Published in final edited form as:

Biometrics. 2014 March ; 70(1): 10–20. doi:10.1111/biom.12110.

Estimation of Time-Dependent Association for Bivariate Failure Times in the Presence of a Competing Risk

Jing Ning^{1,*} and Karen Bandeen-Roche^{2,**}

¹Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030

²Department of Biostatistics, Johns Hopkins University, Baltimore, MD 21025

Summary

This paper targets the estimation of a time-dependent association measure for bivariate failure times, the conditional cause-specific hazards ratio (CCSHR), which is a generalization of the conditional hazards ratio (CHR) to accommodate competing risks data. We model the CCSHR as a parametric regression function of time and event causes and leave all other aspects of the joint distribution of the failure times unspecified. We develop a pseudo-likelihood estimation procedure for model fitting and inference and establish the asymptotic properties of the estimators. We assess the finite-sample properties of the proposed estimators against the estimators obtained from a moment-based estimating equation approach. Data from the Cache County study on dementia are used to illustrate the proposed methodology.

Keywords

Association measure; Competing risk; Conditional cause-specific hazards ratio; Dementia; Multivariate survival; Pseudo-likelihood

1. Introduction

Bivariate failure time data are often encountered in biomedical research (Skytthe et al., 2003; Struwing et al., 1997; Vanobbergen et al., 2000) when study units are paired, such as two observations from the same family in genetics studies. When analyzing such data, it is important to measure the strength of association between two correlated failure times, which, even if it is not of scientific interest, at least ensures the correctness of inferences. Although there is a substantial amount of literature on the association between correlated failure times, considerable methodological gaps remain. For example, previous work has often focused on time-invariant, or “global” dependence measures (Clayton, 1978; Oakes, 1982, 1986; Clayton and Cuzick, 1985). Much of the methodology allowing for time, or “local” variations in dependence has used highly parametric models, such as the Archimedean copula family (Genest and Mackay, 1986; Shih and Louis, 1995). A number of methods have addressed the selection of a copula family (Genest and Rivest, 1993; Viswanathan and Manatunga, 2001; Andersen et al., 2005; Chen and Bandeen-Roche, 2005; Chen and Fan, 2007); however, because copulas reduce time dependence through the survival function, the resulting inferences may be misleading if the reduction in dependence

*jning@mdanderson.org. **kbandeen@jhsph.edu.

Supplementary Materials

The computational code and details of those referenced in Sections 2.4 and 5 are available with this paper at the Biometrics website on Wiley Online Library.

is inconsistent with the data distribution (Bandeen-Roche and Ning, 2008). Methods that describe dependence nonparametrically avoid these problems. Many nonparametric methods have been proposed (Prentice and Cai, 1992; Hsu and Prentice, 1996; Wang and Wells, 2000) over the past twenty years; however, most are subject to complicated estimation or interpretation. Recently, a number of authors (Fan et al., 2000; Bandeen-Roche and Ning, 2008; Cheng and Fine, 2008) have proposed nonparametric estimations of the local conditional hazards ratio (Oakes, 1989), an easily interpreted association measure. However, in this scenario, each association is assumed to be constant over finite time “bins”. More recently, Hu et al. (2011) developed a time-dependent estimation of the local conditional hazards ratio for bivariate survival data, which cannot be easily extended in the presence of competing risks owing to the violation of the independent censoring assumption.

Dementia is a progressive degenerative medical condition and is one of the leading causes of death in the United States and Canada. There is evidence that dementia aggregates in families (Hendrie, 1998), and it is important to assess familial association in dementia onsets. The Cache County Study on Memory Health and Aging is a population study of dementias based in an exceptionally long-lived population (aged 65 and over) residing in northern Utah (Breitner et al., 1999). The study recruited participants in family units and provided information to estimate aggregation of dementia within families. A necessary consideration for the study due to the particular study population is the competing risk of death, which may predate the onset of dementia. An approach that ignores this competing risk assumes independence between death and dementia onset, and could result in biased inference of the familial aggregation of dementia. Additionally, previous studies (Bandeen-Roche and Ning, 2008; Cheng and Fine, 2008) have suggested that the association between familial age at onset of dementia is not a constant, but is particularly strong when the onset of dementia occurs at younger ages. Although time-dependent association has been proposed for bivariate failure times with competing risks, the available estimating methods often make an assumption of a constant or piecewise constant. Our goal in this paper is to provide an estimating procedure to assess the familial association of disease with competing risks without the restriction that the association is a constant or piecewise constant over time.

In this paper, we propose a semiparametric, time-dependent model that accommodates competing risks and does not require binning. The time-dependent measure, the conditional cause-specific hazards ratio (CCSHR), is described via a parametric, easily interpreted regression function, and all other aspects of the joint distribution of bivariate failure times are left unspecified. The remainder of our paper is organized as follows. In Section 2, we introduce the model and develop the pseudo-likelihood approach (Clayton, 1978; Oakes, 1982) for estimating the regression parameters. A moment-based approach is described for comparison. We assess the empirical performance of the proposed estimators via simulation studies in Section 3 and describe an application of the proposed methods to a real data example in Section 4. A discussion is provided in Section 5.

2. Methods

2.1 Notation and Estimand

We consider bivariate failure time data involving competing events $1, \dots, K$. The outcomes are times to events and types of events (X_j, K_j) , $j = 1, 2$; X_j represents the time of the earliest event, including censoring; K_j identifies the type of first-occurring event, where $K_j = 0$ if censoring occurs before any competing event of interest. We assume that censoring occurs independently of the competing events. Observed data are represented by n independent pairs, $(X_{i1}, X_{i2}, K_{i1}, K_{i2})$, $i = 1, \dots, n$. The component-wise minimum $(X_{a1} \wedge X_{b1}, X_{a2} \wedge X_{b2})$

of two pairs of failure times (X_{a1}, X_{a2}) and (X_{b1}, X_{b2}) is denoted by (X_1^{ab}, X_2^{ab}) , and (K_1^{ab}, K_2^{ab}) are the causes corresponding to (X_1^{ab}, X_2^{ab}) . One commonly used association measure of bivariate failure time is the conditional hazards ratio. Here we aim to develop a simple methodology for estimating an extension of the conditional hazards ratio in the context of competing risks, the CCSHR. That measure is defined as (Bandeen-Roche and Liang, 2002)

$$\theta_{cs}(x_1, x_2, k_1, k_2) = \frac{\lambda_{1,k_1}(x_1|X_2=x_2, K_2=k_2)}{\lambda_{1,k_1}(x_1|X_2 \geq x_2)},$$

where $\lambda_{1,k_1}(x_1|X_2=x_2, K_2=k_2) = \lim_{\Delta x \downarrow 0} Pr(x_1 - X_1 < x_1 + \Delta x, K_1 = k_1 | X_1 = x, X_2 = x_2, K_2 = k_2) / \Delta x$ and $\lambda_{1,k_1}(x_1, |X_2 \geq x_2) = \lim_{\Delta x \downarrow 0} Pr(x_1 - X_1 < x_1 + \Delta x, K_1 = k_1 | X_1 = x, X_2 \geq x_2) / \Delta x$ are two conditional cause-specific hazards for X_1 given information on its partner.

2.2 Model

Our strategy uses the fact that $\theta_{cs}(x_1, x_2, k_1, k_2)$ divides the conditional probability of concordance between the pairs' failure times by the conditional probability of discordance, each given (X_1^{ab}, X_2^{ab}) and (K_1^{ab}, K_2^{ab}) , which thus equals the conditional odds of concordance (Clayton, 1978; Oakes, 1982):

$$\theta_{cs}(x_1, x_2, k_1, k_2) = \frac{Pr\{(X_{a1} - X_{b1})(X_{a2} - X_{b2}) > 0 | (X_1^{ab}, X_2^{ab}, K_1^{ab}, K_2^{ab}) = (x_1, x_2, k_1, k_2)\}}{Pr\{(X_{a1} - X_{b1})(X_{a2} - X_{b2}) < 0 | (X_1^{ab}, X_2^{ab}, K_1^{ab}, K_2^{ab}) = (x_1, x_2, k_1, k_2)\}}. \quad (1)$$

We propose to model the conditional odds of concordance by a logistic regression function of time and failure causes,

$$\text{logit}[Pr\{(X_{a1} - X_{b1})(X_{a2} - X_{b2}) > 0 | (X_1^{ab}, X_2^{ab}, K_1^{ab}, K_2^{ab}) = (x_1, x_2, k_1, k_2)\}] = P_{cs}(x_1, x_2, k_1, k_2; \alpha),$$

where P_{cs} is a pre-specified function of times and causes, and α is a finite-dimensional Euclidean parameter. This model implies that

$$\theta_{cs}(x_1, x_2, k_1, k_2; \alpha) = \exp\{P_{cs}(x_1, x_2, k_1, k_2; \alpha)\}.$$

Our formulation accommodates certain existing models for time dependence by choosing appropriate functions of time. For instance, for a partition of $R^+ \times R^+$, $(B_j, j = 1, \dots, J)$, and a total number of M possible failure causes $((k_{j1}, k_{j2}), j = 1, \dots, M)$, the specification for time dependence described by Bandeen-Roche and Ning (2008) is equivalent to

$$\begin{aligned}
& \log \theta_{cs}(x_1, x_2, k_1, k_2) \\
&= \alpha_0 + \sum_{i=1}^{J-1} \alpha_i I\{(x_1, x_2) \in B_i\} \\
&+ \sum_{j=1}^{M-1} \alpha_{J-1+i_j} I\{(k_1, k_2) = (k_{j1}, k_{j2})\} \\
&+ \sum_{i=1}^{J-1} \sum_{j=1}^{M-1} \alpha_{M+J-2+i \times j} I\{(x_1, x_2) \in B_i, (k_1, k_2) \\
&= (k_{j1}, k_{j2})\}.
\end{aligned}$$

2.3 Estimation

Under the assumption of a constant conditional hazards ratio, Clayton (1978) suggested the pseudo-likelihood, a product of conditional probabilities for which the first log derivative yields unbiased estimating equations. In this paper, we extend Clayton's pseudo-likelihood method to a regression framework and allow the CCSHR to vary with time and possibly with other factors.

Denote $\delta(x_1, x_2) = \{i : X_{i1} \leq x_1, X_{i2} \leq x_2\}$ as the bivariate risk set and $R(x_1, x_2)$ as the size of risk set $\delta(x_1, x_2)$. The respective "concordance" and "discordance" events, $e_1(x_1, x_2, k_1, k_2)$ and $e_2(x_1, x_2, k_1, k_2)$, underlie the pseudo-likelihood estimator we propose:

$$\begin{aligned}
e_1(x_1, x_2, k_1, k_2) &= \{\exists i \in \delta(x_1, x_2) : X_{i1} \leq x_1 + dx_1, X_{i2} \leq x_2 + dx_2, \forall dx_1 > 0, dx_2 > 0; K_{i1} = k_1, K_{i2} = k_2\} \\
e_2(x_1, x_2, k_1, k_2) &= \{\exists i, j \in \delta(x_1, x_2), i \neq j : X_{i1} \leq x_1 + dx_1, X_{j2} \leq x_2 + dx_2, \forall dx_1 > 0, dx_2 > 0; K_{i1} = k_1, K_{j2} = k_2\}.
\end{aligned}$$

Event $e_1(x_1, x_2, k_1, k_2)$ occurs if any pair in the risk set $\delta(x_1, x_2)$ has failure events at time (x_1, x_2) with the associated causes (k_1, k_2) . Event $e_2(x_1, x_2, k_1, k_2)$ occurs if k_1 - and k_2 -type failures occur in different pairs in the risk set $\delta(x_1, x_2)$ at the specified times.

Calculating event probabilities conditional on the composition of the risk set yields a quantity involving conditional hazards-the component dividends of the CCSHR:

$$\begin{aligned}
Pr\{e_1(x_1, x_2, k_1, k_2) | \delta(x_1, x_2)\} &= R(x_1, x_2) \lambda_{1,k_1}(x_1 | X_2 \geq x_2) \lambda_{2,k_2}(x_2 | X_1 = x_1, K_1 = k_1) dx_1 dx_2; \\
Pr\{e_2(x_1, x_2, k_1, k_2) | \delta(x_1, x_2)\} &= R(x_1, x_2) \{R(x_1, x_2) - 1\} \lambda_{1,k_1}(x_1 | X_2 \geq x_2) \lambda_{2,k_2}(x_2 | X_1 = x_1) dx_1 dx_2.
\end{aligned}$$

We obtain terms involving ratios of conditional hazards by considering the conditional probability of $e_1(x_1, x_2, k_1, k_2)$, given the occurrence of $e_1(x_1, x_2, k_1, k_2)$ or $e_2(x_1, x_2, k_1, k_2)$ and the risk set

$$Pr\{e_1 | e_1 \cup e_2, \delta(x_1, x_2)\} = \frac{\theta_{cs}(x_1, x_2, k_1, k_2)}{R(x_1, x_2) - 1 + \theta_{cs}(x_1, x_2, k_1, k_2)}$$

and its complement probability

$$Pr\{e_2 | e_1 \cup e_2, \delta(x_1, x_2)\} = \frac{R(x_1, x_2) - 1}{R(x_1, x_2) - 1 + \theta_{cs}(x_1, x_2, k_1, k_2)}.$$

For each bivariate risk set that may be constructed from realized failure time minima (X_1^{ij}, X_2^{ij}) , either $e_1(X_1^{ij}, X_2^{ij}, K_1^{ij}, K_2^{ij})$ or $e_2(X_1^{ij}, X_2^{ij}, K_1^{ij}, K_2^{ij})$ must occur. The pseudo-likelihood multiplies the probabilities of the events observed to occur over the risk sets. Synthesizing, we have the log-pseudo-likelihood L_p as

$$L_p = \sum_i \delta_i \log \left\{ \frac{\theta_{cs}(O_i)}{R(X_{i1}, X_{i2}) - 1 + \theta_{cs}(O_i)} \right\} + \sum_{i < j} \delta^{ij} (1 - \Delta_{ij}) \log \left\{ \frac{R(X_1^{ij}, X_2^{ij}) - 1}{R(X_1^{ij}, X_2^{ij}) - 1 + \theta_{cs}(O^{ij})} \right\},$$

where $O_i = (X_{i1}, X_{i2}, K_{i1}, K_{i2})$, $O^{ij} = (X_1^{ij}, X_2^{ij}, K_1^{ij}, K_2^{ij})$, $\Delta_{ij} = 1$ if the two pairs (X_{i1}, X_{i2}) and (X_{j1}, X_{j2}) are concordant and $\Delta_{ij} = 0$ otherwise, $\delta_i = I(K_{i1} > 0, K_{i2} > 0)$ and $\delta^{ij} = I(K_1^{ij} > 0, K_2^{ij} > 0)$. Consider a regression model $\theta_{cs}(x_1, x_2, k_1, k_2; \alpha)$, where α is a $P \times 1$ vector of parameters. To obtain parameter estimates, we maximize the log-pseudo-likelihood; that is, we solve for the root of the first derivative,

$$\tilde{U}_p(\alpha) = \sum_i \delta_i \left\{ \frac{\theta'_{cs}(O_i; \alpha)}{\theta_{cs}(O_i; \alpha)} - \frac{\theta'_{cs}(O_i; \alpha)}{R(X_{i1}, X_{i2}) - 1 + \theta_{cs}(O_i; \alpha)} \right\} - \sum_{i < j} \delta^{ij} (1 - \Delta_{ij}) \frac{\theta'_{cs}(O^{ij}; \alpha)}{R(x_1^{ij}, x_2^{ij}) - 1 + \theta_{cs}(O^{ij}; \alpha)}, \quad (2)$$

where $\theta'_{cs}(O_i; \alpha) = \partial \theta_{cs}(O_i; \alpha) / \partial \alpha$.

2.4 Inference: Asymptotic Distributional Properties

Unlike partial likelihoods, pseudo-likelihoods do not require that the conditional events from which components are constructed come from a nested sequence. Therefore, they can be treated as neither partial likelihood functions nor likelihood functions for the purpose of making inferences about parameters. Instead, asymptotic properties of the pseudo-likelihood estimator must be employed to develop Wald inferences. To study these, we will further explore the first derivative of the log-pseudo-likelihood. $\tilde{U}_p(\alpha)$ can be rewritten as

$$\begin{aligned} \tilde{U}_p(\alpha) &= \sum_i \delta_i \frac{\theta'_{cs}(O_i; \alpha)}{\theta_{cs}(O_i; \alpha)} \left\{ 1 - \frac{\theta_{cs}(O_i; \alpha)}{R(X_{i1}, X_{i2}) - 1 + \theta_{cs}(O_i; \alpha)} \right\} - \sum_{i < j} \delta^{ij} (1 - \Delta_{ij}) \frac{\theta'_{cs}(O^{ij}; \alpha)}{R(X_1^{ij}, X_2^{ij}) - 1 + \theta_{cs}(O^{ij}; \alpha)} \\ &= \sum_{i < j} \delta^{ij} \frac{\theta'_{cs}(O^{ij}; \alpha)}{\theta_{cs}(O^{ij}; \alpha)} \frac{\Delta_{ij} \{1 + \theta_{cs}(O^{ij}; \alpha)\} - \theta_{cs}(O^{ij}; \alpha)}{R(X_1^{ij}, X_2^{ij}) - 1 + \theta_{cs}(O^{ij}; \alpha)}. \end{aligned} \quad (3)$$

Note that $\tilde{U}_p(\alpha)$ is an unbiased equation that follows from observing that Δ_{ij} is independent of $R(x_1, x_2)$ conditional on the component-wise minima. We can approximate equation (3) with a simplified estimating equation,

$$U_p(\alpha) = \sum_{i < j} \delta^{ij} \frac{\theta'_{cs}(O^{ij}; \alpha)}{\theta_{cs}(O^{ij}; \alpha)} \frac{\Delta_{ij} \{1 + \theta_{cs}(O^{ij}; \alpha)\} - \theta_{cs}(O^{ij}; \alpha)}{R(X_1^{ij}, X_2^{ij})} \quad (4)$$

Although the solutions to equations (3) and (4) have different values, analogous to the estimator considered by Oakes (1986), they are asymptotically equal due to the fact that the difference between equations (3) and (4) goes to zero when the sample size goes to infinity. We can prove that the solution to equation (4), $\hat{\alpha}$, is consistent and asymptotically normally

distributed under some regularity conditions, which are listed in the Supplementary Materials. Theorem 1 summarizes these asymptotic properties.

Theorem 1: Denote α^* to be the true value of α . Under regularity assumptions A.1–A.4 listed in the Web Supplementary Materials, $\hat{\alpha}$ is a consistent estimator of α^* . Moreover, $\sqrt{n}(\hat{\alpha} - \alpha^*)$ converges weakly to a normal distribution with mean zero and variance-covariance matrix $\Sigma(\alpha^*)$.

See the Supplementary Materials for the detailed proof. Here we list main steps for the proof. First we show that estimating equation (4) yields a unique and consistent estimator. Then, by Taylor series expansion, we have

$$0 = n^{-3/2}U_p(\alpha^*) - n^{-2}I_n(\alpha^*)\sqrt{n}(\hat{\alpha} - \alpha^*) + L_n(\hat{\alpha}, \alpha^*), \quad (5)$$

where α^* is the true value, $I_n(\alpha^*)$ is the first derivative of estimating equation (4), and $L_n(\alpha, \alpha^*) \propto o_p(1)$. To prove the asymptotic normality of $n^{-3/2}U_p(\alpha^*)$, we study the asymptotic normality of the difference between $n^{-3/2}U_p(\alpha^*)$ and $n^{-3/2}U_p^*(\alpha^*)$ defined to have summands as in (4); however, we replace $R(X_1^{ij}, X_2^{ij})$ with $nS(X_1^{ij}, X_2^{ij})$, where $S(\cdot, \cdot)$ is the joint survival function of X_1 and X_2 . Note that $U_p^*(\alpha^*)$ is a U-statistic; hence $n^{-3/2}U_p^*(\alpha^*)$ and thus $n^{-3/2}U_p(\alpha^*)$ are asymptotically normally distributed. This fact, taken together with the consistency of $\hat{\alpha}$ and $n^{-2}I_n(\alpha^*)$, implies that $\sqrt{n}(\hat{\alpha} - \alpha^*)$ converges to a zero-mean normal distribution.

2.5 A Moment-Based Estimator

Our generalized linear modeling form suggests an alternative estimating equation approach based more directly on the concordance data than our pseudo-likelihood approach. To simplify the description, we consider a simple linear additive model

$$\theta_{cs}(x_1, x_2, k_1, k_2) = \exp(\alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 k_1 + \alpha_4 k_2), \quad (6)$$

and propose the following algorithm:

1. For each pair of pairs $(X_{a1}, X_{a2}, K_{a1}, K_{a2})$ and $(X_{b1}, X_{b2}, K_{b1}, K_{b2})$, record

$$\left\{ \delta^{ab} = I(K_1^{ab} > 0, K_2^{ab} > 0), \Delta^{ab} = I\{(X_{a1} - X_{b1})(X_{a2} - X_{b2}) > 0\}, Z^{ab} = (1, X_1^{ab}, X_2^{ab}, K_1^{ab}, K_2^{ab})' \right\}.$$

2. Use the first moments of quantities in (1) to construct unbiased estimating equations for $\alpha' = (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$:

$$\sum_{i < j} \delta^{ij} Z^{ij} \left\{ \Delta^{ij} - \frac{\exp(\alpha' Z^{ij})}{1 + \exp(\alpha' Z^{ij})} \right\} = 0.$$

3. Obtain standard errors of the estimators using a bootstrap procedure.

We propose step (3) because the outcomes in step (1) are not mutually independent, so an analytic expression for the dependence structure will be very complicated. The estimator from the moment-based estimating equation can be proven to be consistent, and a bootstrap procedure can be used to obtain a variance for the estimator. However, only pairs with uncensored pair-wise minimums contribute to this equation, suggesting that, due to censoring, the moment-based equation may lose more information than the pseudo-

likelihood estimator. In Section 3 we compare the small sample accuracy and precision of estimators from the pseudo-likelihood approach and the moment-based estimating equation.

3. Simulation Study

3.1 Design and Data Generation

We generated bivariate data according to the frailty model for subject- and cause-specific failure hazards given by equation (6) from Bandeen-Roche and Liang (2002). That equation hypothesizes two components of within-pair dependence: a “size” frailty modeling dependence among the times to the first failure and a “shape” frailty defining pair-specific propensities to fail from two competing causes, “disease” ($k = 1$) and “death” ($k = 2$). Specifically, the subject-and cause-specific conditional hazard function is

$$\lambda_{ik}(x|A=a, B=b)=ab\lambda^*(x),$$

where “A” is a “size” frailty, “B” is a “shape” frailty by which the overall hazard is allocated to the various causes, and $\lambda^*(.) = 1$. We first generated bivariate time according to the distribution of the “size” frailty, which was assumed to be either gamma or positive stable distribution. For the competing causes, we began with generating the “shape” frailty b from a Beta distribution with mean equal to shape parameter R and scale parameter equal to 1. Then we generated two independent Uniform (0,1) random numbers u_1 and u_2 , and let $k_1 = 1$ if $u_1 < b$, otherwise $k_1 = 2$. Similarly, we set the value of k_2 by comparing u_2 to b . We considered scenarios with and without right censoring. In the scenarios with right censoring, we generated two censoring times from either a exponential or a uniform distribution independently. Each study was composed of 500 runs.

We first considered the performance of pseudo-likelihood estimators with data generated from gamma frailty and independence models. Here the “shape” frailty is a constant $B = 1$ and thus there is no competing risks (Table 1). All estimators were based on the following model:

$$\log\theta_{cs}(x_1, x_2, k_1, k_2)=\alpha_0+\alpha_1x_1+\alpha_2x_2. \quad (7)$$

For independence runs, we generated two independent failure times from the standard exponential distribution, and hence all true values of the coefficients were zero. For gamma frailty runs, we generated the “size” frailty from a gamma distribution with mean and variance of 1, yielding $\alpha_0 = 0.693$, $\alpha_1 = 0$, and $\alpha_2 = 0$.

We then considered competing risk models with size frailties distributed as gamma distributions with mean of 1 and variance of 1, and the “shape” parameter R ranging from 0.2 to 0.8. This produced CCSHR values of 6.0, 3.0, and 2.25 (Bandeen-Roche and Ning, 2008), which corresponded to the regression model

$$\log\theta_{cs}(x_1, x_2, k_1, k_2)=\alpha_0+\alpha_1I(k_1=k_2=1). \quad (8)$$

In this scenario, the true values of α_0 in model (8) are 0.60, 0.51, and 0.45, and the true values of α_1 in model (8) are 1.19, 0.59, and 0.36. To evaluate the performance of the proposed estimators with negative association, we also considered the scenarios with CCSHR values of 1.8, 0.89, and 0.66 when R ranging from 0.2 to 0.8. We first generated bivariate failure times with negative association per Nan et al. (2006) and then used the aforementioned mechanism to generate the cause indicators. Specifically, for bivariate

failure time, we generated x_1 from the standard exponential distribution, and calculated x_2 by solving $S(x_2|x_1) = u_2$, where u_2 is a random number from Uniform(0,1), and the conditional survival distribution $S(x_2|x_1)$ is specified as

$$S(x_2|x_1) = \{(1 - e^{-x_1})^{0.5} + (1 - e^{-x_2})^{0.5} - 1\}^2 / e^{-x_1}.$$

In this case, the true values of α_0 in model (8) are -0.60 , -0.71 , and -0.77 , and the true values of α_1 in model (8) are 1.19 , 0.59 , and 0.36 , respectively. Next we contrasted our proposed estimating procedure using the pseudo-likelihood approach against the moment-based estimating equation procedure with data generated from gamma frailty and independence models (Table 3). We set the conditional hazards ratio at 1.0 for independent runs and 2.0 for gamma frailty runs.

Finally, we specified the model as

$$\log \theta_{cs}(x_1, x_2, k_1, k_2) = \alpha_0 + \sum_{j=1}^3 \alpha_j I\{(x_1, x_2) \in B_j\} \quad (9)$$

to compare the small sample properties of our regression procedure against those of the U-statistic procedure for independence, gamma frailty and positive stable models described by Bandeen-Roche and Ning (2008) (Table 4). For each run, we applied the two methods to estimate the CCSHR as a bivariate function of time over a four-cell grid that bisected each time dimension at the marginal distributional median. For positive stable scenarios, the true values of CCSHR within each bin to be estimated were calculated as described in Bandeen-Roche and Ning (2008). In the presence of censoring, analyses of U-statistics treated censoring as a second “failure” cause. For ease of comparison, we tabulated the results concerning the estimation of the CCSHR for four bins instead of the estimated coefficients for the regression method.

3.2 Simulation Results

We obtained the proposed estimates by maximizing the log-pseudo-likelihood, which is implemented by the “optim” function in R (R Development Core Team, Version 2.14.1) using the quasi-Newton method.

We first considered the performance of the estimated coefficients in the absence of competing risks (see Table 1). Even with small samples (100) and moderate censoring, the mean estimates of the coefficients agreed with the true parameter settings. With sample sizes equal to 500, associated inferences were also accurate for all of the scenarios we considered: the empirical biases were less than 5%, and coverage of the 95% confidence intervals ranged from 91% to 98%. In contrast, with $n = 100$, the empirical biases ranged from 0% to 16% and coverage ranged from 85% to 97%.

In the presence of competing risks, Table 2 lists simulation results for the estimator of $\theta_{cs}(x_1, x_2, 1, 1)$. For ease of presentation, we only tabulated results concerning the estimation of $\theta_{cs}(1, 1)$, with asymptotical standard errors by the delta method. With the exception of the scenarios with small sample sizes ($n = 100$), both the estimator and its associated inference were accurate no matter what the association was positive or negative. For instance, the percentage bias was less than 2%, and coverage of the 95% confidence intervals ranged from 96% to 97% for all scenarios with sample sizes of 500. In contrast, with gamma frailty model with positive association ($n = 100$ and $R = 0.2$), the percentage bias was 14% and coverage was 92% because few samples were observed to have disease in this scenario.

Table 3 compares the performance of the estimators based on the pseudo-likelihood method with those derived from the moment-based estimating equation. Both procedures achieved outstanding accuracy even when the sample size was 100. The pseudo-likelihood method exhibited clearly superior precision, with empirical standard deviations 6%–38% smaller than those for the moment-based estimating equation. The moment-based estimating equation lost even more information in the presence of strong censoring because it only utilized fully observed pairs. For example, under the independence model and $n = 500$, without censoring and then with exponential censoring, the empirical standard deviations resulting from the pseudo-likelihood method were 25% and then 38% smaller, respectively, than those resulting from the moment-based estimating equation. We also evaluated the performance of bootstrap standard deviations of the moment-based estimators, which are based on 200 bootstrap resamples. As expected, the empirical standard deviations and bootstrap standard deviations agree well even in the scenarios with a small sample size ($n=100$).

Table 4 compares the regression method with the U-statistic method. Somewhat surprisingly, the two estimators performed quite similarly even in the case of strong censoring, suggesting that the U-statistic method may not lose too much information due to censoring.

4. Application: Aggregation of Dementia in Families

We applied the proposed methodology to analyze the aggregation of dementia in families who participated in the Cache County Study on Memory Health and Aging (Breitner et al., 1999). The data comprise the same 3635 failure-time pairs that were analyzed by Bandeen-Roche and Liang (2002), Bandeen-Roche and Ning (2008), and Cheng and Fine (2008). Among the 3635 pairs, there were 40 pairs in which both had dementia, 1132 pairs in which both died free of dementia, 259 pairs in which mother and child had different failure causes (dementia or death), 136 pairs in which both outcomes were censored, 2059 pairs in which only the child's outcome was censored, and 9 pairs in which only the mother's outcome was censored. In the analysis, the observed data for the i th family are $(X_{i1}, X_{i2}, K_{i1}, K_{i2})$, where X_{i1} is the event time of the oldest sibling, X_{i2} is the event time of the mother, and K_{ij} indicates the event type corresponding to the failure time X_{ij} . Censoring, dementia, and death are coded as 0, 1, and 2, respectively. With these notations, $\theta_{cs}(x_1, x_2, 1, 1)$ denotes here the multiplicative increase in the risk of dementia onset for children at age x_1 whose mothers have dementia at age x_2 , versus those whose mothers survive without dementia beyond age x_2 .

We began by estimating $\theta_{cs}(x_1, x_2, k_1, k_2)$ on a time grid created by dichotomizing the children's and mothers' time scales approximately at the respective medians (75 for children and 80 for mothers) for the observed time-to-first event; this yielded four bins: $B_0 = (x_1 \leq 75, x_2 \leq 80)$, $B_1 = (x_1 \leq 75, x_2 > 80)$, $B_2 = (x_1 > 75, x_2 \leq 80)$, and $B_3 = (x_1 > 75, x_2 > 80)$. Bandeen-Roche and Ning (2008) computed U-statistic estimates on the same time grids, finding $\hat{\theta}_{cs}(x_1 \leq 75, x_2 \leq 80, 1, 1) = 3.81$, $\hat{\theta}_{cs}(x_1 \leq 75, x_2 > 80, 1, 1) = 0.80$, $\hat{\theta}_{cs}(x_1 > 75, x_2 \leq 80, 1, 1) = 2.41$, and $\hat{\theta}_{cs}(x_1 > 75, x_2 > 80, 1, 1) = 5.89$. The following regression model replicates their specification:

$$\log \theta_{cs}(x_1, x_2, k_1, k_2) = \alpha_0 + \sum_{i=1}^3 \alpha_i I_{B_i} + \sum_{i=1}^3 \alpha_{3+i} I_{K_i} + \sum_{i=1}^3 \alpha_{i+6} I_{B_i} I_{K_3}, \quad (10)$$

where $I_{B_i} = I\{(x_1, x_2) \in B_i\}$, $I_{K1} = I\{(k_1, k_2) = (2, 1)\}$, $I_{K2} = I\{(k_1, k_2) = (1, 2)\}$ and $I_{K3} = I\{(k_1, k_2) = (1, 1)\}$. We fitted three versions of (10); see Table 5. The first version (Model 1) specified the additive effects of the failure times and cause, i.e., set $\alpha_7 = \alpha_8 = \alpha_9 = 0$ in (10).

This analysis indicated a convincing intra-familial association for life span: $\exp(\alpha_0) = 1.43$ with 95% confidence interval (1.29, 1.57); α_1 , α_2 and α_3 were small and lacked statistical significance, suggesting little specificity of the association to the timing bins. It further estimated the dementia-specific association to be meaningfully larger than the death-specific association: $\exp(\alpha_6) = 2.27$ with 95% confidence interval (1.78, 2.76). Overall, the analysis estimated the conditional hazards ratios specific to two dementia onsets as $\theta_{cs}(x_1 = 75, x_2 = 80, 1, 1) = 3.27$ with 95% confidence interval (2.24, 4.20), $\theta_{cs}(x_1 = 75, x_2 > 80, 1, 1) = 2.32$ with 95% confidence interval (1.51, 3.16), $\theta_{cs}(x_1 > 75, x_2 = 80, 1, 1) = 2.09$ with 95% confidence interval (1.40, 2.67), and $\theta_{cs}(x_1 > 75, x_2 > 80, 1, 1) = 2.52$ with 95% confidence interval (1.70, 3.53). Next we allowed the strengths of the failure time association to differ for two dementia onsets compared to any other failure-cause pairing (Model 2); i.e., we fitted (10) in its entirety. Estimates of the first six parameters changed little compared to Model 1; additionally we found $\exp(\alpha_7) = 0.28$ with 95% confidence interval (0.00, 0.57), indicating relative disaggregation between a late dementia onset for a mother and an early dementia onset for her child. The estimated values of cause-specific hazards ratios under this model were similar to those reported by Bandeen-Roche and Ning (2008), but the standard errors of the estimates were lower by as much as 50%. In Model 2, α_8 and α_9 were of sizes comparable to their standard errors, so we revised Model 2 by removing the corresponding terms (Model 3). The resulting fit indicated similar strengths of dementia-specific association in all time bins, except for the one corresponding to late maternal and early child onset: $\theta_{cs}(x_1 = 75, x_2 = 80, 1, 1) = 3.67$ with 95% confidence interval (2.75, 4.84), $\theta_{cs}(x_1 > 75, x_2 > 80, 1, 1) = 4.06$ with 95% confidence interval (3.08, 5.25), $\theta_{cs}(x_1 = 75, x_2 > 80, 1, 1) = 1.08$ with 95% confidence interval (0.39, 2.38), and $\theta_{cs}(x_1 > 75, x_2 = 80, 1, 1) = 3.37$ with 95% confidence interval (2.64, 4.73). Among these values, the point estimate for two late onsets differed most from the one in Model 2, which was equal to 5.13. Confidence intervals were narrower for Model 3 than for Model 2 by approximately 25%–43%.

In Models 1–3, we forced $\theta_{cs}(x_1, x_2, k_1, k_2)$ to be constant within each time grid. To check the piece-wise constant assumption, we added terms that allow $\theta_{cs}(x_1, x_2, k_1, k_2)$ to vary continuously with the event times of the mother and child (x_1, x_2). From Models 1–3 we observe that the strength of the failure time association may depend on the difference between the times of dementia onset for the mother and child ($|x_1 - x_2|$). Hence, we considered adding this time difference to Model 4

$$\log \theta_{cs}(x_1, x_2, k_1, k_2) = \alpha_0 + \sum_{i=1}^3 \alpha_i I_{B_i} + \sum_{i=1}^3 \alpha_{3+i} I_{K_i} + \alpha_7 I_{B_1} I_{K_3} + \sum_{i=1}^2 \alpha_{8+i} x_i + \sum_{i=1}^2 \alpha_{9+i} I_{K_3} x_i + \alpha_{10} |x_1 - x_2| + \alpha_{11} I_{K_3} |x_1 - x_2| \quad (11)$$

The fit of (11) strongly indicated that the intra-familial dementia-specific association decreased with increasing child onset times after controlling for bin effects (see Table 6). The small sizes and lack of statistical significance of θ_{82} and θ_{92} suggested the absence of considerable dependence of the intra-familial associations for dementia onset and death on maternal onset times after controlling for the child onset times and bin effects. As expected, the intra-familial associations for dementia onset and death depended on the difference between maternal and child onset: the larger the distance between onsets, the weaker the association. Interestingly, the small size and nonsignificance of θ_{11} suggested the absence of a considerable difference between dementia onset and death in the attenuation of the association, with a disparity in timing between the mother and the child. Accordingly, we revised Model 4 by removing the weakly supported terms involving continuous time (Model 5), i.e., we set $\alpha_{82} = \alpha_{92} = \alpha_{11} = 0$ in (11). Point estimates of the remaining parameters changed little when compared with those of Model 4. The estimated values of the cause-specific hazards ratios under Model 5 were similar to those under Model 4 but had much lower standard errors. For example, under Model 4, $\theta_{cs}(65, 65, 1, 1) = 5.04$ with standard

error 1.11 and $\hat{\theta}_{cs}(65, 85, 1, 1) = 0.82$ with standard error 0.36; under Model 5, $\hat{\theta}_{cs}(65, 65, 1, 1) = 5.08$ with standard error 0.76 and $\hat{\theta}_{cs}(65, 85, 1, 1) = 0.92$ with standard error 0.25. Finally we set $\alpha_3 = \alpha_5 = 0$ to simplify Model 5 (resulting in Model 6) due to the small size and large variances of $\hat{\alpha}_3$ and $\hat{\alpha}_5$. The estimated coefficients for Model 6 were similar to those for Model 5 with the exception of the intercept, which was equal to 0.10 due to the change of the reference group. The standard errors were smaller for Model 6 than for Model 5 by 0%–25%. Overall, Model 6 resulted in point estimates for the CCSHRs that were similar to those in Model 5 but with slightly smaller standard errors. The fit of Model (6) indicated that the strength of the intra-familial dementia-specific association was much greater than those of the other three intra-familial associations specified in the model. There was disaggregation between a late dementia onset for a mother and an early dementia onset for her child. After controlling for the children's onset times and bin effects, the association between intra-familial dementia onset and death still strongly depended on the difference between maternal and child age at the onset of dementia: the greater the difference in ages at dementia onset, the weaker the association.

To check the accuracy of the asymptotic inferences, we computed bootstrap standard errors. For models with small numbers of parameters, such as Model 3, the inferences agreed well: the differences between two standard errors were less than 6%. In contrast, for analysis with larger numbers of parameters, such as Model 4, there were considerable discrepancies for the terms involving cause (dementia, dementia). The worst case was for the interaction between the timing of onset (early, late) and causes (dementia, dementia), for which the bootstrap standard error was 0.75 and the asymptotic standard error was 0.29. This may have been due to the small sample size ($n = 40$) of families in which both members had dementia.

5. Concluding Remarks

Our methodology estimates the CCSHR within a simple regression framework, providing a direct test of whether there is any specificity of the association to time or causes based on the corresponding coefficients in the regression model. The pseudo-likelihood approach to estimation readily and advantageously handles censoring. The proposed estimator benefits from simplicity relative to alternative bivariate survival function plug-in estimators. Although our methodology employs a parametric regression model for the CCSHR, strong modeling assumptions can be avoided by using sufficiently flexible functions of time, such as polynomial functions or regression splines. We have conducted additional simulations, in which data from Frank family (Fan et al., 2000) were considered to check the adequacy of the proposed parametric functions of time. It is shown that the estimated results by using cubic polynomial models agree well with the true time-dependent association, suggesting that cubic polynomial models can yield satisfactory estimates for complex time-dependent association. See the Supplementary Materials for the details of the additional simulations.

Following Oakes (1986, 1989) and Bandeen-Roche and Liang (2002), the relationship between the CCSHR and the joint survival function of failure times and disease types can be described by a second-order partial differential equation,

$$\theta_{cs}(x_1, x_2, k_1, k_2) = \frac{f(x_1, x_2, k_1, k_2)S(x_1, x_2)}{\int_{x_2}^{\infty} \sum_{k=1}^K f(x_1, x, k_1, k) dx \int_{x_1}^{\infty} \sum_{k=1}^K f(x, x_2, k, k_2) dx}, \quad (12)$$

where $f(x_1, x_2, k_1, k_2) = \lim_{(\Delta x_1, \Delta x_2) \downarrow 0} \Pr(X_1 \in x_1 + \Delta x_1, X_2 \in x_2 + \Delta x_2, K_1 = k_1, K_2 = k_2) / (\Delta x_1 \Delta x_2)$, and $S(x_1, x_2) = \int_{x_1}^{\infty} \int_{x_2}^{\infty} \sum_{k=1}^K \sum_{k'=1}^K f(s, t, k, k') dt ds$. Without competing risks, the CCSHR will reduce to the CHR, and the constant CHR implies that the partial differential equation has a unique solution for the joint distribution, which belongs to

Clayton copula models. Furthermore, Nan et al. (2006) indicates that if the CHR is a piecewise constant on a grid of sample space, the above partial differential equation is also solvable. However, for bivariate data with competing risks, we can not have analogous results derived from equation (12).

An alternative approach is a composite likelihood approach (Shih and Albert, 2010), in which a two-stage estimation procedure is adopted. This approach can be used to estimate the CCSHR by the fact that the CCSHR is a function of the overall conditional hazard ratio for the bivariate first event times and the odds ratio for two competing causes. One assumption required in the composite likelihood approach is the bivariate distribution assumption for the bivariate first event times, for which appropriate model checking techniques are lacking. For example, Shih and Albert (2010) assumed a joint model in which the overall CHR for the bivariate first event times is a piecewise constant. In contrast, the proposed pseudo-likelihood approach only depends on the model assumptions of the CCSHR and provides a robust way to model the association without specifying the joint survival distributions. On the other hand, when the true association is a piecewise constant, the proposed pseudo-likelihood approach is likely to be inefficient compared with the composite likelihood approach, since the pseudo-likelihood approach does not fully use the distribution information.

The dependency structure of bivariate competing risk data may depend not only on time and causes, but also on other covariates. Extending the proposed approach to accommodate covariate adjustments is nontrivial, and requires modifications including additional model specifications on the conditional hazard functions. The methods of incorporating covariate into the CCSHR are worthy of future research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Jing Ning's research was partially supported by a start-up fund from the University of Texas MD Anderson Cancer Center. The authors are grateful to Dr. Peter Zandi for providing the Cache County data and for the support of the National Institutes of Health.

References

- Andersen PK, Ekstrom CT, Klein JP, Shu Y, Zhang MJ. A class of goodness of fit tests for a copula based on bivariate right-censored data. *Biometrical Journal*. 2005; 47:815–824. [PubMed: 16450854]
- Bandeen-Roche K, Liang KY. Modelling multivariate failure time associations in the presence of a competing risk. *Biometrika*. 2002; 89:299–314.
- Bandeen-Roche K, Ning J. Nonparametric estimation of bivariate failure time association in the presence of a competing risk. *Biometrika*. 2008; 95:221–232. [PubMed: 20305739]
- Breitner JC, Wyse BW, Anthony JC, Welsh-Bohmer KA, Steffens DC, Norton MC, Tschanz JT, Plassman BL, Meyer MR, Skoog I, Khachaturian A. Apoe-epsilon4 count predicts age when prevalence of ad increases, then declines: The cache county study. *Neurology*. 1999; 53:321–331. [PubMed: 10430421]
- Chen MC, Bandeen-Roche K. A diagnostic for association in bivariate survival models. *Lifetime Data Analysis*. 2005; 11:245–264. [PubMed: 15940823]
- Chen X, Fan Y. A model selection test for bivariate failure time data. *Econometric Theory*. 2007; 23:414–439.

- Cheng Y, Fine JP. Nonparametric estimation of cause-specific cross hazard ratio with bivariate competing risks data. *Biometrika*. 2008; 95:233–240.
- 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; 65:141–151.
- Clayton DG, Cuzick J. Multivariate generalizations of the proportional hazards model. *Journal of the Royal Statistical Society, Series A*. 1985; 148:82–108.
- Fan J, Hsu L, Prentice RL. Dependence estimation over a finite bivariate failure time region. *Lifetime Data Anal*. 2000; 6:343–355. [PubMed: 11190604]
- Genest C, Mackay J. Copules archimediennes et familles de lois bidimensionnelles dont les marges sont donnees. *The Canadian Journal of Statistics*. 1986; 14:145–159.
- Genest C, Rivest LP. Statistical inference procedures for bivariate archimedean copulas. *Journal of the American Statistical Association*. 1993; 88:1034–1043.
- Hendrie HC. Epidemiology of dementia and alzheimers disease. *Am. J. Geriatr. Psychol*. 1998; 6:3–18.
- Hsu L, Prentice RL. On assessing the strength of dependency between failure time variates. *Biometrika*. 1996; 83:491–506.
- Hu T, Nan B, Lin X, Robins J. Time dependent cross-ratio estimation for bivariate failure times. *Biometrika*. 2011; 98:341–354. [PubMed: 22822258]
- Nan B, Lin X, Lisabeth LD, Harlow SD. Piecewise constant cross-ratio estimation for association of age at a marker event and age at menopause. *Journal of the American Statistical Association*. 2006; 101:65–77.
- Oakes D. A model for association in bivariate survival data. *Journal of the Royal Statistical Society, Series B*. 1982; 44:414–422.
- Oakes D. Semiparametric inference in a model for association in bivariate survival data. *Biometrika*. 1986; 73:353–361.
- Oakes D. Bivariate survival models induced by frailties. *Journal of the American Statistical Association*. 1989; 84:487–493.
- Prentice RL, Cai J. Covariance and survivor function estimation using censored multivariate failure time data. *Biometrika*. 1992; 79:495–512.
- Shih JH, Albert PS. Modeling familial association of ages at onset of disease in the presence of competing risk. *Biometrics*. 2010; 66:1012–1023. [PubMed: 20002400]
- Shih JH, Louis TA. Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*. 1995; 51:1584–1599.
- Skytthe A, Pedersen N, Kaprio J, Stazi M, Hjelmborg J, Iachine I, Vaupel J, Christensen K. Longevity studies in genomeutwin. *Twin Res*. 2003; 6(5):448–455. [PubMed: 14624729]
- Struwing JP, Hartge P, Wacholder S, Baker SM, Berlin M, McAdams M, Timmerman MM, Lawrence BC, Tucker MA. The risk of cancer associated with specific mutations of brca1 and brca2 among ashkenazi jews. *The New England Journal of Medicine*. 1997; 336:1401–1408. [PubMed: 9145676]
- Vanobbergen J, Martens L, Lesaffre E, Declerck D. The signal-tandmobiel project, a longitudinal intervention heath promotion study in flanders (belgium): baseline and first year results. *European Journal of Paediatrics and Dentistry*. 2000; 2:87–96.
- Viswanathan B, Manatunga AK. Diagnostic plots for assessing the frailty distribution in multivariate survival data. *Lifetime Data Analysis*. 2001; 7:143–155. [PubMed: 11458654]
- Wang W, Wells MT. Estimation of kendall's tau under censoring. *Statistica Sinica*. 2000; 10:1199–1215.

Table 1

Simulation Study: Estimation of Coefficients in the Absence of Competing Risks

	n	Censoring %	Mean	SD_A	SD_M	95% CP
Independent Model: $(\alpha_0, \alpha_1, \alpha_2) = (0, 0, 0)$						
No censoring	100	0	(.00, .00, .00)	(.23, .16, .16)	(.20, .15, .15)	(.97, .95, .96)
	500	0	(.00, .00, .00)	(.09, .06, .06)	(.08, .05, .05)	(.97, .96, .96)
C ~ U(0,3)	100	32%	(-.05, .05, .06)	(.33, .33, .35)	(.33, .37, .36)	(.95, .91, .93)
	500	32%	(-.01, .01, .02)	(.13, .13, .12)	(.13, .14, .14)	(.94, .92, .93)
C ~ exp(1)	100	50%	(-.04, .07, .09)	(.45, .63, .61)	(.43, .71, .68)	(.94, .87, .89)
	500	50%	(-.02, -.02, -.05)	(.16, .19, .19)	(.17, .21, .21)	(.94, .93, .93)
Gamma Model: $(\alpha_0, \alpha_1, \alpha_2) = (0.69, 0, 0)$						
No censoring	100	0	(.71, .00, .00)	(.16, .01, .03)	(.16, .03, .03)	(.93, .91, .89)
	500	0	(.69, .00, .00)	(.07, .01, .01)	(.06, .01, .01)	(.98, .91, .93)
C ~ U(0,3)	100	46%	(.67, .07, .04)	(.34, .44, .46)	(.33, .43, .41)	(.94, .89, .91)
	500	46%	(.69, .01, .01)	(.13, .15, .15)	(.14, .16, .17)	(.94, .92, .91)
C ~ exp(1)	100	59%	(.58, .19, .16)	(.44, .82, .83)	(.45, .84, .82)	(.91, .85, .85)
	500	59%	(.69, .04, .00)	(.16, .22, .22)	(.17, .25, .26)	(.94, .91, .92)

Note: SD_A is the square root of the average of asymptotic variance estimates over 500 runs; SD_M is the empirical standard deviation of estimates over 500 runs; and 95% CP is the empirical coverage rates of 95% confidence intervals over 500 runs.

Table 2
Simulation Study: Estimation of Cause-Specific Hazards Ratio in the Gamma Frailty Model

	$\theta_{cs}(1,1)$	Sample Size	Mean	SD_A	SD_M	95% CP
Gamma Frailty Model with Positive Association						
R=0.8	2.25	100	2.31	0.37	0.39	0.93
		500	2.26	0.17	0.15	0.96
		1000	2.25	0.12	0.11	0.96
R=0.5	3	100	3.17	0.68	0.69	0.94
		500	3.04	0.30	0.28	0.96
		1000	3.00	0.21	0.20	0.96
R=0.2	6	100	6.82	2.68	2.81	0.92
		500	6.11	1.02	0.90	0.97
		1000	6.03	0.72	0.65	0.97
Gamma Frailty Model with Negative Association						
R=0.8	0.66	100	0.69	0.09	0.08	0.98
		500	0.67	0.04	0.03	0.97
		1000	0.66	0.02	0.02	0.96
R=0.5	0.89	100	0.92	0.20	0.18	0.96
		500	0.90	0.08	0.08	0.96
		1000	0.89	0.05	0.05	0.96
R=0.2	1.80	100	1.89	0.85	0.86	0.93
		500	1.80	0.32	0.28	0.96
		1000	1.80	0.20	0.19	0.96

Note: SD_A is the square root of the average of asymptotic variance estimates over 500 runs; SD_M is the empirical standard deviation of estimates over 500 runs; and 95% CP is the empirical coverage rates of 95% confidence intervals over 500 runs.

Table 3
Simulation Study: Comparison of Pseudo-Likelihood and Moment-Based Methods

Censoring Distribution	Sample Size	Censoring %	Pseudo-likelihood		Moment-based EE	
			Estimator	SD	Estimator	SD
Independent Model: $\theta_{cs}(x_1, x_2, 1, 1) = 1$						
No censoring	100	0	1.02	0.23	1.03	0.28
	500	0	1.01	0.09	1.03	0.12
$C \sim U(0,3)$	100	32%	1.04	0.33	1.06	0.41
	500	32%	1.00	0.10	1.01	0.13
$C \sim \exp(1)$	100	50%	1.09	0.53	1.10	0.66
	500	50%	1.00	0.16	1.01	0.26
Gamma Model: $\theta_{cs}(x_1, x_2, 1, 1) = 2$						
No censoring	100	0	2.05	0.31	2.04	0.35
	500	0	2.01	0.11	2.01	0.16
$C \sim U(0,3)$	100	46%	2.02	0.51	2.04	0.56
	500	46%	2.01	0.28	2.02	0.31
$C \sim \exp(1)$	100	59%	2.04	0.83	2.10	0.97
	500	59%	2.02	0.32	2.05	0.41

Note: Estimator is the mean of estimates over 500 runs; SD is the empirical standard deviation of estimates over 500 runs; SD_B is the mean of Bootstrap standard deviation over 500 runs.

Table 4

Simulation Study: Comparison of the Regression Model Using the Pseudo-Likelihood Method and U-Statistic Method

Censoring	n	Pseudo-likelihood method				U-statistic method			
		Upper left	Upper right	Lower left	Lower right	Upper left	Upper right	Lower left	Lower right
		Est(SD)	Est(SD)	Est(SD)	Est(SD)	Est(SD)	Est(SD)	Est(SD)	Est(SD)
Independent Model: $\theta_{cs}(B_i, 1, 1) = 1, i = 1, 2, 3, 4$									
No Censoring	500	1.01 (.09)	1.01 (.10)	1.01 (.09)	1.01 (.10)	1.01 (.10)	1.01 (.10)	1.01 (.11)	0.99 (.13)
	1000	1.00 (.07)	1.01 (.07)	1.00 (.06)	1.00 (.06)	0.99 (.07)	1.01 (.07)	0.99 (.07)	0.99 (.08)
C ~ exp(1)	500	1.00 (.11)	1.01 (.15)	1.02 (.19)	1.03 (.19)	1.01 (.12)	1.00 (.16)	1.01 (.15)	1.01 (.16)
	1000	1.01 (.15)	1.01 (.14)	1.00 (.13)	1.02 (.14)	1.01 (.16)	1.00 (.15)	0.99 (.15)	0.99 (.16)
Gamma Model: $\theta_{cs}(B_i, 1, 1) = 2, i = 1, 2, 3, 4$									
No Censoring	500	2.00 (.19)	2.04 (.26)	2.03 (.23)	2.00 (.18)	1.99 (.19)	2.02 (.26)	2.01 (.25)	1.96 (.19)
	1000	2.01 (.14)	2.01 (.16)	2.02 (.15)	2.01 (.11)	2.01 (.14)	2.01 (.18)	2.01 (.16)	2.00 (.12)
C ~ exp(1)	500	2.05 (.34)	2.11 (.48)	2.05 (.46)	2.10 (.53)	2.03 (.32)	2.09 (.51)	2.02 (.48)	2.04 (.59)
	1000	2.03 (.23)	2.11 (.30)	2.03 (.32)	2.04 (.40)	2.01 (.23)	2.10 (.33)	2.03 (.34)	2.02 (.46)
Positive Stable Model: $\theta_{cs}(B_i, 1, 1) = 1.77, \theta_{cs}(B_i, 1, 1) = 1.20, i = 2, 3, \theta_{cs}(B_i, 1, 1) = 1.13$									
No Censoring	500	1.66 (.15)	1.17 (.13)	1.19 (.13)	1.11 (.10)	1.78 (.16)	1.20 (.14)	1.21 (.15)	1.12 (.12)
	1000	1.66 (.10)	1.18 (.08)	1.18 (.08)	1.10 (.07)	1.78 (.11)	1.21 (.10)	1.21 (.10)	1.12 (.09)
C ~ exp(1)	500	1.83 (.23)	1.28 (.31)	1.26 (.33)	1.29 (.56)	2.19 (.30)	1.29 (.37)	1.28 (.39)	1.19 (.63)
	1000	1.80 (.15)	1.24 (.22)	1.22 (.21)	1.18 (.36)	2.17 (.20)	1.27 (.25)	1.23 (.24)	1.15 (.42)

Note: Mean is the mean of estimates over 500 runs; SD is the empirical standard deviation of estimates over 500 runs. Upper left, upper right, lower left, and lower right are the upper left cell, the upper right cell, the lower left cell, and the lower right cell within the 4-cell grid, respectively.

Table 5

Analysis of Cache County Data to the Replicate Model of Bandeen-Roche and Ning (2008) Using Regression Equation (10)

Parameter	Model 1 Est. (SE)	Model 2 Est. (SE)	Model 3 Est. (SE)
Onset timing			
α_0 : (early, early)	0.36 (0.05)	0.36 (0.05)	0.36 (0.04)
α_1 : (late, early) vs. (early, early)	0.02 (0.07)	0.03 (0.07)	0.03 (0.06)
α_2 : (early, late) vs. (early, early)	-0.09 (0.06)	-0.08 (0.07)	-0.09 (0.05)
α_3 : (late, late) vs. (early, early)	0.10 (0.08)	0.09 (0.07)	0.10 (0.07)
Causes			
α_4 : (death, dementia) vs. reference	-0.26 (0.08)	-0.26 (0.08)	-0.26 (0.08)
α_5 : (dementia, death) vs. reference	-0.07 (0.07)	-0.06 (0.07)	-0.06 (0.05)
α_6 : (dementia, dementia) vs. reference	0.82 (0.11)	0.97 (0.19)	0.94 (0.10)
Onset timing interactions with causes = (dementia, dementia)			
α_7 : onset timing = (late, early)		-1.28 (0.53)	-1.26 (0.27)
α_8 : onset timing = (early, late)		-0.22 (0.25)	
α_9 : onset timing = (late, late)		0.21 (0.22)	

Note: Pairings are ordered as (child, mother). The reference pairing of failure causes is (death, death). Model fitting was done with the pseudo-maximum likelihood method. Est. is the parameter estimate; SE is the estimated asymptotic standard error of the parameter estimator.

Table 6

Analysis of Cache County Data Using Regression Equation (11)

Parameter	Model 4 Est. (SE)	Model 5 Est. (SE)	Model 6 Est. (SE)
Onset timing			
α_0 : (early, early)	0.45 (0.43)	0.20 (0.28)	0.10 (0.25)
α_1 : (late, early) vs. (early, early)	0.18 (0.10)	0.13 (0.07)	0.11 (0.06)
α_2 : (early, late) vs. (early, early)	-0.10 (0.08)	-0.10 (0.08)	-0.14 (0.06)
α_3 : (late, late) vs. (early, early)	0.10 (0.11)	0.04 (0.08)	
$10 \times \alpha_{81}$: child onset time	0.03 (0.04)	0.03 (0.04)	0.04 (0.03)
$10 \times \alpha_{82}$: maternal onset times	-0.04 (0.04)		
$10 \times \alpha_{10}$: difference between two onset times	-0.06 (0.03)	-0.07 (0.03)	-0.06 (0.03)
Causes			
α_4 : (death, dementia) vs. reference	-0.25 (0.08)	-0.25 (0.08)	-0.25 (0.08)
α_5 : (dementia, death) vs. reference	-0.08 (0.05)	-0.08 (0.05)	
α_6 : (dementia, dementia) vs. reference	2.08 (1.09)	2.80 (0.79)	2.74 (0.79)
Onset timing interactions with causes = (dementia, dementia)			
α_7 : onset timing = (late, early)	-1.39 (0.29)	-1.47 (0.27)	-1.45 (0.27)
$10 \times \alpha_{91}$: child onset time	-0.20 (0.11)	-0.24 (0.10)	-0.23 (0.10)
$10 \times \alpha_{92}$: maternal onset times	0.06 (0.11)		
$10 \times \alpha_{11}$: difference between two onset times	-0.13 (0.12)		

Note: Pairings are ordered as (child, mother). The reference pairing of failure causes is (death, death). Model fitting was done with the pseudo-maximum likelihood method. Est. is the parameter estimate; SE is the estimated asymptotic standard error of the parameter estimator.