11-17-2005

NONPARAMETRIC ESTIMATION OF BIVARIATE FAILURE TIME ASSOCIATIONS IN THE PRESENCE OF A COMPETING RISK

Karen Bandeen-Roche
Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, kbandeen@jhsph.edu

Jing Ning
Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, njstat@gmail.com

NONPARAMETRIC ESTIMATION OF BIVARIATE FAILURE TIME ASSOCIATIONS IN THE PRESENCE OF A COMPETING RISK

BY KAREN BANDEEN-ROCHE, JING NING

Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health,

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

e-mail: kbandeen@jhsph.edu jning@jhsph.edu

SUMMARY

There has been much research on the study of associations among paired failure times. Most has either assumed time invariance of association or been based on complex measures or estimators. Little has accommodated failures arising amid competing risks. This paper targets the conditional cause specific hazard ratio, a recent modification of the conditional hazard ratio to accommodate competing risks data. Estimation is accomplished by an intuitive, nonparametric method that localizes Kendall's tau. Time variance is accommodated through a partitioning of space into "bins" between which the strength of association may differ. Inferential procedures are researched, small sample performance evaluated, and methods applied to investigate familial association in dementia onset. The proposed methodology augments existing methodology with an approach that may be more readily applied and interpreted, thus facilitate dissemination of methodology addressing failure time associations into the substantive literature.

Some key words: Cause-specific; Kendall's tau; Multivariate; Paired; Survival; U-statistic

1. INTRODUCTION

Methodology to analyze correlated failure time data has potentially wide import for biomedical research. With the proliferation of genetics studies and outcomes research, studies must account for time-to-event clustering within families or providers of care. Examples abound where health is quantified as multiple occurrences per individual, be it recurrent events such as serial falls (Stel et al., 2003), or times to "comorbid" disease onset (Camp et al., 2005), or repeated assessments of episode duration (e.g. wakefulness; Punjabi et al., 1999). When statistical analyses involve such data, there must be accounting of failure time correlations to ensure correctness of inferences, at the least. Further, strength of dependence among related failure times may be of scientific interest. This paper concerns this latter case. We both propose methodology to estimate strength of failure time dependence and apply it to estimate familial association in ages of dementia onset.

There has been considerable work on the assessment of failure time associations. Among the earliest-proposed measures was the cross-, or conditional hazard, ratio (Clayton, 1978; Clayton & Cuzick, 1985; Oakes, 1982; 1986). Clayton's (1978) measure provides a single, time-invariant summary of dependence. The cross-ratio function has a parametric representation with a direct link to two well discussed families for modeling multivariate failure times, parametric copula (Genest & MacKay, 1986) and frailty (Oakes, 1989) models. Two primary approaches have been proposed for the estimation of failure time associations using these models: full or approximate maximum likelihood (Nielsen et al., 1992; Ripatti et al., 2002; Ripatti & Palmgren 2000), and two-stage, "pseudo"-maximum likelihood (Genest et al, 1995; Shih & Louis, 1995; Glidden, 2000).

A number of nonparametric association measures have also been proposed, including the general formulation of the conditional hazard ratio function. There have also been two primary approaches to estimation in this case. A first plugs into the association measure a nonparametric estimator of the multivariate survival (i.e. Dabrowska, 1988; Prentice et. al, 2004) or cumulative hazard function (Prentice & Cai, 1992; Hsu & Prentice, 1996; Fan et al., 2000; Wang & Wells, 2000). A second employs one-dimensional empirical processes whose expectations relate conveniently to the measure of interest, affording method-of-moments estimation (Oakes, 1982; Oakes, 1989; Genest & Rivest, 1993; Barbe et al., 1996 Viswanathan & Manatunga, 2001; Chen & Bandeen-Roche, 2005). Regression models relating such association measures to covariates have also been proposed (e.g. Prentice & Hsu 1997; Fine & Jiang, 2000).

This considerable body of research notwithstanding, measures of failure time associations have been slow to find utilization in biomedical studies. A Web of Science search carried out on June 6, 2005 identified the vast majority of citations to articles just referenced to be by quantitative methodology articles, with scarcely any excepting in review articles appearing in the biomedical literature. Among potential explanations, two are relevant to the present work. First, the complexity of estimation involved for most existing approaches, and in some cases, of interpretation, may be off-putting. Second, little of the existing work accounts explicitly for competing or semi-competing risks. Yet, these are unavoidable in applications involving conditions that may lead to death or affect only a fraction of individuals within their lifetimes. Bandeen-Roche and Liang (2002) studied the estimation of failure time associations accounting for competing risks; at that paper's

completion, no other papers on the topic could be found. We are aware of one subsequent paper, whose measures of association are based on bivariate cause-specific hazard and cumulative-incidence functions and are of a combined empirical process, survival function estimator plug-in type (J. Fine, personal communication, December 16, 2004).

Here we aim to progress toward filling the gap we have just argued, by developing inference procedures for a simple, nonparametric estimator of an easily interpreted measure of bivariate failure time association, the conditional cause-specific hazard ratio (CCSHR, Bandeen-Roche & Liang, 2002). Ignoring censoring for the present, let X_1,X_2 be failure times to be observed for two family members; K_1,K_2, the respective causes of failure, with $k_j = 1$ indicating dementia onset and $k_j = 2$ indicating death before dementia; and λ_k, the hazard function for failure specific to cause k. The CCSHR defines the multiplicative increase in risk of dementia onset for family members whose relatives are diagnosed as cases at, say, age x_1 versus those whose relatives survive without disease beyond that age:

$$\lambda_{X_2,1}(x_2|X_1=x_1,K_1=1)/\lambda_{X_2,1}(x_2|X_1>x_1). \quad \textbf{(1)}$$

Our 2002 paper was primarily focused on a parametric, copula-based formulation of this quantity, whose estimation proved highly sensitive to modeling assumptions. In contrast, this paper studies estimation by a localized version of Kendall's tau to which we have made previous allusion (Bandeen-Roche & Liang 1996, 2002) and studied in a paper not focused on competing risks (Chen & Bandeen-Roche, 2005). Its idea dates to the seminal papers on the cross-ratio function and has been prominent in the unidimensional empirical process-type association measures identified above. However, to our knowledge,

asymptotic inference has only been developed for versions of the estimator that are global and ignore competing risks, which simply involve the standard Kendall's tau (e.g., Kendall, 1948, p. 67; Oakes 1982) or a weighted version thereof (Oakes, 1986). Here, in contrast, we localize to time as well as causes, using an easily applicable procedure. We develop inferences for the resulting estimator; evaluate small sample performance in a simulation study; and apply our methodology to analyze familial data on dementia from the Cache County Study (Breitner et al., 1999). Inference does not follow from existing theory on Kendall's tau (e.g., Shieh, 1998), because our localization procedure weights concordances according to observed failure times and causes, hence the weights and data defining concordance may be stochastically dependent. Rather, we obtain inference directly through representation of our estimator as a U-statistic. As a by-product, we gain insight into the convergence behavior of time-invariant estimators of the cross-ratio when in fact the ratio is time-varying, as well as distributional and operational features that affect precision.

We now define our association measure, describe estimation and develop associated inferences. Section 3 reports on our simulation study. Section 4 details application of our methodology. Section 5 briefly addresses bin choice and study design. We conclude with discussion.

2. METHODS

2.1. Notation and Estimand

We first formally introduce the CCSHR. Suppose there are competing events 1,...,C such that interest is in the time, X^*, to the first of the events, and $K^* \in \{1,...,C\}$, a

code identifying the first-occurring event. We consider situations where the observable data for an individual are: X, the minimum of X^* and the time at which there is censoring of X^* for non-competing reasons; and K, a code equaling 0 if failure is censored altogether and K^* if the earliest competing event occurs prior to censoring. Then the individual's cause-specific hazard for the occurrence of each kth event is:

$$\lambda_k(x)=\lim_{\Delta x\downarrow 0} \Pr(x\le X\le x+\Delta x, K=k|X\ge x)/\Delta x \qquad (2)$$

(Prentice et al., 1978; Benichou & Gail, 1990).

With correlated failure processes, observable data are times $(X_{i1},...,X_{im_i})$ and associated causes $(K_{i1},...,K_{im_i})$ jointly sampled in 'clusters,' i=1,...,n. Specifically X_{ij} is the time of the earliest event (including censoring) occurring for member j of cluster i, and K_{ij} codes the event that occurs. As the CCSHR is bivariate, we henceforth assume m_i=2 and $(X_{i1},X_{i2},K_{i1},K_{i2})$ as independently and identically distributed so that cause-specific densities

$$f(x,k)=\lim_{(\Delta x_1,\Delta x_2)\downarrow 0} \Pr(x_1\le X_1\le x_1+\Delta x_1, x_2\le X_2\le x_2+\Delta x_2, K_1=k_1, K_2=k_2)/(\prod_{m=1}^{2}\Delta x_m) \qquad (3)$$

exist for each combination of failure causes $k = (k_1,k_2)$. Then, (X_1,X_2), has an absolutely continuous joint survival function $S(x_1,x_2) = \int_{x_1}^{\infty}\int_{x_2}^{\infty}\sum_{k_1=1}^{K}\sum_{k_2=2}^{K} f(x,y,k_1,k_2)dydx$. Here, subscripted variables denote scalars, and unsubscripted, vectors, e.g. $S(x_1,x_2) = S(x)$.

Employing the quantities defined in (3) ff, Bandeen-Roche and Liang (2002)

defined the CCSHR as in (1), that is:

$$\theta_{CS}(x;k_1,k_2)=(1)=\frac{\lambda_{1,k_1}(x_1|X_2=x_2,K_2=k_2)}{\lambda_{1,k_1}(x_1|X_2>x_2)}=\left\{\frac{f(x_1,x_2,k_1,k_2)}{\int_{x_1}^{\infty}\sum_{k=1}^{K}f(x,x_2,k,k_2)dx}\right\}\left\{\frac{S(x_1,x_2)}{\int_{x_2}^{\infty}\sum_{k=1}^{K}f(x_1,x,k_1,k)dx}\right\}. \quad \textbf{(4)}$$

2·2. Estimator

While (4) is a recently proposed measure, a transformed, localized Kendall's tau serves to estimate it nonparametrically. In brief, it can be shown that $\theta_{CS}(x;k)$ 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 $(K_1^{(ab)},K_2^{(ab)})$:

$$\theta_{CS}(x_1,x_2;k_1,k_2)=\frac{\text{pr}\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})>0 \mid (X_1^{(ab)},X_2^{(ab)})=(x_1,x_2),(K_1^{(ab)},K_2^{(ab)})=(k_1,k_2)\}}{\text{pr}\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})<0 \mid (X_1^{(ab)},X_2^{(ab)})=(x_1,x_2),(K_1^{(ab)},K_2^{(ab)})=(k_1,k_2)\}},$$

where $(X_1^{(a)},X_2^{(a)})$ and $(X_1^{(b)},X_2^{(b)})$ are two independently drawn failure time pairs, $(X_1^{(ab)},X_2^{(ab)})$ are the componentwise minima $(X_1^{(a)}\wedge X_1^{(b)},X_2^{(a)}\wedge X_2^{(b)})$, and $(K_1^{(ab)},K_2^{(ab)})$ are the causes corresponding to $(X_1^{(ab)},X_2^{(ab)})$. Thus, a simple estimator determines the concordance status for every two pairs with $(K_1^{(ab)},K_2^{(ab)})=(k_1,k_2)$ and then divides the number of concordances by the number of discordances. Here we must be mindful that $\theta_{CS}(x;k)$ is potentially a continuous function of (x_1,x_2) on $\{x: x>0\}$ (henceforth, $\mathbb{R}^{2+}$). If so, samples from a continuous-time failure time process will yield at most one pairing of pairs with $(X_1^{(ab)},X_2^{(ab)})$ equal to any given (x_1,x_2). To obtain stable ratios of concordance and discordance counts, then, one must bin or smooth the counts and/or ratios.

Here we propose to bin in two dimensional space. Let $\mathcal{B}=\{B_1,...,B_J\}$ be a partition

of $\mathbb{R}^{2+}$, with $\mathscr{B}$ established a priori and J finite. Our estimator is

$$\theta_{CS}(x_1,x_2;k_1,k_2) = \frac{\sum_{a<b} \mathbb{I}\{K^{(ab)}=(k_1,k_2)\}\mathbb{I}\{X^{(ab)}\in B_{j(x_1,x_2)}\}\mathbb{I}\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})>0\}}{\sum_{a<b} \mathbb{I}\{K^{(ab)}=(k_1,k_2)\}\mathbb{I}\{X^{(ab)}\in B_{j(x_1,x_2)}\}\mathbb{I}\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})<0\}}. \quad \textbf{(5)}$$

Here, $\mathbb{I}$ is the standard indicator function, $\mathbb{I}\{A\}=1$ if A is true and 0 otherwise, and $j(x_1,x_2)$ indexes the cell of the partition that includes (x_1,x_2). This is the same estimator employed by Bandeen-Roche & Liang (2002) and, ignoring causes, Chen & Bandeen-Roche (2005), but with $\mathscr{B}$ defined on $\mathbb{R}^{2+}$ rather than $\{S(x), x \in \mathbb{R}^{2+}\}$.

2.3. Distributional Properties

At each (x_1,x_2,k_1,k_2) the numerator of (5) is a U-statistic with kernel

$$h_{1(x1,x2,k1,k2)}\{(x^{(a)},k^{(a)}),(x^{(b)},k^{(b)})\} = \mathbb{I}\{k^{(ab)}=(k_1,k_2)\}\mathbb{I}\{x^{(ab)}\in B_{j(x_1,x_2)}\}\mathbb{I}\{(x_1^{(a)}-x_1^{(b)})(x_2^{(a)}-x_2^{(b)})>0\},$$

and similarly for the denominator with kernel we label $h_{2(x1,x2,k1,k2)}\{(x^{(a)},k^{(a)}),(x^{(b)},k^{(b)})\}$. Thus, inferences follow directly from U-statistic theory (e.g. Serfling, 1980). Beginning with point convergence: replacing sums by averages, (5) converges almost surely to

$$\theta^*_{CS}(x_1,x_2,k_1,k_2) = \frac{pr\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})>0 | X^{(ab)}\in B_{j(x_1,x_2)},\ K^{(ab)}=(k_1,k_2)\}}{pr\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})<0 | X^{(ab)}\in B_{j(x_1,x_2)},\ K^{(ab)}=(k_1,k_2)\}} \quad \textbf{(6)}$$

under weak conditions, so long as the denominator exceeds 0. More interesting is to consider interpretation if, further, $pr\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})<0, | X^{(ab)}=y,\ K^{(ab)}=(k_1,k_2)\}$ is

bounded above 0 almost everywhere y on $B_{j(x_1,x_2)}$. Then, (6) equals

$$E\{\theta_{CS}(X_1^{(ab)},X_2^{(ab)};k_1,k_2)|X^{(ab)}\in B_j,(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})<0,\ K^{(ab)}=(k_1,k_2)\} \qquad (7)$$

$= \theta^*_{CS}(j;k_1,k_2)$ (Appendix 1), where we now suppress the (x_1,x_2) notation indexing bins, retaining only the subscript j. If the CCSHR is constant over B_j, (7) equals that constant value; otherwise, it is the expectation over potential time realizations within B_j, weighted with respect to probabilities of discordance in pairs (a) and (b). Interestingly, then, the average conservatively up-weights regions of less strongly positive association, thus dampens the magnitude of association relative to a straight expectation over (x_1,x_2).

We proceed to derive asymptotic distributions of the dividends that define our estimator. As a first step, it is useful to write the respective means in a different format than given preceding (6). We begin with the concordance (numerator) term. Note that the compound event $\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})>0,\ X^{(ab)}\in B_j,\ K^{(ab)}=(k_1,k_2)\}$ occurs if and only if

$\{(X^{(a)}<X^{(b)},\ X^{(a)}\in B_j,\ K^{(a)}=(k_1,k_2))\} \cup \{(X^{(a)}>X^{(b)},\ X^{(b)}\in B_j,K^{(b)}=(k_1,k_2))\}$ occurs. Then,

$$\begin{aligned}
&E[\mathbb{I}\{k^{(ab)}=(k_1,k_2)\}\mathbb{I}\{X^{(ab)}\in B_j\}\mathbb{I}\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})>0\}]\\
&= E_{(a)}E_{(b)|(a)}[\mathbb{I}\{X^{(a)}<X^{(b)},\ X^{(a)}\in B_j,\ K^{(a)}=(k_1,k_2)\}\\
&\qquad +\mathbb{I}\{X^{(a)}>X^{(b)},\ X^{(b)}\in B_j,\ K^{(b)}=(k_1,k_2)\}]\\
&= E_{(a)}[\mathbb{I}\{X^{(a)}\in B_j,\ K^{(a)}=(k_1,k_2)\}S(X^{(a)})]\\
&\qquad +E_{(b)}E_{(a)|(b)}[\mathbb{I}\{X^{(a)}>X^{(b)},\ X^{(b)}\in B_j,\ K^{(b)}=(k_1,k_2)\}]\\
&= 2E_{(a)}[\mathbb{I}\{X^{(a)}\in B_j,\ K^{(a)}=(k_1,k_2)\}S(X^{(a)})]
\end{aligned} \qquad (8)$$

where $E_{(b)|(a)}$ denotes expectation with respect to $(X^{(b)},K^{(b)})$ conditioning on $(X^{(a)},K^{(a)})$, etc.

Henceforth we denote this expression as $\bar{E}_C$.

Proceeding, the concordance term variance depends on the quantities

$$\zeta_{1C} = Var_{(a)}[E_{(b)|(a)}\{h_1(X^{(a)},X^{(b)},K^{(a)},K^{(b)})\}];\ \zeta_{2C} = Var\{h_1(X^{(a)},X^{(b)},K^{(a)},K^{(b)})\}.$$

Term $h_1(X^{(a)},X^{(b)},K^{(a)},K^{(b)})$ is an indicator function with mean $\bar{E}_C$, thus has variance $\zeta_{2C} = \bar{E}_C(1-\bar{E}_C)$. Term ζ_{1C} follows from line 2 of (8), replacing $E_{(a)}$ with $Var_{(a)}$. If we define $\mu_{(b)}\{B_j \cap (0,x^{(a)});k_1,k_2\} = E_{(b)}[\mathbb{I}\{x^{(a)} > X^{(b)},\ X^{(b)} \in B_j,\ K^{(b)} = (k_1,k_2)\}]$, the probability that $X^{(b)}$ is a (k_1,k_2)-type failure occurring in the intersection of B_j and the quadrant $\{0 < x < X^{(a)}\}$,

$$\zeta_{1C} = E_{(a)}[\mathbb{I}\{X^{(a)} \in B_j,\ K^{(a)} = (k_1,k_2)\}S(X^{(a)}) + \mu_{(b)}\{B_j \cap (0,X^{(a)});k_1,k_2\} - \bar{E}_C]^2. \qquad (9)$$

Finally, the numerator variance is given by $[4(n-2)/\{n(n-1)\}]\zeta_{1C} + [2/\{n(n-1)\}]\zeta_{2C}$.

The asymptotic distribution of the numerator is normal provided $\zeta_{1C} > 0$. Trivially, then, bins and failure causes must be such that failures of type (k_1,k_2) may occur, excluding $pr\{X^{(a)} \in B_j, K^{(a)} = (k_1,k_2)\} = 0$. A more interesting case arises when $S(x)$ is restricted to one dimension such that $S(x) = 1-F(x)$. If there is only one bin (the positive real quadrant) and failure cause, then (9) evidently equals 0. However $\zeta_{1C} > 0$ if there are multiple causes or bins with well-defined measure > 0.

The denominator has variance = $[4(n-2)/\{n(n-1)\}]\zeta_{1D} + [2/\{n(n-1)\}]\zeta_{2D}$; $\zeta_{1D} = Var_{(a)}[E_{(b)|(a)}\{h_2(X^{(a)},X^{(b)},K^{(a)},X^{(b)})\}]$; $\zeta_{2D} = Var\{h_2(X^{(a)},X^{(b)},K^{(a)},X^{(b)})\}$. Elucidation of ζ_{1D} and ζ_{2D} is analogous as for the concordance terms, albeit more unwieldy. We relegate details to

Appendix 2. Finally, under conditions already noted,

$$\sqrt{n}\left\{\begin{array}{l}\binom{n}{2}^{-1}\sum_{a<b}h_1(X^{(a)},X^{(b)},K^{(a)},K^{(b)})-\bar{E}_C\\ \binom{n}{2}^{-1}\sum_{a<b}h_2(X^{(a)},X^{(b)},K^{(a)},K^{(b)})-\bar{E}_D\end{array}\right\}\xrightarrow{d} N\left\{\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}4\zeta_{1C} & 4\eta_1\\ 4\eta_1 & 4\zeta_{1D}\end{pmatrix}\right\}$$

where $\eta_1 = cov_{(a)}[E_{(b)|(a)}\{h_1(X^{(a)},X^{(b)},K^{(a)},K^{(b)})\},E_{(b)|(a)}\{h_2(X^{(a)},X^{(b)},K^{(a)},K^{(b)})\}]$ and has expansion similar to those for ζ_{1C} and ζ_{1D}.

Applying the delta method, it follows that the proposed estimator is asymptotically normal with limiting mean (7) and variance equal to

$$\frac{4}{n\bar{E}_D^2}\left(\zeta_{1C}-2\frac{\bar{E}_C\eta_1}{\bar{E}_D}+\frac{\zeta_{1D}\,\bar{E}_C^2}{\bar{E}_D^2}\right). \qquad (10)$$

2·4. Variance estimation

We hoped that equations (8)-(9) and Appendix 2 would afford a time-saving strategy for approximating our estimators' variability. Due to the complexity of the discordance-associated terms, however, they seem not to. Therefore, we merely estimate quantities defining the limiting variance of our CCSHR estimator by their sample counterparts. Estimates for $\bar{E}_C$ and $\bar{E}_D$ are given by numerator and denominator of the CCSHR calculation (5); those for ζ_{1C}, ζ_{1D}, and η_1 are calculated similarly, for instance

$$\hat{\zeta}_{1C}=\sum_{a=1}^{n}\left\{\sum_{b\neq a}h_1(X^{(a)},X^{(b)},K^{(a)},K^{(b)})/(n-1)\right\}^2/n-\bar{E}_C^2.$$

Computations involve nested sums thus are intensive at n^2 complexity, but simple in form.

3. SIMULATION STUDY

3·1. Design

We conducted studies of the small sample accuracy and precision of our estimator and associated inferences. Our design mimicked that of Bandeen-Roche and Liang (henceforth "BRL"; 2002), to afford comparison with previous findings. Each scenario we studied envisioned two failure causes, "disease" (k=1) and "death" (k=2), without censoring, and comprised 500 runs. We considered two sample sizes: n=500, 1000.

We generated bivariate data according to the frailty model for subject- and cause-specific failure hazards given by equation (6) in BRL. In brief, let $\lambda(t) = \sum_{k=1}^{2} \lambda_k(t)$ denote the overall failure hazard, and $R(t) = \lambda_1(t)/\lambda(t)=R$, the proportional contribution of the disease-specific hazard to the overall hazard. Then, the model at issue is

$$\lambda_k(x|A=a,B=b)=ab_k\lambda^*(x), \qquad (11)$$

where "A" is a scalar "size" frailty and "B" is a compositional vector "shape" frailty shared by the members of a given "familial" pair. The frailties allow heterogeneity both in overall failure propensity and proportional allocation of the overall hazard to component causes. Per equation (13) of BRL, this formulation induces a CCSHR that multiplies the standard conditional hazard ratio (CHR) for a scalar frailty model by a factor involving R. As in BRL we assumed *B* distributed as Beta with mean R and scale parameter=1 and set $\lambda^*(t)=1$.

Our design varied the size frailty distribution to be either gamma or positive stable. The former leads to a time invariant CCSHR; the latter, to one that decreases in each time

dimension. We also varied the magnitude of the CCSHR. For runs with gamma frailty, we replicated the BRL design, yielding $\theta_{CS}(x;1,1) = \theta_{CS}(1,1)$ values of 6.0, 3.0, and 2.25. Positive stable distributions have Laplace transform $exp\{-u^{\alpha}\}$. For runs with such frailties, we fixed R=.5 and varied α over values .4, .6, .8. Both the global CCSHR and the rate of CCSHR decline over time increase as α decreases. Data were generated per Appendix 2 of BRL except that in runs with positive stable frailty, we generated frailties per Lee (1979).

For each run, we applied (5) 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. Standard errors and 95% confidence intervals (CIs) were constructed per section 3.2. For each set of runs, we (i) evaluated bias *vis à vis* CCSHR values defined by the data-generating distribution; (ii) compared average of estimated CCSHR variances to the empirical variances of estimates over runs; and (iii) calculated coverage of Wald 95% CIs. Concerning (i): with gamma size frailties, the CCSHR is an easily defined constant value. However with positive stable size frailties, the CCSHR varies continuously with time, and the per-bin targets of estimation are given by equation (7). To estimate these, we generated 20,000 pairs from each distributional scenario and replaced the expectation in (7) by a sample average over the $\binom{20,000}{2}$ pairings of pairs. To assess adequacy of this sample size we generated estimates over a range of sample sizes; by n=20,000, the series of estimates had leveled to an approximate asymptote.

Finally, we conducted studies to compare the small sample accuracy and precision of our estimator and its associated inferences to that of the cumulative hazard plug-in

estimator proposed by Fan et al. (2000; henceforth, "Fan estimator"). The Fan estimator is not designed to accommodate competing risks; thus data for these studies were generated from distributions assuming a single failure cause, without censoring (setting b_k in equation 11 equal to 1). Moreover it estimates the inverse CHR over lower (t_1,t_2) quadrants; therefore we applied the inverse of our estimator and derived findings accordingly, over quadrants with $t_1=t_2$=(a) lower quartile; (b) median; and (c) upper quartile of the marginal survival function. We studied models assuming independence within pairs; gamma frailty with CHR=2; and positive stable frailty with Laplace transform parameter α=0.4, 0.8. For positive stable models, the per-quadrant targets of estimation were approximated as described in the previous paragraph, employing n=20,000 pairs; for the Fan estimator, we averaged the inverse CHR over the quadrant in question–that is, computed the empirical cumulative distribution function version of Fan et al. (2000) equation (2). There were 500 replicates per simulation run; sample sizes of n=100 and n=1000 pairs were compared.

3·2. Results

We first consider estimator performance on data generated with gamma size frailties, thus having time-invariant CCSHR (Table 1). In scenarios with n=1000 pairs, both the estimator and its associated inferences were very accurate on all time quadrants, with slight upward biases ranging from 1% to 7% for both point and standard error estimation as the underlying CCSHR ranged from 2.25 to 6.00. Simulations with the n=500 exhibited similar, moderately exacerbated patterns, with percentage biases primarily ranging between 3% and 20% as the underlying CCSHR increased. In both cases, and

particularly for n=500, estimator distributions were somewhat right-skewed. They also included outliers whose severity increased with the underlying CCSHR and were largely responsible for the associated increase in percentage biases of estimators. Coverage of 95% confidence intervals ranged within 93%-97% for all sample sizes and scenarios.

Estimator performance was even better with positive stable size frailties than their gamma counterparts (Table 2). CCSHR distributions appeared considerably more symmetric, and this was reflected advantageously both in accuracy of estimation and its robustness to size of the underlying CCSHR.

Table 3 compares performance of our estimator and the Fan estimator. For independence and gamma frailty runs, each achieved outstanding accuracy in all cases except for a 20% upward bias in the small-sample independence case; the Fan estimator exhibited modestly superior precision, to a degree increasing with the quadrant size, with empirical standard deviations 1%-20% lower than ours. In contrast, for positive stable scenarios, our estimator exhibited modestly superior accuracy and precision. Our estimator's bias was negligible relative to its target of estimation; that of the Fan estimator increased with the quadrant size, topping at 25% upward bias for α=.4 and the quadrant bounded by the upper quartiles. Empirical standard deviations for our estimator were as much as 24% lower than those for the Fan estimator, with discrepancy increasing with strength of association and quadrant size. In all, the estimators performed quite similarly.

4. APPLICATION: AGGREGATION OF DEMENTIA IN FAMILIES

There is evidence that dementia aggregates in families (Hendrie, 1998) with greater

heritability for early-onset than later-onset dementia (Silverman et al., 2005). If so, we would anticipate dementia onset ages to be associated within families, with particularly strong association in a lower left quadrant of ages. Additionally, death is a competing risk that very often predates a dementia diagnosis. Thus, analysis of the aggregation of dementia in families is well suited to illustrating the methodology we propose.

We now analyze the same data, provided by the Cache County Study on Memory in Aging (Breitner et al., 1999), as were analyzed in the BRL (2002) paper; there, readers may find a comprehensive description. In brief, the study sampling frame was the entire 65-and-older permanent resident population of Cache County, Utah, U.S.A. Study participants were diagnosed for dementia; information about all the participant's immediate family members was collected by interview, and relatives were designated as dementia cases if interview information met set criteria. Pairs we analyzed comprise the participant's mother and oldest sibling inclusive of self. We denote children's event times by X_1, and mothers', by X_2. Five-hundred and 70 pairs with missing data were excluded from analysis, as were another 887 pairs for which either member died or became demented prior to age 55, leaving 3635 pairs for analysis. There were 40 pairs in which both members had a dementia, 1132 in which both members died free of dementia, 259 in which members were observed to fail of different causes, 145 in which the mother's outcome was censored and an additional 2059 in which the eldest child's outcome was censored. Analyses treated censoring as a third "failure" cause, along with dementia onset and death.

We began by estimating $\theta_{CS}(\text{dementia,dementia}) = \theta_{CS}(1,1)$ on a four-bin time grid created by dichotomizing children's and mothers' time scales approximately at the

respective medians for time-to first event (dementia, death, or censoring; Table 3), yielding bins: ($x_1 \leq 75$, $x_2 \leq 80$), ($x_1 \leq 75$, $x_2 > 80$), ($x_1 > 75$, $x_2 \leq 80$), and ($x_1 > 75$, $x_2 > 80$). To reference analytic findings: In BRL (2002), the estimated CCSHR was 8.86 for times with joint (first-event) survival probability greater than .80; and, on the order of 2.5 for times with joint survival probability no greater than .80. In our current analysis, we also found early maternal onset and early child onset to be strongly associated, with $\hat{\theta}_{CS}(1,1)$=3.81 for ($x_1 \leq 75$, $x_2 \leq 80$); 95% CI= (1.48,6.14). Somewhat surprisingly, however, the estimated strength of association was not less for late maternal and child onset: $\hat{\theta}_{CS}(1,1)$=5.89 on ($x_1 > 75$, $x_2 > 80$); 95% CI= (1.67,10.1). Only the association for early child onset in combination with late maternal onset was notably weaker: $\hat{\theta}_{CS}(1,1)$=0.80 on ($x_1 \leq 75$, $x_2 > 80$); 95% CI= (-0.27,1.86).

Before exploring this finding further, let us consider the accuracy of asymptotic inferences reported above. In addition to inferences derived as described in §3.2, we also computed bootstrap standard errors and confidence intervals, taking 1000 bootstrap samples as in BRL (2002). With the exception of the (late, late) onset quadrant, bootstrap standard errors closely matched the respective asymptotic approximations; in that quadrant, the approximation was about 10% smaller than the bootstrap estimate. Asymptotic confidence intervals were shifted somewhat to the left of their (bias-corrected percentile-based) bootstrap counterparts, with lower limits decreased by 10%-20%, and upper limits, considerably more modestly.

To further explore the unexpected strength of association found for late child, with

late mother, dementia onsets, we conducted analyses trichotomizing to age ranges ≤ 70, 70-80, and >80 in each dimension (Table 4). As expected, with an "early" onset cutoff that more closely approximated the earliest-combined-onset category in the 2002 paper, the strength of estimated association for two early onsets was increased: $\hat{\theta}_{CS}(1,1)$=5.35. Associations were weakest in the bins representing maximally disparate children's and mothers' dementia onset ages: $\hat{\theta}_{CS}(1,1)$=1.09 for ($x_1 \leq 70$, $x_2 > 80$) and =0.81 for ($x_1 > 80$, $x_2 \leq 70$). However, there was little suggesting against a comparable strength of association for two late onsets as for two early onsets, and one higher than estimated for later onsets in the 2002 paper. With small sample sizes in most cells, few of the estimated associations differed significantly from the null of $\theta_{CS}(1,1)=1$.

In summary, analysis of familial associations in time-to-dementia onset in the bivariate time domain has clarified analysis that considered strength of association as a function of joint survival probability. The latter analysis may have understated the heritability of late-onset dementia, likely because the region of lower joint survival probability mixes regions of comparably late onset times with regions of very disparate onset times. Accordingly, this appears a good example where assumptions made by copula-based association analysis may be inadequate.

5. BIN CHOICE AND STUDY DESIGN

As a practical matter, our methodology requires choices on the number and cut points defining "bins" of failure time space. For the dementia analysis and simulation runs

with n=1000 we originally attempted to estimate associations by a 5x5 grid of failure times. This proved too sparse a partitioning, with several 0-count cells: When failure-time association is strong, the number of (a), (b) pairings with componentwise minima falling in (early failure, late failure) regions of space decreases with the degree of discrepancy between "early" and "late." Moreover the number of at-risk pairs, hence (a), (b) pairings with componentwise minima falling in later-time regions of space, declines with time. Ultimately a partitioning by equally spaced marginal quantiles is not an optimal approach.

Beyond such ad hoc considerations, it will sometimes be necessary to design studies assuring that strength of association is estimated with suitable precision in given regions of space. While full elaboration is beyond the scope of this paper, a tractable formula emerges if we multiply and divide the asymptotic variance expression (10) for our estimator by $\bar{E}_C^2$. Noting that $\bar{E}_C/\bar{E}_D = \theta^*_{CS}(j;k_1,k_2)$, one obtains equality of (10) to

$$\frac{4}{n}\{\theta^*_{CS}(j;k_1,k_2)\}^2[\zeta_{1C}-2\eta_1\{\theta^*_{CS}(j;k_1,k_2)\}+\zeta_{1D}\{\theta^*_{CS}(j;k_1,k_2)\}^2]/\bar{E}_C^2. \qquad (12)$$

Candidates for $\theta^*_{CS}(j;k_1,k_2)$ determine candidates for $\bar{E}_C$. Then, to complete (12), one must obtain candidates for ζ_{1C}, ζ_{1D}, and η_1. While these will be both complicated and unknown, equations (9) and Appendix 2 provide a template for their approximation with pilot data on bivariate failure location and cause frequencies, and marginal failure time distributions.

6. DISCUSSION

Our methodology estimates failure-time associations accurately and, to within the evaluation we provided, comparably precisely to estimation as proposed by Fan et al.

(2000). A strength is the ready interpretation of the measure we estimate. Another is the simplicity of our estimator relative to survival function and cumulative hazard plug-in counterparts. We do not intend our estimator as a replacement for such methods, but as a potentially more easily interpreted and readily implemented complement to them.

Among limitations of our approach, the CCSHR will not always capture association features of clinical interest. However, our methods could easily be elaborated to estimate other ratios involving the concordance probabilities we studied, i.e. comparing offspring with mothers diagnosed with dementia versus dying free of dementia at a given age: $\lambda_{X_2,k}(x_2|X_1=x_1,K_1=1)/\lambda_{X_2,k}(x_2|X_1=x_1,K_1=2)$. Moreover, clinical interest may require comparing strength of association by familial or individual characteristics. Our findings easily elaborate to comparisons across strata; extension to accommodate regression of CCSHRs on covariates would be valuable. Finally, while our strategy easily handles censoring as a distinct failure cause, it is desirable to accommodate such more efficiently relative to estimation absent censoring. Doing so is an advantage of methods like that proposed by Fan et al. (2000). Accommodation is complex for our method, because censoring introduces uncertainty into not only the determination of concordance but also the value of (x_1,x_2) to which a given determination should be assigned.

As equation (7) reveals, our estimator up-weights regions of less strongly positive association when strength of association varies within a bin. The reason for this is unrelated to the presence of competing causes; therefore, the effect prevails for CHR estimators grounded in Kendall's tau construction, as well. This suggests the worthiness of delineating estimation targets for maximum likelihood and pseudo-maximum likelihood

approaches assuming constant CHR when the assumption is mistaken.

As an alternative to binning, one might kernel-smooth counts defining our estimator's numerator and denominator, to obtain pointwise CCSHR estimates. So long as one chooses bandwidth *a priori*, inferences go through as in the current paper. However for time-varying CCSHR, we found kernel-smoothed estimators to be quite biased. Given this, as well as the inferential complexity of procedures with automated bandwidth choice, we prefer binning in conjunction with clarity on the target of estimation, per (7).

Our failure to find strongly for greater familial aggregation of early-onset, than late-onset, dementia contrasts with the findings of Silverman et al. (2005). To the credit of the Silverman et al. (2005) study, there was enrichment to include a substantially larger number and proportion of persons with dementia as index cases than did the data we analyzed. To our credit, data were population-based hence did not entail selection of probands or controls indexing relatives to be compared. Differences in methodologies employed were substantial and further complicate comparison. One must be mindful that our analyses' "youngest" bin included 70-year olds, and a stricter early-event definition would have resulted in a larger CCSHR (per BRL, 2002). However our analysis cautions against too strongly downplaying familial aggregation in later-onset dementia.

ACKNOWLEDGEMENT

The authors acknowledge the support of the National Institutes of Health. They are grateful to Dr. Peter Zandi for providing the Cache County data.

APPENDIX 1

Equality of (6) and (7)

Let $A_j^{(ab)}$ stand for the event $\{X^{(ab)} \in B_j, K^{(ab)} = (k_1,k_2)\}$. Then, equation (6) =

$$\frac{pr\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})>0, A_j^{(ab)}\}}{pr\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})<0, A_j^{(ab)}\}},$$

which in turn equals

$$\frac{E[pr\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})>0|A_j^{(ab)}\}\mathrm{I}\{A_j^{(ab)}\}]}{E[pr\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})<0|A_j^{(ab)}\}\mathrm{I}\{A_j^{(ab)}\}]}$$

$$=\frac{\int_{y\in B_j} pr\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})>0|X^{(ab)}=y,K^{(ab)}=k\}f^{(ab)}(y,k)dy}{\int_{y\in B_j} pr\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})<0|X^{(ab)}=y,K^{(ab)}=k\}f^{(ab)}(y,k)dy}.$$

The second expression is well defined because $P\{X^{(ab)}\le x, K^{(ab)}=k\} = P\{(X^{(a)}\le x, K^{(a)}=k) \cup (X^{(b)}\le x, K^{(b)}=k) \cup (X_1^{(a)}\le x_1, K_1^{(a)}=k_1, X_2^{(b)}\le x_2, K_2^{(b)}=k_2) \cup (X_1^{(b)}\le x_1, K_1^{(b)}=k_1, X_2^{(a)}\le x_2, K_2^{(a)}=k_2)\}$ and thus the cause-specific distribution of the componentwise minimum decomposes as sums and products of cause-specific marginal and the joint failure time distributions and, per (3), has a valid density $f^{(a,b)}(x,k)$. Denote the probability integrand in the denominator as $d^{(ab)}(y,k)$. Provided this probability is bounded above 0 almost everywhere (y) with respect to $f^{(a,b)}$, the expression may further be rewritten as

$$\frac{\int_{y\in B_j} \frac{pr\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})>0|X^{(ab)}=y,K^{(ab)}=k\}}{pr\{(X_1^{(a)}-X_1^{(b)})(X_2^{(a)}-X_2^{(b)})<0|X^{(ab)}=y,K^{(ab)}=k\}} d^{(ab)}(y,k)f^{(ab)}(y,k)dy}{\int_{y\in B_j} d^{(ab)}(y,k)f^{(ab)}(y,k)dy}.$$

If the ratio integrand is a constant, $\theta_j(k)$, almost everywhere (y) on B_j, then the expression equals $\theta_j(k)$; otherwise it equals the weighted average

$$\int_{y\in B_j} \theta_j(y,k)w^{(ab)}(y,k)f^{(ab)}(y,k)dy;$$

$$w^{(ab)}(y,k)=\frac{d^{(ab)}(y,k)}{\int_{y\in B_j} d^{(ab)}(y,k)f^{(ab)}(y,k)dy}.$$

APPENDIX 2

Elucidation of Discordance Variance Terms

To accomplish this concisely, it is useful to define one-dimensional analogs of a few already-defined quantities. First, it is convenient to denote "slices" of B_j: let $B_{j1}(y)$ be the set of x-axis value s such that $(x,y) \in B_j$, and conversely for $B_{j2}(x)$. Let the version without an argument, B_{j1} (B_{j2}), be the set of x-axis (y-axis) values such that $(x,y) \in B_j$ for at least one y (x). Second, denote one-dimensional regions where bin slices intersect (0,z) line segments by $D_1\{(z,y)\}=B_{j1}(y)\cap(0,z)$ and $D_2\{(x,z)\}=B_{j2}(x)\cap(0,z)$. Then, assuming that all regions in question are measurable, the discordance analogs of the concordance-related quantities (8) and (9) follow:

$$
\begin{aligned}
\bar{E}_D &= E_{(a)}[\mathbb{I}\{X_1^{(a)}\in B_{j1},\ K_1^{(a)}=k_1\}\mu_{(b)}\{(X_1^{(a)},\infty)\times D_2(X^{(a)});.,k_2\} \\
&\qquad +\mathbb{I}\{X_2^{(a)}\in B_{j2},\ K_2^{(a)}=k_2\}\mu_{(b)}\{D_1(X^{(a)})\times(X_2^{(a)},\infty);k_1,.\}] \\
&= 2E_{(a)}[\mathbb{I}\{X_1^{(a)}\in B_{j1},\ K_1^{(a)}=k_1\}\mu_{(b)}\{(X_1^{(a)},\infty)\times D_2(X^{(a)});.,k_2\}] \\
&= 2E_{(a)}[\mathbb{I}\{X_2^{(a)}\in B_{j2},\ K_2^{(a)}=k_2\}\mu_{(b)}\{D_1(X^{(a)})\times(X_2^{(a)},\infty);k_1,.\}]; \qquad \text{(A1)} \\
\zeta_{1D} &= E_{(a)}[\mathbb{I}\{X_1^{(a)}\in B_{j1},\ K_1^{(a)}=k_1\}\mu_{(b)}\{(X_1^{(a)},\infty)\times D_2(X^{(a)});.,k_2\} \\
&\qquad +\mathbb{I}\{X_2^{(a)}\in B_{j2},\ K_2^{(a)}=k_2\}\mu_{(b)}\{D_1(X^{(a)})\times(X_2^{(a)},\infty);k_1,.\}-\bar{E}_D]^2; \\
\zeta_{2D} &= \bar{E}_D(1-\bar{E}_D),
\end{aligned}
$$

where "." denotes the usual sum over all possibilities with respect to the argument at issue. If bins are defined on a rectangular grid, $D_1\{(z,y)\}$ simplifies to $B_{j1}\cap(0,z)$, and similarly for $D_2\{(x,z)\}$. To summarize, the denominator of (5) has normal asymptotic distribution provided $\zeta_{1D} > 0$. This condition is satisfied for reasonable bin choices and distributions, analogously as for ζ_{1C}.

REFERENCES

BANDEEN-ROCHE, K. & LIANG, K.Y. (1996). Modelling failure-time associations in data with multiple levels of clustering. *Biometrika* **83**, 29-39.

BANDEEN-ROCHE, K. & LIANG, K.Y. (2002). Modelling multivariate failure time associations in the presence of a competing risk. *Biometrika* **89**, 299-314.

BARBE, P., GENEST, C., GHOUDI, K. & RÉMILLARD, B. (1996). On Kendall's process. *J Multivar. Anal.* **58**, 197-229.

BENICHOU, J. & GAIL, M.H. (1990). Estimates of absolute cause-specific risk in cohort studies. *Biometrics* **46**, 813-26.

BREITNER, J. C. S., WYSE, B.W., ANTHONY, J. C., WELSH-BOHMER, K. A., STEFFENS, D. C., NORTON, M. C., TSCHANZ, J. T., PLASSMAN, B. L., MEYER, M. R., SKOOG, I., & KHACHATURIAN, A. (1999). APOE-ϵ4 count predicts age when prevalence of AD increases, then declines: The Cache County Study. *Neurology* **53**, 321-31.

CAMP, N.J., LOWRY, M.R., RICHARDS, R.L., PLENK, A.M., CARTER, C., HENSEL, C.H., ABKEVICH, V., SKOLNICK, M.H., SHATTUCK, D., ROWE, K.G., HUGHES, D.C., & CANNON-ALBRIGHT, L.A. (2005). Genome-wide linkage analyses of extended Utah pedigrees identifies loci that influence recurrent, early-onset major depression and anxiety disorders. *Am. J. Med. Genet. B: Neuropsychiatr. Gen.* **135(1)**, 85-93.

CHEN, M.C. & BANDEEN-ROCHE, K. (2005). A diagnostic for association in bivariate survival models. *Lifetime Data Anal.* **11**, 245-64.

CLAYTON, D.G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141-51.

CLAYTON, D., & CUZICK, J. (1985). Multivariate generalizations of the proportional hazards model. *J. R. Statist. Soc. A* **148**, 82-108.

FAN, J., PRENTICE, R.L. & HSU, L. (2000). A class of weighted dependence measures for bivariate failure time data. *J. R. Statist. Soc. B* **62**, 181-90.

FINE, J.P. & JIANG, H. (2000). On association in a copula with time transformations. *Biometrika* **87**, 559-71.

GENEST, C. & MACKAY, J. (1986). Copules archimediennes et familles de lois bidimensionelles dont les marges sont donnees. *Can. J. Statist.* **14**, 145-59.

GENEST, C., & RIVEST, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *J. Am. Statist. Assoc.* **88**, 1034-43.

GENEST, C., GHOUDI, K. & RIVEST, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543-52.

GLIDDEN, D.V. (2000). A two-stage estimator of the dependence parameter for the Clayton-Oakes model. *Lifetime Data Anal.* **6**, 141-56.

HENDRIE, H.C. (1998). Epidemiology of dementia and Alzheimer's disease. *Am. J. Geriatr. Psych.* **6**, S3-18.

HSU, L. & PRENTICE, R.L. (1996). On assessing the strength of dependency between failure time variates. *Biometrika* **83**, 491-506.

KENDALL, M.G. (1955). *Rank Correlation Methods,* 2nd ed. London: Griffin.

LEE, L. (1979). Multivariate distributions having Weibull properties. *J. Multivar. Anal.* **9**, 267-77.

NIELSEN, G.G., GILL, R.D., ANDERSEN, P.K. & SORENSEN, T.I.A. (1992). A counting process approach to maximum likelihood estimation in frailty models. *Scand. J. Statist.* **19**, 25-43.

OAKES, D. (1982). A model for association in bivariate survival data. *J. R. Statist. Soc. B* **44**, 414-22.

OAKES, D. (1986). Semiparametric inference in a model for association in bivariate survival data. *Biometrika* 73, 353-61.

OAKES, D. (1989). Bivariate survival models induced by frailties. *J. Am. Statist. Assoc.* **84**,

487-93.

PRENTICE, R.L., KALBFLEISCH, J.D., PETERSON, A.V. JR., FLOURNOY, N., FAREWELL, V.T. & BRESLOW, N.E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics* **34**, 541-54.

PRENTICE, R.L. & CAI, J. (1992). Covariance and survivor function estimation using censored multivariate failure time data. *Biometrika* **79**, 495-512.

PRENTICE, R.L. & HSU, L. (1997). Regression on hazard ratios and cross ratios in multivariate failure time analysis. *Biometrika* **84**, 349-363.

PRENTICE, R.L., ZOE, M.F. & WU, J. (2004). Hazard-based nonparametric survivor function estimation. *J. R. Statist. Soc. B* **66**, 305-19.

PUNJABI, N.M., O'HEARN, D.J., NEUBAUER, D.N., NIETO, F.J., SCHWARTZ, A.R., SMITH, P.L, & BANDEEN-ROCHE, K. (1999). Modeling hypersomnolence in sleep-disordered breathing. A novel approach using survival analysis. *Am. J. Resp. Crit. Care Med.* **159(6)**, 1703-9.

RIPATTI, S. & PALMGREN, J. (2000). Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics* **56**, 1016-22.

RIPATTI, S., LARSEN, K. & PALMGREN, J. (2002). Maximum likelihood inference for multivariate frailty models using an automated Monte Carlo EM algorithm. *Lifetime Data Anal.* **8**, 349-60.

SERFLING, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: Wiley, Chapter 5.

SHIEH, G.S. (1998). A weighted Kendall's tau statistic. *Statist. Prob. Let.* **39**, 17-24.

SHIH, J.H. & LOUIS, T. A. (1995). Inferences on the association parameter in copula models

for bivariate survival data. *Biometrics* **51**, 1584-99.

SILVERMAN, J.M., CIRESI, G., SMITH, C.J., MARIN, D.B., & SCHAIDER-BEERI, M. (2005). Variability of familial risk of Alzheimer disease across the late life span. *Arch. Gen. Psych.* **62**, 565-73.

STEL, V.S., PLUIJM, S.M., DEEG, D.J., SMIT, J.H., BOUTER, L.M., LIPS, P. (2003). A classification tree for predicting recurrent falling in community-dwelling older persons. *J. Am. Geriatr. Soc.* **51(10)**,1356-64.

VISWANATHAN, B. & MANATUNGA, A.K. (2001). Diagnostic plots for assessing the frailty distribution in multivariate survival data. *Lifetime Data Anal.* **7**, 143-55.

WANG, W. & WELLS, M.T. (2000). Estimation of Kendall's tau under censoring. *Statistica Sinica* **10**, 1199-1215.

Table 1 - Simulation Study Findings
Association estimator distributions, Gamma/Beta frailty data

Quadrant	n	R_1=.2 — $\theta_{CS}(1,1) = 6$				R_1=.5 — $\theta_{CS}(1,1) = 3$				R_1=.8 — $\theta_{CS}(1,1) = 2.25$			
		Mean	**SD_E**	**SD_M**	**Cov**	**Mean**	**SD_E**	**SD_M**	**Cov**	**Mean**	**SD_E**	**SD_M**	**Cov**
1: $(t_1,t_2) \leq$ medians	1000	6.11	1.06	1.00	0.96	3.03	0.30	0.28	0.96	2.24	0.16	0.16	0.95
	500	6.30	2.16	1.93	0.96	3.05	0.56	0.53	0.97	2.31	0.31	0.32	0.95
2: $t_1 \leq$ median, $t_2 >$ median	1000	6.12	1.64	1.60	0.93	3.04	0.44	0.44	0.96	2.27	0.24	0.24	0.94
	500	7.18	4.39	4.16	0.94	3.15	0.91	0.92	0.94	2.34	0.47	0.47	0.96
3: $t_1 >$median, $t_2 \leq$ median	1000	6.36	1.70	1.65	0.94	3.04	0.44	0.43	0.95	2.25	0.24	0.24	0.94
	500	7.06	6.59	4.96	0.93	3.12	0.89	0.88	0.93	2.29	0.46	0.46	0.93
4: (t_1,t_2) >medians	1000	6.15	1.38	1.28	0.96	3.03	0.38	0.37	0.95	2.26	0.20	0.20	0.96
	500	7.16	4.27	3.85	0.93	3.12	0.79	072	0.95	2.31	0.40	0.39	0.95

Data generated as described in Section 3.1, equation (11): Gamma size copula (A), $\theta(t)=2$; Beta shape frailty (B) with mean R_1 and scale=1; conditional baseline distributions exponential(1); bivariate data with sample size n per each of 500 runs; no non-competing censoring.
SD_E = square root of the average of variance estimates over 500 runs.
SD_M = the empirical standard deviation of estimates over 500 runs.

Table 2 - Simulation Study Findings
Association estimator distributions, Positive stable/Beta frailty data

Quadrant	n	α=0.4; $\{\theta_{CS}(1,1;Q1),...,\theta_{CS}(1,1;Q4)\}$ = {9.24, 3.68, 3.66,2.89}				α=0.6; $\{\theta_{CS}(1,1;Q1),...,\theta_{CS}(1,1;Q4)$ = {4.72,2.36,2.37,2.05}				α=0.8; $\{\theta_{CS}(1,1;Q1),...,\theta_{CS}(1,1;Q4)$ = {2.66,1.80,1.80,1.69}			
		Mean	**SD_E**	**SD_M**	**Cov**	**Mean**	**SD_E**	**SD_M**	**Cov**	**Mean**	**SD_E**	**SD_M**	**Cov**
1: (t_1,t_2) ≤ medians	1000	9.24	1.00	1.04	0.93	4.77	0.51	0.53	0.95	2.67	0.29	0.27	0.96
	500	9.41	1.92	1.99	0.94	4.91	0.98	0.94	0.95	2.79	0.56	0.59	0.94
2: t_1≤median, t_2 >median	1000	3.76	0.73	0.75	0.94	2.39	0.38	0.38	0.95	1.81	0.25	0.25	0.95
	500	3.93	1.52	1.49	0.94	2.55	0.78	0.73	0.97	1.89	0.51	0.52	0.94
3: t_1>median, t_2≤ median	1000	3.77	0.74	0.76	0.94	2.38	0.38	0.38	0.95	1.82	0.25	0.23	0.96
	500	4.02	1.61	1.60	0.95	2.52	0.79	0.76	0.96	1.85	0.49	0.50	0.93
4: (t_1,t_2) > medians	1000	2.89	0.33	0.35	0.93	2.07	0.26	0.25	0.95	1.72	0.23	0.23	0.95
	500	2.98	0.66	0.64	0.96	2.15	0.53	0.51	0.95	1.75	0.47	0.46	0.95

Data generated as described in Section 3.1, equation (11): Positive stable size copula (A); Beta shape frailty (B) with mean R=.5 and scale=1; conditional baseline distributions exponential(1); bivariate data with sample size n per each of 500 runs; no non-competing censoring.
SD_E = square root of the average of variance estimates over 500 runs.
SD_M = the empirical standard deviation of estimates over 500 runs.

Table 3 - Simulation Study Findings
Association estimator distributions, frailty data without competing risks[1]

Association Model	n	*(t_1,t_2) ≤ lower quartiles*				*2: (t_1,t_2) ≤ medians*				*3: (t_1,t_2) ≤ upper quartiles*			
		Our estimator		Fan estimator		Our estimator		Fan estimator		Our estimator		Fan estimator	
		Mean	SD_E	Mean	SD_E	Mean	SD_E	Mean	SD_E	Mean	SD_E	Mean	SD_E
Independent CHR^{-1}=1	1000	1.02	0.14	1.02	0.14	1.01	0.07	1.00	0.07	1.01	0.05	1.00	0.04
	100	1.21	0.90	1.20	0.89	1.02	0.23	1.01	0.21	1.00	0.16	1.00	0.13
Gamma CHR^{-1}=0.5	1000	0.50	0.06	0.50	0.06	0.49	0.03	0.50	0.03	0.50	0.02	0.50	0.02
	100	0.53	0.19	0.54	0.18	0.51	0.10	0.52	0.10	0.51	0.08	0.52	0.08
Pos. stable α=.4	1000	0.09	0.01	0.09	0.01	0.16	0.01	0.18	0.01	0.22	0.01	0.28	0.03
	100	0.09	0.03	0.10	0.03	0.16	0.04	0.20	0.04	0.22	0.04	0.30	0.05
	limit	0.09	NA	0.09	NA	0.16	NA	0.17	NA	0.22	NA	0.24	NA
Pos. stable α=.8	1000	0.40	0.05	0.42	0.05	0.56	0.04	0.60	0.04	0.64	0.04	0.72	0.04
	100	0.43	0.22	0.45	0.23	0.57	0.13	0.62	0.13	0.65	0.11	0.73	0.11
	limit	0.40	NA	0.42	NA	0.56	NA	0.59	NA	0.64	NA	0.67	NA

Conditional baseline distributions exponential(1); bivariate data with sample size n per each of 500 runs; no non-competing censoring.
SD_E = square root of the average of variance estimates over 500 runs.

Table 4 - Cache County Data (n=3635)
Association estimator distributions, $\theta_{CS}(1,1)$ (dementia, dementia); 2x2 time grid

Quadrant	Mean	SD_E	SD_M	SD_B	95% CI - asymptotic	95% CI - bootstrap
1: $t_1 \le 75, t_2 \le 80$	3.81	1.19	1.23	1.23	(1.48, 6.14)	(1.68, 6.22)
2: $t_1 \le 75, t_2 > 80$	0.80	0.54	0.53	0.57	(-0.27, 1.86)	(0.00, 1.87)
3: $t_1 > 75, t_2 \le 80$	2.41	0.73	0.76	0.77	(0.97, 3.84)	(1.10, 3.92)
4: $t_1 > 75, t_2 > 80$	5.89	2.15	2.41	2.38	(1.67,10.11)	(2.08,10.39)

SDe = Asymptotic standard deviation approximation
SDm = Square-root of the average of asymptotic variance approximation estimates over 1000 bootstrap replicates
SDb Bootstrap standard deviation estimate, 1000 replicates
Bootstrap CI is bias-corrected and percentile-based

Table 5 - Cache County Data (n=3635)
Association estimator distributions, $\theta_{CS}(1,1)$ (dementia, dementia); 3x3 time grid

Cell	Mean	SD_E	SD_M	SD_B	95% CI - asymptotic	95% CI - bootstrap
1: $(t_1,t_2) \le 70$	5.35	3.15	3.32	3.35	(-0.82,11.53)	(0.00,12.18)
2: $t_1 \le 70$, $70 < t_2 \le 80$	1.72	1.21	1.27	1.25	(-0.65, 4.10)	(0.00, 4.10)
3: $t_1 \le 70$, $t_2 > 80$	1.09	0.99	0.89	1.05	(-0.84, 3.03)	(0.00, 3.21)
4: $70 < t_1 \le 80$, $t_2 \le 70$	4.03	1.94	2.12	2.11	(0.23, 7.84)	(0.65, 8.36)
5: $70 < t_1 \le 80$, $70 < t_2 \le 80$	3.39	1.14	1.18	1.16	(1.16, 5.61)	(1.25, 5.62)
6: $70 < t_1 \le 80$, $t_2 > 80$	2.99	1.13	1.21	1.15	(0.77, 5.20)	(0.88, 5.21)
7: $t_1 > 80$, $t_2 \le 70$	0.81	0.80	0.89	0.88	(-0.77, 2.39)	(0.00, 2.54)
8: $t_1 > 80$, $70 < t_2 \le 80$	2.50	1.28	1.43	1.40	(-0.02, 5.01)	(0.03, 5.28)
9: $(t_1,t_2) > 80$	3.62	1.91	2.36	2.33	(-0.13, 7.36)	(0.00, 7.65)

SDe = Asymptotic standard deviation approximation
SDm = Square-root of the average of asymptotic variance approximation estimates over 1000 bootstrap replicates
SDb Bootstrap standard deviation estimate, 1000 replicates
Bootstrap CI is bias-corrected and percentile-based