



Published in final edited form as:

*Biometrika*. 2008 March 1; 95(1): 221–232. doi:10.1093/biomet/asm091.

## Non-parametric estimation of bivariate failure time associations in the presence of a competing risk

**Karen Bandeen-Roche and Jing Ning**

Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, Maryland 21205, U.S.A.

Karen Bandeen-Roche: kbandeen@jhsph.edu; Jing Ning: jning@jhsph.edu

### Abstract

Most research on the study of associations among paired failure times has either assumed time invariance or been based on complex measures or estimators. Little has accommodated competing risks. This paper targets the conditional cause-specific hazard ratio, henceforth called the cause-specific cross ratio, a recent modification of the conditional hazard ratio designed to accommodate competing risks data. Estimation is accomplished by an intuitive, non-parametric 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 developed, small-sample performance is evaluated and the methods are applied to the investigation of familial association in dementia onset.

### Keywords

Cause-specific; Kendall's tau; Multivariate; Paired; Survival; *U*-statistic

## 1. INTRODUCTION

Methodology for analyzing correlated failure-time data must ensure correctness of inferences, accounting for failure-time dependence within clusters. We propose methodology for estimating the strength of dependence for paired failure times.

Much work assessing failure time associations has been focused on the cross, or conditional hazard, ratio function (Clayton, 1978; Oakes, 1982, 1986; Clayton & Cuzick, 1985). Various researchers have exploited parametric models for the cross ratio (Genest & MacKay, 1986; Oakes, 1989; Nielsen et al., 1992; Genest et al., 1995; Shih & Louis, 1995; Glidden, 2000; Ripatti & Palmgren 2000; Ripatti et al., 2002). Measures allowing non-parametrically time-varying association have also been proposed, with common estimation strategies being to 'plug into' one's measure a non-parametric estimator of the multivariate survival or cumulative hazard function (Prentice & Cai, 1992; Hsu & Prentice, 1996; Fan et al., 2000; Wang & Wells, 2000) or solve method of moments equations involving appropriately chosen empirical processes for the measure of interest (Oakes, 1982, 1989; Genest & Rivest, 1993; Barbe et al., 1996; Viswanathan & Manatunga, 2001; Chen & Bandeen-Roche, 2005). Related regression models have also been proposed (Prentice & Hsu, 1997; Fine & Jiang, 2000).

However, measures of failure time association have had slow uptake in biomedical studies, and we know of only three papers addressing the (semi-)competing risks that are unavoidable when conditions may lead to death or affect only a fraction of individuals; see Bandeen-Roche & Liang (2002), a University of Wisconsin technical report by Y. Cheng, J. P. Fine & M. R. Kosorok (2004), and a paper soon to appear by Cheng & Fine (2008). To narrow these gaps,

this paper studies a simple, non-parametric estimator of an easily interpreted measure of failure time association, namely the cause-specific cross ratio (Bandeem-Roche & Liang, 2002). Cheng & Fine (2008) have studied an alternative approach to estimating the same measure. Our derivation yields insight into the convergence behaviour of cross-ratio estimators when time invariance of the cross ratio is wrongly assumed, and into factors affecting the estimator's precision.

## 2. METHODS

### 2.1. Notation and estimand

Consider competing events  $1, \dots, m$ . Let  $X^*$  be the time to the first event, and let  $K \in \{1, \dots, m\}$  be a code identifying that event. We observe, for individuals,  $X$ , the minimum of  $X^*$  and the time at which  $X^*$  is censored for non-competing reasons, together with  $K$ , a code equalling 0 if failure is censored and  $K^*$ , if the earliest competing event precedes censoring. Then, the individual's cause-specific hazard for the  $k$ th event is

$$\lambda_{x,k}(x) = \lim_{\Delta x \downarrow 0} \text{pr}(x \leq X \leq x + \Delta x, K = k | X \geq x) / \Delta x$$

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

With bivariate failure processes, observable data are times  $(X_i, Y_i)$  and associated causes  $(K_i, L_i)$  samples in pairs, for  $i = 1, \dots, n$ . We assume  $(X_i, Y_i, K_i, L_i)$  to be independently and identically distributed so that, with  $i$  suppressed, cause-specific densities

$$f(x, y, k, l) = \lim_{(\Delta x, \Delta y) \downarrow 0} \text{pr}(x \leq X \leq x + \Delta x, y \leq Y \leq y + \Delta y, K = k, L = l) / (\Delta x \Delta y) \tag{1}$$

exist for each combination of failure causes  $(k, l)$ . Then,  $(X, Y)$  has an absolutely continuous joint survival function

$$S(x, y) = \int_x^\infty \int_x^\infty \sum_{k=1}^m \sum_{l=1}^m f(u, v, k, l) dv du.$$

Bandeem-Roche & Liang (2002) defined the cause-specific conditional hazard, i.e. cross ratio, as

$$\theta_{cs}(x, y; k, l) = \left( \frac{f(x, y, k, l)}{\int_x^\infty \sum_{h=1}^m f(u, y, h, l) du} \right) \left( \frac{S(x, y)}{\int_y^\infty \sum_{j=1}^m f(x, v, k, j) dv} \right). \tag{2}$$

If  $X$  and  $Y$  are failure times for two family members, and causes  $K$  and  $L$  indicate dementia onset, if equal to 1, or death before dementia, if equal to 2, then the cause-specific cross ratio for  $k = 1, l = 1$  gives the factor by which dementia onset at age  $y$  is increased for family members whose relatives are diagnosed as cases at age  $x$  versus those whose relatives survive, disease-free, beyond age  $x$ , or vice versa:

$$\lambda_{y,1}(y | X = x, K = 1) / \lambda_{y,1}(y | X > x).$$

**2.2. Estimator**

A transformed, localized Kendall’s tau serves to estimate (2) non-parametrically. Let  $(X_a, Y_2)$  and  $(X_b, Y_a)$  be independently drawn failure time pairs; let  $(X_{(ab)}, Y_{(ab)})$  be the componentwise minima  $(X_a \wedge X_b, Y_a \wedge Y_b)$ ; and let  $(K_{(ab)}, L_{(ab)})$  be causes corresponding to  $(X_{(ab)}, Y_{(ab)})$ . It can be shown that

$$\theta_{CS}(x, y; k, l) = \frac{\text{pr}\{(X_a - X_b)(Y_a - Y_b) > 0 | (X_{(ab)}, Y_{(ab)}) = (x, y), (K_{(ab)}, L_{(ab)}) = (k, l)\}}{\text{pr}\{(X_a - X_b)(Y_a - Y_b) < 0 | (X_{(ab)}, Y_{(ab)}) = (x, y), (K_{(ab)}, L_{(ab)}) = (k, l)\}}$$

that is, the quotient of conditional probabilities of concordance and discordance between the pairs’ failure times, given  $(X_{(ab)}, Y_{(ab)})$  and  $(K_{(ab)}, L_{(ab)})$ . Thus, a simple estimator determines concordance status for every two pairs with  $(K_{(ab)}, L_{(ab)}) = (k, l)$  and then divides the number of concordances by the number of discordances. However, if  $\theta_{CS}(x, y; k, l)$  is a continuous function of  $(x, y)$  on  $\{(u, v) : u > 0, v > 0\} = \mathcal{R}^{2+}$ , the numbers or ratios must be binned or smoothed if stable ratios are to be obtained. Let  $\mathcal{B} = \{B_1, \dots, B_j\}$  be a partition of  $\mathcal{R}^{2+}$ , with  $\mathcal{B}$  set a priori and  $J$  finite, and  $\mathcal{Q}\{A\} = 1$  if  $A$  is true and 0 otherwise. Our estimator is

$$\widehat{\theta}_{CS}(x, y; k, l) = \frac{\sum_{a < b} \mathcal{I}\{(K_{(ab)}, L_{(ab)}) = (k, l)\} \cdot \mathcal{I}\{(X_{(ab)}, Y_{(ab)}) \in \mathcal{B}_{j(x,y)}\} \cdot \mathcal{I}\{(X_a - X_b)(Y_a - Y_b) > 0\}}{\sum_{a < b} \mathcal{I}\{(K_{(ab)}, L_{(ab)}) = (k, l)\} \cdot \mathcal{I}\{(X_{(ab)}, Y_{(ab)}) \in \mathcal{B}_{j(x,y)}\} \cdot \mathcal{I}\{(X_a - X_b)(Y_a - Y_b) > 0\}} \tag{3}$$

with  $j(x, y)$  indexing the partition cell including  $(x, y)$ . It parallels estimators employed by Bandeen-Roche & Liang (2002) and, with causes ignored, Chen & Bandeen-Roche (2005), but defines  $\mathcal{B}$  on  $\mathcal{R}^{2+}$  rather than on  $\{S(x, y), (x, y) \in \mathcal{R}^{2+}\}$ .

**2.3. Distributional properties**

Asymptotic inference employing (3) seems only to have been developed for like estimators ignoring competing risks and involving the global Kendall’s tau (Kendall, 1948, p. 67; Oakes 1982) or a weighted analogue (Oakes, 1986). Existing theory for localized versions of tau (Shieh, 1998) does not apply, because our localization procedure incorporates weights that may stochastically relate to the terms being weighted. Rather, let  $r_i$  represent a generic four-tuple realization  $(x_i, y_i, k_i, l_i)$  for pair  $i$ , and  $R_i$ , the analogous random variable. At each  $(x, y, k, l)$ , the numerator of (3) is a  $U$ -statistic with kernel

$$h_{C(x,y,k,l)}(r_a, r_b) = \mathcal{I}\{(k_{(ab)}, l_{(ab)}) = (k, l)\} \cdot \mathcal{I}\{(x_{(ab)}, y_{(ab)}) \in B_{j(x,y)}\} \cdot \mathcal{I}\{(x_a - x_b)(y_a - y_b) > 0\},$$

and similarly for the denominator with kernel, we label  $h_{D(x,y,k,l)}(r_a, r_b)$ . Thus, inferences follow directly from  $U$ -statistic theory (Serfling, 1980, Ch. 5). With sums replaced by averages, (3) converges almost surely to

$$\theta_{CS}^*(x, y; k, l) = \frac{\text{pr}\{(X_a - X_b)(Y_a - Y_b) > 0 | (X_{(ab)}, Y_{(ab)}) \in B_{j(x,y)}, (K_{(ab)}, L_{(ab)}) = (k, l)\}}{\text{pr}\{(X_a - X_b)(Y_a - Y_b) < 0 | (X_{(ab)}, Y_{(ab)}) \in B_{j(x,y)}, (K_{(ab)}, L_{(ab)}) = (k, l)\}} \tag{4}$$

under weak conditions, so long as the denominator exceeds 0. More interesting is to consider interpretation if  $\text{pr}\{(X_a - X_b)(Y_a - Y_b) < 0 | (X_{(ab)}, Y_{(ab)}) = (x, y), (K_{(ab)}, L_{(ab)}) = (k, l)\}$  is bounded above 0 almost everywhere on  $B_{j(x,y)}$ . Then, by an argument in a longer version of this paper available from the authors, (4) equals

$$E\{\theta_{CS}(X_{(ab)}, Y_{(ab)}; k, l) | (X_{(ab)}, Y_{(ab)}) \in B_j, (X_a - X_b)(Y_a - Y_b) < 0, (K_{(ab)}, L_{(ab)}) = (k, l)\} = \theta_{CS}^*(j; k, l), \tag{5}$$

where we now suppress the  $(x, y)$  notation indexing bins, retaining only  $j$ . If the cause-specific cross ratio is constant over  $B_j$ , (5) equals that constant value. Otherwise, (5) averages over potential time realizations within  $B_j$ , weighting with respect to probabilities of discordance in pairs  $a$  and  $b$ . Thus, with time-varying cause-specific cross ratio, (5) summarizes up-weighting regions of less strongly positive association over the bin, and the limiting target of (3) is closer to the null than a straight expectation of (2) over  $B_j$ .

We report asymptotic distributions of the dividends defining (3). Notation to index properly each ‘C’ and ‘D’ subscripted term in terms of  $(j, k, l)$  is suppressed. The dividends depend on the numerator and denominator kernel means,

$$\bar{E}_C = E\{h_C(R_a, R_b)\}, \bar{E}_D = E\{h_D(R_a, R_b)\},$$

and the corresponding variances and covariances. The numerator variance depends on

$$\zeta_{1C} = \text{var}_a[E_{b|a}\{h_C(R_a, R_b)\}], \zeta_{2C} = \text{var}\{h_C(R_a, R_b)\},$$

and the denominator variance on

$$\zeta_{1D} = \text{var}_a[E_{b|a}\{h_D(R_a, R_b)\}], \zeta_{2D} = \text{var}\{h_D(R_a, R_b)\}.$$

The terms  $h_C(R_a, R_b)$  and  $h_D(R_a, R_b)$  are indicator functions, and thus have variances  $\zeta_{2C} = \bar{E}_C(1 - \bar{E}_C)$  and  $\zeta_{2D} = \bar{E}_D(1 - \bar{E}_D)$ ; the terms  $\zeta_{1C}$  and  $\zeta_{1D}$  are elucidated in Appendix 1. The numerator and denominator variances are obtained, respectively from  $U$ -statistic theory as,

$$[4(n - 2)/\{n(n - 1)\}]\zeta_{1C} + [2/\{n(n - 1)\}]\zeta_{2C}, [4(n - 2)/\{n(n - 1)\}]\zeta_{1D} + [2/\{n(n - 1)\}]\zeta_{2D}.$$

Moreover, the dividends asymptotically have normal distributions, provided  $\zeta_{1C} > 0$  and  $\zeta_{1D} > 0$ . Under these and previous conditions,

$$\sqrt{n} \left\{ \begin{array}{l} \left( \binom{n}{2}^{-1} \sum_{a < b} h_C(R_a, R_b) - \bar{E}_C \right) \\ \left( \binom{n}{2}^{-1} \sum_{a < b} h_D(R_a, R_b) - \bar{E}_D \right) \end{array} \right\} \rightarrow N \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4\zeta_{1C} & 4\eta_1 \\ 4\eta_1 & \zeta_{1D} \end{pmatrix} \right\},$$

in distribution, where  $\eta_1 = \text{cov}_a[E_{b|a}\{h_C(R_a, R_b)\}, E_{b|a}\{h_D(R_a, R_b)\}]$ . By the delta method, the proposed estimator is asymptotically normal with limiting mean (5) and variance

$$\frac{4}{n\bar{E}^2 D} \left( \zeta_{1C} - 2 \frac{\bar{E}_{C\eta_1}}{\bar{E}_D} + \frac{\zeta_{1D} \bar{E}_C^2}{\bar{E}_D^2} \right). \tag{6}$$

To see what  $\zeta_{1C} > 0, \zeta_{1D} > 0$  entail, consider the form of  $\zeta_{1C}$ , from Appendix 1:

$$\zeta_{1c} = E_a[\mathcal{I}\{(X_a, Y_a) \in B_j, (K_a, L_a) = (k, l)\} S(X_a, Y_a) + \mu_b\{B_j \cap (0, X_a) \times (0, Y_a); k, l\} - \bar{E}_c]^2 \tag{7}$$

Trivially, bins and failure causes must be such that failures of type  $(k, l)$  may occur, excluding  $\text{pr}\{(X_a, Y_a) \in B_j, (K_a, L_a) = (k, l)\} = 0$ . Furthermore, if  $S(x, y)$  is restricted to one dimension,  $S(x, y) = 1 - F(x, y)$ , (7) equals 0 if there are one bin,  $\mathcal{R}^{2+}$  and failure cause. However, even in this case, all  $\zeta_{1c} > 0$  if there are multiple causes or bins having well-defined positive measure.

Equation (3) estimates  $\theta_{CS}(x, y; k, l)$  for each bin and pair of failure causes  $(k, l)$ . There may also be interest in estimators involving more than one of these, such as the ratio comparing individuals whose relative is diagnosed with dementia, versus who dies dementia-free, at a given age:

$$\lambda_{y,1}(y|X=x, K=1) / \lambda_{y,1}(y|X=x, K=2).$$

It is relatively simple so to extend inferences, because the cause-specific cross ratios for different cause combinations and bins are asymptotically uncorrelated; see Appendix 2.

**2.4. Variance estimation**

Since the discordance-associated terms are complicated, we estimate quantities defining the limiting variance of our estimator by their sample counterparts. Estimators for  $\bar{E}_C$  and  $\bar{E}_D$  are given by numerator and denominator of (3), and those for  $\zeta_{1C}$ ,  $\zeta_{1D}$  and  $\eta_1$  are calculated similarly; for instance,

$$\widehat{\zeta}_{1c} = \sum_{a=1}^n \left\{ \sum_{b \neq a} h_c(R_a, R_b) / (n - 1) \right\}^2 / n - \bar{E}_c^2.$$

Computations involve nested sums and thus, are  $O(n^2)$  in complexity, but simple in form.

**3. SIMULATION STUDY**

**3.1. Design**

Each study scenario envisaged two failure causes, ‘disease’ ( $k = 1$ ) and ‘death’ ( $k = 2$ ), and comprised 500 runs. Bivariate data were generated according to the frailty model for subject- and cause-specific failure hazards given by equation (6) of Bandeen-Roche & Liang (2002), as described in Appendix 2 of that article. For a given hazard function  $\lambda^*(t) = 1$ , that model is

$$\lambda k(x|G=g, W=w) = gw_k \lambda^*(x), \tag{8}$$

with  $G$  as a scalar ‘size’ frailty, and  $W$ , a compositional ‘shape’ frailty, shared by a given pair. Throughout, we generated  $W$  to be beta-distributed with scale parameter 1 and set  $\lambda^*(t) = 1$ .

We allowed the distribution of  $G$  to be gamma or positive stable. The cause-specific cross ratio is time-invariant with gamma frailties; we varied it to be 2.25 or 6.0. In positive stable runs, we generated size frailties as in Lee (1979) and fixed  $E(W) = (0.5, 0.5)$ . Positive stable distributions have Laplace transform  $\exp(-t^\alpha)$ ; we varied  $\alpha$  to be 0.4 or 0.8.

For each run, we applied (3) over a four-cell grid, bisecting each time, dimension at its median. We considered two sample sizes,  $n = 500, 1000$ . Standard errors were constructed as in § 2.4.

For each scenario, we evaluated estimator bias, compared asymptotic and empirical estimator variance calculations and determined 95% confidence interval coverage. With positive stable-size frailties, the cause-specific cross ratio decreases with time, and the bin-specific association parameters are given by equation (5). To estimate these, we generated  $n_s = 20\,000$  pairs for each scenario and replaced the expectation in (5) by a sample average over the 20 000-choose-2 pairings of pairs. To select this sample size, we generated estimates over a range of  $n_s$ ; by  $n_s = 20\,000$ , an approximate asymptote was achieved.

Secondary studies compared the performance of our estimator and the cumulative hazard plug-in estimator proposed by Fan et al., (2000). This estimator addresses single-cause failure outcomes; data were generated accordingly, setting  $w_k$  in equation (8) equal to 1. It estimates the cross ratio,  $\theta(t)$ , over lower  $(t_1, t_2)$  quadrants; we report findings over quadrants with  $t_1$  and  $t_2$  equal to (a) the lower quartile and (b) the upper quartile of the marginal survival function. We studied models of independence within pairs, gamma frailty with cross ratio equal to 2; and positive stable frailty with  $\alpha = 0.4, 0.8$ . For positive stable scenarios, the values of parameters to be estimated were approximated, as described in the previous paragraph. For the Fan estimator, equation (6) of Fan et al., (2000) was averaged, with weights as given at the top of page 183 of that article. We compared sample sizes of  $n = 100$  and  $n = 1000$  pairs as well as scenarios without censoring and with independent standard-exponential-distributed random censoring of individual failure times, yielding censoring probability of approximately 0.5.

### 3.2. Results

We first consider performance with gamma scenarios in our primary study; see Table 1. Both our estimator and its associated inferences were accurate on all time quadrants. With  $n = 1000$ , upward biases ranged from 1% to 7% for both point and standard error estimation; with  $n = 500$ , these were moderately exacerbated. Confidence interval coverage ranged from 93% to 97%.

Estimator performance was even better for positive stable scenarios, see Table 2, yet upward biases persisted for the smaller sample sizes and less populated bins. As suggested by a referee, we explored the impact of estimator distribution skewness on these biases by comparing geometric means and arithmetic means for  $n = 500$  and  $\alpha = 0.4$ . Geometric means were generally closer to limiting target values than arithmetic means, with biases of 0.05, 0.01,  $< 0.01$  and  $-0.03$  as compared to 0.17, 0.16, 0.17 and 0.01 per bins ordered as the rows of Table 2.

Table 3 compares our estimator and the Fan estimator. In independence and gamma frailty runs, both estimators achieved outstanding accuracy, and they exhibited nearly identical precision. In positive stable scenarios, our estimator's limiting values exceeded those of the Fan estimator, increasingly with the strength of association and quadrant size. Equation (5) predicts this. Our estimator's standard error was higher in absolute terms, but was only modestly higher as a percentage of its limiting target. In the absence of censoring, both estimators exhibited little bias. In censored scenarios, both estimators exhibited appreciable bias, ranging approximately from 10% to 50%. 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- than later-onset dementia (Silverman et al., 2005). If so, dementia onset ages should be associated within families, most strongly in a lower left quadrant of ages. In addition, death very often predates a dementia diagnosis.

We analyzed the same 3635 failure time pairs, from the Cache County Study on Memory in Aging (Breitner et al., 1999), as were analyzed by Bandeen-Roche & Liang (2002). Event times were ascertained for the oldest sibling inclusive of self ( $X$ ) and mother ( $Y$ ). Analyses treated censoring as a third ‘failure’ cause, besides dementia onset and death.

We first estimated  $\theta_{CS}(j; 1, 2)$  on a grid dichotomizing children’s and mothers’ ages approximately at the respective medians for time-to-first event, yielding bins ( $x \leq 75, y > 80$ ), ( $x > 75, y \geq 80$ ), ( $x > 75, y > 80$ ) and ( $x > 75, y > 80$ ); see Table 4. Early maternal onset of dementia was strongly associated with early child onset:  $\hat{\theta}_{CS}(j; 1, 1) = 3.81$  for ( $x \leq 75, y \leq 80$ ) with 95% confidence interval (1.48, 6.14). Surprisingly, the estimated strength of association was greater for paired late onsets:  $\hat{\theta}_{CS}(j; 1, 1) = 5.89$  on ( $x > 75, y > 80$ ) with 95% confidence interval (1.67, 10.1). It was only notably weaker for early child onset and late maternal onset:  $\hat{\theta}_{CS}(j; 1, 1) = 0.80$  on ( $x \leq 75, y > 80$ ) with 95% confidence interval (−0.27, 1.86). Here, asymptotic inferences were derived as described in § 2.4. We also computed bootstrap standard errors and confidence intervals, following Bandeen-Roche & Liang (2002), taking 1000 bootstrap samples. Bootstrap standard errors closely matched asymptotic approximations except in the (late, late) quadrant, where the approximation was roughly 10% smaller than the bootstrap estimate. Asymptotic confidence intervals had lower limits decreased by 10–20%, and upper limits considerably more modestly, than their bias-corrected percentile-based bootstrap counterparts.

To explore further the unexpected strength of association found for paired late dementia onsets, we conducted analyses trichotomizing to age ranges  $\leq 70$ ,  $70 - 80$  and  $> 80$  in each dimension. The strength of estimated association for two early onsets was increased:  $\theta_{CS}(j; 1, 1) = 5.35$ . Associations were the weakest in the bins with maximally disparate children’s and mothers’ dementia onset ages. However, there was little evidence against a comparable strength of association for two late onsets as for two early onsets. With small sample sizes in most cells, few of the estimated associations differed significantly from the null of  $\theta_{CS}(j; 1, 1) = 1$ .

Bandeem-Roche & Liang (2002) estimated the cause-specific cross ratio as 8.86 for times with joint first-event survival probability greater than 0.80, and of the order of 2.5 for times with joint survival probability no greater than 0.80. Their analysis may have understated heritability of late-onset dementia, because comparably, late times and disparate times may both have low joint survival probability. The assumptions made by copula-based association analysis appear inadequate for the current example.

## 5. BIN CHOICE AND STUDY DESIGN

Our methodology requires the choice of cut-points defining failure time ‘bins.’ For our dementia analysis and larger simulation runs, we originally attempted estimation on a  $5 \times 5$  time grid. This partitioning yielded zero-count cells. The number of ( $a, b$ ) pairings with (early failure, late failure) componentwise minima decreases with degree of discrepancy between ‘early’ and ‘late’, and the number of ( $a, b$ ) pairings with (late, late) componentwise minima declines with time. Ultimately, a partitioning by equally spaced marginal quantiles is not optimal.

Beyond ad hoc considerations, it will sometimes be necessary to design studies, ensuring that strength of association is estimated with suitable precision in given regions of space. While full elaboration is beyond this paper’s scope, a tractable design basis emerges if we multiply and divide the asymptotic variance (6) by  $\bar{E}_c^2$ . Given  $\bar{E}_c / \bar{E}_D = \theta_{CS}^*(j; k, l)$ , (6) equals

$$\frac{4}{n} \{\theta_{CS}^*(j;k, l)\}^2 [\zeta_{1C} - 2\eta_1 \{\theta_{CS}^*(j;k, l)\} + \zeta_{1D} \{\theta_{CS}^*(j;k, l)\}^2] / \bar{E}_C^2. \quad (9)$$

Candidates for  $\theta_{CS}^*(j;k, l)$  determine candidates for  $\bar{E}_C$ . Then, to complete (9), one must obtain candidates for  $\zeta_{1C}$ ,  $\zeta_{1D}$  and  $\eta_1$ . While these may be both complicated and unknown, Appendix 1 provides a template for their approximation with pilot data on bivariate failure location and cause frequencies, and marginal failure time distributions.

## 6. DISCUSSION

Our methodology accurately estimates failure time associations over a range of sample sizes and associations encountered in practice. Its accuracy appears comparable, and its precision nearly so, to that of estimation as proposed by Fan et al., (2000). Relative to survival function and cumulative hazard plug-in counterparts, our estimator's strengths include its ready interpretation and simplicity. We do not intend it as a replacement for alternative methods, but as a potentially more easily interpreted and readily implemented complement to them.

Our strategy handles censoring as a distinct failure cause. Doing this more efficiently is an advantage of competing methods, such as that of Fan et al., (2000). Doing so for our method is complicated, because censoring introduces uncertainty into both the determination of concordance and the value of  $(x, y)$  to which a given determination should be assigned.

Through (5), our estimator up-weights regions of less strongly positive association in summarizing time-varying strength of association over a bin. This effect is not due to competing risks, and thus, also prevails for Kendall's tau-based cross-ratio estimators. That there may be a need to delineate estimation targets when cross-ratio constancy is mistakenly assumed, was highlighted in our positive stable simulations, where the Fan estimator appeared to down-weight the conditional hazard ratio more strongly than ours relative to an unweighted average over a bin.

Rather than binning, one might kernel-smooth counts defining our estimator's dividends, and thus, obtain pointwise association estimates. So long as bandwidth choice is a priori, inferences go through, as in the current paper. However, for non-constant association, we find kernel-smoothed estimators to be quite biased.

Our failure to find strongly for greater familial aggregation of early-, than late-onset dementia contrasts with a recent study (Silverman et al., 2005). Comparison is complicated by substantial differences in the sampling strategies and analytical methodologies used in the two studies. However, our analysis cautions against too strongly downplaying familial aggregation in later-onset dementia.

## Acknowledgments

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

## APPENDIX 1

### Elucidation of variance terms

As a first step, it is useful to write the respective means in a different format from that given preceding (4). We begin with the concordance, i.e. numerator term. Note that, the compound



event  $\{(X_a - X_b)(Y_a - Y_b) > 0, (X_{(ab)}, Y_{(ab)}) \in B_j, (K_{(ab)}, L_{(ab)}) = (k, l)\}$  occurs, if and only if there occurs  $\{(X_a < X_b), (Y_a < Y_b), (X_{(ab)}, Y_{(ab)}) \in B_j, (K_{(ab)}, L_{(ab)}) = (k, l)\}$  and/or  $\{(X_a > X_b), (Y_a > Y_b), (X_{(ab)}, Y_{(ab)}) \in B_j, (K_{(ab)}, L_{(ab)}) = (k, l)\}$ . Then,

$$\begin{aligned} & E[\mathcal{S}\{(K_{(ab)}, L_{(ab)})=(k, l)\} \cdot \mathcal{S}\{(X_{(ab)}, Y_{(ab)}) \in B_j\} \cdot \mathcal{S}\{(X_a - X_b)(Y_a - Y_b) > 0\}] \\ = & E_a E_{b|a}[\mathcal{S}\{(X_a < X_b), (Y_a < Y_b), (X_{(ab)}, Y_{(ab)}) \in B_j, (K_{(ab)}, L_{(ab)})=(k, l)\} + \mathcal{S}\{(X_a > X_b), (Y_a > Y_b), (X_{(ab)}, Y_{(ab)}) \in B_j, (K_{(ab)}, L_{(ab)})=(k, l)\}] \\ = & E_a[\mathcal{S}\{(X_a, Y_a) \in B_j, (K_a, L_a)=(k, l)\} S(X_a, Y_a)] + E_b E_{a|b}[\mathcal{S}\{(X_a > X_b), (Y_a > Y_b), (X_{(ab)}, Y_{(ab)}) \in B_j, (K_{(ab)}, L_{(ab)})=(k, l)\}] \\ = & 2E_a[\mathcal{S}\{X_a, Y_a \in B_j, (K_a, L_a)=(k, l)\} S(X_a, Y_a)], \end{aligned} \tag{A1}$$

where  $E_{b|a}$  denotes expectation with respect to  $R_b$  conditioning on  $R_a$ , and so on. Term  $\zeta_{1C}$  follows from the second line of (A1), replacing  $E_a$  with  $\text{var}_a$ . If we define

$$\mu_b\{B_j \cap (0, x_a) \otimes (0, y_a); k, l\} = E_b[\mathcal{S}\{x_a > X_b, y_a > Y_b, (X_a, Y_b) \in B_j, (K_b, L_b)=(k, l)\}],$$

i.e. the probability that  $(X_b, Y_b)$  is a  $(k, l)$ -type failure occurring in the intersection of  $B_j$  and the quadrant  $\{0 < x < X_a, 0 < y < Y_a\}$ , then,

$$\zeta_{1C} = E_a[\mathcal{S}\{(X_a, Y_a) \in B_j, (K_a, L_a)=(k, l)\} S(X_a, Y_a) + \mu_b\{B_j \cap (0, X_a) \otimes (0, Y_a); k, l\} - \bar{E}_c]^2.$$

To elucidate discordance terms concisely, we define one-dimensional analogues of a few already defined quantities. First, we define ‘slices’ of  $B_j$ : let  $B_{j1}(y)$ , be the set of  $x$ -axis values, such that  $(x, y) \in B_j$ , and analogously 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  $y$  ( $x$ ). Secondly, denote one-dimensional regions where bin slices, on the real line, intersect  $0 : z$  interval line segments  $\tau_{0:z}$  by  $\delta_1\{(z, y)\} = B_{j1}(y) \cap \tau_{0:z}$  and  $\delta_2\{(x, z)\} = B_{j2}(x) \cap \tau_{0:z}$ . Then, if we assume that all regions in question are measurable, the discordance analogues of the concordance-related quantities (6) and (7) follow:

$$\begin{aligned} \bar{E}_D &= E_a[\mathcal{S}\{X_a \in B_{j1}, K_a=k\} \mu_b\{(X_a, \infty) \otimes \delta_2(X_a, Y_a); \cdot, l\} + \mathcal{S}\{Y_a \in B_{j2}, L_a=l\} \mu_b\{\delta_1(X_a, X_b) \otimes (Y_a, \infty); k, \cdot\}]; \\ &= 2E_a[\mathcal{S}\{X_a \in B_{j1}, K_a=k\} \mu_b\{(X_a, \infty) \otimes \delta_2(X_a, Y_a); \cdot, l\}]; \\ &= 2E_a[\mathcal{S}\{Y_a \in B_{j2}, L_a=l\} \mu_b\{\delta_1(X_a, Y_a) \otimes (Y_a, \infty); k, \cdot\}]; \\ \zeta_{1D} &= E_{(a)}[\mathcal{S}\{X_a \in B_{j1}, K_a=k\} \mu_b\{(X_a, \infty) \otimes \delta_2(X_a, Y_a); \cdot, l\} + \mathcal{S}\{Y_a \in B_{j2}, L_a=l\} \mu_b\{\delta_1(X_a, Y_a) \otimes (Y_a, \infty); k, \cdot\} - \bar{E}_D]^2; \\ \zeta_{2D} &= \bar{E}_D(1 - \bar{E}_D), \end{aligned}$$

where ‘ $\cdot$ ’ 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  $B_{j1} \cap \tau_{0:z}$ , and similarly for  $D_2\{(x, z)\}$ . The expansion for  $\eta_1$  is similar to that for the variance terms, and we omit it.

## APPENDIX 2

### Asymptotic independence across causes and bins

Consider cause-specific cross ratios for the same, ‘ $j$ th’ bin but different cause combinations; without loss of generality,  $(k, l) = (1, 1)$  and  $(1, 2)$ . We have,

$$E \left\{ \frac{\sum_{a < b} h_{C(j,1,1)}(R_a, R_b) \sum_{a < b} h_{C(j,1,2)}(R_a, R_b)}{\sum_{a < b} h_{D(j,1,1)}(R_a, R_b) \sum_{a < b} h_{D(j,1,2)}(R_a, R_b)} \right\} - E\{\widehat{\theta}_{CS}(j; 1, 1)\}E\{\widehat{\theta}_{CS}(j; 1, 2)\}$$

$$= \frac{\binom{n}{2}^{-1} \{2(n-2)\rho_{C1} + \rho_{C2}\} + E\{h_{C(j,1,1)}(R_a, R_b)\}E\{h_{C(j,1,2)}(R_a, R_b)\}}{\binom{n}{2}^{-1} \{2(n-2)\rho_{D1} + \rho_{D2}\} + E\{h_{D(j,1,1)}(R_a, R_b)\}E\{h_{D(j,1,2)}(R_a, R_b)\}} - E\{\widehat{\theta}_{CS}(j; 1, 1)\}E\{\widehat{\theta}_{CS}(j; 1, 2)\},$$

where  $\rho_{C1} = \text{cov}\{h_{C(j,1,1)}(R_a, R_b), h_{C(j,1,2)}(R_a, R_b)\}$ ,  $a < b \neq i$ ;  $\rho_{C2} = \text{cov}\{h_{C(j,1,1)}(R_a, R_b), h_{C(j,1,2)}(R_a, R_b)\}$ , and similarly for  $\rho_{D1}$  and  $\rho_{D2}$  with respect to discordance kernels. The last of these terms has summands that converge to equal limits, equal to

$$\frac{E\{h_{C(j,1,1)}(R_a, R_b)\}E\{h_{C(j,1,2)}(R_a, R_b)\}}{E\{h_{D(j,1,1)}(R_a, R_b)\}E\{h_{D(j,1,2)}(R_a, R_b)\}}.$$

Hence,  $\text{cov}\{\widehat{\theta}_{CS}(j; 1, 1), \widehat{\theta}_{CS}(j; 1, 2)\}$  converges to 0. Asymptotic independence follows with asymptotic normality. The proof for independence of cause-specific cross ratios across bins follows sufficiently similarly, that we omit it.

## REFERENCES

- Bandeen-roche K, Liang KY. Modelling multivariate failure time associations in the presence of a competing risk. *Biometrika* 2002;89:299–314.
- Barbe P, Genest C, Ghoudi K, RÁillard B. On Kendall's process. *J. Multivariate Anal* 1996;58:197–229.
- Benichou J, Gail MH. Estimates of absolute cause-specific risk in cohort studies. *Biometrics* 1990;46:813–826. [PubMed: 2242416]
- Breitner JCS, Wyse BW, Anthony JC, Welsh-Bohmer KA, Steffens DC, Norton MC, Tschanz JT, Plassman BL, Meyer MR, Skoog I, Khachaturian A. APOE-ε4 count predicts age when prevalence of AD increases, then declines: The Cache County Study. *Neurology* 1999;53:321–331. [PubMed: 10430421]
- Chen MC, Bandeem-Roche K. A diagnostic for association in bivariate survival models. *Lifetime Data Anal* 2005;11:245–264. [PubMed: 15940823]
- Cheng Y, Fine JP. Nonparametric estimation of cause-specific cross hazard ratio with bivariate competing risks data. *Biometrika*. 2008 to appear.
- 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 D, Cuzick J. Multivariate generalizations of the proportional hazards model. *J. Roy. Statist. Soc. A* 1985;148:82–108.
- Fan J, Prentice RL, Hsu L. A class of weighted dependence measures for bivariate failure time data. *J. R. Stat. Soc. B* 2000;62:181–190.
- Fine JP, Jiang H. On association in a copula with time transformations. *Biometrika* 2000;87:559–571.
- Genest C, Ghoudi K, Rivest L-P. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* 1995;82:543–552.
- Genest C, Mackay J. Copules archimediennes et familles de lois bidimensionnelles dont les marges sont données. *Canad. J. Statist* 1986;14:145–159.
- Genest C, Rivest L-P. Statistical inference procedures for bivariate Archimedean copulas. *J. Amer. Statist. Assoc* 1993;88:1034–1043.
- Glidden DV. A two-stage estimator of the dependence parameter for the Clayton–Oakes model. *Lifetime Data Anal* 2000;6:141–156. [PubMed: 10851839]

- Hendrie HC. Epidemiology of dementia and Alzheimer's disease. *Am. J. Geriatr. Psychol* 1998;6:S3–S18.
- Hsu L, Prentice RL. On assessing the strength of dependency between failure time variates. *Biometrika* 1996;83:491–506.
- Kendall, MG. *Rank Correlation Methods*. 2nd ed.. London: Griffin; 1955.
- Lee L. Multivariate distributions having Weibull properties. *J. Multivariate Anal* 1979;9:267–277.
- Nielsen GG, Gill RD, Andersen PK, Sorensen TIA. A counting process approach to maximum likelihood estimation in frailty models. *Scand. J. Statist* 1992;19:25–43.
- Oakes D. A model for association in bivariate survival data. *J. R. Stat. Soc. 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. *J. Am. Statist. Assoc* 1989;84:487–493.
- Prentice RL, Cai J. Covariance and survivor function estimation using censored multivariate failure time data. *Biometrika* 1992;79:495–512.
- Prentice RL, Hsu L. Regression on hazard ratios and cross ratios in multivariate failure time analysis. *Biometrika* 1997;84:349–363.
- Prentice RL, Kalbfleisch JD, Peterson AVPJR, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. *Biometrics* 1978;34:541–554. [PubMed: 373811]
- Ripatti S, Larsen K, Palmgren J. Maximum likelihood inference for multivariate frailty models using an automated Monte Carlo EM algorithm. *Lifetime Data Anal* 2002;8:349–360. [PubMed: 12471944]
- Ripatti S, Palmgren J. Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics* 2000;56:1016–1022. [PubMed: 11129456]
- Serfling, RJ. *Approximation Theorems of Mathematical Statistics*. New York: Wiley; 1980.
- Shieh GS. A weighted Kendall's tau statistic. *Statist. Probab. Lett* 1998;39:17–24.
- Shih JH, Louis TA. Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* 1995;51:1584–1599.
- Silverman JM, Cirese G, Smith CJ, Marin DB, Schaidt-beeri M. Variability of familial risk of Alzheimer disease across the late life span. *Arch. Gener. Psychol* 2005;62:565–573.
- Viswanathan B, Manatunga AK. Diagnostic plots for assessing the frailty distribution in multivariate survival data. *Lifetime Data Anal* 2001;7:143–155. [PubMed: 11458654]
- Wang W, Wells MT. Estimation of Kendall's tau under censoring. *Statist. Sinica* 2000;10:1199–1215.

*Simulation study. Estimation of  $\theta_{CS}(t_1, t_2; 1, 1)$  from data corresponding to gamma size copula, with  $\theta(t) = 2$ , beta shape frailty, with mean equal to  $R_1$  and scale equal to 1, and  $\text{Ex}(1)$  as conditional baseline distribution. Sample size is  $n$  in each of the 500 runs. No non-competing censoring*

**Table 1**

Quadrant	$n$	$R_1 = 0.2; \theta_{CS}(1, 1) = 6$			$R_1 = 0.8; \theta_{CS}(1, 1) = 2.25$				
		Mean	SD <sub>E</sub>	SD <sub>M</sub>	Cov	Mean	SD <sub>E</sub>	SD <sub>M</sub>	Cov
1 : $(t_1, t_2) \leq$ medians	1000	6.11	1.06	1.00	0.96	2.24	0.16	0.16	0.95
	500	6.30	2.16	1.93	0.96	2.31	0.31	0.31	0.95
2 : $t_1 \leq$ median $t_2 >$ median	1000	6.12	1.64	1.60	0.93	2.27	0.24	0.24	0.94
	500	7.18	4.39	4.16	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	2.25	0.24	0.24	0.94
	500	7.06	6.59	4.96	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	2.26	0.20	0.20	0.96
	500	7.16	4.27	3.85	0.93	2.31	0.40	0.39	0.95

$\theta_{CS}(t_1, t_2; 1, 1)$ , cause-specific cross ratio function as in equation (2);  $\theta(t)$ , cross ratio function; SD<sub>E</sub>, square root of the average of variance estimates over 500 runs; SD<sub>M</sub>, empirical standard deviation of estimates over 500 runs; Cov, empirical confidence interval coverage.

**Table 2**

*Simulation study. Estimation of  $\theta_{CS}(t_1, t_2; 1, 1)$  from data corresponding to positive stable size copula, with parameter equal to  $\alpha$ , beta shape frailty, with mean equal to 0.5 and scale equal to 1, and  $Ex(1)$  as conditional baseline distribution. Sample size is  $n$  in each of the 500 runs. No non-competing censoring*

Quadrant	$n$	$\alpha = 0.4;$ $\{\theta_{CS}(Q; 1; 1, 1), \dots, \theta_{CS}(Q4; 1, 1)\}$ $= \{9.24, 3.68, 3.66, 2.89\}$		$\alpha = 0.8;$ $\{\theta_{CS}(Q1; 1, 1), \dots, \theta_{CS}(Q4; 1, 1)\}$ $= \{2.66, 1.80, 1.80, 1.69\}$					
		Mean	SD <sub>E</sub>	SD <sub>M</sub>	Cov	Mean	SD <sub>E</sub>	SD <sub>M</sub>	Cov
1 : $(t_1, t_2) \leq$	1000	9.24	1.00	1.04	0.93	2.67	0.29	0.27	0.96
medians	500	9.41	1.92	1.99	0.94	2.79	0.56	0.59	0.94
2 : $t_1 \leq$ median $t_2 >$ median	1000	3.76	0.73	0.75	0.94	1.81	0.25	0.25	0.95
3 : $t_1$ median, $t_2 \leq$ median	500	3.93	1.52	1.49	0.94	1.89	0.51	0.52	0.94
4 : $(t_1, t_2)$	1000	2.89	0.33	0.35	0.93	1.72	0.23	0.23	0.95
> medians	500	2.98	0.66	0.64	0.96	1.75	0.47	0.46	0.95

$\theta_{CS}(t_1, t_2; 1, 1)$ , cause-specific cross ratio function as in equation (2); SD<sub>E</sub>, square root of the average of variance estimates over 500 runs; SD<sub>M</sub>, empirical standard deviation of estimates over 500 runs; Cov, empirical confidence interval coverage.

**Table 3**

*Simulation study. Estimation of  $\theta(t)$  from data corresponding to copula models in the absence of competing risks, with  $\text{Ex}(1)$  as conditional baseline distribution. Sample size is  $n$  in each of the 500 runs. In censored data runs, censoring is random according to elementwise, independent  $\text{Ex}(1)$  distributions.*

Association Model	$n$	Censored	$(t_1, t_2) \leq$ lower quartiles			$(t_1, t_2) \leq$ upper quartiles			
			Our estimator	Fan estimator	SD <sub>E</sub>	Our estimator	Fan estimator	SD <sub>E</sub>	
Independent	1000	N	Mean	1.00	0.13	1.00	0.13	1.00	0.05
			SD <sub>E</sub>	0.13	1.00	0.15	1.00	0.08	1.00
$\theta(t) = 1$	100	Y	Mean	1.05	0.43	1.05	0.42	1.01	0.16
			SD <sub>E</sub>	0.43	1.05	0.51	1.01	0.19	1.02
Gamma	1000	N	Mean	2.02	0.20	2.02	0.20	2.00	0.09
			SD <sub>E</sub>	0.20	2.02	0.23	2.00	0.15	1.99
$\theta(t) = 2.0$	100	Y	Mean	2.11	0.73	2.08	0.69	2.01	0.31
			SD <sub>E</sub>	0.73	2.11	0.82	2.07	0.52	2.05
Pos-stable	1000	N	Mean	11.3	1.10	10.1	1.00	4.53	0.24
			SD <sub>E</sub>	1.10	11.3	1.26	6.31	0.49	3.95
$\alpha = 0.4$	100	Y	Mean	12.2	4.27	9.80	3.05	4.56	0.75
			SD <sub>E</sub>	4.27	12.2	3.97	6.69	1.64	4.11
Pos-stable	limit	N	Mean	11.2	NA	10.1	NA	4.51	NA
			SD <sub>E</sub>	NA	11.2	0.26	1.56	0.08	1.28
$\alpha = 0.8$	1000	Y	Mean	2.53	0.28	2.32	0.28	1.82	0.13
			SD <sub>E</sub>	0.28	2.53	0.80	1.57	0.23	1.29
Pos-stable	100	Y	Mean	2.62	0.90	2.38	0.80	1.86	0.46
			SD <sub>E</sub>	0.90	2.62	1.00	1.86	0.46	1.46
Pos-stable	limit	N	Mean	2.52	NA	2.33	NA	1.55	NA
			SD <sub>E</sub>	NA	2.52	NA	1.55	NA	1.28

$\theta(t)$ , cross ratio function; SD<sub>E</sub>, square root of the average of variance estimates over 500 runs.

**Table 4**

*Data application. Estimation of  $\theta_{CS}(t_1, t_2; 1, 1)$  from data collected in the Cache County Study, over a  $2 \times 2$  time grid. Bootstrap is based on 1000 replicates. Bootstrap confidence interval is bias-corrected and percentile-based.*

Quadrant	Mean	SD <sub>E</sub>	SD <sub>M</sub>	SD <sub>B</sub>	95% CI-asymptotic	95% CI-bootstrap
1 : $t_1 \leq 75, t_2 \leq 80$	3.81	1.19	1.23	1.23	(1.48, 6.14)	(1.68, 6.22)
2 : $t_1 \leq 75, t_2 > 80$	0.80	0.54	0.53	0.57	(-0.27, 1.86)	(0.00, 1.87)
3 : $t_1 > 75, t_2 \leq 80$	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)

SD<sub>E</sub>, asymptotic standard deviation approximation; SD<sub>M</sub>, square-root of the average of asymptotic variance approximation estimates over bootstrap replicates; SD<sub>B</sub>, bootstrap standard deviation estimate; CI, confidence interval.