	2.84	3.24	3.15	2.98	2.89	3.32	3.24	2.99
		MSE	0.21	0.26	0.26	0.23	0.30	0.36	0.38	0.32
		ESE	0.21	0.26	0.27	0.22	0.33	0.40	0.41	0.31
		COV	0.87	0.88	0.93	0.94	0.89	0.89	0.92	0.94
0.8	2.25	EST	2.14	2.49	2.37	2.23	2.13	2.49	2.36	2.23
		MSE	0.12	0.15	0.15	0.13	0.16	0.20	0.20	0.17
		ESE	0.12	0.15	0.15	0.13	0.17	0.21	0.21	0.18
		COV	0.82	0.67	0.91	0.93	0.84	0.81	0.92	0.94
			$n = 200, CP = 35\%$				$n = 200, CP = 35\%$			
0.2	6	EST	5.63	6.72	6.66	6.21	5.71	6.73	6.77	6.20
		MSE	1.40	2.14	1.91	1.70	1.34	1.69	1.80	1.56
		ESE	1.45	1.79	0.89	1.80	1.45	1.82	2.04	1.64
		COV	0.87	0.98	0.94	0.92	0.87	0.96	0.95	0.91
0.5	3	EST	2.83	3.33	3.23	3.01	2.85	3.36	3.26	3.05
		MSE	0.40	0.52	0.54	0.47	0.39	0.47	0.51	0.45
		ESE	0.41	0.50	0.55	0.48	0.33	0.49	0.55	0.45
		COV	0.90	0.96	0.96	0.94	0.89	0.94	0.95	0.93
0.8	2.25	EST	2.12	2.54	2.40	2.27	2.13	2.55	2.41	2.27
		MSE	0.22	0.29	0.30	0.26	0.21	0.26	0.28	0.24
		ESE	0.22	0.26	0.29	0.26	0.23	0.29	0.31	0.24
		COV	0.87	0.90	0.95	0.95	0.86	0.85	0.92	0.95
			$n = 100, CP = 35\%$				$n = 100, CP = 35\%$			
0.2	6	EST	5.74	7.45	7.43	6.69	5.65	7.21	7.15	6.42
		MSE	1.96	4.24	3.07	2.57	1.83	2.89	2.70	2.26
		ESE	2.51	4.01	4.26	3.09	2.02	2.87	3.09	2.75
		COV	0.84	0.98	0.93	0.89	0.84	0.97	0.94	0.88
0.5	3	EST	2.82	3.40	3.30	3.15	2.85	3.43	3.34	3.11
		MSE	0.57	0.84	0.79	0.70	0.54	0.70	0.75	0.65
		ESE	0.57	0.76	0.80	0.75	0.57	0.73	0.81	0.69
		COV	0.89	0.97	0.96	0.95	0.89	0.98	0.97	0.93
0.8	2.25	EST	2.16	2.66	2.49	2.29	2.14	2.61	2.45	2.31
		MSE	0.32	0.45	0.44	0.37	0.30	0.38	0.41	0.35
		ESE	0.31	0.43	0.43	0.38	0.31	0.41	0.43	0.37
		COV	0.90	0.94	0.97	0.94	0.88	0.91	0.96	0.93

Note: Mean of estimates (EST), mean of standard error/bootstrap standard error (MSE), empirical standard error (ESE), coverage rate of 95% confidence interval (COV), cluster size (n), and censoring proportion (CP).

estimator $\hat{\theta}^L$ and the U-statistic $\hat{\theta}^U$ tend to overestimate it. We also evaluated the performance of the three estimators for sibship data with tied observations. The results are given in the upper right panels of Table 3. We observed similar trends as in the mother–multiple children data. The plug-in estimators are likely to underestimate the association while the other two are prone to overestimate the association. As a comparison, we also evaluated the performance of the modified U-statistics, denoted by $\hat{\theta}_{CM}^{MU}$ and $\hat{\theta}_{CC}^{MU}$, for the rounded mother–children and sibship data. With the presence of rounding errors, the modified U-estimates

clearly outperform the other three, as they are closest to the true values. There is also some efficiency gain over the original U-statistic because the latter loses information by discarding all tied observations. When there are fewer clusters ($n = 200$ or $n = 100$) and higher censoring rate ($CP = 35\%$), we observed more bias in all the four estimators (bottom panel of Table 3). However, the modified U-statistic still clearly outperforms the other three.

3.4 Negative association

The proposed plug-in and pseudo-likelihood estimators, as well as the original and modified U-estimators, can be naturally applied to the clustered data where the paired events have a negative association. To evaluate their practical performance under negative association, we simulated 500 or 100 clusters of mother–multiple children data where the mother–child paired event times satisfy a Clayton model with copula parameter $\alpha = 0.5$ and have the same marginal exponential distributions with mean 1. For each subject, the cause indicator was generated from a $\text{Beta}(0.5, 0.5)$. The other settings are the same as in Section 3.1 except that there is no independent censoring. The corresponding mother–child association is $\theta_{CM} = 0.75$, estimated based on 100 replications of 1,000 mother–child pairs. The results are summarized in Table 4. Under this simulation setting, all the three estimators have satisfactory performance. We were also interested in examining the effect of rounding errors on the three types of estimators when there is a negative association. Hence, we rounded the simulated data to the first decimal place and computed the three estimators and the modified U-statistic as listed in the bottom panel of Table 4. Interestingly, when $\theta_{CM} = 0.75$ and there are rounding errors, the plug-in estimator and pseudo-likelihood estimator overestimated the true value with bias toward independence, while the original U-estimator underestimated the true value over-evaluating the negative association. We also observe some improvements in the coverage rates when the cluster size decreases from 500 to 100, because the estimators are roughly as biased as before but the standard errors increase in the smaller sample. In contrast to the performance of the three

Table 4 Simulation results for mother–multiple children data (CM) for negative association

$\theta(\zeta)_{CM}$	Stat	$\hat{\zeta}_{CM}$	$\hat{\theta}_{CM}^L$	$\hat{\theta}_{CM}^U$	$\hat{\theta}_{CM}^{MU}$	$\theta(\zeta)_{CM}$	Stat	$\hat{\zeta}_{CM}$	$\hat{\theta}_{CM}^L$	$\hat{\theta}_{CM}^U$	$\hat{\theta}_{CM}^{MU}$
Untied data $n = 500$											
0.75	EST	0.75	0.75	0.75	–	0.45	EST	0.45	0.45	0.45	–
	MSE	0.06	0.06	0.04	–		MSE	0.04	0.03	0.04	–
	ESE	0.06	0.06	0.04	–		ESE	0.04	0.03	0.04	–
	COV	0.95	0.95	0.95	–		COV	0.94	0.94	0.94	–
Untied data $n = 100$											
	EST	0.78	0.77	0.77	–		EST	0.48	0.46	0.46	–
	MSE	0.13	0.12	0.13	–		MSE	0.09	0.08	0.09	–
	ESE	0.14	0.12	0.15	–		ESE	0.09	0.07	0.10	–
	COV	0.93	0.95	0.92	–		COV	0.94	0.97	0.92	–
Tied data $n = 500$											
	EST	0.81	1.02	0.70	0.79		EST	0.53	0.87	0.39	0.50
	MSE	0.06	0.05	0.06	0.06		MSE	0.04	0.04	0.04	0.04
	ESE	0.06	0.06	0.05	0.06		ESE	0.05	0.04	0.04	0.04
	COV	0.86	0	0.82	0.92		COV	0.56	0	0.6	0.80
Tied data $n = 100$											
	EST	0.82	0.99	0.70	0.79		EST	0.54	0.77	0.40	0.50
	MSE	0.13	0.14	0.13	0.14		MSE	0.09	0.10	0.08	0.08
	ESE	0.13	0.14	0.13	0.14		ESE	0.09	0.09	0.09	0.08
	COV	0.95	0.66	0.88	0.95		COV	0.90	0.04	0.83	0.91

Note: Refer to Table 1 for the definitions of EST, MSE, ESE, and COV.

estimators, the modified U-estimator produced the most reliable result with the presence of negative association for both cluster sizes.

4 Cache County Study

The data from the Cache County Study on Memory in Aging have been analyzed in previous works. Bandeen-Roche and Liang [4] and Bandeen-Roche and Ning [6] applied their U-type estimator, and Cheng and Fine [7] used a plug-in estimator to investigate the child–mother association in dementia based on 3,635 pairs of mothers and their eldest children. In order to utilize the information on other siblings, Cheng et al. [10] adapted the U-estimator for bivariate competing risks data to multivariate cases, assuming exchangeability among all siblings. In this paper, we were interested in both the child–mother and the sibship associations. Hence we extended the plug-in estimator from bivariate competing risks data and the pseudo-likelihood estimator from a typical bivariate survival setting to more complicated data structures such as mother–multiple children data and sibship data. The two extended estimators were applied to 3,635 families containing mothers and all her children for the mother–child association and to the sibship data containing 4,770 families of siblings for the sibship association. There are ~60% censored cases among mothers and 70% censored cases among children. The number of children in each family ranges from 1 to 14 in the mother–children data with mean and median both around 5. The distribution is similar for sibship data before we excluded single child families from the sibship analysis.

More importantly, the data had tied observations due to ages being rounded to the nearest integers, which has been ignored in the previous works. We computed the modified U-statistic along with the two extended estimators and the original U-statistic. To facilitate comparison with previous work, we examined the strengths of association over the same time regions as reported in Cheng et al. [10]. We first divided the age ranges into three intervals: ≤ 70 , $71-80$, and > 80 for both dimensions. For each age region, we computed the plug-in estimate ($\hat{\zeta}_{CM}$), the pseudo-likelihood estimate ($\tilde{\theta}_{CM}^L$), the original U-estimate ($\hat{\theta}_{CM}^U$), and the modified U-estimate ($\hat{\theta}_{CM}^{MU}$) of the child–mother association according to eqs (2), (7), (4) and an equation similar to eq. (12) with uniform weight. The model-based standard errors (MSE) were available for the plug-in and U-type estimators and calculated based on their influence functions. We also computed the bootstrap standard errors (BSE) for the four estimators based on 500 bootstrap samples. The results are summarized in the top panel of Table 5. Similarly, for the sibship association, we computed the four estimates $\hat{\zeta}_{CC}$, $\tilde{\theta}_{CC}^L$, $\hat{\theta}_{CC}^U$, and $\hat{\theta}_{CC}^{MU}$. The model-based standard errors were given for the plug-in estimator and U-estimators. The bootstrap standard errors were also calculated for all four estimators based on 500 bootstrap samples, each containing 2,000 clusters of siblings drawn from 4,770 families with replacement. The reported BSEs were adjusted by a factor of $\sqrt{\frac{2,000}{4,770}}$. The results on sibship association are given in the lower panel of Table 5. The same analyses were performed on four regions $\leq (>)75 \times \leq (>)80$, and the results are shown in Table 6.

The original U-estimates for mother–child associations are identical to those reported in Cheng et al. [10]. However, there are some that differ in the sibship analysis because we are now using not only the pairs whose event times belong to the region of interest but also the pairs whose switched times under exchangeability belong to the region. For mother–child associations, the four estimates are coherent on most regions. The modified U-estimates almost always lie between the plug-in estimates and the original U-estimates except for the region $s > 80, t \leq 70$ where there are only three mother–child pairs both having dementia ($N_{11} = 3$). We observe a similar trend for sibship associations in the lower panel of Table 5. This is in agreement with our simulation studies in Section 2.5 where we have found that the plug-in estimator tends to underestimate and the original U-statistic tends to overestimate, while the modified U-statistic is closest to the true positive association. This is the first time that tied events have been taken into account in estimating the constant association over a time region, which produces more reliable results as suggested by our simulation studies.

Table 5 Child–mother association and sibship association in the Cache County Study estimated within 3×3 time regions Ω_{qr} , number of pairs at risk in each time region (N), number of cause 1 pairs in each time region (N_{11}), estimate (EST), model-based standard error (MSE), bootstrap standard error (BSE), s for child, and t for mother

Child–mother association													
$(s, t) \in \Omega_{qr}$	Obs		$\hat{\zeta}_{CM}$			$\tilde{\theta}_{CM}^L$		$\hat{\theta}_{CM}^U$			$\hat{\theta}_{CM}^{MU}$		
	N	N_{11}	EST	MSE	BSE	EST	BSE	EST	MSE	BSE	EST	MSE	BSE
$s \leq 70, t \leq 70$	1,338	10	4.03	1.33	1.40	2.57	1.22	4.21	1.42	1.51	4.06	1.36	1.34
$s \leq 70, 70 < t \leq 80$	2,171	12	2.27	0.69	0.72	3.13	1.16	2.34	0.73	0.76	2.29	0.70	0.69
$s \leq 70, t > 80$	3,907	7	1.24	0.46	0.48	1.31	0.56	1.30	0.49	0.51	1.25	0.47	0.49
$70 < s \leq 80, t \leq 70$	1,258	18	4.86	1.26	1.30	4.83	1.51	5.14	1.39	1.44	4.93	1.30	1.36
$70 < s \leq 80, 70 < t \leq 80$	2,271	23	2.64	0.58	0.60	3.00	0.84	2.76	0.64	0.66	2.68	0.60	0.58
$70 < s \leq 80, t > 80$	3,757	15	2.13	0.63	0.64	1.85	0.63	2.27	0.70	0.72	2.15	0.65	0.67
$s > 80, t \leq 70$	650	3	0.96	0.49	0.48	1.83	0.71	0.94	0.49	0.49	0.96	0.49	0.51
$s > 80, 70 < t \leq 80$	961	10	2.44	0.89	0.95	2.55	0.89	2.55	0.98	1.08	2.47	0.92	0.96
$s > 80, t > 80$	1,715	9	2.40	1.04	1.12	1.88	0.91	2.66	1.19	1.15	2.43	1.08	1.07

Sibship association													
$(s, t) \in \Omega_{qr}$	Obs		$\hat{\zeta}_{CC}$			$\tilde{\theta}_{CC}^L$		$\hat{\theta}_{CC}^U$			$\hat{\theta}_{CC}^{MU}$		
	N	N_{11}	EST	MSE	BSE	EST	BSE	EST	MSE	BSE	EST	MSE	BSE
$s \leq 70, t \leq 70$	22,478	26	3.40	1.01	1.08	3.16	1.23	3.42	1.03	1.12	3.44	1.03	1.09
$s \leq 70, 70 < t \leq 80$	17,525	44	3.35	0.58	0.56	3.87	0.76	3.45	0.61	0.59	3.37	0.59	0.57
$s \leq 70, t > 80$	5,256	27	2.42	0.60	0.60	2.42	0.63	2.56	0.66	0.68	2.44	0.62	0.66
$70 < s \leq 80, 70 < t \leq 80$	21,500	104	2.95	0.49	0.48	3.16	0.55	3.11	0.53	0.53	3.00	0.51	0.52
$70 < s \leq 80, t > 80$	10,531	56	2.58	0.44	0.40	2.41	0.42	2.80	0.50	0.47	2.61	0.45	0.44
$s > 80, t > 80$	9,984	60	1.65	0.45	0.42	1.43	0.27	1.81	0.52	0.52	1.67	0.45	0.46

Table 6 Child–mother association and sibship association in the Cache County Study estimated within 2×2 time regions Ω_{qr} , number of pairs at risk in each time region (N), number of cause 1 pairs in each time region (N_{11}), estimate (EST), model-based standard error (MSE), and bootstrap standard error (BSE)

Child–mother association													
$(s, t) \in \Omega_{qr}$	Obs		$\hat{\zeta}_{CM}$			$\tilde{\theta}_{CM}^L$		$\hat{\theta}_{CM}^U$			$\hat{\theta}_{CM}^{MU}$		
	N	N_{11}	EST	MSE	BSE	EST	BSE	EST	MSE	BSE	EST	MSE	BSE
$s \leq 75, t \leq 80$	5,319	41	3.41	0.60	0.59	3.57	0.68	3.55	0.65	0.65	3.44	0.61	0.62
$s \leq 75, t > 80$	5,891	12	1.30	0.39	0.40	1.25	0.40	1.36	0.42	0.44	1.30	0.39	0.45
$s > 75, t \leq 80$	3,330	35	2.40	0.48	0.52	2.45	0.49	2.49	0.53	0.58	2.43	0.49	0.50
$s > 75, t > 80$	3,488	19	2.71	0.87	0.55	1.80	0.41	2.90	0.99	0.63	2.75	0.90	0.95

Sibship association													
$(s, t) \in \Omega_{qr}$	Obs		$\hat{\zeta}_{CC}$			$\tilde{\theta}_{CC}^L$		$\hat{\theta}_{CC}^U$			$\hat{\theta}_{CC}^{MU}$		
	N	N_{11}	EST	MSE	BSE	EST	BSE	EST	MSE	BSE	EST	MSE	BSE
$s \leq 75, t \leq 80$	60,922	128	3.28	0.52	0.51	3.17	0.50	3.36	0.55	0.54	3.33	0.53	0.52
$s \leq 75, t > 80$	9,663	44	2.61	0.47	0.52	2.41	0.45	2.76	0.52	0.58	2.63	0.48	0.49
$s > 75, t \leq 80$	33,893	168	2.94	0.35	0.34	2.77	0.34	3.11	0.39	0.39	2.97	0.36	0.36
$s > 75, t > 80$	16,108	94	2.03	0.34	0.33	1.67	0.25	2.24	0.40	0.40	2.06	0.35	0.34

However, when the number of double cause 1 events N_{11} is small (i.e. $N_{11} = 2$), there is a noticeable discrepancy between the pseudo-likelihood estimate and the other three. For sibship associations, because of the exchangeability assumption, the estimators are the same over two symmetric regions, for example, $s \leq 70, 70 < t \leq 80$ and $70 < s \leq 80, t \leq 70$. We only report one of the symmetric regions. The four sets of estimates are generally close, and the standard errors are smaller than those for mother–child associations. As more sibship pairs were used in the sibship analysis than child–mother pairs in the child–mother analysis, the sibship estimates tend to be more robust than the child–mother estimates. This suggests that it would be important to extend original estimators for bivariate data to clustered data and to incorporate as much information as possible in quantifying mother–child associations and sibship associations when the event is rare.

When we examine the mother–child associations on broader regions, as given in Table 6, the plug-in estimates, the modified U-estimates, and the original U become closer and always follow the same order. A similar trend is observed in the sibship analysis. However, the discrepancy between the likelihood-based estimates and the other three still exists, especially on the region $s > 75, t > 80$, for both mother–child and sibship associations, in spite of the fact that the N_{11} s get bigger in the four-region analysis. One possible explanation is that the association may be time-varying on these broader regions, and the estimated association measure can be thought of as some weighted average over each region. The likelihood-based method weights the time-varying association differently as compared to the plug-in and U-statistics, since the denominator of eq. (11) is in general different from 1.

In this application, we consider and contrast the associations over a 3×3 grid and a 2×2 grid. The selection of endpoints requires careful consideration. In the Cache County Study, the number of family is large. However, the incidence of dementia is relatively low. Cheng et al. [20] reported the cumulative incidence of dementia for a mother or her eldest child is about 0.1 at age 90. Since we are quantifying association in dementia onset among family members, we pay particular attention to N_{11} . One of the 3×3 grids $s > 80, t \leq 70$ only contains three double cause 1 events, and the result is likely unreliable. The N_{11} s over the 2×2 grid are 12 or above in the mother–children analysis, and the analysis tends to be more reliable. On the other hand, the varying association is likely to be smoothed out. Our experience suggests starting with large grids and then refining the analysis depending on the change in association and the number of double cause 1 events in each cell.

5 Remarks

In this paper we discuss three types of nonparametric approaches to estimating two equivalent and easily interpreted association measures for multivariate competing risks data. The pseudo-likelihood estimator lacks an explicit variance estimator and relies on bootstrapping, hence it is most computationally intensive. For example, it took about half a second to obtain the plug-in estimate or the U-estimate as well as the corresponding standard error and about $0.5(500) = 250$ seconds to obtain the pseudo-likelihood estimate and its bootstrap standard error based on 500 bootstrap samples, for the case of $n = 100$, $CP = 35\%$, and $R_1 = 0.5$ as reported in Table 1, on a PC with intel i7 1.6 GHz CPU. When the sample size increases to 200, the computing time for obtaining the plug-in estimate or the U-estimate and its corresponding standard error increases to be around 2 seconds, and that for the pseudo-likelihood estimate increases to be $5(500) = 2,500$ seconds. Therefore, we recommend using the plug-in estimator or the U-estimator for their simplicity and good practical performance as illustrated in our simulations, when the assumption of constant association holds within each grid and there are no or few tied events in a practical application.

On the other hand, the pseudo-likelihood estimator has the potential to be most efficient. In this paper we use a stratified approach which always matches up children with the same sibling order. It is possible to check concordance status for two pairs, for example, a mother and her eldest child from one family and another mother and her second child from another family. There are actually many more combinations than

those used in eqs (7) and (8). Based on our simulation studies, the stratified pseudo-likelihood estimator has comparable efficiency to the other two estimators. If we were able to use all possible combinations, which of course would be very computationally intensive, we would further improve the efficiency of the pseudo-likelihood estimator. Moreover, the pseudo-likelihood estimator weights the time-varying associations differently from the plug-in estimator and the U-estimator. We hence recommend using both the pseudo-likelihood estimator and the plug-in (or U-) estimator to capture different aspects of the underlying association, when we suspect the association to vary over time but cannot refine the grids due to the limited sample size. In addition, the likelihood-based approach is more adaptive to a regression setting or to a parametric model. For example, one may extend the regression model for bivariate competing risks data [19] to multivariate competing risks data based on the pseudo-likelihood functions in eqs (7) and (8). Recently, Hu et al. [23] used a polynomial function to approximate cross hazard ratio and estimated the unknown parameters based on a pseudo-partial likelihood. Similar parametric approximation may be used for the cause-specific cross hazard ratio, and the pseudo-likelihood functions considered in this paper can again be used as the base of estimation. These will be future research topics.

More importantly, we investigate how rounding errors affect the performance of the three piecewise estimators and propose a modified U-statistic to handle tied multivariate competing risks data. For the modified U-statistic, we simply adjust for concordant and discordant pairs for the tied events. It is also fast to compute the modified U-estimate, e.g. taking about a second to obtain the modified U-estimate and standard error for the case of $n = 100$, $CP = 35\%$, $R_1 = 0.5$ as reported in Table 3. Based on our current simulation studies and real data analysis, we suggest using the modified U-statistic when the data are subject to rounding errors.

It is worth pointing out that our method can also handle other-cause association and across-cause association. However, the estimators may not have satisfactory performance under settings in which it is relatively rare to observe two different types of outcomes for two relatives. This is also true for the primary cause. In the Cache County Study, the onset of dementia is relatively rare. Thus, a large sample is needed to warrant reliable estimation of the familial association in dementia, and it is more crucial to use the information from all children in the estimation. If a large sample is not available, some parametric or semiparametric models such as the frailty model considered in Gorfine and Hsu [11] may be more efficient in estimating the familial association of a rare event among family members.

Appendix

Asymptotic properties of $\hat{\zeta}_{CM}$

Recall the extended association measure for mother–children data within region Ω_{qr}

$$\zeta_{CM}(\Omega_{qr}) = \frac{\int \int_{\Omega_{qr}} w(s, t) EN_{11}^*(ds, dt) EH^*(s, t)}{\int \int_{\Omega_{qr}} w(s, t) EN_{10}^*(ds, t) EN_{01}^*(s, dt)},$$

where N_{11}^* , N_{10}^* , N_{01}^* , and H^* are the double or single event processes and at-risk process defined over all mother–child pairs, for instance, $N_{11}^*(s, t) = \sum_j I(T_j \leq s, \varepsilon_j = 1, T_0 \leq t, \varepsilon_0 = 1)$. We now consider asymptotic properties of the nonparametric plug-in estimator

$$\hat{\zeta}_{CM}(\Omega_{qr}) = \frac{\int \int_{\Omega_{qr}} \hat{w}(s, t) \mathbb{P}N_{11}^*(ds, dt) \mathbb{P}H^*(s, t)}{\int \int_{\Omega_{qr}} \hat{w}(s, t) \mathbb{P}N_{10}^*(ds, t) \mathbb{P}N_{01}^*(s, dt)},$$

as given in eq. (6).

Following the arguments in Cheng and Fine [7], we establish consistency and asymptotic normality as below. For simplicity, in the sequel, we omit the integration region Ω_{qr} . Observe that

$$\hat{\zeta}_{CM} - \zeta_{CM} = \frac{A}{\left\{ \int \int \hat{w}(s, t) \mathbb{P}_n N_{10}^*(ds, t) \mathbb{P}_n N_{01}^*(s, dt) \right\} \times \left\{ \int \int w(s, t) EN_{10}^*(ds, t) EN_{01}^*(s, dt) \right\}},$$

where

$$\begin{aligned} A &= \int \int w(s, t) EN_{10}^*(ds, t) EN_{01}^*(s, dt) \int \int \hat{w}(s, t) \mathbb{P}_n N_{11}^*(ds, dt) \mathbb{P}_n H^*(s, t) \\ &\quad - \int \int w(s, t) EN_{11}^*(ds, dt) EH^*(s, t) \int \int \hat{w}(s, t) \mathbb{P}_n N_{10}^*(ds, t) \mathbb{P}_n N_{01}^*(s, dt) \\ &= \int \int w(s, t) EN_{10}^*(ds, t) EN_{01}^*(s, dt) \left\{ \int \int \hat{w}(s, t) \mathbb{P}_n N_{11}^*(ds, dt) \mathbb{P}_n H^*(s, t) \right. \\ &\quad \left. - \int \int w(s, t) EN_{11}^*(ds, dt) EH^*(s, t) \right\} \\ &\quad - \int \int w(s, t) EN_{11}^*(ds, dt) EH^*(s, t) \left\{ \int \int \hat{w}(s, t) \mathbb{P}_n N_{10}^*(ds, t) \mathbb{P}_n N_{01}^*(s, dt) \right. \\ &\quad \left. - \int \int w(s, t) EN_{10}^*(ds, t) EN_{01}^*(s, dt) \right\}. \end{aligned}$$

Consistency of $\hat{\zeta}_{CM}$: Since $N_{11}^*(s, t)$ and $H^*(s, t)$ are summations of bivariate indicator functions with $EM < \infty$, we have $\sqrt{n} \begin{pmatrix} \mathbb{P}_n N_{11}^* - EN_{11}^* \\ \mathbb{P}_n H^* - EH^* \end{pmatrix} \rightsquigarrow \begin{pmatrix} Z_{N_{11}^*} \\ Z_{H^*} \end{pmatrix}$, where $Z_{N_{11}^*}$ and Z_{H^*} are mean zero Gaussian processes, and $\rightsquigarrow$ denotes convergence in distribution. By the functional δ -method,

$$\begin{aligned} &\sqrt{n} \left\{ \int \int \mathbb{P}_n N_{11}^*(ds, dt) \mathbb{P}_n H^*(s, t) - \int \int EN_{11}^*(ds, dt) EH^*(s, t) \right\} \\ &\xrightarrow{P} \int \int Z_{N_{11}^*}(ds, dt) EH^*(s, t) + \int \int EN_{11}^*(ds, dt) Z_{H^*}(s, t), \end{aligned}$$

where $\xrightarrow{P}$ denotes convergence in probability. Then,

$$\begin{aligned} &\left| \int \int \hat{w} \mathbb{P}_n N_{11}^*(ds, dt) \mathbb{P}_n H^*(s, t) - \int \int w EN_{11}^*(ds, dt) EH^*(s, t) \right| \\ &\leq \left| \int \int \mathbb{P}_n N_{11}^*(ds, dt) \mathbb{P}_n H^*(s, t) - \int \int EN_{11}^*(ds, dt) EH^*(s, t) \right| \\ &\quad + \|\hat{w} - w\|_\infty \cdot |\tau_{1,q+1} - \tau_{1,q}| \cdot |\tau_{2,r+1} - \tau_{2,r}| \xrightarrow{P} 0, \end{aligned}$$

as $n \rightarrow \infty$, where $\|\cdot\|_\infty$ denotes supremum over region Ω_{qr} . Similarly, we obtain

$$\left| \int \int \hat{w} \mathbb{P}_n N_{10}^*(ds, t) \mathbb{P}_n N_{01}^*(s, dt) - \int \int w EN_{10}^*(ds, t) EN_{01}^*(s, dt) \right| \xrightarrow{P} 0,$$

as $n \rightarrow \infty$. The consistency of $\hat{\zeta}_{CM}^{qr}$ is immediate.

Asymptotic normality of $n^{1/2}(\hat{\zeta}_{CM} - \zeta_{CM})$: By some simple algebra,

$$\begin{aligned} \sqrt{n}A &= \int \int w EN_{10}^*(ds, t) EN_{01}^*(s, dt) \left\{ \int \int \sqrt{n}(\hat{w} - w) \mathbb{P}_n N_{11}^*(ds, dt) \mathbb{P}_n H^*(s, t) \right. \\ &\quad \left. + \int \int w EN_{11}^*(ds, dt) \sqrt{n}(\mathbb{P}_n - E)H^*(s, t) + \int \int w \sqrt{n}(\mathbb{P}_n - E)N_{11}^*(ds, dt) \mathbb{P}_n H^*(s, t) \right\} \\ &\quad - \int \int w EN_{11}^*(ds, dt) EH^*(s, t) \left\{ \int \int \sqrt{n}(\hat{w} - w) \mathbb{P}_n N_{10}^*(ds, t) \mathbb{P}_n N_{01}^*(s, dt) \right. \\ &\quad \left. + \int \int w EN_{10}^*(ds, t) \sqrt{n}(\mathbb{P}_n - E)N_{01}^*(s, dt) + \int \int w \sqrt{n}(\mathbb{P}_n - E)N_{10}^*(ds, t) \mathbb{P}_n N_{01}^*(s, dt) \right\}. \end{aligned}$$

Note that

$$\begin{aligned}
 & \int \int wEN_{10}^*(ds, t)EN_{01}^*(s, dt) \int \int \sqrt{n}(\hat{w} - w)\mathbb{P}_n N_{11}^*(ds, dt)\mathbb{P}_n H^*(s, t) \\
 & - \int \int wEN_{11}^*(ds, dt)EH^*(s, t) \int \int \sqrt{n}(\hat{w} - w)\mathbb{P}_n N_{10}^*(ds, t)\mathbb{P}_n N_{01}^*(s, dt) \\
 & = \int \int wEN_{10}^*(ds, t)EN_{01}^*(s, dt) \left\{ \int \int \sqrt{n}(\hat{w} - w)EN_{11}^*(ds, dt)EH^*(s, t) + o_P(1) \right\} \\
 & - \int \int wEN_{11}^*(ds, dt)EH^*(s, t) \left\{ \int \int \sqrt{n}(\hat{w} - w)EN_{10}^*(ds, t)EN_{01}^*(s, dt) + o_P(1) \right\} \\
 & = o_P(1),
 \end{aligned} \tag{9}$$

since we assume constant ζ in the integration region. It follows that under time-invariance, $\sqrt{n}(\hat{\zeta}_{CM} - \zeta_{CM}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n I_i + o_P(1)$, where the influence function

$$\begin{aligned}
 I_i = & \frac{1}{\int \int wEN_{10}^*(ds, t)EN_{01}^*(s, dt)} \left[\int \int wEN_{11}^*(ds, dt) \left\{ \sum_{j \leq m_i} I(Y_{ij} \geq s, Y_{i0} \geq t) \right. \right. \\
 & \left. \left. - E(m_i)P(Y_{i1} \geq s, Y_{i0} \geq t) \right\} + \int \int w \left\{ \sum_{j \leq m_i} I(Y_{ij} = s, \eta_{ij} = 1, Y_{i0} = t, \eta_{i0} = 1) \right. \right. \\
 & \left. \left. - E(m_i)P(Y_{i1} = s, \eta_{i1} = 1, Y_{i0} = t, \eta_{i0} = 1) \right\} EH^*(s, t) \right] - \frac{\int \int wEN_{11}^*(ds, dt)EH^*(s, t)}{\left\{ \int \int wEN_{10}^*(ds, t)EN_{01}^*(s, dt) \right\}^2} \\
 & \left[\int \int wEN_{10}^*(ds, t) \left\{ \sum_{j \leq m_i} I(Y_{ij} \geq s, Y_{i0} = t, \eta_{i0} = 1) - E(m_i)P(Y_{i1} \geq s, Y_{i0} = t, \eta_{i0} = 1) \right\} \right. \\
 & \left. + \int \int w \left\{ \sum_{j \leq m_i} I(Y_{ij} = s, \eta_{ij} = 1, Y_{i0} \geq t) - E(m_i)P(Y_{i1} = s, \eta_{i1} = 1, Y_{i0} \geq t) \right\} EN_{01}^*(s, dt) \right].
 \end{aligned} \tag{10}$$

By the asymptotic linearity, and assuming $EM^2 < \infty$, we have the normality of $n^{1/2}(\hat{\zeta}_{CM} - \zeta_{CM})$ via standard central limit theorem.

Asymptotic properties of $\tilde{\theta}_{CM}^L$

Without confusion we simply denote $\theta_{CM}(\Omega_{qr})$ as θ . Taking the first derivative of L_n in θ and multiplying by $\frac{\theta}{\binom{n}{2}}$, we have the following estimating equation:

$$\begin{aligned}
 \Psi_n(\theta) = & \frac{1}{\binom{n}{2}} \left[\sum_{1 \leq i \leq n} \sum_{d=1}^{m_{\max}} I\{i \in D_d\} \frac{\phi_{i,d0}^{qr} \{a(Y_{id}, Y_{i0}; k, l) - 1\}}{a(Y_{id}, Y_{i0}; k, l) - 1 + \theta} \right. \\
 & \left. - \sum_{1 \leq i < j \leq n} \sum_{d=1}^{m_{\max}} I\{i, j \in D_d\} \frac{\psi_{ij,dd00}^{qr} \theta}{a(Y_{(idjd)}, Y_{(iojo)}; k, l) - 1 + \theta} \right] \\
 = & \frac{1}{\binom{n}{2}} \left[\sum_{1 \leq i < j \leq n} \sum_{d=1}^{m_{\max}} I\{i, j \in D_d\} \frac{\phi_{ij,dd00}^{qr}}{a(Y_{(idjd)}, Y_{(iojo)}; k, l) - 1 + \theta} \right. \\
 & \left. - \sum_{1 \leq i < j \leq n} \sum_{d=1}^{m_{\max}} I\{i, j \in D_d\} \frac{\psi_{ij,dd00}^{qr} \theta}{a(Y_{(idjd)}, Y_{(iojo)}; k, l) - 1 + \theta} \right] = 0,
 \end{aligned}$$

where $\phi_{ij,dd00}^{qr}$ is the concordant indicator defined for U-statistics.

Define $X_i = (Y_{i0}, Y_{i1}, \dots, Y_{im_i}, \eta_{i0}, \eta_{i1}, \dots, \eta_{im_i}, m_i), i = 1, \dots, n$, and denote

$$h(X_i, X_j; \theta) = \sum_{d=1}^{m_{\max}} I\{i, j \in D_d\} \left[\frac{\phi_{ij,dd00}^{qr} - \psi_{ij,dd00}^{qr} \theta}{a(Y_{(idj)}, Y_{(iojo)}; k, l) - 1 + \theta} \right]. \quad (11)$$

Then $\Psi_n(\theta)$ is a U-statistic of order 2 with the kernel function $h(X_i, X_j; \theta)$.

Let $\Psi(\theta) = E\Psi_n(\theta)$ and θ_0 be the solution to $\Psi(\theta) = 0$. Define $\hat{\Psi}(\theta) = \frac{2}{n} \sum_{i=1}^n h_1(X_i; \theta)$, where $h_1(X_i; \theta) = Eh(X_i, X_j; \theta | X_i) - \Psi(\theta)$. Since $\left| \frac{\phi_{ij,dd00}^{qr} - \psi_{ij,dd00}^{qr} \theta}{a(Y_{(idj)}, Y_{(iojo)}; k, l) - 1 + \theta} \right| \leq \frac{1+\theta}{\theta}$, for any $\theta > 0$, and $m_{\max}$ is finite for a family study, we have $Eh^2(X_i, X_j; \theta) < \infty$. By Theorem 12.3 of van der Vaart [16], $\sqrt{n}(\Psi_n(\theta) - \Psi(\theta) - \hat{\Psi}(\theta)) \xrightarrow{P} 0$ and $\sqrt{n}(\Psi_n(\theta) - \Psi(\theta))$ converge to a mean zero random variable Z for any $\theta > 0$. Note that the map $\theta \rightarrow \Psi_n(\theta)$ is non-increasing. By Lemma 5.10 of van der Vaart [16], we have $\tilde{\theta}_{CM}^L \rightarrow \theta_0$ in probability, as $n \rightarrow \infty$.

Next, we will establish the asymptotic normality of the estimator using a master theorem of Z-estimators. For any $\theta_1, \theta_2 > 0$, note that $|h(X_i, X_j; \theta_1) - h(X_i, X_j; \theta_2)| = \left| \sum_{d=1}^{m_{\max}} I\{i, j \in D_d\} \frac{\{\phi_{ij,dd00}^{qr} - \psi_{ij,dd00}^{qr} \theta_1\}(\theta_2 - \theta_1)}{(a - 1 + \theta_1)(a - 1 + \theta_2)} \right| \leq \frac{m_{\max}}{a(Y_{(idj)}, Y_{(iojo)}; k, l) - 1} |\theta_1 - \theta_2|$. Therefore,

$$\begin{aligned} \sqrt{n}(\Psi_n - \Psi)(\tilde{\theta}_{CM}^L) - \sqrt{n}(\Psi_n - \Psi)(\theta_0) &= \sqrt{n}\{\hat{\Psi}(\tilde{\theta}_{CM}^L) - \hat{\Psi}(\theta_0)\} + o_P(1) \\ &= \sqrt{n} \cdot \frac{1}{n} \sum_{i=1}^n \{h_1(X_i; \tilde{\theta}_{CM}^L) - h_1(X_i; \theta_0)\} + o_P(1) \\ &\leq \sqrt{n}|\tilde{\theta}_{CM}^L - \theta_0| \cdot E \left\{ \frac{m_{\max}}{a(Y_{(idj)}, Y_{(iojo)}; k, l) - 1} \right\} + o_P(1) = o_P(1 + \sqrt{n}|\tilde{\theta}_{CM}^L - \theta_0|), \end{aligned}$$

since $a(Y_{(idj)}, Y_{(iojo)}; k, l)$ is the at-risk set of order n . Let $\dot{\Psi}(\theta_0) = E\left\{\frac{\partial h_1(X_i; \theta)}{\partial \theta} \Big|_{\theta=\theta_0}\right\}$. $\tilde{\theta}_{CM}^L$ is chosen such that $|\Psi_n(\tilde{\theta}_{CM}^L)| = o_P(n^{-1/2})$. By Theorem 13.4 of Kosorok [15], we have $\sqrt{n}(\tilde{\theta}_{CM}^L - \theta_0) \rightarrow -\dot{\Psi}_{\theta_0}^{-1}Z$ in distribution. The bootstrap validity also follows.

Formula for corrected U

We now define concordant and discordant pairs for tied sibship observations as follows:

$$\begin{aligned} \tilde{\phi}_{ij,abcd}^{qr} &= I\{Y_{ia} = Y_{jc}, Y_{ib} > Y_{jd}, (Y_{ia}, Y_{jd}) \in \Omega_{qr}, \eta_{ia} \neq 1, \eta_{jc} = 1, \eta_{jd} = 1\} \\ &\quad + I\{Y_{ia} = Y_{jc}, Y_{ib} < Y_{jd}, (Y_{ia}, Y_{ib}) \in \Omega_{qr}, \eta_{ia} = 1, \eta_{jc} \neq 1, \eta_{ib} = 1\} \\ &\quad + I\{Y_{ia} > Y_{jc}, Y_{ib} = Y_{jd}, (Y_{jc}, Y_{ib}) \in \Omega_{qr}, \eta_{jc} = 1, \eta_{ib} \neq 1, \eta_{jd} = 1\} \\ &\quad + I\{Y_{ia} < Y_{jc}, Y_{ib} = Y_{jd}, (Y_{ia}, Y_{ib}) \in \Omega_{qr}, \eta_{ia} = 1, \eta_{ib} = 1, \eta_{jd} \neq 1\} \\ &\quad + 0.5 \times \left\{ I\{Y_{ia} \neq Y_{jc}, Y_{ib} = Y_{jd}, (Y_{iajc}, Y_{ib}) \in \Omega_{qr}, \eta_{iajc} = 1, \eta_{ib} = \eta_{jd} = 1\} \right. \\ &\quad \quad \left. I\{Y_{ia} = Y_{jc}, Y_{ib} \neq Y_{jd}, (Y_{ia}, Y_{ibjd}) \in \Omega_{qr}, \eta_{ia} = \eta_{jc} = 1, \eta_{ibjd} = 1\} \right. \\ &\quad \quad \left. I\{Y_{ia} = Y_{jc}, Y_{ib} = Y_{jd}, (Y_{ia}, Y_{ib}) \in \Omega_{qr}, \right. \\ &\quad \quad \quad \left. \eta_{ia} = \eta_{jc} = 1, \{\eta_{ib} = 1, \eta_{jd} \neq 1\} \text{ or } \{\eta_{ib} \neq 1, \eta_{jd} = 1\}\} \right. \\ &\quad \quad \left. I\{Y_{ia} = Y_{jc}, Y_{ib} = Y_{jd}, (Y_{ia}, Y_{ib}) \in \Omega_{qr}, \right. \\ &\quad \quad \quad \left. \{\eta_{ia} = 1, \eta_{jc} \neq 1\} \text{ or } \{\eta_{ia} \neq 1, \eta_{jc} = 1\}, \eta_{ib} = \eta_{jd} = 1\} \right\} \\ &\quad + 0.5 \times I\{Y_{ia} = Y_{jc}, Y_{ib} = Y_{jd}, (Y_{ia}, Y_{ib}) \in \Omega_{qr}, \eta_{ia} = \eta_{jc} = \eta_{ib} = \eta_{jd} = 1\}, \end{aligned}$$

$$\begin{aligned}
\tilde{\psi}_{ij,acbd}^{qr} = & I\{Y_{ia} = Y_{jc}, Y_{ib} > Y_{jd}, (Y_{ia}, Y_{jd}) \in \Omega_{qr}, \eta_{ia} = 1, \eta_{jc} \neq 1, \eta_{jd} = 1\} \\
& + I\{Y_{ia} = Y_{jc}, Y_{ib} < Y_{jd}, (Y_{ia}, Y_{ib}) \in \Omega_{qr}, \eta_{ia} \neq 1, \eta_{jc} = 1, \eta_{ib} = 1\} \\
& + I\{Y_{ia} > Y_{jc}, Y_{ib} = Y_{jd}, (Y_{jc}, Y_{ib}) \in \Omega_{qr}, \eta_{jc} = 1, \eta_{ib} = 1, \eta_{jd} \neq 1\} \\
& + I\{Y_{ia} < Y_{jc}, Y_{ib} = Y_{jd}, (Y_{ia}, Y_{ib}) \in \Omega_{qr}, \eta_{ia} = 1, \eta_{ib} \neq 1, \eta_{jd} = 1\} \\
& + 0.5 \times \{I\{Y_{ia} \neq Y_{jc}, Y_{ib} = Y_{jd}, (Y_{iajc}, Y_{ib}) \in \Omega_{qr}, \eta_{(iajc)} = 1, \eta_{ib} = \eta_{jd} = 1 \\
& \quad I\{Y_{ia} = Y_{jc}, Y_{ib} \neq Y_{jd}, (Y_{ia}, Y_{ibjd}) \in \Omega_{qr}, \eta_{ia} = \eta_{jc} = 1, \eta_{(ibjd)} = 1\} \\
& \quad I\{Y_{ia} = Y_{jc}, Y_{ib} = Y_{jd}, (Y_{ia}, Y_{ib}) \in \Omega_{qr}, \\
& \quad \eta_{ia} = \eta_{jc} = 1, \{\eta_{ib} = 1, \eta_{jd} \neq 1\} \text{ or } \{\eta_{ib} \neq 1, \eta_{jd} = 1\}\} \\
& \quad I\{Y_{ia} = Y_{jc}, Y_{ib} = Y_{jd}, (Y_{ia}, Y_{ib}) \in \Omega_{qr}, \\
& \quad \eta_{ia} = 1, \eta_{jc} \neq 1\} \text{ or } \{\eta_{ia} \neq 1, \eta_{jc} = 1\}, \eta_{ib} = \eta_{jd} = 1\} \\
& + 0.5 \times I\{Y_{ia} = Y_{jc}, Y_{ib} = Y_{jd}, (Y_{ia}, Y_{ib}) \in \Omega_{qr}, \eta_{ia} = \eta_{jc} = \eta_{ib} = \eta_{jd} = 1\}.
\end{aligned}$$

Similarly we define concordant and discordant pairs for the reversed tied observations under exchangeability $\tilde{\phi}_{ij,adbc}^{qr}$ and $\tilde{\psi}_{ij,adbc}^{qr}$. Then we have the modified U-statistic for the sibship association

$$\hat{\theta}_{CC}^{MU}(\Omega_{qr}; k, l) = \frac{\binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \sum_{1 \leq a < b \leq m_i} \sum_{1 \leq c < d \leq m_j} (\phi_{ij,acbd}^{qr} + \tilde{\phi}_{ij,acbd}^{qr} + \phi_{ij,adbc}^{qr} + \tilde{\phi}_{ij,adbc}^{qr})}{\binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \sum_{1 \leq a < b \leq m_i} \sum_{1 \leq c < d \leq m_j} (\psi_{ij,acbd}^{qr} + \tilde{\psi}_{ij,acbd}^{qr} + \psi_{ij,adbc}^{qr} + \tilde{\psi}_{ij,adbc}^{qr})}. \quad (12)$$

Acknowledgments: The authors would like to thank the two referees and Dr. Nancy Pfenning for their constructive comments. We are grateful to the Cache County Study Steering Committee for providing the dementia dataset. This research was supported by the National Science Foundation grants DMS-0906449 and DMS-1207711 to Cheng.

References

1. Hougaard P. Analysis of multivariate survival data. New York: Springer, 2000.
2. Pruitt RC. Identifiability of bivariate survival curves from censored data. J Am Stat Assoc 1993;88:573–9.
3. Tsiatis A. A nonidentifiability aspect of the problem of competing risks. Proc Natl Acad Sci 1975;72:20–2.
4. Bandeen-Roche K, Liang K. Modelling multivariate failure time associations in the presence of a competing risk. Biometrika 2002;89:299–314.
5. Oakes D. Bivariate survival models induced by frailties. J Am Stat Assoc 1989;84:487–93.
6. Bandeen-Roche K, Ning J. Nonparametric estimation of bivariate failure time associations in the presence of a competing risk. Biometrika 2008;95:221–32.
7. Cheng Y, Fine JP. Nonparametric estimation of cause-specific cross hazard ratio with bivariate competing risks data. Biometrika 2008;95:233–40.
8. Shih JH, Albert PS. Modeling familial association of ages at onset of disease in the presence of competing risk. Biometrics 2010;66:1012–923.
9. Shih JH, Louis TA. Inferences on the association parameter in copula models for bivariate survival data. Biometrics 1995;51:1384–99.

10. Cheng Y, Fine JP, Bandeen-Roche K. Association analyses of clustered competing risks data via cross hazard ratio. *Biostatistics* 2010;11:82–92.
11. Gorfine M, Hsu L. Frailty-based competing risks model for multivariate survival data. *Biometrics* 2011;67:415–26.
12. Katki HA, Blackford A, Chen S, Parmigiani G. Multiple diseases in carrier probability estimation: accounting for surviving all cancers other than breast and ovary in BRCA PRO. *Stat Med* 2008;27:4532–48.
13. Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 1978;65:141–52.
14. Oakes D. A model for association in bivariate survival data. *J R Stat Soc Ser B Methodological* 1982;44:414–22.
15. Kosorok MR. Introduction to empirical processes and semiparametric inference. New York: Springer, 2007.
16. van der Vaart AW. Asymptotic statistics. Cambridge: Cambridge University Press, 1998.
17. Fine JP, Jiang H. On association in a copula with time transformations. *Biometrika* 2000;87:559–71.
18. Lindsay BG. Composite likelihood methods. In: Prabhu NU, editor. Statistical inference from stochastic processes. American Mathematical Society, 1988:221–39.
19. Ning J, Bandeen-Roche K. Estimation of time-dependent association for bivariate failure times in the presence of a competing risk. *Biometrics* 2014;70:10–20.
20. Cheng Y, Fine JP, Kosorok MR. Nonparametric analysis of bivariate competing risks data. *J Am Stat Assoc* 2007;102:1407–16.
21. Cheng Y, Fine JP. Cumulative incidence association models for bivariate competing risks data. *J R Stat Soc Ser B* 2012. DOI: 10.1111/j.1467-9868.2011.01012.x.
22. Cheng Y. Modeling cumulative incidences of dementia and dementia-free death using a novel three-parameter logistic function. *Int J Biostat* 2009. DOI: 10.2202/1557-4679.1183.
23. Hu T, Nan B, Lin X, Robins JM. Time-dependent cross ratio estimation for bivariate failure times. *Biometrika* 2011;98:341–54.