STATISTICAL METHODS FOR MULTIVARIATE FAILURE-TIME DATA UNDER COMPETING RISKS

by

Jeongyong Kim

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

March, 2017

© Jeongyong Kim 2017

All rights reserved

Abstract

Traditional research on survival analysis often centered on univariate data where the observations are mutually independent. In many modern studies, however, data of interest are observed in clusters, so may be associated. Primary scientific interest often centers on the effect of a treatment on the individuals' outcomes in studies involving multivariate failure time data, but this thesis is mainly concerned with analyses in which the estimation of association between failure times is of interest. A considerable body of literature has addressed this topic, but they have been limited in many ways. They may depend on parametric assumptions that may easily be violated, they may not be flexbile enough, or their interpretations are not intuitive. The primary purpose of this thesis is to investigate the drawbacks of existing methods, and suggest an alternative measure of association that is flexible and interpretable, especially under the competing risks setting. This thesis consists of three main chapters. Chapter 2 discusses a nonparametric estimation of the local version of Kendall's τ. The performance of several smoothing methods are compared, and new methods to deal with censored data are also proposed and assessed. Chapter 3 studies the sensitivity of the Bandeen-Roche and Liang (2002) estimator of the CCSHR to the imposed statistical assumptions and investigate the source of a bias reported in its foundational work. In Chapter 4, novel parametric and nonparametric estimators for the association between failure causes are proposed. Various combinations of existing and new methods for the association between failure times and between failure causes are assessed.

Advisor:

Karen Bandeen-Roche, Ph.D.

Committee:

Terri Beaty, Ph.D. (chair); Mei-Cheng Wang, Ph.D.; Emily Agree, Ph.D.

Alternates:

Daniel O. Scharfstein, Sc.D.; George W. Rebok, Ph.D.

Acknowledgements

During the long arduous journey of Ph.D. program, I have been helped and motivated by many people. It was a great honor for me to have an opportunity to study with world-class faculty and classmates. I was also very lucky to have emotionally supportive friends and family in the United States and my home country, Korea. Most of all, I would like to express my deep appreciation to my advisor, Dr. Bandeen-Roche, who has always been inspiring and patient during my entire Ph.D. life.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Until recent decades, research on survival analysis mainly addressed univariate data in which a main statistical assumption is the mutual independence of the observations, conditional on covariates to be considered. In many modern studies, however, the data of interest are observed in clusters, so may be mutually dependent. Such multivariate failure time data arise in many fields. In biomedical studies, for example, one may have interest in ages at onset of schizophrenia in relatives, times to the occurrence of blindness in the left and right eyes in patients with diabetic retinopathy, or time to coronary heart disease and time to cerebrovascular accident. In other disciplines, times-to-default for closely connected companies in economics and times-to-failure of multiple components in a system in reliability engineering are often of interest.

In studies involving multivariate failure times, primary scientific interest often centers on the effect of a treatment on the individuals' outcomes, and the association within clusters is considered as a technicality that should be addressed in analyses. The study of within-cluster association among failure times, however, also may have importance in its own right. This may provide clues, for example, into genetic heritability, geographical trends, or environmental risk factors of health outcomes. This thesis primarily is concerned with analyses in which the estimation of association is of primary interest.

This thesis also addresses a second common challenge for multivariate survival analysis in health research: persons may be at risk for multiple types of failure, of which only one type that occurs first is observable. In studies of dementia in older adults, for example, participants may become demented, but many die without experiencing dementia. In such competing risks data, it typically is not natural to view the failure type of interest as being independent of the others. Then, multivariate survival analysis has an advantage over univariate analysis. In univariate analysis of competing risks, only one type of failure can be observed per sampling unit, hence no information pertaining to association between risks is available. With multivariate data, multiple failures of either the same or different types can be observed in a cluster, providing empirical information about the associations between causes.

A considerable body of literature has addressed association in multivariate failure time analysis. Available methods, however, have been limited in important ways. Parametric methods have been shown to be sensitive to model assumptions such as form of the time dependence of the association or distribution of failure times. Nonparametric estimators of association often have had complex interpretation. This thesis aims to address these gaps. In the following subsections, I will provide key overarching background and identify specific topics to be studied.

1.1 Prior work on the estimation of association among clustered failure times

The most general characterization of association among multivariate failure times is provided by the multivariate survival function $S(t_1, t_2) = \Pr(T_1 > t_1, T_2 > t_2)$. Estimators have been proposed and studied by Dabrowska (1988), Pruitt (1991), Prentice and Cai (1992), van der Laan (1996), Prentice (2014), and others. However, their implementation and interpretation may be complex because they do not distinguish marginal incidence information and association information.

Employing a simple summary measure can ameliorate this problem. The conditional hazard ratio (CHR), $\theta(t_1, t_2) = \dfrac{\lambda(t_2|T_1 = t_1)}{\lambda(t_2|T_1 > t_1)}$, first proposed by Clayton (1978), has become one of the most popular measures of association for multivariate failure-time data. It can be interpreted as the ratio of an individual's hazard of failure at t_2 given failure of his paired partner at t_1 to the hazard given that the partner has not yet failed by time t_1. It also has been shown to follow a one-to-one relationship to a local version of Kendall's tau. One of the main advantages of the CHR is that, for Archimedean copula distributions, it depends on time t_1 and t_2 only through the joint survival function (Oakes, 1989). Many subsequent authors have presented approaches to model the CHR parametrically (e.g., Clayton & Cuzick, 1985; Hougaard, 1986). Parametric measures of association, however, constrain the form of the time dependence. In some instances, a more general time-dependent characterization of association may be needed. In a genetic study, for example, a researcher may believe that a certain gene influences risk only in old age. Prentice and Cai (1992), Hsu and Prentice (1996), Fan, Prentice, and Hsu (2000), and others have provided measures of association flexibly indexed by time to detect such an effect. However, their interpretation may be complex or the measures may address time-dependence only in discrete 'bins' of failure time. Thus, I aimed to develop a nonparametric estimator of association between bivariate failure times to address these limitations.

1.2 Prior work on the estimation of failure time associations in the presence of competing risks

The decade 2000-2010 was an active period for research to develop measures of association in the competing risks setting. For example, Cheng, Fine, and Kosorok (2007, 2009) proposed methods by which to nonparametrically estimate the bivariate cumulative cause-specific hazard function and the bivariate cumulative incidence function from bivariate failure-time data with competing risks; they developed two measures of association based on these quantities. Scheike et al. (2010) proposed a cross-odds ratio function as a measure of association between cause-specific failure times, which they defined as the ratio of the conditional odds of occurrence of one cause-specific event for one cluster member given occurrence of the same or different cause-specific event for another cluster member, over the unconditional odds of occurrence of the cause-specific event.

My dissertation focuses on the conditional cause-specific hazard ratio (CCSHR) proposed by Bandeen-Roche and Liang (2002), which is a modified version of the conditional hazard ratio. One unique and appealing feature of the Bandeen-Roche and Liang paper is the decomposition of the association between cause-specific failure times into two elements – the association between the causes of failure, and the association between the times to failure. However, application of the original work proved too restrictive in two ways. First, the proposed decomposition was implemented via a fully parametric model which appeared overly restrictive for a number of plausible data generation scenarios. I considered it worthwhile to more extensively study estimator performance if the defining distributional assumptions are violated to inform development of alternative models as needed. Secondly, the strength of association between failure causes may vary with time, but the implementation of such time-dependence was limited in the original work. Thus, I aimed to develop methods to flexibly estimate the time-variation in this association.

1.3 Organization of the thesis

The thesis consists of three research chapters addressing the issues discussed above. Chapter 2 discusses a nonparametric approach to the estimation of the local version of Kendall's τ, hence with a simple transformation, the CHR. The proposed method makes use of all available parings of bivariate failure times and operates by smoothing data on concordance and discordance. Several smoothing methods are applied, and their performances are compared. New methods to deal with censored data are also proposed and assessed. Chapter 3 studies the sensitivity of the Bandeen-Roche and Liang (2002) estimator of the CCHSR to the imposed statistical assumptions and investigates the source of a bias reported in its foundational work. In Chapter 4, I use the method in Chapter 2 to estimate the times-to-failure component of the CCSHR nonparametrically, and propose novel parametric and nonparametric estimators for the causes-of-failure association. The performance of the new and existing methods are compared. Chapter 5 summarizes what each chapter achieved, discusses their strength and weakness, and suggests future work .

Chapter 2

Nonparametric Estimation of Association in Bivariate Failure-time Data

2.1 Introduction

In studies focused on familial or geographic determinants of health, multimorbid diseases or settings nesting patients within health care providers, failure-time data may occur as clustered observations. In such studies, primary interest often centers on the effect of a treatment or exposure on individuals' outcomes, and association within clusters is treated as a technicality that must be addressed in analyses. In contrast, our concern is with data for which estimation of association is of interest in its own right. Then, within-cluster association may provide clues into genetic heritability, geographical and environmental risk factors, or practice-level determinants of health outcomes.

Among methods characterizing failure-time association, bivariate survival function estimators have been developed and studied by many authors (Dabrowska, 1988; Pruitt, 1991; Prentice & Cai, 1992; van der Laan, 1996; Prentice, 2014). However, their implementation and interpretation may be complex, and they blend marginal incidence information and association information. To address these issues, a number of summary measures of association have been proposed. The conditional hazard ratio, $\theta(t_1, t_2) = \dfrac{\lambda(t_1|T_2 = t_2)}{\lambda(t_1|T_2 > t_2)}$, has become one of the most popular, because for Archimedean copula distributions it depends on time t_1 and t_2 only through the bivariate survival function. (Oakes, 1989). Clayton's (1978) model for association provides the earliest, time-invariant proposal of this measure. Many subsequent authors have presented approaches to model the conditional hazard ratio parametrically (e.g., Clayton & Cuzick, 1985; Hougaard, 1986) but these constrain the form of the time dependence. However, more general time-dependent association measures may be of interest. In a genetic study, for example, researchers may believe that a certain gene influences risk only in old age. Association measures flexibly indexed by time could provide means of detecting such an effect.

Considerable work to describe failure-time association non- or semi-parametrically has been reported. Oakes (1989) provided a definition of the conditional hazard ratio as a local version of Kendall's τ. Anderson et al. (1992) proposed three time-dependent measures of association, including the conditional hazard ratio of Oakes (1989), conditional expected residual life, $\phi(t_1, t_2) = \dfrac{E(T_1|T_1 > t_1, T_2 > t_2) - t_1}{E(T_1|T_1 > t_1) - t_1}$, and conditional probability, $\psi(t_1, t_2) = \dfrac{\Pr(T_1 > t_1|T_2 > t_2)}{\Pr(T_1 > t_1)}$, where T_1 and T_2 are failure times for a randomly sampled pair. Sankaran, Abraham and Antony (2006) suggested a dependence measure based on a covariance residual life function, $C(t_1, t_2) = M(t_1, t_2) - r_1(t_1, t_2)r_2(t_1, t_2)$ where $M(t_1, t_2) = E[(T_1 - t_1)(T_2 - t_2)|T_1 > t_1, T_2 > t_2]$ and $r_i(t_1, t_2) = E[T_i - t_i|T_1 > t_1, T_2 > t_2]$, and a method of its nonparametric estimation. Nair and Sankaran (2010) suggested another measure of association, $\alpha(t_1, t_2) = \dfrac{M(t_1, t_2)}{r_1(t_1, t_2)r_2(t_1, t_2)}$, using the same definition of $M(t_1, t_2)$ and $r_i(t_1, t_2)$. Hsu and Prentice (1996) suggested the correlation between marginal martingales, $\rho^*(t_1, t_2) = \mathrm{Corr}(M_1(t_1), M_2(t_2))$ as a local measure of association where $M_i(t_i) = N_i(t_i) - \Lambda_i(t_i \wedge T_i)$, $N_i(t_i)$ is a binary variate of failure and $\Lambda_i(\cdot)$ is a cumulative hazard. These quantities generally are estimated by plugging nonparametric estimates of joint and marginal survival or cumulative hazard functions into these expressions.

Whereas the above measures quantify the strength of association at a specific time pair (T_1, T_2), other measures deal with the strength of association for a specific ‘region.’ Chen and Bandeen-Roche (2005) exploited Oakes’ (1989) idea to estimate the conditional hazard ratio in ‘bins’ of (bivariate) survival probability. Bandeen-Roche and Ning (2008) generalized such estimation to bins of bivariate time and to allow for competing risks; Cheng and Fine (2008) and Cheng, Fine, and Bandeen-Roche (2010) addressed the same estimand with alternative estimators. Fan, Prentice, and Hsu (2000) and Fan, Hsu, and Prentice (2000) proposed a new class of weighted dependence measures for bivariate failure time data. These measures characterize the strength of association in a bivariate failure time region $[0, t_1] \times [0, t_2]$; versions of them reduce to the reciprocal of conditional hazard ratio or of Kendall's τ when $t_1, t_2 \to \infty$ in the absence of censoring. Fan and Prentice (2002) generalized these measures to accommodate regression effects on marginal hazard functions.

Though many approaches have been proposed, these methods show only a fixed value of the association measure at a specific time point or region. However, researchers may wish to visualize time-dependent association for the entire time domain. We aimed to develop a method that displays association flexibly and interpretably. We propose smoothing to produce a 'map' of the local Kendall's τ over time to achieve this goal in bivariate data. The estimand has a ready interpretation described below, which transforms easily to the conditional hazard ratio. We evaluated candidate smoothing methods to create a map of association.

The remainder of this paper proceeds as follows. Section 2 introduces notation and relevant background. Section 3 introduces the smoothing methods and describes how we obtain estimates for evaluation in simulation studies. Section 4 reports a series of simulation studies to compare the smoothing methods, and Section 5 presents the application of these methods to data on dementia onset in families. Section 6 concludes.

2.2 Background

2.2.1 Notation

Let T_{ij} denote the failure time and C_{ij}, the time of censoring, for subjects $j = 1, 2$ in pair i. Then the observed event time is the minimum of failure and censoring times, $X_{ij} = T_{ij} \wedge C_{ij}$, and the failure-censoring indicator δ_{ij} is 1 if $T_{ij} < C_{ij}$ and 0 otherwise. $S(t_1, t_2)$ is the joint survival function of T_1 and T_2, and $S_1(t_1)$ and $S_2(t_2)$ are the marginal survival functions of T_1 and T_2, respectively. We assume the data are independently and identically distributed across pairs $i = 1, \cdots, n$ and censoring is independent of failure time.

2.2.2 Definitions

The conditional (cross) hazard ratio (CHR) is defined by $\theta(t_1, t_2) = \dfrac{\lambda(t_2 | T_1 = t_1)}{\lambda(t_2 | T_1 > t_1)} = \dfrac{f(t_1, t_2) S(t_1, t_2)}{\dfrac{\partial S(s_1, t_2)}{\partial s_1}\Big|_{s_1 = t_1} \cdot \dfrac{\partial S(t_1, s_2)}{\partial s_2}\Big|_{s_2 = t_2}}$. Oakes (1989) showed this measure can be viewed as a ratio of conditional probabilities that two bivariate failure time pairs are concordant or discordant given the componentwise minimum failure times. For two bivariate observations $T^{(a)} = (T_1^{(a)}, T_2^{(a)})$ and $T^{(b)} = (T_1^{(b)}, T_2^{(b)})$, denote the corresponding componentwise minimum $(T_1^{(a)} \wedge T_1^{(b)}, T_2^{(a)} \wedge T_2^{(b)})$ by $(T_1^{(ab)}, T_2^{(ab)})$. We say $T^{(a)}$ and $T^{(b)}$ are concordant if $(T_1^{(a)} - T_1^{(b)})(T_2^{(a)} - T_2^{(b)}) > 0$ and discordant if $(T_1^{(a)} - T_1^{(b)})(T_2^{(a)} - T_2^{(b)}) < 0$. Then, it can be shown that $\theta(t_1, t_2)$ equals $\dfrac{\Pr\{(T_1^{(a)} - T_1^{(b)})(T_2^{(a)} - T_2^{(b)}) > 0 | (T_1^{(ab)}, T_2^{(ab)}) = (t_1, t_2)\}}{\Pr\{(T_1^{(a)} - T_1^{(b)})(T_2^{(a)} - T_2^{(b)}) < 0 | (T_1^{(ab)}, T_2^{(ab)}) = (t_1, t_2)\}}$.

If we calculate this ratio without conditioning on the componentwise minimum, we obtain a 'global' CHR, θ. It then is easily seen that θ has a one-to-one relationship with Kendall's τ by $\tau = \dfrac{\theta - 1}{\theta + 1}$. Thus, these two measures share similar interpretation using the concept of concordance: Kendall's τ is defined as the difference of the probabilities that two identically distributed bivariate random vectors are concordant versus discordant, while the CHR is defined as the ratio of these probabilities. We can then think of 'local' Kendall's τ by applying conditioning on the componentwise minimum as above.

2.3 Estimation

To 'map' the local Kendall's τ over a failure time domain, we propose to directly smooth concordance and discordance data to produce a Kendall's τ 'surface' over the time domain. This then may be transformed to estimate the CHR by plugging into $\theta(t_1, t_2) = \dfrac{1 + \tau(t_1, t_2)}{1 - \tau(t_1, t_2)}$. Specifically, given any bivariate failure time data, we can obtain all available pairs of bivariate observations, and concordance status of the pairs. If we assign $+1$ to a concordant pair and -1 to a discordant pair, these data provide 'raw data' for smoothing. A smoothed function from these data can then be interpreted as a function of a 'local' version of Kendall's τ.

As a first step to creating association maps as we propose, a dataset of concordances and discordances must be created. We begin with a dataset consisting of observations of paired variables – the 1st and 2nd components of bivariate failure times. From these data, we can construct a new dataset which consists of pairwise minima between all available pairings of failure time pairs and concordance status of the pairing. For observations $T^{(i)} = (T_1^{(i)}, T_2^{(i)})$ and $T^{(j)} = (T_1^{(j)}, T_2^{(j)})$, the pairwise minimum is $(\min(T_1^{(i)}, T_1^{(j)}), \min(T_2^{(i)}, T_2^{(j)}))$, and the concordance status is $+1$ if $(T_1^{(i)} - T_1^{(j)})(T_2^{(i)} - T_2^{(j)}) > 0$ and -1 if $(T_1^{(i)} - T_1^{(j)})(T_2^{(i)} - T_2^{(j)}) < 0$. If there is a tie in either component of the pair, we may assign 0 as a concordance indicator for the tied pair, or if a binary outcome is desired, we may randomly assign $+1$ or -1 to break the tie. Then, smoothing methods can be applied to this data set to estimate an association function in terms of the failure times.

As the domain for our association function, we propose to use one minus Kaplan-Meier estimates of each failure time coordinate instead of the original failure times. By so doing, we standardize the bivariate failure times into a $[0, 1] \times [0, 1]$ space. This enables us to compare association structures of bivariate failure times with different ranges and to explore how well Archimedean copulas can fit the data. These estimates may be back-transformed to the raw time scale by a relabeling of axes with times corresponding to survival probability values. In summary, our smoothing procedure uses the concordance status indicators for pairs of bivariate observations as a response variable, and the pairwise minima of standardized failure times as explanatory variables.

2.3.1 Candidate smoothing methods

In a simulation study reported in the next section of this paper, we compared estimator performance among seven methods: Loess, logistic regression, four types of generalized additive models (GAM), and multivariate adaptive regression splines (MARS). For Loess and MARS, we propose to directly apply the method with concordance status (-1, 0, or 1) as a response variable and the pairwise minimum of standardized failure times of two pairs as explanatory variables. For other methods, tweaking is needed or variations are possible, as described below. We begin by describing application absent censoring, and then propose strategies to accommodate censoring.

In each case, we describe the method and then report implementations evaluated in simulation studies. All methods were implemented using R packages.

2.3.1.1 Loess

Loess is a local regression method that was developed as a flexible means for modeling central tendency of a response distribution conditional on covariates (Cleveland, 1979, 1988). It does this by fitting simple linear or quadratic regression models for a small subset, or nearest neighbors, of each response point, for all the x values where the Loess curve should be evaluated. Thus, there is a separate local regression for each value of x and the fitted values from these regressions are connected to produce the regression curve. Typically, a locally weighted linear regression or a locally weighted quadratic regression is used, but higher order polynomials or methods targeting measures of central tendency other than means may be used. In this paper, T_1 and T_2 are the explanatory variables. In simulation studies, the size of the neighborhood was fixed to 0.75 – the R default setting, and locally quadratic regression was used.

2.3.1.2 Logistic regression

Logistic regression deals with situations where the observed outcome for a response variable can have only two possible types, for example, true vs. false. It assumes that the logit of the probability (log odds) of success is linearly associated with the predictors. In this paper, a case (success) represents the concordance between two pairs of bivariate failure times and a non-case (failure) the discordance. Thus, the model is expressed as $\log\left(\frac{p(t_1, t_2)}{1 - p(t_1, t_2)}\right) = b_0 + b_1 t_1 + b_2 t_2$, where $p(t_1, t_2)$ is the probability that a pairing of pairs with componentwise minimum failure times (t_1, t_2) is concordant.

Whereas logistic regression requires the response variable should be zero and one, our response variables have values $+1$ for concordance and -1 for discordance. We transform our $\{-1, 1\}$ response variable to zero-one scale by $y' = \frac{1}{2}y + \frac{1}{2}$. To obtain the smoothed Kendall's τ estimate, the predicted values of the logistic regression are back-transformed to $[-1, 1]$ by $y = 2y' - 1$.

2.3.1.3 Generalized additive models

Generalized additive models (GAM) were originally developed to blend properties of generalized linear models (GLM) and smoothing (Hastie & Tibshirani, 1986). Recall that GLMs model a mean response by $g(E[Y]) = b_0 + b_1 x_1 + \cdots + b_m x_m$. The GAM replaces the simple linear terms $b_k x_k$ by $f_k(x_k)$ where f_k is an unspecified function, yielding $g(E(Y)) = f_1(x_1) + \cdots + f_m(x_m)$. This f_k may be a function with a specific parametric form or may be specified nonparametrically. In this paper, we use smoothing splines to model and estimate f_k. In simulation studies to follow, we used $k = 8$ as the dimension of the basis to control smoothness.

GAMs allow for various choices of ‘family’ (distribution) for the response variable and link function ($g(x)$) as GLMs do. In our simulation studies, we compared two choices: binomial family with $g(x) = \log\left(\frac{x}{1-x}\right)$ using the same conversion as for logistic regression, and Gaussian family with $g(x) = x$ (as if the values of the response variable $-1, 0$, and 1 were continuous). Moreover, we compared two ways to model the additive function of our two standardized time variables (x_1, x_2) in the GAM formula: specification in terms of univariate functions, $g(E(y)) = f_1(x_1)+f_2(x_2)$, and specification as a bivariate function (estimated by a bivariate smoothing spline), $g(E(y)) = f(x_1, x_2)$. In summary, we compared four ways to model these data using GAM (two ‘families’ by two linear predictor specifications). In the following sections, GAM with Gaussian family and two univariate functions will be labeled as GAM1, the one with Gaussian family and a bivariate functions as GAM2, the one with binomial family and two univariate functions as GAM3 and the one with binomial family and a bivariate functions as GAM4.

2.3.1.4 Multivariate Adaptive Regression Splines (MARS)

Multivariate adaptive regression splines (MARS) is a nonparametric regression method that makes no assumption about the underlying functional relationship between the response and predictor variables (Friedman, 1991). MARS builds this relation from a set of coefficients and basis functions derived from the data. The MARS description of the response variable mean typically has the form: $f(x) = \sum_{k=1}^{m} c_k B_k(x)$ where each c_k is a constant coefficient and $B_k(x)$ is a basis function. The basis function $B_k(x)$ has one of the three forms: a constant 1, a hinge function with the form $\max(0, x - c)$ or $\max(0, c - x)$ and a product of two or more hinge functions. Then the MARS algorithm automatically selects the variables and the location of knots, c in the hinge function, using a two-stage approach consisting of a forward and backward pass. The forward pass starts with a model which has only the intercept term. Then MARS repeatedly adds a pair of basis functions to the model that most decreases the residual sum of squares. This process continues until the decrease of residual sum of squares is sufficiently small or it reaches the maximum number of terms which is pre-specified by the user. In the forward pass, the maximum degree of interaction was chosen as 1, which equates to building an additive model.

The forward pass usually builds an overfitted model that has a good fit to the data used to build the model, but will not generalize well to new data. The backward pass is performed to build a model that generalizes better to new data. This procedure removes the least effective basis functions one-by-one from the model so that whose removal will lead to the least decrease in the goodness-of-fit. The backward pass continues until it finds the best submodel which is compared using a generalized cross validation (GCV) criterion.

2.3.2 Boundary of reliable estimation

When the two components of bivariate failure times are strongly correlated either positively or negatively, data may occur very sparsely in some portions of the bivariate time domain, and hence yield nonsensical estimates. So we will suggest a method to set a boundary where the estimation can be considered as valid. The basic idea of this method is to consider the estimation in a specific region to be trustworthy when the density of the data is greater than a specific criterion. This method is applied to the observed pairwise minima rather than the failure times themselves. The procedures are as follows:

First, suppose we already have the pairwise minima of the failure times as described in the Section 3.1. We recommend estimating the bivariate density function of these times, and then restricting estimation of failure time association within the region with density exceeding a criterion value. For this paper, we implemented using two-dimensional kernel density estimation in the function 'kde2d' in the R 'MASS' package. The criterion can be decided by various methods, for example, by thresholding at a multiple of the mean density for the entire time domain or the maximum of density estimates. In evaluating the best smoothing method in the next section, for example, we obtained a mean of bivariate density estimates from 300 replicates for each simulation scenario, and set an eighth of the maximum mean density as the criterion. This multiplying coefficient $1/8$ was decided by browsing scatterplots of simulated datasets: As an example, we present a scatterplot of a dataset of sample size 200 with the inside of the region marked with 'o' and the outside with 'x,' (Figure 2.1).

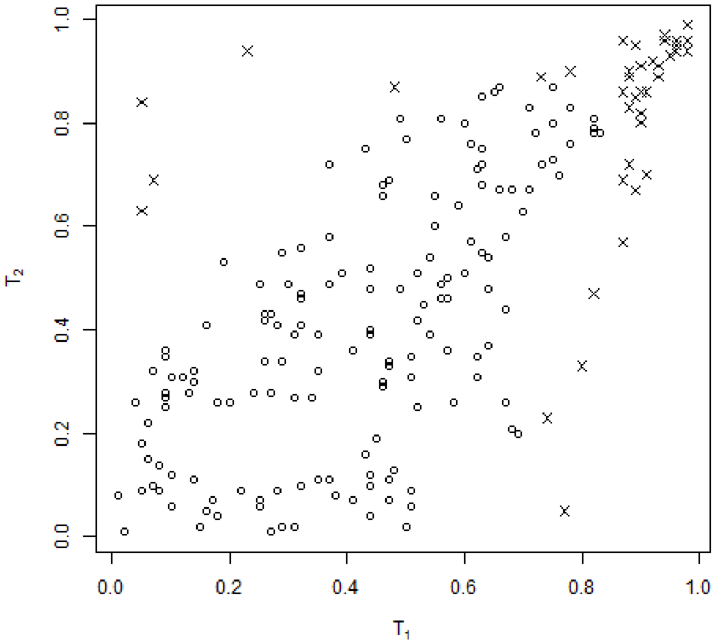


Figure 2.1: A scatterplot of an example dataset with the area of reliable estimation marked with ‘o’

2.3.3 Censoring

The method described so far assumes that there are no censored observations. To address censored data, we evaluated various methods. The first adapted the Brown et al. (1974) estimator of Kendall's τ. The second was multiple imputation using conditional bivariate density estimates derived from the Dabrowska (1988) estimator. The third was to utilize concordance information which can be decided from censored observations. More detailed description of these three methods are as follows.

The basic idea of the first method is to consider a censored observation may have had a larger failure time if it had not been censored and adjust the difference using a Kaplan-Meier estimator. Brown et al. (1974) regarded the concordance indicator as a product of two scores $a_{ij} = \Pr(T_{1i} > T_{1j}) - \Pr(T_{1i} < T_{1j}) = 2 \times \Pr(T_{1i} > T_{1j}) - 1$ and $b_{ij} = 2 \times \Pr(T_{2i} > T_{2j}) - 1$. For example, if $X_{1i} > X_{1j}$ and $\delta_{1j} = 1$, then we can tell that $T_{1i} > T_{1j}$ for certain which makes $a_{ij} = 1$. If $\delta_{1j} = 0$, however, we can utilize Kaplan-Meier estimates of X_i and X_j to obtain the expected value of the indicator $I(T_{1i} > T_{1j}) - I(T_{1i} < T_{1j})$. Then, the scores a_{ij} are defined as follows (Table 2.1); b_{ij} are defined similarly.

Table 2.1: Definition of a_{ij} from Brown et al. (1974)

$(\delta_{1i}, \delta_{1j})$	$X_{1i} > X_{1j}$	$X_{1i} = X_{1j}$	$X_{1i} < X_{1j}$
(1,1)	1	0	-1
(0,1)	1	1	$2 \times S_1(X_j)/S_1(X_i) - 1$
(1,0)	$1 - 2 \times S_1(X_i)/S_1(X_j)$	-1	-1
(0,0)	$1 - S_1(X_i)/S_1(X_j)$	$1 - S_1(X_i)/S_1(X_j)$	$S_1(X_j)/S_1(X_i) - 1$

The product of a_{ij} and b_{ij} matches the definition of the concordance indicator if there is no censoring and has a real value between $+1$ and -1 for censored observations. Brown et al.'s paper took an average of concordance scores $a_{ij} \times b_{ij}$ for all available pairings of bivariate failure times to obtain global Kendall's τ estimates. We used the score $a_{ij} \times b_{ij}$ for each pairing as a response variable and corresponding pairwise minimum of failure times as explanatory variables. This approach is intuitive, but may be biased because joint information is ignored.

In the second method, multiple imputation replaces each singly and doubly censored observation by random numbers and calculates local Kendall's τ using this imputed dataset. First, we obtain bivariate probability mass estimates for a rectangular grid, $B_{pq} = \left[\frac{p-1}{m}, \frac{p}{m}\right] \times \left[\frac{q-1}{m}, \frac{q}{m}\right],\ p, q = 1, 2, \cdots, m$, where B_{pq} is a unit rectangle for which the mass is estimated and m is an appropriately chosen positive integer considering the smoothness of bivariate mass estimates. The mass is calculated using Dabrowska estimates, i.e., $\Pr(B_{pq}) = \Pr\left(\frac{p-1}{m} < T_1 < \frac{p}{m}, \frac{q-1}{m} < T_2 < \frac{q}{m}\right) = S\left(\frac{p-1}{m}, \frac{q-1}{m}\right) - S\left(\frac{p}{m}, \frac{q-1}{m}\right) - S\left(\frac{p-1}{m}, \frac{q}{m}\right) + S\left(\frac{p}{m}, \frac{q}{m}\right)$. It is well known that Dabrowska's estimator may not be a proper survival function and have negative mass (Pruitt, 1991). Since negative and very small positive density may cause an error in the following processes, we replaced non-positive numbers by 1.0×10^{-10} and very small positive numbers (less than 1.0×10^{-5}) by 1.0×10^{-5}. The next step is to randomly choose a failure time for censored observations based on the bivariate mass estimates from the previous step. For a doubly censored observation (X_1, X_2), we pick a grid $B_{p'q'}$ $(p' \geq p, q' \geq q)$ with conditional bivariate density

$\Pr\left(t_1, t_2 \middle| t_1 > \frac{p-1}{m}, t_2 > \frac{q-1}{m}\right)$ where $\frac{p-1}{m} < X_1 < \frac{p}{m}, \frac{q-1}{m} < X_2 < \frac{q}{m}$. Then the imputed value for this observation is the $(X_1', X_2') = \left(\frac{p'-0.5}{m}, \frac{q'-0.5}{m}\right)$. For a singly censored observation (X_1, X_2) where X_1 is censored and X_2 is an observed failure, we pick a cell $B_{p'q}$ $(p' \geq p)$ with conditional density $\Pr\left(t_1, t_2 \middle| t_1 > \frac{p-1}{m}, \frac{q-1}{m} < t_2 < \frac{q}{m}\right)$ where $\frac{p-1}{m} < X_1 < \frac{p}{m}, \frac{q-1}{m} < X_2 < \frac{q}{m}$. The imputed value for this observation is then $(X_1', X_2') = \left(\frac{p-0.5}{m}, X_2\right)$. If the second component of the pair is singly censored, the imputed value is defined similarly. When all censored observations are replaced by imputed values, we obtain concordance indicators for all pairings for this dataset, then apply a selected smoothing method. We can repeat these procedures $10 \sim 20$ times and take an average of predicted values from smoothing.

The third method has been proposed in Chen and Bandeen-Roche (2005): It aims to obtain concordance indicators not only from complete data, but also from censored data if concordance status can be confirmed. If the smaller observation of each component of the pair is an observed failure, the concordance status can be fully determined: For example, an uncensored time pair (20,30) and a censored time pair (30+,45+) are surely concordant whatever the censored values are. However, for a pair (20+,30+) and (30,40), concordance status is undeterminable. Such undeterminable pairs are excluded and the other pairs are used as input data of smoothing.

2.4 Simulation study

To assess the performance of our local Kendall's τ estimator, we designed two sets of simulation studies. The first set compares the seven smoothing methods introduced in Section 3.1, and the second set compares methods to deal with censored data.

2.4.1 Methods

We created Clayton, Frank, and Gumbel Archimedean copulas with parameters generating equal correlation coefficients. For example, a Clayton copula with parameter -0.53 and Frank copula with parameter -3.5 both have correlation -0.5. We also created three copulas with correlation 0.3 (Clayton with parameter 0.5, Frank with parameter 1.9, and Gumbel with parameter 1.26), and three with correlation 0.7 (Clayton with parameter 2.15, Frank with parameter 5.8, and Gumbel with parameter 2.07). We also generated independent bivariate data.

For each copula, corresponding true values of local Kendall's τ were obtained as follows. Firstly, note that the CHR functions are given by $\theta(t_1,t_2) = \dfrac{f(t_1,t_2)S(t_1,t_2)}{\left.\dfrac{\partial S(s_1,t_2)}{\partial s_1}\right|_{s_1=t_1} \left.\dfrac{\partial S(t_1,s_2)}{\partial s_2}\right|_{s_2=t_2}}$ where $S(t_1,t_2) = C(S_1(t_1), S_2(t_2))$ is a bivariate survival function. We can replace the joint survival function by $C(1-u_1, 1-u_2)$ upon transforming the two arguments t_1 and t_2 by their survival functions to be between 0 and 1. These CHR functions were evaluated on a grid of points defined as the Cartesian product of $(0.01, 0.02, \cdots, 0.99)$ and $(0.01, 0.02, \cdots, 0.99)$ and then transformed to local Kendall's τ by $\tau(u_1, u_2) = \dfrac{\theta(u_1, u_2) - 1}{\theta(u_1, u_2) + 1}$. Local Kendall's τs for the Clayton copula and independence scenarios are constant. Those for the Frank copula are monotonically increasing with (u_1, u_2) for negative association, and those for the Frank and Gumbel copulas are decreasing for positive association. For the Gumbel copula, $\tau(u_1, u_2)$ is steeply decreasing in the early failure time region.

To obtain the estimates of local Kendall's τ using the candidate smoothing methods, we generated 300 dataset replicates per each type of association structure. For each replicate, bivariate random numbers between 0 and 1 with sample size 300 were generated, using the function 'rCopula' in the R 'copula' package, and the complement of those numbers were taken as survival times; for independent data, two uniform-distributed vectors with sample size 300 were separately generated. To create outcome times (T_j rather than U_j), these bivariate random numbers were transformed to quantiles of the Weibull distribution with scale parameter 1.0 and shape parameter 1.5. Each smoothing method was applied to these times following the procedures described in Section 3.1. Before smoothing, times were transformed to $[0, 1]$ as $\hat{U}_j = \hat{S}_j(T_j)$, where $\hat{S}_j$ denotes the Kaplan-Meier estimator. The resulting predicted values were evaluated at the same grid of points as true values of the local Kendall's τ.

To assess the quality of the fit, we calculated the root-mean-squared-deviation (RMSD) between the true values and the estimates of local Kendall's τ:
RMSD $\approx \left\{\frac{1}{99^2}\sum_{i=1}^{99}\sum_{j=1}^{99}(\tau(u_i, u_j) - \hat{\tau}(u_i, u_j))^2\right\}^{0.5}$, where $u_k = \dfrac{k}{100}$. We evaluated RMSDs locally by splitting the bivariate domain into three regions by tertiles of the joint

survival function, $C(u_1, u_2) \in [0, 0.333), [0.333, 0.667)$, and $[0.667, 1]$, and then calculating RMSDs separately for each region. We also obtained RMSDs for the subset region where the density exceeded the criterion defined in Section 3.2.

We also evaluated the performance of each smoothing method by visual representation. True and estimated values of local Kendall's τ were displayed on a grid $(0.01, 0.02, \cdots, 0.99) \times (0.01, 0.02, \cdots, 0.99)$ using 3D scatterplots, with each component of bivariate standardized failure times along the X- and Y- axes and the local Kendall's τ values along the Z-axis. For each specific type of copula and each smoothing method, we displayed a 3D scatterplot of means of 300 replicates of estimates overlaid with their true values. To visualize the range of better and worse performance for a specific association structure and smoothing method, we compared true and estimated values of local Kendall's τ estimates for single datasets whose overall RMSDs were at the 5th and 95th percentiles among the 300 replicates. Overlaying true and estimated local Kendall's τ in these ways aims to elucidate overall performance as well as identify portions of the domain in which the association is correctly estimated and portions in which estimation is largely biased.

The second set of simulation studies aimed to compare methods for their accuracy in estimating the local Kendall's τ with censored data. Here we chose one best smoothing method from the previous simulation study and adhered to it for the entire procedure. We generated bivariate failure times with sample size 200 from four association structures, one from independence scenario and three from copulas with correlation coefficient 0.7 (Clayton with parameter 2.15, Frank with parameter 5.8, and Gumbel with parameter 2.07). The method was as described for the previous simulation study. We also generated bivariate censoring times with sample size 200 from three association structures: Gumbel with parameter 2.07 for positive association, Clayton with parameter -0.5 for negative association, and true independence. The data generated from these copulas have marginally uniform distribution, which assumes the proportion censored in the dataset to be about 50%. To change the proportion censored, we convert the uniformly-distributed numbers to quantiles of a beta distribution $B(p, 1-p)$, where $1-p$ is the desired proportion censored. We generated scenarios with proportions censored of 30% and 50%, respectively. We defined observed times as the pairwise minimum of the failure times and censoring times. Thus, we had 24 types of datasets – four types of failure-time association by three types of censoring-time association by two censoring proportions.

In addition to comparing the three methods of treating censored data described in Section 3.3, we implemented a naïve method of handling censoring: We excluded singly or doubly censored pairs and analyzed only fully observed data. This serves as the least criterion that any censoring-tackling approach should achieve with independent censoring. We also analyzed the actual failure times, which usually would be unknown, but are known in a simulation study. The fit of this simulated data was used as a criterion of best performance that any method can achieve. Two hundred replicates were generated and analyzed each, for 24 data generation scenarios and five methods.

2.4.2 Results

Mean RMSDs over 300 repetitions from the 1st simulation study are presented in Table 2.2, over the whole time domain (first row) as well as the subset region defined in Section 3.2. These compare true values of local Kendall's τ and their estimates. When evaluating over the whole time domain, logistic regression showed the best performance for independence and all Clayton scenarios, where the local Kendall's τ is flat, as well as for the Frank scenario with a modest association gradient. For these scenarios, the GAM estimators followed, with RMSDs that were similar in the Frank scenarios and higher by roughly 50% in independence and the Clayton scenarios. For the other Frank scenarios and the Gumbel scenarios, the normal-distribution GAM estimator modeled as a bivariate function of time (GAM2) generally performed best, followed by its binomial family/logit link counterpart (GAM4). The one exception was the Frank family with parameter -3.5, where Loess performed best. Otherwise, Loess performance generally was mediocre to poor. MARS performance was clearly inferior to the other methods, with the largest or next-to-largest RMSD for all scenarios.

The 2nd row of each cell in Table 2.2 shows the RMSDs estimated over the region of reliable estimation according to the criterion proposed in Section 3.2. The RMSDs were significantly smaller than those for the entire domain for all association structures and smoothing methods. Differences were striking for the Frank and Gumbel copulas with strong association. Performance rankings among the seven smoothing methods were preserved in many scenarios, but logistic regression became the top performer in all Frank

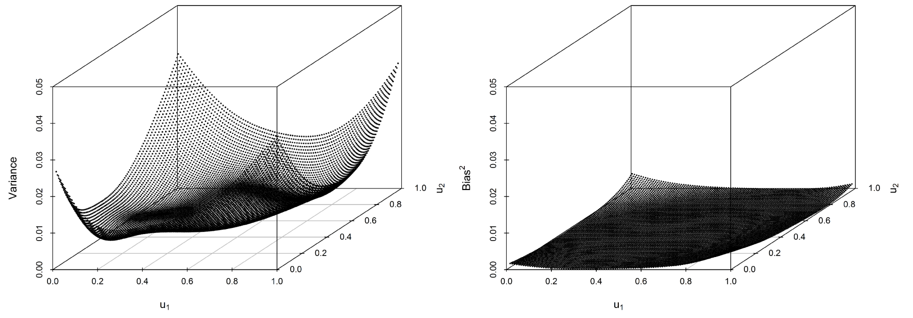


Figure 2.2: Variance (left) and squared bias (right) of RMSDs of local Kendall's τ for data generated from Frank (1.9) copula estimated by GAM2

scenarios.

The RMSDs we have reported can be decomposed into variance across the 300 replicates and squared bias of the estimators. We present 3D scatterplots of variance and squared bias for the GAM2 estimator of the Frank copula with parameter 1.9 in Figure 2.2: it can be seen in this scenario, the variance dominates. Plots for other smoothing methods are presented in the Supplementary Material Section 1. Boxplots showing RMSDs over 300 replicates can also be found in the Supplementary Material Section 1.

Tables 2.3 displays the local RMSDs for three sub-regions of bivariate time domain as determined by the joint survival function. The 1st row in each cell represents the region where $S(t_1, t_2) \in [0, 0.333)$ (late failures), the 2nd row is for $S(t_1, t_2) \in [0.333, 0.667)$, and the 3rd row is for $S(t_1, t_2) \in [0.667, 1]$ (early failures). Generally, the middle region had smaller RMSDs than the other two regions. Exceptions were apparent with Frank and Gumbel copulas with correlation 0.3 and the Clayton copula with correlation 0.7, where the late failure time region had high RMSDs and those for the other two regions were

Table 2.2: RMSDs of local Kendall's τ for the entire space (first row per scenario) and region of valid estimation (second row per scenario), for seven smoothing methods and various association structures (mean of 300 repetitions)

Corr.	Copula	Loess	logistic	GAM1	GAM2	GAM3	GAM4	MARS
-0.5	Clayton	0.152	0.073	0.106	0.108	0.107	0.106	0.162
	(-0.53)	0.101	0.061	0.089	0.096	0.089	0.096	0.157
	Frank	0.129	0.188	0.185	0.141	0.206	0.142	0.239
	(-3.5)	0.098	0.085	0.095	0.095	0.108	0.098	0.165
0	Indep.	0.112	0.072	0.102	0.106	0.101	0.105	0.150
		0.101	0.063	0.088	0.095	0.088	0.095	0.146
0.3	Clayton	0.113	0.068	0.097	0.099	0.097	0.098	0.134
	(0.5)	0.096	0.060	0.082	0.088	0.082	0.088	0.125
	Frank	0.118	0.080	0.103	0.102	0.104	0.103	0.142
	(1.9)	0.097	0.065	0.085	0.090	0.086	0.090	0.130
	Gumbel	0.118	0.120	0.124	0.108	0.128	0.112	0.155
	(1.26)	0.098	0.095	0.105	0.093	0.109	0.095	0.142
0.7	Clayton	0.150	0.057	0.082	0.083	0.082	0.082	0.095
	(2.15)	0.079	0.047	0.063	0.070	0.063	0.069	0.082
	Frank	0.169	0.127	0.123	0.117	0.133	0.120	0.153
	(5.8)	0.082	0.062	0.075	0.076	0.077	0.077	0.109
	Gumbel	0.189	0.146	0.142	0.111	0.165	0.117	0.167
	(2.07)	0.092	0.106	0.091	0.079	0.104	0.084	0.121

GAM1: univariate functions, Gaussian family
GAM2: bivariate function, Gaussian family
GAM3: univariate functions, binomial family
GAM4: bivariate function, binomial family

similar. For the Frank and Gumbel copulas with correlation 0.7, the RMSDs monotonically increased as the failure times increased from early to late. RMSDs calculated over regions of reliable estimation (Table 2.4) were much less distinguished between the late failure time region and the other two regions, whereas RMSDs for the middle and early failure time regions were unchanged or slightly changed. This makes sense because the data grow sparse in late failure time region.

To better understand differences in how these seven smoothing methods estimate local Kendall's τ, we selected Frank copula with parameter 1.9 and then compared 3D scatterplots of mean estimates of 300 replicates, best 5% (5th percentile in terms of RMSD across 300 replicates) and worst 5% (95th percentile) estimates from each of the seven methods. Logistic regression produces planar estimates because of the parametric assumptions used, and MARS produces 'piecewise' flat surfaces by its nature. All the other methods produce smooth and curved surfaces, but we observed subtle distinctions. Since the additive form of two univariate functions in GAM is more restrictive than the bivariate function form, the estimates from the former look more 'parametric' than the latter. There was little difference between estimates from binomial family and Gaussian family. The true values (red) and the GAM2 estimates (black) for Frank (1.9) and Gumbel (1.26) scenarios are displayed in Figure 2.3. For the Frank scenario, the estimator was highly accurate except at the far edges, and particularly the anti-diagonal edges, of the time quadrant. For the Gumbel scenario, the estimator exhibited a more notable bias in the middle of the time range and also was severely biased in the upper-right region (close to (1,1)), which was outside the boundary of

Table 2.3: RMSDs of local Kendall's τ for three sub-regions split by joint survival function from seven smoothing methods for various association structures (mean of 300 repetitions): 1st, 2nd, and 3rd rows for $S(t_1, t_2) \in [0, 0.333)$, $[0.333, 0.667)$, and $[0.667, 1]$, respectively

Corr.	Copula	Loess	logistic	GAM1	GAM2	GAM3	GAM4	MARS
-0.5	Clayton (-0.53)	0.159	0.076	0.108	0.106	0.109	0.105	0.158
		0.114	0.056	0.094	0.105	0.094	0.105	0.169
		0.181	0.089	0.116	0.141	0.115	0.140	0.193
	Frank (-3.5)	0.131	0.213	0.206	0.149	0.231	0.152	0.258
		0.108	0.060	0.088	0.100	0.089	0.103	0.163
		0.164	0.089	0.111	0.137	0.103	0.122	0.172
0	Indep.	0.109	0.076	0.106	0.108	0.105	0.107	0.146
		0.107	0.055	0.088	0.094	0.088	0.094	0.153
		0.155	0.077	0.107	0.124	0.107	0.124	0.177
0.3	Clayton (0.5)	0.117	0.074	0.104	0.103	0.103	0.102	0.135
		0.099	0.052	0.080	0.086	0.080	0.085	0.129
		0.130	0.067	0.095	0.108	0.095	0.108	0.149
	Frank (1.9)	0.125	0.090	0.112	0.108	0.112	0.109	0.145
		0.100	0.054	0.082	0.086	0.082	0.086	0.132
		0.118	0.064	0.091	0.099	0.090	0.098	0.142
	Gumbel (1.26)	0.123	0.133	0.130	0.114	0.134	0.118	0.155
		0.109	0.086	0.113	0.096	0.118	0.099	0.153
		0.103	0.103	0.097	0.096	0.098	0.099	0.162
0.7	Clayton (2.15)	0.181	0.065	0.094	0.090	0.093	0.089	0.102
		0.085	0.044	0.060	0.070	0.060	0.069	0.082
		0.084	0.048	0.067	0.074	0.067	0.074	0.094
	Frank (5.8)	0.205	0.157	0.148	0.139	0.162	0.143	0.179
		0.096	0.061	0.073	0.076	0.076	0.076	0.107
		0.070	0.039	0.057	0.060	0.055	0.057	0.085
	Gumbel (2.07)	0.224	0.167	0.161	0.126	0.185	0.133	0.186
		0.123	0.115	0.116	0.088	0.141	0.095	0.142
		0.078	0.067	0.070	0.060	0.074	0.066	0.102

Table 2.4: RMSDs of local Kendall's τ for three sub-regions (confined to the region of reliable estimation) split by joint survival function from seven smoothing methods for various association structures (mean of 300 repetitions): 1st, 2nd, and 3rd rows for $S(t_1, t_2) \in [0, 0.333)$, $[0.333, 0.667)$, and $[0.667, 1]$, respectively

Corr.	Copula	Loess	logistic	GAM1	GAM2	GAM3	GAM4	MARS
-0.5	Clayton (-0.53)	0.080	0.058	0.083	0.085	0.084	0.086	0.147
		0.114	0.056	0.094	0.105	0.094	0.105	0.169
		0.181	0.089	0.116	0.141	0.115	0.140	0.193
	Frank (-3.5)	0.083	0.092	0.096	0.086	0.115	0.093	0.165
		0.108	0.060	0.088	0.100	0.089	0.103	0.163
		0.164	0.089	0.111	0.137	0.103	0.122	0.172
0	Indep.	0.088	0.064	0.086	0.090	0.086	0.090	0.136
		0.107	0.055	0.088	0.094	0.088	0.094	0.153
		0.155	0.077	0.107	0.124	0.107	0.124	0.177
0.3	Clayton (0.5)	0.087	0.063	0.080	0.086	0.080	0.086	0.118
		0.099	0.052	0.080	0.086	0.080	0.085	0.129
		0.130	0.067	0.095	0.108	0.095	0.108	0.149
	Frank (1.9)	0.091	0.072	0.087	0.091	0.088	0.092	0.126
		0.100	0.054	0.082	0.086	0.082	0.086	0.132
		0.118	0.064	0.091	0.099	0.090	0.098	0.142
	Gumbel (1.26)	0.089	0.099	0.100	0.090	0.104	0.092	0.131
		0.109	0.086	0.113	0.096	0.118	0.099	0.153
		0.103	0.103	0.097	0.096	0.098	0.099	0.162
0.7	Clayton (2.15)	0.072	0.051	0.066	0.069	0.065	0.069	0.078
		0.082	0.043	0.060	0.069	0.060	0.069	0.081
		0.084	0.048	0.067	0.074	0.067	0.074	0.094
	Frank (5.8)	0.079	0.073	0.085	0.085	0.089	0.087	0.120
		0.087	0.056	0.070	0.073	0.072	0.073	0.104
		0.070	0.039	0.057	0.060	0.055	0.057	0.085
	Gumbel (2.07)	0.081	0.117	0.085	0.082	0.091	0.086	0.117
		0.104	0.105	0.101	0.082	0.122	0.088	0.130
		0.078	0.067	0.070	0.060	0.074	0.066	0.102

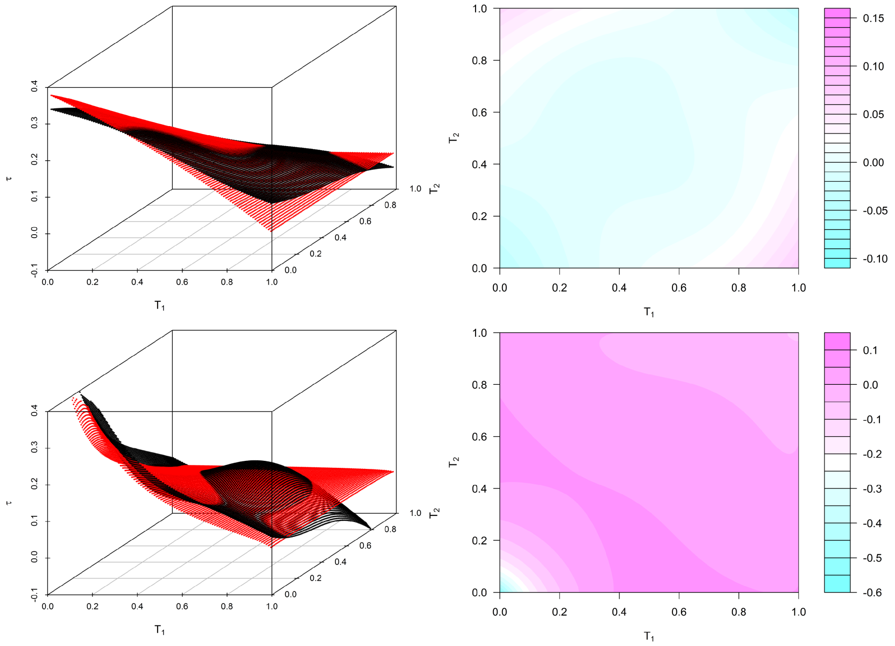


Figure 2.3: 3D scatterplots of local Kendall's τ (true in red and estimates in black) and contour plots of their differences: Frank (1.9) in upper row and Gumbel (1.26) in lower row

reliable estimation. Those for other smoothing methods can be found in the Supplementary Material Section 1.

For each copula type, smoothing methods overestimated in some regions and underestimated in others. For the same copula type, the bias patterns were very similar across seven smoothing methods. In general, severest bias was explained by data sparsity, but this was not the only source of bias. To visualize how the estimates of local Kendall's τ differed from true values for various association structures, 3D scatterplots are provided in the Supplementary Material Section 2. In brief: For Clayton copula with negative correlation (parameter -0.53), all seven smoothing methods underestimated local Kendall's τ at the lower-left corner (early failure times) and overestimated at the 'off-diagonal' region.

For this copula, there is no or little data at the upper-right corner (late failure times), so the estimates in this region do not make sense. Kendall's τs for Clayton copulas with positive correlation (parameters 0.5 and 2.15) were severely underestimated in the lower-left region, severely overestimated at the off-diagonal region and moderately overestimated at the upper-right region. The biases at the off-diagonal region were more severe when the correlation is stronger because the data are sparser. The Frank copula with negative correlation (parameter -3.5) was accurately estimated for all regions except for the upper-right region, where it is nonsensical to estimate because of the data sparsity. Frank copulas with positive correlation (parameters 1.9 and 5.8) were slightly underestimated in the lower-left corner and overestimated at the off-diagonal region. Gumbel copulas with positive correlation (parameters 1.26 and 2.07) were accurately estimated in the lower-left region, but were overestimated in the off-diagonal regions. The biases in the off-diagonal region when the correlation is very strong were extreme for Clayton copula but moderate for Frank copula and Gumbel copula. The region of reliable estimation excluded area in the 'off-diagonal' region in positive association scenarios and 'upper-right' region in negative association scenarios.

Based on the performance of these smoothing methods discussed in this section, we recommend generalized additive models with bivariate function form (referred to as GAM2 and GAM4) for the smoothing of local Kendall's τ. Gaussian and binomial families performed equally well. GAMs showed worse performance than logistic regression for association structures that were time-invariant or only gently time-varying, but the difference was small, and the nonparametric capacity to capture substantially time-varying association structures is a great advantage of GAM over logistic regression.

With the selected smoothing method, GAM2, we assessed the performance of censoring-treatment techniques. In Table 2.5, we present RMSDs for 12 failure-censoring association type pairs and five methods described in Section 4.1. We also present the efficacy of each method, defined as the difference of RMSD of using only fully observed data and each censoring-treatment method divided by the difference of RMSD using only fully observed data and assuming no censoring: This is the amount of RMSD that each method reduced from the worst case divided by the amount of RMSD increased by censoring from ideal case. The following results are for 50% censored datasets if not otherwise specified.

Excluding all the censored observations increased the RMSDs significantly. Negatively associated censoring times increased the RMSDs most, and positively associated censoring times, least. In general, MI outperformed the other methods. It yielded the smallest RMSD in most scenarios. Both Chen's and Brown's methods had a niche of superior performance – Chen's method performed best for the Clayton model and Brown's method, for the independence model where MI method's RMSD exceeded each of them by $10 \sim 30\%$. Each seemed inferior to the MI method in other scenarios.

In Table 2.6, we decomposed RMSDs presented in Table 2.5 into variance and bias squared. Chen's method exhibited least bias for most scenarios; not surprisingly, Brown's method was least biased for independent data. For Clayton copula estimation, MI had significantly larger bias than using fully observed data only; its bias was considerably less

Table 2.5: RMSDs for censoring-treatment methods and their efficacy

30% censoring

Failure	Censoring	NC	CC	Brown		MI		Chen	
Clayton	Positive	0.101	0.162	0.189	-44.2%	0.167	-7.6%	0.138	39.3%
Clayton	Negative	0.101	0.263	0.284	-12.6%	0.215	29.5%	0.168	58.7%
Clayton	Indep.	0.101	0.222	0.246	-19.3%	0.190	26.9%	0.152	58.4%
Frank	Positive	0.124	0.187	0.172	24.7%	0.130	90.6%	0.171	25.3%
Frank	Negative	0.124	0.245	0.205	32.9%	0.142	84.9%	0.199	37.9%
Frank	Indep.	0.124	0.211	0.191	22.9%	0.139	83.4%	0.186	29.4%
Gumbel	Positive	0.124	0.197	0.181	21.4%	0.139	80.3%	0.177	27.4%
Gumbel	Negative	0.124	0.265	0.211	38.6%	0.156	77.3%	0.217	34.6%
Gumbel	Indep.	0.124	0.229	0.197	30.2%	0.149	76.7%	0.197	30.4%
Indep.	Positive	0.129	0.172	0.118	125.0%	0.131	94.1%	0.179	-16.1%
Indep.	Negative	0.129	0.202	0.105	132.5%	0.134	93.2%	0.200	3.3%
Indep.	Indep.	0.129	0.187	0.108	136.4%	0.132	93.9%	0.191	-6.1%

50% censoring

Failure	Censoring	NC	CC	Brown		MI		Chen	
Clayton	Positive	0.101	0.230	0.243	-10.3%	0.241	-8.5%	0.179	39.3%
Clayton	Negative	0.101	0.465	0.365	27.6%	0.334	36.0%	0.331	37.0%
Clayton	Indep.	0.101	0.370	0.330	14.6%	0.294	28.3%	0.249	44.9%
Frank	Positive	0.124	0.254	0.225	22.3%	0.167	67.3%	0.222	25.2%
Frank	Negative	0.124	0.427	0.264	53.9%	0.200	75.0%	0.341	28.2%
Frank	Indep.	0.124	0.342	0.256	39.5%	0.183	73.1%	0.279	28.8%
Gumbel	Positive	0.124	0.272	0.239	22.1%	0.178	63.4%	0.250	15.0%
Gumbel	Negative	0.124	0.476	0.253	63.3%	0.213	74.9%	0.369	30.4%
Gumbel	Indep.	0.124	0.372	0.258	46.1%	0.191	73.2%	0.306	26.8%
Indep.	Positive	0.129	0.245	0.122	105.6%	0.146	84.9%	0.248	-2.5%
Indep.	Negative	0.129	0.390	0.112	106.6%	0.147	93.2%	0.384	2.4%
Indep.	Indep.	0.129	0.304	0.115	108.1%	0.147	89.5%	0.312	-4.7%

NC: Assuming there is no censoring, CC: Using only complete case pairs
The efficacy for Brown, MI, and Chen method is defined as

$$\frac{(\text{RMSD of CC}) - (\text{RMSD of each method})}{(\text{RMSD of CC}) - (\text{RMSD of NC})}$$

in estimation of Frank and Gumbel copulas, and was much less for independent data. In terms of variance, Brown's method and MI were significantly lower than Chen's method–however, in neither of these cases did the variance estimate account for imputation uncertainty. Synthesizing the above discussion, we recommend using Chen's method to handle censored data.

Table 2.6: Variance (1st row) and bias squared (2nd row), 50% censored

Failure	Censoring	NC	CC	Brown	MI	Chen
Clayton	Positive	0.011 0.017	0.046 0.096	0.013 0.219	0.013 0.233	0.034 0.031
	Negative	0.011 0.017	0.214 0.169	0.013 0.352	0.014 0.338	0.124 0.048
	Indep.	0.011 0.017	0.113 0.188	0.014 0.316	0.014 0.296	0.069 0.039
Frank	Positive	0.015 0.060	0.063 0.099	0.018 0.192	0.017 0.113	0.051 0.090
	Negative	0.015 0.060	0.179 0.144	0.015 0.247	0.015 0.165	0.129 0.107
	Indep.	0.015 0.060	0.110 0.127	0.019 0.226	0.016 0.136	0.085 0.096
Gumbel	Positive	0.017 0.034	0.079 0.068	0.021 0.207	0.020 0.119	0.073 0.054
	Negative	0.017 0.034	0.231 0.163	0.017 0.235	0.018 0.170	0.161 0.054
	Indep.	0.017 0.034	0.141 0.127	0.018 0.233	0.018 0.143	0.112 0.053
Indep.	Positive	0.018 0.016	0.063 0.046	0.017 0.005	0.023 0.013	0.068 0.020
	Negative	0.018 0.016	0.172 0.034	0.015 0.006	0.022 0.014	0.167 0.038
	Indep.	0.018 0.016	0.101 0.028	0.015 0.004	0.022 0.015	0.108 0.019

2.5 Data analysis

We applied our method to data from the Cache County Study on Memory Health and Aging. This study aimed to investigate the prevalence of dementia in terms of age, education, sex, and *APOE* genotype (Breitner et al., 1999). To its end, the study recruited participants from the entire population of Cache County, Utah, U.S.A. aged 65 and over. Data were collected on each participant as well as all their first-degree relatives. These data have previously been used to illustrate multivariate failure time methods related to those developed here (Bendeen-Roche & Liang, 2002; Bandeen-Roche & Ning, 2008; Cheng & Fine, 2008; Cheng, Fine, & Bandeen-Roche, 2010).

We analyzed a subset of Cache County data comprising the eldest sibling in each participant's family (inclusive of self) and the participant's mother. This subset has 4,522 pairs of observations, $(X_{i1}, X_{i2}, K_{i1}, K_{i2})$, where X_{i1} is the age of event occurrence of the oldest sibling, X_{i2} is the event time of the mother, and K_{ij} is the event type corresponding to $X_{ij},\ j = 1, 2$. Censoring, dementia onset, and death without dementia were coded as 0, 1, and 2, respectively. We included 3,635 pairs of observations for which some data were observed and who had not yet failed due to either cause by age 55. Among these, 1,431 pairs had no censored component, that is, both pair members either were demented or died. Since our method is supposed to be used for failure times of a single cause, and we are interested in the association between onset ages of dementia, we regarded dementia onset as a failure cause of interest and death as censoring.

Firstly, we excluded any pairs which were singly or doubly censored, with failure type

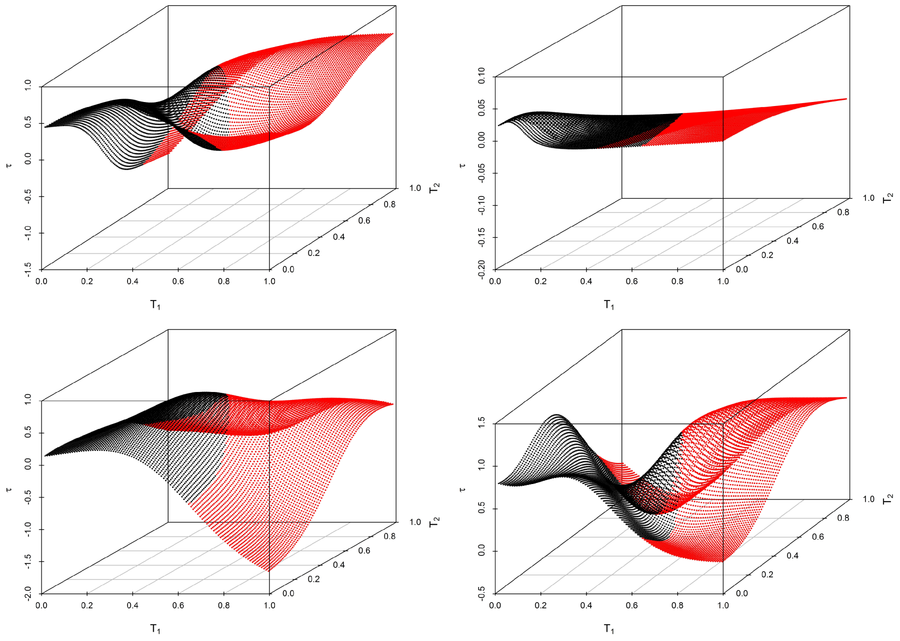


Figure 2.4: 3D scatterplots of local Kendall's τ estimated from Cache County study data: Complete cases (upper-left), Brown (upper-right), MI (lower-left), Chen (lower-right); Red implies outside of the region of reliable estimation.

code 0, leaving a dataset with 1,431 pairs. In this dataset, there were 196 pairs in which the eldest child experienced dementia and 143 pairs in which the mother experienced dementia, of which there were 40 pairs in which both individuals in a pair were demented. We used the dataset with these 40 pairs as a way of investigating the effect of using only fully observed pairs (Method 1). Secondly, using the dataset of 1,431 pairs with censoring, we applied three methods for dealing with censoring: Brown's method, multiple imputation, and Chen's method.

Using only fully observed data and Chen's method showed similar tendency in the change of association as a function of time. The association was very strong for similar times of dementia onset for child and mother – both early or both late which is the only region within the boundary of reliable estimation. In this region, using fully observed pairs only and multiple imputation gave similar estimates which seems most trustworthy while estimates from Chen's method seems overestimated. For early child onset and late mother onset, the association was strongly negative, and for late child onset and early mother onset, the association was weakly positive. For the multiple imputation method, the association was strongly positive for shared early onset only; it was weakly positive for a late maternal onset and modestly negative for early maternal onset together with late child onset. For Brown's method, the estimates ranged from -0.12 to 0.04. The response variables for this method are not binary, but take continuous values between -1 and 1, hence the estimate surface fluctuates less than for the other methods.

We compared this result with previous analyses of the same data by Bandeen-Roche and Ning (2008). This paper calculated cause-specific CHR by counting concordances and discordances with specified failure causes in specific regions. Children's and mothers' ages of dementia onset were dichotomized at 75 and 80 years, and cause-specific CHR was estimated for the resulting quadrants of the bivariate time domain, $(x \leq 75, y \leq 80), (x \leq 75, y > 80), (x > 75, y \leq 80)$ and $(x > 75, y > 80)$. For purposes of comparison, we partitioned our zero-one-scale standardized bivariate failure times at 0.305 and 0.505, corresponding to ages 75 and 80, and obtained the means of our local Kendall's τ estimates for each region. Table 2.7 displays the cause-specific CHRs from the 2008 paper, with corresponding Kendall's τ values in parentheses, side-by-side with estimates from the methods studied in this paper. Estimates from the 2008 paper, fully observed data analysis,

Table 2.7: Cause-specific CHRs for Cache County data in Bandeen-Roche and Ning (2008): failure times dichotomized

(child, mother)	2008: CHR (τ)	CC	Brown	MI	Chen
(early, early)	3.81 (0.58)	0.377	-0.003	0.264	0.721
(early, late)	0.80 (-0.11)	-0.305	-0.076	-0.099	-0.020
(late, early)	2.41 (0.41)	0.126	-0.003	-0.371	0.245
(late, late)	5.89 (0.71)	0.454	-0.038	-0.009	0.528

CC: Using only complete case pairs

and Chen's method coincided in their signs in all four regions, whereas multiple imputation estimates were of opposite sign in the (late,early) and (late,late) regions. Silverman et al. (2005) reported the dementia aggregation in families is stronger in early ages than late ages; the multiple imputation findings are most consistent with this report.

2.6 Conclusion

In this paper, we showed we can visualize the association structure as a nonparametric function of bivariate failure times or inverse quantiles of them by smoothing concordance indicators as a response variable. We compared the performances of various smoothing methods in terms of RMSDs between true and estimated values of local Kendall's τ; we recommend using GAM with a bivariate function with Gaussian or binomial family because it is fully capable of describing complex association functions of time while exhibiting reasonably comparable performance to logistic regression when the association structure is planar.

We evaluated several methods to deal with censored data. We adapted methods suggested by Brown et al. (1974) and Chen and Bandeen-Roche (2005) which were applied for global Kendall's τ and CHR estimation, respectively, to be suited for estimating local Kendall's τ. A new method was also suggested which replaces censored observations by imputed values based on bivariate mass estimates and calculates local Kendall's τ from these. The performance of these methods in terms of RMSDs were compared by a simulation study. Brown's method proved unsatisfactory. Between multiple imputation and Chen's method, however, we could not conclude which was generally better. Multiple imputation showed relatively higher bias than Chen's method, whereas Chen's method was more variable.

An appealing feature of our method is the intuitive interpretation of the strength of time-dependent association – the difference of the probabilities of concordance and discordance. Another strength is the capacity to describe and visualize association in the entire failure time domain, and not only in bins or regions.

One limitation of our method is that the estimates have large bias when the data are sparse at a specific region of the bivariate time domain. We observed especially large bias with censored data addressed by multiple imputation and failure times generated from a Clayton copula with parameter 2.15. We believe the seriousness of this bias can be mitigated because the bias was inflated mainly due to the extreme bias in the 'off-diagonal' region where the data are very sparse, and there would be little need to estimate local association in such a region. Commensurately, we recommend our method be applied in conjunction with a method for identifying a sub-region where estimation can be considered valid, such as our proposed method for thresholding in terms of the bivariate density of observed failure times. Finally, it remains difficult to characterize variability of the proposed estimators. The development of pointwise and simultaneous confidence bands for association function is still a necessary topic of future research.

Chapter 3

Parametric Estimation of Association in Bivariate Failure-time Data Subject to Competing Risks: Sensitivity to Underlying Assumptions

3.1 Introduction

Until recent decades, research on survival analysis mostly concerned univariate data, with observations assumed to be independent. In many modern studies, however, data of interest contain observations that are clustered, and so may be associated. Characterizing failure time associations may sometimes then be of direct interest. Addressing this, multivariate survival function estimators have been developed by Dabrowska (1988), Pruitt (1991), Prentice and Cai (1992), van der Laan (1996), and Prentice (2014). Their implementation and the functions' interpretation, however, may be complex. Employing a simple summary measure of dependence structure can ameliorate this problem. Along these lines, Clayton (1978) suggested representing the dependence structure as a 'cross' (or conditional) hazard ratio. When generalized to vary with time, this quantity is defined as follows:

$$\theta(t_1,t_2) = \frac{\lambda(t_2|T_1 = t_1)}{\lambda(t_2|T_1 > t_1)} = \frac{f(t_1,t_2)\cdot S(t_1,t_2)}{\left.\dfrac{\partial S(s_1,t_2)}{\partial s_1}\right|_{s_1=t_1}\cdot\left.\dfrac{\partial S(t_1,s_2)}{\partial s_2}\right|_{s_2=t_2}}. \tag{3.1}$$

It can be interpreted as the ratio comparing an individual's hazard of failure at t_2 given failiure of his pair partner at t_1 to the hazard given that the partner has not yet failed by t_1.

Multivariate survival analysis may have particular benefits to offer in research involving competing risks. Most such research has focused on the univariate setting in which only one type of failure may be observed per sampling unit. Multivariate survival analysis with competing risks informs the study of relationships among failure types in ways univariate analysis cannot, because multiple failure types may be observed in a cluster. Among many available measures of association in the competing risks setting (e.g. Cheng, Fine, & Kosorok, 2007, 2009; Scheike et al., 2010), this paper focuses on the modified conditional hazard ratio, and a parametric model and estimator for this, proposed by Bandeen-Roche and Liang (2002). Bandeen-Roche and Ning (2008) developed a nonparametric estimator of the modified conditional hazard ratio and proved its distributional properties; Cheng, Fine, and Bandeen-Roche (2010) extended it to exchangeable data in which the cluster size may be greater than two. Gorfine and Hsu (2011) suggested a frailty-based conditional regression model in which the frailty processes have general distributional structure, and which subsumes the Bandeen-Roche and Liang parametric model as a special case.

The parametric model of Bandeen-Roche and Liang (2002) has an appealing feature that is not shared by the nonparametric approaches to estimation of the modified conditional hazard ratio, nor is retained in the Gorfine and Hsu (2011) formulation: a conceptually intuitive decomposition of failure time associations into ‘size’ and ‘shape’ components. To explicate the idea, consider two failure causes: onset of a given disease, or death. The ‘size’ component governs clustering between times to earliest failure from any cause - either disease onset or death. It does this through cluster-specific frailties that multiply the overall, population failure hazard. The ‘shape’ component governs clustering in the tendency to fail preferentially from certain causes as opposed to others. It does this through cluster-specific compositional frailty processes (time-varying vectors of proportions) that generate cause-specific hazards by multiplying the overall cluster hazard. Such a decomposition opens prospects for distinguishing shared genetic or environmental influences that predispose faster overall health declines from those that speed or delay some diseases as opposed to others. The methodology was never pursued beyond the 2002 paper, however, because it performed badly in simulation scenarios in which its underlying assumptions were replaced by alternative reasonable assumptions. Our goal herein is to better understand the source of this sensitivity, with an eye to correcting it.

The remainder of this paper proceeds as follows. Section 2 introduces notation and relevant background. Section 3 investigates sensitivity to one of the methodology's major assumptions: Dirichlet distribution of the shape frailty. We study the behavior of the estimator when the data are generated from a logit-normal distribution and also investigate the potential influence of mis-specified size frailty. Section 4 investigates the second major assumption: that size and shape frailty variables are statistically independent. Both investigations employed simulation studies. Section 5 concludes.

3.2 Background and Motivation

3.2.1 Notation

We consider a simple setting in which the data are independently and identically distributed across clusters, there are two types of competing risks, and there are two units per cluster (pairs). For members $j = 1, 2$ of a given pair (subscript i tracking pairs suppressed for the time being), let T_{j1} denote the failure time of interest and T_{j2} the failure time for the competing risk, each with hazard function $\lambda_j(t)$. Then the time of the first failure is $X_j = T_{j1} \wedge T_{j2}$; if events truly are competing, only X_j is observable, whereas for semicompeting risks T_{j1} and T_{j2} both may be observed in certain instances. The data also includes a failure type indicator K_j which is 1 when $X_j = T_{j1}$ and 2 when $X_j = T_{j2}$. For now we treat the data as fully observed; later we introduce the possibility of censoring independent of the occurrence of both types of risks.

3.2.2 The conditional cause-specific hazard ratio (CCSHR)

The CCSHR compares two instances of the cause-specific hazard - a fundamental quantity estimable from observed data in the competing risks setting. In the univariate setting, the cause-specific hazard is defined as $\lambda_k(x) = \lim_{h_1 \downarrow 0} \Pr(x \leq X < x + h_1, K = k | X \geq x)/h_1$. Its generalization to the bivariate setting is given by $\lambda_{(k_1,k_2)}(x_1, x_2) = \lim_{(h_1,h_2)\downarrow 0} \Pr(x_1 \leq X_1 < x_1 + h_1, K_1 = k_1, x_2 \leq X_2 < x_2 + h_2, K_2 = k_2 | X_1 \geq x_1, X_2 \geq x_2)/(h_1 h_2)$; Bandeen-Roche and Liang (2002) considered a corresponding joint density for the failure times and causes, given by $f(x, k) = \lim_{(h_1,h_2)\downarrow 0} \Pr(x_1 \leq X_1 \leq x_1 + h_1, x_2 \leq X_2 \leq x_2 + h_2, K_1 = k_1, K_2 = k_2)/(h_1 h_2)$. Then, the conditional cause-specific hazard ratio (CCSHR) may be defined as

$$\begin{aligned}\theta_{CS}(x_1, x_2; k_1, k_2) &= \frac{\lambda_{1,k_1}(x_1 | X_2 = x_2, K_2 = k_2)}{\lambda_{1,k_1}(x_1 | X_2 > x_2)} \\ &= \frac{S(x_1, x_2) f(x_1, x_2; k_1, k_2)}{\{\int_{x_2}^{\infty} \sum_{k=1}^{2} f(x_1, x, k_1, k) dx\}\{\int_{x_1}^{\infty} \sum_{k=1}^{2} f(x, x_2, k, k_2) dx\}}. \qquad (3.2)\end{aligned}$$

Roughly it is the factor by which an individual's risk of failure at x_1 due to cause k_1 is changed if his pair partner fails at x_2 due to cause k_2 versus has not yet failed at all by x_2. It generalizes the conditional hazard ratio which has similar definition as in (3.2), only omitting all references to causes k.

3.2.3 A parametric model for the CCSHR

The model we seek to study is grounded in the frailty modeling (Vaupel et al., 1979). A frailty variable, A, is an unobserved random effect that multiplicatively modifies the hazard function of an individual, or of related individuals. Taking G as the frailty distribution and a as a generic realization, the bivariate survival function can be expressed as follows: $S(x_1, x_2) = \int \{\prod_{m=1}^{2} S_m^*(x_m)\}^a dG(a) = \int \exp\{-a \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(x)dx\} dG(a)$, where $S_m^*(x_m)$ are survival functions and $\lambda_m^*(x_m)$ are corresponding hazard functions conditional on $A = 1$ (henceforth, 'reference' survival or hazard functions). The conditional hazard ratio then can be represented in terms of A and λ_m^* as

$$\theta(x_1, x_2) = \frac{E[A^2 \exp\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t)dt\}]E[\exp\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t)dt\}]}{E^2[A \exp\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t)dt\}]} \tag{3.3}$$

.

Importantly for what follows, the survival function for each m-th pair member conditional on $A = a$ is $S_m^*(x_m)^a$, and the corresponding hazard function is

$$\lambda_m(x_m | A = a) = a\lambda_m^*(x_m) \tag{3.4}$$

.

To represent the CCSHR, Bandeen-Roche and Liang observed that because the overall failure hazard is the sum of cause-specific hazards, $\lambda(x) = \lambda_1(x)+\lambda_2(x)$, the cause-specific hazard can be written as a proportion $R_k(x)$ of the overall hazard, $\lambda_k(x) = R_k(x)\lambda(x)$, $k = 1, 2$. To characterize a hazard specific to both pair and cause k, then, they proposed to modify the right-hand side of (3.4) by multiplying the frailty for overall failure, A, by a proportional shape frailty vector $B(x) = \{B_1(x), B_2(x)\}$ having mean function $\{R_1(x), R_2(x)\}$. This yields

$$\lambda_{mk}(x_m|A = a, B(x_m) = b(x_m)) = ab_k(x_m)\lambda^*_m(x_m) \tag{3.5}$$

where $\sum_k b_k(x_m) = 1$. Conceptually, A amplifies or diminishes a pair's tendency to fail early, regardless of cause, and $B(x)$ tailors the pair's allocation of the overall hazard to the respective causes.

To develop an estimator for the CCSHR, Bandeen-Roche and Liang imposed two assumptions upon (3.5): Dirichlet distribution of the shape frailty $B(x)$, and independence between the size frailty A and the shape frailty $B(x)$. With the independence assumption, the CCSHR for causes k_1 and k_2 becomes

$$\frac{E\{B_{k_1}(x_1)B_{k_2}(x_2)\}}{E\{B_{k_1}(x_1)\}E\{B_{k_2}(x_2)\}} \times \theta(x_1, x_2). \tag{3.6}$$

If $B(x)$ has Dirichlet distribution with parameter $\delta(x)$ and mean function $R(x)$ and we set $\delta(x) = \Delta R(x)$, the first multiplicand becomes

$$1 - \frac{1}{\Delta + 1} \left\{ \frac{R_{k_1}(x_1 \wedge x_2) - 1}{R_{k_1}(x_1 \wedge x_2)} \right\}^{I(k_1 = k_2)}. \tag{3.7}$$

The second multiplicand is the conditional hazard ratio for the frailty model without competing risks. The first and second multiplicands have interpretations as association in failure causes and in times to first failure, respectively. The distributional assumptions yield convenient estimators.

Notwithstanding these advantages, prior studies have suggested that estimators employing (3.6) and (3.7) may be sensitive to assumptions made. In the next two sections we study this issue seeking means to ameliorate the sensitivity.

3.3 Sensitivity to assumption 1: Dirichlet distribution of shape frailty

To evaluate the sensitivity of the Bandeen-Roche and Liang (2002) parametric estimator (henceforth, BRL estimator) to the Dirichlet distribution assumption, a natural comparator is one incorporating a logit-normal distribution instead. In this section, we propose an estimator based on logit-normal-distributed shape frailty, and then compare the performance of the two estimators for simulated data sets in which the shape frailty has Dirichlet versus logit-normal distribution. Additionally, we repeated simulation scenarios in which the underlying assumptions of the BRL framework were replaced by alternative reasonable assumptions, but revisited estimation not only of the shape frailty component of our model but also the size frailty component – a source of sensitivity not considered in the original 2002 paper.

3.3.1 Introduction of distributions to be studied

The Dirichlet distribution is frequently used to model vectors of multivariate proportions, W, which sum to one (i.e. 'compositional' data). Thus it is suited to allocate proportions of hazards of the various failure types to the overall hazard. It has density $\frac{\Gamma(\alpha)}{\prod_{k=1}^{K}\Gamma(\alpha_k)}\prod_{k=1}^{K} w_k^{\alpha_k-1}$ where $\alpha = \sum_{k=1}^{K}\alpha_k$, $E(W_k) = \frac{\alpha_k}{\alpha}$ and $Var(W_k) = \frac{\alpha_k(\alpha-\alpha_k)}{\alpha^2(\alpha+1)}$ (Aitchison, 1982). It arises intuitively by dividing a collection of 'amounts' by their sum when the amounts are mutually independent, and the proportions resulting from dividing the amounts by their sum are independent of the sum, or when the amounts are independent gamma random variables with common scale, or in certain cases when amounts are positively correlated (Bandeen-Roche & Ruppert, 1991). In the failure time context, if disease A and disease B arise independently within families, and the type of failure occurring first is independent of the total propensity to fail, then the assumptions of the Dirichlet distribution are satisfied. If diseases A and B have a common cause, these assumptions are likely to be violated because the propensities to fail from two diseases are correlated. Moreover, the Dirichlet distribution constrains the covariance between any pair of proportions to be negative. If there are only two types of failures (i.e. a single proportion and its difference from one to be modeled), the Dirichlet distribution reduces to the beta distribution.

The logit-normal distribution is a primary alternative to the Dirichlet for modeling compositional data (Aitchison & Shen, 1980). Suppose a $(K-1)$-dimensional random vector Y follows a multivariate normal distribution $N_{K-1}(\mu, \Sigma)$ over $\mathbb{R}^{K-1}$. Then W with $W_j = \dfrac{\exp(Y_j)}{1+\sum_{k=1}^{K-1}\exp(Y_k)}, j = 1, \cdots, K-1$ and $W_K = 1 - \sum_{k=1}^{K-1} W_k$ defines the logit-normal distribution of dimension K. The associated density function is given by $|2\pi\Sigma|^{-\frac{1}{2}}(\prod_{k=1}^{K} W_k)^{-1}\exp[-\frac{1}{2}\{\log(W_{-K}/W_K)-\mu\}^T\Sigma^{-1}\{\log(W_{-K}/W_K)-\mu\}]$ where $W_{-K} = (W_1, \cdots, W_{K-1})$. The logit-normal distribution has $\dfrac{1}{2}(K-1)(K+2)$ parameters compared with only K parameters for the Dirichlet distribution; in fact, a suitably chosen logit-normal can closely approximate any Dirichlet. It relaxes some of the assumptions underlying the Dirichlet class, for example independence of the bases, making it a worthwhile choice for further study.

Following on the 2002 paper by Bandeen-Roche and Liang, we proceed to study the case of two competing causes.

3.3.2 Methods

We began by implementing a maximum likelihood estimator for the parameters of a logit-normal shape distribution in the BRL framework, assuming that $B_j(x) = B_j$ for all x. The likelihood function for hazard and frailty quantities based on a sample of pairs

$i = 1, \cdots, n$ is given by

$$\prod_{i=1}^{n} E\{B_{K_{i1}}(x_{i1})B_{K_{i2}}(x_{i2})\}E[A^2\lambda_1^*(x_{i1})\lambda_2^*(x_{i2})\exp\{-A\sum_{m=1}^{2}\int_0^{x_{im}}\lambda_m^*(t)dt\}] \tag{3.8}$$

(Bandeen-Roche & Liang, 2002). Additionally assuming size and shape independence factorizes this into quantities involvinng only the shape frailty distribution versus only the reference hazard and size frailty distribution. Inference for the pair-specific hazards and size frailty can be accomplished by existing methods such as Shih and Louis (1995). Inference for the shape frailty involves only the first multiplicand of Equation (3.8), taking the likelihood function for the logit-normal parameters proprotional to

$$\begin{aligned}
&\prod_{i\in I_1} E(B_1(x)B_1(x)) \prod_{i\in I_3} E(B_1(x)B_2(x)) \prod_{i\in I_2} E(B_2(x)B_2(x)) \\
&\qquad = \prod_{i\in I_1} E(B^2) \prod_{i\in I_3} E(B(1-B)) \prod_{i\in I_2} E((1-B)^2) \\
&= \prod_{i\in I_1} E\left(\frac{\exp(2Y)}{(1+\exp(Y))^2}\right) \prod_{i\in I_3} E\left(\frac{\exp(Y)}{(1+\exp(Y))^2}\right) \prod_{i\in I_2} E\left(\frac{1}{(1+\exp(Y))^2}\right) \qquad (3.9)
\end{aligned}$$

where Y is a normal, and B, a logit-normal, random variable, and I_1, I_2, and I_3 refer respectively to sets of pairs whose members both fail of cause 1, both fail of cause 2, and fail of different causes. Then the log-likelihood is

$$
\begin{aligned}
n_1 \log \int_{-\infty}^{\infty} \frac{\exp(2y)}{\{1+\exp(y)\}^2} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) dy \\
+ n_3 \log \int_{-\infty}^{\infty} \frac{\exp(y)}{\{1+\exp(y)\}^2} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) dy \\
+ n_2 \log \int_{-\infty}^{\infty} \frac{1}{\{1+\exp(y)\}^2} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) dy \quad (3.10)
\end{aligned}
$$

where μ and σ are the mean and standard deviation of the logit, n_1 is the number of pairs whose members both fail due to cause 1, n_2 is the number of pairs whose members both fail due to cause 2, and n_3 is the pairs whose members fail of different causes. For improved numerical stability, we replaced the standard deviation σ with $\exp(\log(\sigma))$ and then estimated $\log(\sigma)$. The values of μ and $\log(\sigma)$ that maximize the log-likelihood function were obtained using the 'optim' function with L-BFGS-B method in the R Statistical Software package.

When we have censored observations, we can still use the same likelihood function to estimate μ and σ. First, we count the number of pairs whose members both fail due to cause 1, both fail due to cause 2 and fail of different causes among pairs in which both members were observed to fail. Using the proportional frequencies of these three groups of pairs, we can get imputed frequencies of three groups for singly and doubly censored pairs. Then adding the observed and imputed frequencies of pairs gives us n_1, n_2, and n_3. This method is described in more detail in Step 1$\sim$3 in the Appendix 1 of Bandeen-Roche and Liang (2002).

A simulation study was conducted to assess the performance of the estimator with logit-normal shape frailty assumption, and the sensitivity of both it and the previously proposed Dirichlet-based estimator to violations of their respective distributional assumptions. The simulation settings and procedures mimicked those of Bandeen-Roche and Liang (2002). A first set of studies assessed the accuracy of the logit-normal parameter estimation. It assumed the pair members' earliest failure times regardless of cause followed a Clayton copula distribution. To create such failure times, we first generated 1,000 replicates of $n = 100$ or $n = 500$ size frailties 'A' drawn independently from a gamma distribution with mean = 1 and variance = 1. Per replicate and pair i, we generated two failure times drawn independently from an exponential distribution with rate parameter A_i. Next, we allocated 'causes' of failure. Per replicate and pair, we drew shape frailties 'B_i' independently from a logit-normal distribution with mean of the logit equal to μ and standard deviation of the logit equal to σ. Parameters $\mu = 0, 0.75$, and 1.5 and $\sigma = 1$ and 3 were varied as true values of the logit-normal parameters. The resulting distribution is symmetric when $\mu = 0$ and increasingly left skewed as μ is larger; $\sigma = 1$ results in unimodal distributions and $\sigma = 3$ results in a bimodal (U-shaped) distribution. In each, to decide the failure type for each failure time in a pair, we generated independent uniformly distributed random numbers and compared these to the shape frailties B_i; if an individual's uniform realization was less than or equal to B_i, we assigned cause 1, and otherwise, cause 2. Finally, we estimated μ and σ as the values maximizing the log-likelihood equation (3.10) and then the CCSHR according to (3.6). In the first multiplicand of $\text{CCSHR}_{1,1}$ (between cause 1 and cause 1), $E(B^2)$ and $E(B)$ were calculated using numerical integration, plugging in the estimated logit-normal parameters. The numerical integration was implemented using the 'integrate' function in R with default settings (R Core Team, 2013). The first multiplicand of $\mathrm{CCSHR}_{1,2}$ and $\mathrm{CCSHR}_{2,2}$ can be obtained by numerical integration of $\dfrac{E(B(1-B))}{E(B)E(1-B)}$ and $\dfrac{E((1-B)^2)}{\{E(1-B)\}^2}$, respectively. The second CCSHR multiplicand is the conditional hazard ratio without competing risks: it was obtained using two-stage semiparametric estimation of Shih and Louis (1995) assuming Clayton's copula.

A next set of studies assessed sensitivity of estimators to mis-specified shape distribution, within the BRL framework. To assess sensitivity of the original, Dirichlet-based estimator to violation of its assumption of distribution for the shape frailty, we applied an estimator assuming beta shape distribution (detailed in Section 4.1, Bandeen-Roche and Liang, 2002) to the same data as described above. Here, we used maximum likelihood method to estimate Dirichlet parameters instead of closed-form formula in their paper. Conversely, to assess performance of the logit-normal estimator under a Dirichlet shape assumption, we fit both estimators to data generated as described above except replacing logit-normal shape frailties with beta frailties, varying the beta parameters as $(\alpha, \beta) = (0.2, 0.8), (1, 4), (0.5, 0.5)$, and $(2, 2)$.

A third set of studies employed a generating mechanism outside of the BRL framework. This mechanism imagines a 'latent' failure time for each cause of which only the first is observed. For each of 500 replicates, we first generated $n = 500$ pairs of 'cause 1' (say, 'disease') failure times as exponential conditional on gamma frailties, A_{i1}, exactly as in

	Cause 1			Cause 2		
Scenario	l_1	l_2	t_1	l_3	l_4	t_2
1	2	2	2	2	2	2
2	2	2	2	2	2	4
3	2	2	2	2	3	2
4	2	2	2	2	3	4
5	2	2	2	3	3	2
6	2	2	2	3	3	4

Table 3.1: Exponential rate (l) and association (gamma shape-defining; t) parameters for the 3rd and 4th sets of simulation studies

the first step of the first set of studies. Then, we independently generated $n = 500$ pairs of 'cause 2' (say, 'death') failure times, also exponential conditional on gamma frailties, A_{i2}; for each individual, we considered the pairwise minimum of the 'disease' and 'death' failure times as the failure time (with their associated cause). For both causes the gamma scale parameter was set equal to 1. The gamma shape parameter was set equal to $1/(t_1 - 1)$ for 'disease' (yielding 'marginal' CHR of t_1) and to $1/(t_2 - 1)$ for 'death', varying of t_1 and t_2 as in Table 3.1 below. For 'disease' the exponential rate parameters were set to $l_1 \times A_{i1}$ and $l_2 \times A_{i1}$ for the respective members of the pair; for 'death', they were set equal to $l_3 \times A_{i2}$ and $l_4 \times A_{i2}$. Values of l_1, l_2, l_3, and l_4 also were varied as in Table 3.1, for a total of six scenarios. CCSHRs were estimated through the same estimation procedures as in the second simulation study (falsely assuming data generated according to the Bandeen-Roche and Liang framework).

A fianl set of simulation studies was similar to the third one in all ways with one exception: rather than generating cause-specific failure times as exponential conditional on the pairwise frailty, we generated them to be marginally exponential. Details are provided in the Appendix. The gamma, exponential, normal, beta, and uniform random numbers needed for the studies just described were generated using standard R functions.

3.3.3 Results

The first and second sets of simulation studies addressed the estimation of the logit-normal parameters μ and σ (Table 3.2) and resulting CCSHR (Table 3.3). The estimator of μ exhibited bias at most 5.3% for completely observed data and at most 7.6% for 30% censored data. The estimator of σ exhibited bias which increased in absolute value with σ, but bias as a percentage of the estimand decreased. For both estimators based on beta and logit-normal shape distributions, biases decreased considerably comparing $n = 500$ to $n = 100$ and increased for 30% censored data compared to complete data. Precision of estimation improved substantially for $n = 500$ compared to $n = 100$, with standard errors in estimation generally smaller by 50% to 60% for both μ and σ. Standard errors for the censored data were greater than those for complete data by 35$\sim$60% for both μ and σ.

Table 3.3 compares performance in estimating $\mathrm{CCSHR}_{1,1}$ between procedures based on a logit-normal shape distribution and on a beta distribution when the true failure types are generated by various parameters of these two distributions. Each column displays mean and standard deviation of the CCSHR estimates using estimators based on logit-normal and beta distribution respectively. The upper and lower parts of the table show the results when the true failure type distribution was beta and logit-normal, respectively. Two estimators exhibited bias no greater than 1.2% for complete data and 2.0% for censored data in all

Table 3.2: Simulation study findings: Performance of ML estimation of logit-normal distribution parameters (Equation (3.10)). Data were generated according to the Bandeen-Roche and Liang parametric model with gamma size frailty and logit-normal shape frailty.

	Estimates of μ							
	True values		No censoring			30% censoring		
	μ	σ	Mean	SD	Bias	Mean	SD	Bias
n=100	0	1	-0.002	0.188	-0.002	-0.001	0.278	-0.001
	0	3	-0.003	0.400	-0.003	0.017	0.548	0.017
	0.75	1	0.770	0.221	0.020	0.789	0.342	0.039
	0.75	3	0.778	0.437	0.028	0.801	0.629	0.051
	1.5	1	1.537	0.294	0.037	1.585	0.472	0.085
	1.5	3	1.579	0.537	0.079	1.614	0.775	0.114
	μ	σ	Mean	SD	Bias	Mean	SD	Bias
n=500	0	1	0.002	0.082	0.002	0.005	0.120	0.005
	0	3	0.003	0.170	0.003	0.004	0.245	0.004
	0.75	1	0.756	0.096	0.006	0.762	0.142	0.012
	0.75	3	0.751	0.179	0.001	0.748	0.261	-0.002
	1.5	1	1.511	0.127	0.011	1.519	0.186	0.019
	1.5	3	1.506	0.218	0.006	1.512	0.306	0.012
	Estimates of σ							
	True values		No censoring			30% censoring		
	μ	σ	Mean	SD	Bias	Mean	SD	Bias
n=100	0	1	0.998	0.406	-0.002	1.029	0.567	0.029
	0	3	3.025	0.657	0.025	3.010	0.903	0.010
	0.75	1	0.997	0.457	-0.003	1.010	0.685	0.010
	0.75	3	3.083	0.686	0.083	3.126	0.993	0.126
	1.5	1	0.951	0.515	-0.049	0.963	0.737	-0.037
	1.5	3	3.120	0.757	0.120	3.146	1.078	0.146
	μ	σ	Mean	SD	Bias	Mean	SD	Bias
n=500	0	1	0.990	0.177	-0.010	0.985	0.260	-0.015
	0	3	2.987	0.300	-0.013	2.994	0.430	-0.006
	0.75	1	0.990	0.191	-0.010	0.972	0.292	-0.028
	0.75	3	2.991	0.315	-0.009	2.991	0.451	-0.009
	1.5	1	0.999	0.214	-0.001	0.980	0.334	-0.020
	1.5	3	2.993	0.327	-0.007	2.998	0.477	-0.002

scenarios except for $(\alpha, \beta) = (0.2, 0.8)$ scenario where logit-normal based estimator had biases of 2.8% and 4.0%, respectively. The coefficient of variation (CV) for $\mathrm{CCSHR}_{1,1}$ was no greater than 14.3% for complete data, and no greater than 18.9% for censored data, and there were little differences between beta-based and logit-normal-based estimators for most scenarios. For the most highly skewed scenario, $(\alpha, \beta) = (0.2, 0.8)$, logit-normal-based estimator was less accurate than beta-based one, but also less variable. For all the other scenarios, both estimators were highly accurate.

Table 3.3: Comparison of CCSHR estimators based on beta and logit-normal distributions when the true failure types respectively are generated from logit-normal and beta distributions. Estimators are those detailed in Section 3.1.

			No censoring				30% censoring			
			Beta		Logit-normal		Beta		Logit-normal	
μ	σ	TRUE	Mean	SD	Mean	SD	Mean	SD	Mean	SD
0	1	2.347	2.336	0.163	2.336	0.163	2.304	0.220	2.305	0.220
0	3	3.081	3.075	0.208	3.068	0.207	3.038	0.291	3.029	0.288
0.75	1	2.177	2.169	0.135	2.169	0.135	2.138	0.182	2.138	0.182
0.75	3	2.762	2.755	0.178	2.751	0.177	2.721	0.244	2.717	0.243
1.5	1	2.080	2.075	0.123	2.076	0.122	2.048	0.165	2.049	0.165
1.5	3	2.529	2.521	0.158	2.519	0.157	2.489	0.216	2.487	0.216
α	β	TRUE	Mean	SD	Mean	SD	Mean	SD	Mean	SD
0.2	0.8	6.000	6.000	0.614	5.830	0.527	5.959	0.843	5.757	0.709
1	4	3.333	3.299	0.466	3.293	0.470	3.218	0.607	3.223	0.602
0.5	0.5	3.000	2.990	0.198	2.987	0.198	2.950	0.275	2.946	0.273
2	2	2.400	2.388	0.163	2.388	0.163	2.352	0.216	2.352	0.216

The third set of simulation studies addressed the estimation of CCSHR for failure times arising as the pairwise minimum of cause-specific failure times (Table 3.4). When failure rates due to cause 1 equaled those for cause 2 for both pair members ($l_1 = l_3$ and $l_2 = l_4$; Scenarios 1 and 2), the bias of CCSHR estimator was very small (<1%). When the cause-specific failure rates differed across causes for only one member of the pair ($l_1 = l_3$ and $l_2 \neq l_4$; Scenarios 3 and 4), the estimator was moderately biased (up to 2.7%). When the cause-specific hazard rates differed for both pair members ($l_1 \neq l_3$ and $l_2 \neq l_4$; Scenarios 5 and 6), the biases inflated further (up to 8.9%). In most of the scenarios, biases in estimating $\mathrm{CCSHR}_{1,2}$ were smaller than those of $\mathrm{CCSHR}_{1,1}$. The coefficients of variation of $\mathrm{CCSHR}_{1,2}$ estimates, however, were greater than those of $\mathrm{CCSHR}_{1,1}$. Beta-based and logit-normal-based estimators performed similarly in all scenarios.

Table 3.4: Comparison of the CCSHR estimators based on beta and logit-normal distributions (3rd simulation study); Data generated from distributions with $\mathrm{CCSHR}_{1,1} = 2$ and $\mathrm{CCSHR}_{1,2} = 1$

	$\mathrm{CCSHR}_{1,1}$		$\mathrm{CCSHR}_{1,2}$	
Scenario	Beta	Logit-normal	Beta	Logit-normal
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
1	2.006 (0.138)	2.001 (0.151)	1.000 (0.084)	1.006 (0.103)
2	2.001 (0.117)	1.993 (0.115)	0.997 (0.099)	1.023 (0.085)
3	2.054 (0.131)	2.054 (0.131)	1.018 (0.092)	1.018 (0.092)
4	2.036 (0.108)	2.029 (0.106)	0.996 (0.099)	1.015 (0.086)
5	2.177 (0.150)	2.177 (0.150)	1.021 (0.086)	1.021 (0.085)
6	2.085 (0.121)	2.077 (0.119)	0.958 (0.094)	0.980 (0.080)

The fourth set of simulation studies differed from the third set only in the distributions of T_1 and T_2 (exponential marginally versus conditionally on the pair frailty; Table 3.5). We observed a pattern of findings quite similar to the third set of studies, however with biases that were much more severe. For scenarios in which the strength of association was equal across causes (Scenarios 1, 3, and 5), the bias increased with increasing differentiation in the cause-specific marginal distributions (0$\sim$10%). For scenarios in which the strength of association differed across causes (Scenarios 2, 4, and 6), the estimators were severely biased regardless of the marginal distributions (30$\sim$ 60%). For each estimand, beta- and logit-normal-based estimators performed similarly. For the beta estimator, this finding replicates that in the Bandeen-Roche and Liang (2002) paper.

Table 3.5: Comparison of the CCSHR estimators based on beta and logit-normal distributions (4th simulation study); Data generated from distributions with $\mathrm{CCSHR}_{1,1} = 2$ and $\mathrm{CCSHR}_{1,2} = 1$

	$\mathrm{CCSHR}_{1,1}$		$\mathrm{CCSHR}_{1,2}$	
Scenario	Beta	Logit-normal	Beta	Logit-normal
1	2.006 (0.139)	1.999 (0.155)	1.002 (0.080)	1.010 (0.105)
2	2.689 (0.200)	2.682 (0.206)	1.009 (0.097)	1.017 (0.111)
3	2.061 (0.131)	2.061 (0.131)	1.038 (0.084)	1.038 (0.084)
4	2.808 (0.218)	2.806 (0.216)	1.082 (0.099)	1.083 (0.097)
5	2.273 (0.168)	2.273 (0.168)	1.029 (0.085)	1.029 (0.085)
6	3.200 (0.257)	3.194 (0.253)	1.107 (0.099)	1.111 (0.095)

In seeking to understand biases in estimating the CCSHR observed in the 3rd and particularly the 4th set of simulations, estimation of the size (second) multiplicand (Equation (3.6)) and not only the shape (first) multiplicand of the CCSHR must be considered. Specifically, even though the dependence between bivariate failure times for each cause follows a gamma frailty model where the strength of association does not change over time, the dependence in observed failure times (generated as the minimum of cause-specific failure times) may not. This was a possibility not considered by Bandeen-Roche and Liang in their 2002 paper. We used the diagnostic method of Chen and Bandeen-Roche (2005) to assess whether the pairwise minimum retained gamma frailty dependence structure. If so, the 'size'-associated conditional hazard ratio, $\theta^*(S(t_1, t_2))$, should be constant considered as a function of the survival function. Results of this diagnostic are displayed in Table

3.6. The numbers in the table are the mean and standard deviation of the CHR (for time to first failure) over 200 replicates of simulation studies when the joint survival function is $0, 1/6, 2/6, 3/6, 4/6, 5/6$, and 1, respectively. For Scenarios 2, 4, and 6 of the 4th simulation study in which the bias in estimating the CCSHR was most severe, the ratios were strikingly non-constant. This implies the association between the first failure times of a pair regardless of cause may not follow gamma frailty dependence structure, even though the association for the cause-specific failure time does. Thus, both herein and in the 2002 paper by Bandeen-Roche and Liang, the bias in CCSHR estimation may reflect mis-specification in estimating its size multiplicand rather than undue sensitivity to the shape distributional assumption.

3.4 Sensitivity to assumption 2: Independence of size and shape frailty

The simplicity of the Bandeen-Roche and Liang method becomes possible by assuming the size frailty A and the shape frailty $B(x)$ are statistically independent. This means the overall tendency to fail early or late should not relate to the propensity to fail from a specific cause at any time. This assumption allows the CCSHR (Equation (3.6)) to be decomposed into multiplicands which respectively characterize the propensity to fail from a particular cause and dependence in the timing of one's earliest failure regardless of cause.

In this section, we evaluate the effect which the dependence structure between the size

Table 3.6: CHR as a function of joint survival function $S(t)$. Non-constant trends with $S(t)$ indicate departure from gamma frailty dependence structure in paired times to first failure (Chen and Bandeen-Roche, 2005).

Simulation 2:		S(t)=0						S(t)=1	
Scenario 1	Mean	1.491	1.561	1.519	1.506	1.469	1.505	1.456	
	SD	0.230	0.238	0.252	0.257	0.323	0.412	0.639	
Scenario 2	Mean	1.805	1.760	1.801	1.788	1.780	1.812	1.834	
	SD	0.267	0.238	0.269	0.293	0.319	0.439	0.771	
Scenario 3	Mean	1.494	1.508	1.493	1.524	1.481	1.548	1.531	
	SD	0.213	0.214	0.244	0.265	0.330	0.413	0.718	
Scenario 4	Mean	1.805	1.736	1.773	1.742	1.817	1.706	1.824	
	SD	0.317	0.243	0.267	0.296	0.350	0.454	0.721	
Scenario 5	Mean	1.491	1.496	1.504	1.480	1.552	1.531	1.416	
	SD	0.219	0.212	0.220	0.252	0.327	0.414	0.692	
Scenario 6	Mean	1.812	1.805	1.765	1.724	1.751	1.689	1.823	
	SD	0.259	0.242	0.251	0.295	0.306	0.419	0.875	
Simulation 3:		S(t)=0						S(t)=1	
Scenario 1	Mean	1.547	1.517	1.498	1.524	1.512	1.498	1.526	
	SD	0.245	0.232	0.206	0.242	0.320	0.389	0.646	
Scenario 2	Mean	1.752	1.857	1.848	1.914	1.928	1.995	2.034	*
	SD	0.267	0.241	0.282	0.353	0.322	0.434	0.821	
Scenario 3	Mean	1.524	1.507	1.512	1.520	1.487	1.464	1.517	
	SD	0.208	0.232	0.226	0.232	0.333	0.369	0.672	
Scenario 4	Mean	1.800	1.876	1.890	2.003	1.935	1.974	2.162	*
	SD	0.294	0.252	0.276	0.366	0.371	0.475	0.829	
Scenario 5	Mean	1.519	1.484	1.545	1.522	1.537	1.526	1.452	
	SD	0.250	0.197	0.227	0.246	0.294	0.406	0.661	
Scenario 6	Mean	1.967	1.920	1.961	2.036	2.091	2.146	2.233	*
	SD	0.310	0.306	0.290	0.344	0.410	0.548	0.880	

and the shape frailty has on estimation of the CCSHR when the size frailty A is gamma distributed and the shape frailty $B(x)$ is beta distributed. Assuming only the 'size-shape frailty' framework and not the independence of A and $B(x)$,

$$\theta_{CS}(x_1, x_2; k_1, k_2) = \frac{E[A^2 B_{K_1}(x_1) B_{K_2}(x_2) \Lambda^*(x_1, x_2)] E[\Lambda^*(x_1, x_2)]}{E[A B_{K_1}(x_1) \Lambda^*(x_1, x_2)] E[A B_{K_2}(x_2) \Lambda^*(x_1, x_2)]}, \tag{3.11}$$

where $\Lambda^*(x_1, x_2) = \exp\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t) dt\}$. This is Equation (9) in Bandeen-Roche and Liang (2002). Equation (10) in this paper,

$$\frac{E\{B_{K_1}(x_1) B_{K_2}(x_2)\}}{E\{B_{K_1}(x_1)\} E\{B_{K_2}(x_2)\}} \times \frac{E[A^2 \Lambda^*(x_1, x_2)] E[\Lambda^*(x_1, x_2)]}{E^2[A \Lambda^*(x_1, x_2)]}, \tag{3.12}$$

on the other hand, decomposes $\theta_{CS}(x_1, x_2; k_1, k_2)$ based on the assumption that A and $B(x)$ are independent. Thus the effect of the assumption of independence between A and $B(x)$ can be assessed by directly comparing the CCSHR calculated by Equation (3.11) and parametrically estimated using Equation (3.12).

In this section, we will approximate true values of the CCSHR for various degrees of dependence between size and shape frailty. Then we compare them with parametric and nonparametric estimates of CCSHR.

3.4.1 Methods

First, we studied the difference between the CCSHR surfaces as functions of t_1 and t_2 when A and $B(x)$ are independent versus dependent. To approximate these surfaces, we generated a random sample of 2000 realizations of size frailty A and shape frailty B, with scenario-specific details to follow shortly. $\mathrm{CCSHR}_{1,1}$, $\mathrm{CCSHR}_{1,2}$, and $\mathrm{CCSHR}_{2,2}$ were obtained using Equation (3.11), replacing expectations by sample means and using $\lambda_m^*(t) = 1$. These CCSHRs were evaluated on a grid consisting of Cartesian products of 1st to 99th percentiles of failure time points generated from an exponential distribution as in the first set of simulation studies in Section 3.

For an independent case, we generated gamma-distributed size frailty A with mean 1 and variance 1 and time-invariant, beta-distributed shape frailty B with parameters 0.2 and 0.8 sampled independently from A. To construct a dependent sample (A^*, B^*) from A and B, we generated a bivariate standard normal-distributed sample with a pre-specified correlation value. We obtained ranks within the first components of the bivariate sample and ranks within the second components; then we re-ordered A and B yielding A^* with the same ranks as the first components of the bivariate sample and B^* with the second components.

After studying the effect of varying the joint distribution of A and B on the true CCSHR values, we evaluated the performance of the Bandeen-Roche and Liang's parametric and nonparametric estimator when A and B are not independent. The parametric estimator of CCSHR was obtained by plugging in maximum likelihood estimates of R_1 and Δ from the beta distribution model into Equation (3.7). We implemented the time-invariant nonparametric estimator of CCSHR described by Bandeen-Roche and Liang (2002) for the CCSHR between cause 1 and cause 1. This estimator compares concordances and discordances for parings of pairs, where a concordance occurs if both failure times of cause 1 for one pair in the pairing are greater than both failure times of cause 1 in the other pair in the pairing, and a discordance occurs otherwise. If all four members of a pairing were observed to fail from cause 1, then a concordance or discordance can be confirmed. If the smaller observation among the first components of the two pairs and the smaller one among the second components were observed to fail from cause 1, then we can confirm concordance/discordance status since the concordance/discordance among observed or latent cause 1 failure times coincides with that among observed (minimum) failure times. On the other hand, either in the first components or the second components, if the smaller observation failed of cause 2, then we cannot decide whether it is concordant or discordant.

3.4.2 Results

When A and B are statistically independent, $\theta_{CS}(x_1, x_2; 1, 1) = 6$ for all (x_1, x_2), and indeed our approximation of this function using the method described in the first paragraph of the previous section was virtually constant (near 6). As the correlation of the bivariate normal distribution used to generate dependence between A and B increased, we observed the CCSHR to increase throughout the $(x_1$, $x_2)$ space (Figure 3.1), particularly rapidly in the upperright region. Conversely, the CCSHR decreased throughout the $(x_1$, $x_2)$ space

Correlation	$\mathrm{CCSHR}_{1,1}(0.2, 0.2)$	$\mathrm{CCSHR}_{1,1}(0.5, 0.5)$	$\mathrm{CCSHR}_{1,1}(0.8, 0.8)$
1	7.909	20.857	115.207
0.7	7.275	11.992	26.169
0.4	6.640	8.234	12.061
0	5.977	6.074	6.421
-0.5	4.514	3.882	3.213
-1	1.614	1.295	1.315

Table 3.7: $\mathrm{CCSHR}_{1,1}$ values resulting for three values of (x, x) and various degrees of correlation between A and B

as the correlation decreased below 0. Table 3.7 displays CCSHR values at three diagonal (x, x) points. Our work further indicates that the CCSHR increases with x_1 and x_2 when A and B are positively correlated and decreases with x_1 and x_2 when A and B are negatively correlated (Figure 3.1). As electronic supplementary material, we present CCSHR contour plots for various degrees of dependence between A and B.

Global (time-invariant) estimates of $\mathrm{CCSHR}_{1,1}$ are presented for comparison: the 2nd column of Table 3.8 presents parametric estimates of CCSHR as in Bandeen-Roche and Liang (2002), and the 3rd column presents nonparametric estimates of $\mathrm{CCSHR}_{1,1}$. Since the parametric estimation does not take the dependence between A and B into consideration, the estimates for all correlation values were close to 6, the true value under independence. In contrast, the nonparametric estimates of the CCSHR increase as the correlation increases, resembling the pattern of underlying values of the CCSHR.

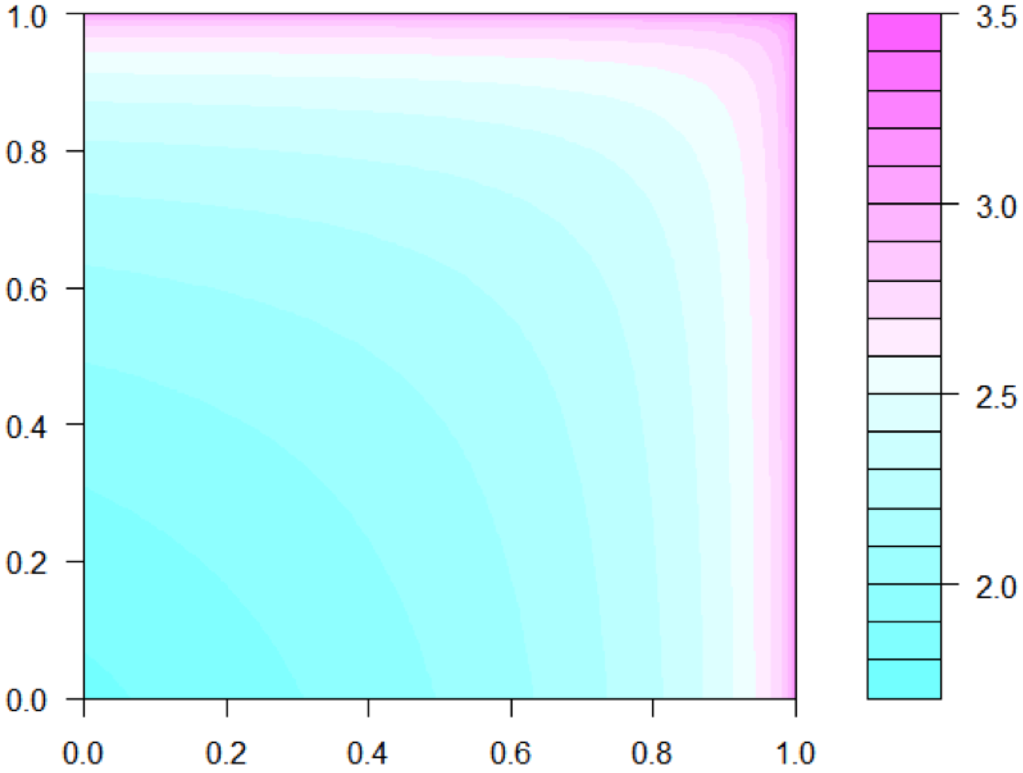


Figure 3.1: Contour plot of $\log(\mathrm{CCSHR}_{1,1})$ when the correlation of the bivariate normal distribution inducing dependence between A and B (Section 4.1) is 0.4. The horizontal and vertical axes are failure time percentiles.

Correlation	Parametric	Nonparametric
	Mean (SD)	Mean (SD)
1	5.992 (0.296)	8.150 (0.566)
0.7	6.001 (0.290)	7.871 (0.552)
0.4	6.005 (0.283)	7.266 (0.559)
0	6.000 (0.296)	6.040 (0.505)
-0.5	6.008 (0.299)	3.892 (0.355)
-1	5.994 (0.301)	1.344 (0.14)

Table 3.8: Comparison of parametric and nonparametric estimates of $\mathrm{CCSHR}_{1,1}$ for various degrees of dependence between A and B

3.5 Discussion

This paper addressed association among paired failure times subject to a competing risk, as defined by conditional cause-specific hazard ratios (CCSHRs), and estimated in the parametric framework proposed by Bandeen-Roche and Liang (2002). This framework partitions the CCSHR into two factors–one reflecting association between times to earliest failure regardless of cause (overall hazard ‘size’), and a second reflecting association between the causes of failure (cause allocation ‘shape’). We implemented a new estimator in this framework based on a logit-normal shape frailty distribution and compared its performance with an existing one based on a beta shape frailty distribution, in data scenarios generated from each distribution within the framework as well as scenarios outside the framework. We also studied the effect of dependence between overall failure propensity and the allocation of this among causes on the CCSHR magnitude and temporal variation, and we evaluated the robustness of the Bandeen-Roche and Liang estimator of the CCHSR to such dependence. We found little difference in estimator performance between the two shape-generating distributions, but large implications of dependence between size and shape frailty for the magnitude and temporal variation of failure time associations hence for estimator performance.

When size and shape were generated independently, both beta- and logit-normal-based estimators estimated the CCSHR accurately when data were generated according to the Bandeen-Roche and Liang framework, regardless of the underlying shape distribution. When data were generated according to models outside the Bandeen-Roche and Liang framework, both estimators exhibited biases comparable to those observed in the 2002 paper; however, based on our application of diagnostics for model fit, we suspect this owes primarily to mis-modeling of the association in first failure times ('size' association) rather than sensitivity to the 'shape' distributional assumption. We conclude parametric estimation of shape component of CCSHR will adequately estimate the CCSHR in many circumstances, provided that association in first failure times is characterized carefully as a function of time. The estimator employing a beta distribution assumption was comparably accurate and precise as the logit-normal-based estimator, hence we recommend both for paired failure-time data.

Independence between the size and shape frailty is a key feature enabling the simplified likelihood formulation in the Bandeen-Roche and Liang framework. Dependence in size and shape induces a mathematically complex likelihood form as well as a complicated time dependence of the resulting CCSHR. It remains to be seen whether a simply estimable, and interpretable, method can be developed to accommodate this scenario. We conjecture that this model is only weakly identifiable from one in which independence of size and shape frailty is maintained but the overall ('size'-dependent) association is allowed to vary arbitrarily with time. If so, one might retain a parametric estimation of shape component of CCSHR together with nonparametric estimation of size component, a flexibly time-varying conditional hazard ratio with CCSHR estimator as a multiplication of shape and size components. In such an approach, methods which accommodate estimation of a time varying ratio of cause-specific to overall hazard, $R(t)$, within the shape component of CCSHR (Equation (3.7)) may well be needed.

A limitation of our work is that we have only evaluated scenarios with two competing causes and two candidate shape distributions. The similarity we observed in estimator performance comparing beta and logit-normal models, as well as for generalized beta distributions (data not shown), is not surprising because logit-normal and Dirichlet distributions closely approximate each other when there are only two categories to be modeled. More substantive differences likely would emerge for 3 or more competing causes, because the logit-normal distribution admits more flexible correlation structures in this case.

We believe there is merit in distinguishing contributions to associations among clustered failures of multiple types into shared overall failure risk and shared failure cause propensities. Multimorbidity–an important and common setting in which clustered failures of multiple types arises–may reflect, both, individuals' overall vulnerability to physiological declines and disease-specific mechanisms (Varadhan et al., 2014). Partitioned disease heritability into these two components, methodology as discussed in this paper could inform the etiology of psychiatric disorders, metabolic syndrome, frailty in aging, and other medical syndromes. Whether the partitioning proposed here does address this goal, alternative means and measures for achieving the goal are still needed, representing another area of needed work.

Chapter 4

Nonparametric and Semiparametric Estimation of Association in Bivariate Failure-time Data under Competing Risks

4.1 Introduction

A large literature on failure-time analysis has addressed univariate data where the observations are independently and identically distributed, and there is only one cause of failure. However, a considerable body of work in recent decades has extended this traditional survival analysis approach. Multivariate failure-time data analysis accommodates multiple observations in a single sampling unit, for example, times to onset of a disease for multiple family members, wherein the observations may be correlated. A number of researchers have aimed to assess the association among multivariate failure times and suggested various measures to this end (Hougaard, 2000). Among the proposed measures, the conditional (or cross) hazard ratio (CHR) is an easily interpreted description of association (Clayton & Cuzick, 1985; Clayton, 1978; Oakes, 1982, 1986, 1989), and parametric models of it have been studied by many researchers (Genest et al., 1995; Genest & MacKay, 1986; Glidden, 2000; Nielsen et al., 1992; Oakes, 1989; Ripatti et al., 2002; Ripatti & Palmgren, 2000; Shih & Louis, 1995). Nonparametric measures of time-varying association have also been proposed (Fan, Prentice, & Hsu, 2000; Hsu & Prentice, 1996). Another direction of extension of traditional survival analysis has been to accommodate competing risks. Competing risks data are frequently encountered in biomedical studies, where subjects may experience failure from one of multiple causes. In such data, we may only observe the time to the first failure experienced and the cause of this failure. In this paper, we address the situation where the failure-time data are both multivariate and reflect competing risks using our measure of association, the conditional cause-specific hazard ratio (CCSHR), an extension of the CHR proposed by Bandeen-Roche and Liang (2002).

A considerable body of literature has considered the estimation of association among multivariate failure times subject to competing risks (Bandeen-Roche & Ning, 2008; Cheng & Fine, 2008; Cheng & Fine, 2012; Cheng, Fine, & Bandeen-Roche, 2010; Cheng, Fine, & Kosorok, 2007, 2009; Ning & Bandeen-Roche, 2014; Scheike & Sun, 2012; Scheike et al., 2010). A few researchers have observed that associations among multivariate failure times subject to competing risks can be decomposed into two elements: the association between the probabilities that the individuals experience a specific failure type, and the association between times to first failure experienced regardless of cause. Bandeen-Roche and Liang (2002) demonstrated that under certain assumptions, the CCSHR can be expressed as a multiplication of the ordinary CHR for association between times to first failure, and a factor representing the association between failure causes. Shih and Albert (2010) also adopted the framework where the overall association between cause-specific failure times could be decomposed into the association between failure times and the association between failure causes. For the former, they also utilized the ordinary CHR for association between times to first failure; for the latter, they defined an odds ratio of having a specific pair of failure causes conditional on the first-failure times. The previous approaches to estimating the decompositions described above either have been parametric (Bandeen-Roche & Liang, 2002) or modeled time dependence as piecewise constant (Shih & Albert, 2010).

The decompositions described above distinguish associations between causes of failure from associations between overall propensity to fail, which may yield insights into shared determinants of disease onset in people (e.g. comorbidity), families (e.g. genetics), or communities (e.g. environment). In this paper, we aimed to augment the available methodology to achieve such insights by enabling completely nonparametric estimation of each component in the CCSHR decomposition proposed by Bandeen-Roche and Liang (2002). To this end, we developed methods to estimate both failure-time and failure-cause components of the CCSHR by smoothing. In Section 2, we detail these methods. In Section 3, we study the performance of our methods through simulation studies. In Section 4, we present an application of our method to study familial associations in times to onset of dementia, using data from the Cache County Study on Memory Health and Aging. Section 5 concludes.

4.2 Methods

4.2.1 Definition and notation

We assume data are independently and identically distributed across pairs $i = 1, \cdots, n$ and censoring is independent of failure time. $S(t_1, t_2)$ is the joint survival function of T_1 and T_2, and $S_1(t_1)$ and $S_2(t_2)$ are the marginal survival functions of T_1 and T_2, respectively. Hazard functions corresponding to $S_m(t_m)$ are denoted by $\lambda_m(t_m)$, and cause-specific hazard functions for the m-th individual is defined as $\lambda_{m,k}(t_m)$.

The conditional hazard ratio (CHR) is defined by

$$\theta(t_1, t_2) = \frac{\lambda(t_2|T_1 = t_1)}{\lambda(t_2|T_1 > t_1)} = \frac{f(t_1, t_2)S(t_1, t_2)}{\left.\frac{\partial S(s_1, t_2)}{\partial s_1}\right|_{s_1=t_1} \cdot \left.\frac{\partial S(t_1, s_2)}{\partial s_2}\right|_{s_2=t_2}}, \tag{4.1}$$

the ratio of an individual's hazard of failure at t_2 given failure of his pair partner at t_1 to the hazard given that the partner has not yet failed by t_1. Oakes (1989) showed that the CHR can be alternatively expressed as

$$\frac{\Pr\{(T_1^{(a)} - T_1^{(b)})(T_2^{(a)} - T_2^{(b)}) > 0 | (T_1^{(ab)}, T_2^{(ab)}) = (t_1, t_2)\}}{\Pr\{(T_1^{(a)} - T_1^{(b)})(T_2^{(a)} - T_2^{(b)}) < 0 | (T_1^{(ab)}, T_2^{(ab)}) = (t_1, t_2)\}} \tag{4.2}$$

where $T^{(a)} = (T_1^{(a)}, T_2^{(a)})$ and $T^{(b)} = (T_1^{(b)}, T_2^{(b)})$ are two randomly chosen bivariate observations and $(T_1^{(ab)}, T_2^{(ab)})$ is the componentwise minimum of $T^{(a)}$ and $T^{(b)}$.

Our approach to augmenting the CHR can be best motivated through a frailty formulation. A frailty variable is a positive random effect that multiplicatively modifies the hazard function of both individuals in a pair. That is, the hazard function for each m-th individual in the i-th pair is $\lambda_{im}(t) = a_i \lambda_m^*(t)$ where a_i is the realization of the frailty variable for the i-th pair and $\lambda_m^*(t)$ is the 'reference' hazard function conditional on $a_i = 1$. The bivariate survival function with frailty variable A can be expressed as follows:

$$S(x_1, x_2) = \int \exp\Big\{-a \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t)dt\Big\} dG(a) = E\Big[\exp\Big\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t)dt\Big\}\Big], \tag{4.3}$$

where G is the distribution of the frailty variable. Taking partial and second derivatives here to compute all the terms needed in Equation (4.1), then the bivariate density can be expressed in terms of the frailty variable A and λ_m^* as

$$f(x_1, x_2) = \lambda_1^*(x_1)\lambda_2^*(x_2) E[A^2 \exp\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t)dt\}], \tag{4.4}$$

and the conditional hazard ratio can be expressed as

$$\theta(x_1, x_2) = \frac{E[A^2 \exp\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t)dt\}]E[\exp\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t)dt\}]}{E^2[A \exp\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t)dt\}]}. \tag{4.5}$$

(Liang et al., 1995)

4.2.2 Introduction to CCSHR and shape-size decomposition

Bandeen-Roche and Liang (2002) introduced the concept of CCSHR as a measure of association between bivariate failure times with competing risks, defined as

$$\begin{aligned}\theta_{CS}(x_1, x_2; k_1, k_2) &= \frac{\lambda_{1,k_1}(x_1|X_2 = x_2, K_2 = k_2)}{\lambda_{1,k_1}(x_1|X_2 > x_2)} \\ &= \frac{S(x_1, x_2)f(x_1, x_2; k_1, k_2)}{\{\int_{x_2}^{\infty} \sum_{k=1}^{2} f(x_1, x, k_1, k)dx\}\{\int_{x_1}^{\infty} \sum_{k=1}^{2} f(x, x_2, k, k_2)dx\}}. \end{aligned} \tag{4.6}$$

It is interpreted as a multiplicative factor by which one's hazard of failing at time x_1 due to failure cause k_1 is inflated when his pair partner fails at time x_2 due to cause k_2 compared to when the partner has not yet failed by time x_2 due to any cause. Bandeen-Roche and Liang pointed out it is possible to decompose the CCSHR into the association between failure *times* regardless of the failure cause (which they termed the hazard 'size' association) and the association between the propensities of failing due to specific *causes* (which they termed the hazard 'shape' association).

In competing risks analysis, the overall hazard is considered as the sum of cause-specific hazard functions. Bandeen-Roche and Liang generalized this idea to the multivariate failure time setting by envisioning that the proportions of allocation of the overall hazard function into cause-specific hazards may heterogeneously vary across pairs, and not only the overall hazard. To this end, they defined two types of frailties: A positive random variable A that governs a pair's tendency to fail early or late (regardless of the cause), and a stochastic process $B(x) = (B_1(x), B_2(x))$ that tailors the pair's allocation of the overall hazard to the respective causes. Then, they defined the hazard function conditional on these frailties as $\lambda_m(x_m|A = a, B_1(x_m) = b_1(x_m), B_2(x_m) = b_2(x_m)) = ab_1(x_m)\lambda^*_m(x_m) + ab_2(x_m)\lambda^*_m(x_m)$ where $b_1(x_m) + b_2(x_m) = 1$. The vector $B(x) = (B_1(x), B_2(x))$ has a mean function $R(x) = (R_1(x), R_2(x))$ where $R_k(x) = \lambda_{mk}(x)/\lambda_m(x)$ which for simplicity we take to be equal across components $m = 1, 2$, in our paper. Henceforth, A will be called the size frailty and B, the shape frailty.

Synthesizing, the cause-specific hazard of cause k_m for the m-th individual given the frailties is $\lambda_{mk_m}(x) = AB_{mk_m}(x)\lambda^*_m(x)$. Then, in analogy to Equation (4.4), the corresponding bivariate cause-specific density $f(x_1, x_2; k_1, k_2)$ equals $\lambda^*_1(x_1)\lambda^*_2(x_2)E\Big[A^2 B_{k_1}(x_1)B_{k_2}(x_2)\exp\Big\{-A\sum_{m=1}^{2}\int_0^{x_m}\lambda^*_m(t)dt\Big\}\Big]$ (Bandeen-Roche & Liang, 2002). By plugging it into Equation (4.6), the CCSHR between cause k_1 and k_2 can be expressed as:

$$\theta(x_1, x_2) = \frac{E[A^2 B_{k_1}(x_1) B_{k_2}(x_2) \exp\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t) dt\}] E[\exp\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t) dt\}]}{E[A B_{k_1}(x_1) \exp\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t) dt\}] E[A B_{k_2}(x_2) \exp\{-A \sum_{m=1}^{2} \int_0^{x_m} \lambda_m^*(t) dt\}]}. \tag{4.7}$$

Bandeen-Roche and Liang assumed statistical independence between the size frailty A, and the shape frailty, $B(x)$, simplifying Equation (4.7) into factors

$$\frac{E[B_{k_1}(x_1) B_{k_2}(x_2)]}{E[B_{k_1}(x_1)] E[B_{k_2}(x_2)]} \times \theta(x_1, x_2) \tag{4.8}$$

where $\theta(x_1, x_2)$ is the ordinary conditional hazard ratio. Henceforth, we will call the first multiplicand the 'shape' component of the CCSHR and the second multiplicand, the 'size' component.

The shape factor governs the association between pair members' failure causes, and the size component measures the strength of association between bivariate failure times regardless of failure cause. Our approach is to estimate these two multiplicands separately, each as a function of bivariate failure times. The next two subsections will discuss existing estimators and our new approach for each component.

4.2.3 Estimation of the shape component

In this section, we present an existing method of estimating the shape component of the CCSHR, and then suggest an alternative. The existing method is semiparametric estimation as suggested by Bandeen-Roche and Liang (2002). Our new approach is nonparametric estimation using an alternative representation of the CCSHR proposed by Shih and Albert (2010).

Method 1: Semiparametric Dirichlet Model – This method assumes that $B(t)$ is a two-dimensional beta-distributed process with parameter $(\delta_1(t), \delta_2(t)) = \Delta \times (R_1(t), R_2(t))$ and combines a parametric estimator of Δ with a nonparametric estimator of $R(t)$. For the present, censored observations are excluded in the estimation of $R(t)$ and Δ.

- Estimation of $R(t)$: The function $R(t) = (R_1(t), R_2(t))$ is defined as a pointwise division of a cause-specific hazard function over an overall hazard function, $\lambda_k(t)/\lambda(t),\ k = 1, 2$. Here we do not distinguish the (cause-specific) hazard functions of the 1st and 2nd individuals of each pair, nor $R(t)$, but rather assume these functions are common for both members of a pair. We therefore take as input for estimation univariate failure time data with corresponding failure causes, pooling the 1st and 2nd individuals in pairs without distinguishing them. To estimate $R(t)$ using these data, we obtain nonparametric estimates of the cause-specific hazard functions and overall hazard function using an existing software such as 'muhaz' function in R 'muhaz' package, and then divide them. Here, we constrain the estimates of $R(t)$ to be confined within the range of $[\epsilon, 1 - \epsilon]$ by winsorizing for a small positive number ϵ to prevent nonsensical values in the next steps. This shape component estimator will be called 'Shape1' hereafter.

- Estimation of Δ: Δ is a parameter which controls the strength of failure cause association within a pair. Under conditions delineated in Bandeen-Roche and Liang (2002), it may be estimated as follows. If pairs are sampled independently, the likelihood function for the frailty distribution parameters and reference hazard function is

$$\prod_{i=1}^{n} E\Big\{B_{k_{i1}}(x_{i1})B_{k_{i2}}(x_{i2})\Big\}E\Big[A^2\lambda_1^*(x_{i1})\lambda_2^*(x_{i2})\exp\Big\{-A\sum_{m=1}^{2}\int_0^{x_{im}}\lambda_m^*(t)dt\Big\}\Big]. \tag{4.9}$$

Since this likelihood function factorizes into a shape frailty component versus a size frailty and hazard function component, estimation of shape frailty parameters involves only $\prod_{i=1}^{n} E\Big\{B_{k_{i1}}(x_{i1})B_{k_{i2}}(x_{i2})\Big\}$. We propose to reduce this likelihood to a function of a single variable rather than two by invoking the assumption that $B(x)./R(x)$ is a martingale process: Bandeen-Roche and Liang (2002) showed when this occurs, both time arguments may be evaluated at their minimum, $x = (\min(x_{i1}, x_{i2}))$. Then, under the beta distribution assumption, $E\Big\{\frac{B_{k_1}(x_1\wedge x_2)B_{k_2}(x_1\wedge x_2)}{R_{k_1}(x_1\wedge x_2)R_{k_2}(x_1\wedge x_2)}\Big\} = \frac{\Delta}{\Delta+1}$ for $k_1 \neq k_2$, and equals $\frac{\Delta + R_k^{-1}(x_1\wedge x_2)}{\Delta+1} = \frac{R_k^{-1}(x_1\wedge x_2)-1}{\Delta+1}+1$ for $k_1 = k_2$. The likelihood with respect to Δ, then, is proportional to $\prod_{i\in I_1\cup I_2}\Big[1-\frac{1}{\Delta+1}\{1-R_{k_1}^{-1}(x_{i1}\wedge x_{i2})\}\Big]\prod_{i\in I_3}\Big(1-\frac{1}{\Delta+1}\Big)$ where

$$\begin{aligned} I_k &= \{i : \text{both members of pair } i \text{ fail due to cause } k\} \quad (k=1,2) \\ I_3 &= \{i : \text{members of pair } i \text{ fail due to different causes}\}. \end{aligned} \tag{4.10}$$

With this likelihood we can estimate Δ by solving the corresponding score equation:

$$\frac{n_3}{\Delta(\Delta+1)}-\sum_{k=1}^{2}\sum_{i\in I_k}\Big(\frac{1}{(\Delta+1)^2}\Big\{\frac{1-R_k(x_{i1}\wedge x_{i2})}{R_k(x_{i1}\wedge x_{i2})}\Big\}\Big[\frac{1}{\Delta+1}\Big\{\frac{1-R_k(x_{i1}\wedge x_{i2})}{R_k(x_{i1}\wedge x_{i2})}\Big\}+1\Big]^{-1}\Big)=0, \tag{4.11}$$

plugging in for $R_{k_1}(x_{i1}\wedge x_{i2})$, the nonparametric estimates of $R_{k_1}(t)$ for $t\in(0,1)$ obtained in the previous step. Following this plug-in, the score equation can be solved by a simple R program.

Finally we obtain the shape component by plugging the maximum likelihood estimate of Δ and nonparametric estimates of $R(t)$ into the formula,

$$1-\frac{1}{\Delta+1}\Big\{\frac{R_{k_1}(x_1\wedge x_2)-1}{R_{k_1}(x_1\wedge x_2)}\Big\}^{I(k_1=k_2)}, \tag{4.12}$$

which is a function of $x_1\wedge x_2$ for the association between the same failure causes and a constant for the association between different causes.

Method 2: Nonparametric Estimator – Shih and Albert (2010) showed that the CCSHR can be alternatively represented as

$$\theta(t_1,t_2)\cdot\frac{\Pr(K_1=k_1,K_2=k_2|T_1=t_1,T_2=t_2)}{\Pr(K_1=k_1|T_1=t_1,T_2>t_2)\Pr(K_2=k_2|T_2=t_2,T_1>t_1)}. \tag{4.13}$$

They proposed quite complicated methodology to estimate the quantities in the right-hand multiplicand. Rather, we propose to directly estimate each component of Equation (4.13) nonparametrically. The numerator and the two components of the denominator can be expressed as functions of t_1 and t_2:

$$g_1(t_1, t_2) = \Pr(K_1 = k_1, K_2 = k_2 | T_1 = t_1, T_2 = t_2) \tag{4.14}$$

$$g_2(t_1, t_2) = \Pr(K_1 = k_1 | T_1 = t_1, T_2 \geq t_2) \tag{4.15}$$

$$g_3(t_1, t_2) = \Pr(K_2 = k_2 | T_2 = t_2, T_1 \geq t_1) \tag{4.16}$$

To estimate g_1, g_2, and g_3, a smoothing method may be applied to the data fully observed for the event of interest. That is, for the estimation of g_1, we need a subset of the full dataset with failures observed for both the 1st and 2nd individuals, and for the estimation of g_2, we need another subset where the 1st individuals are not censored, and similarly for g_3. Thus, the input data for the smoothing of g_1 are two explanatory variables, t_{i1} and t_{i2}, and the response variable, $I(K_{i1} = k_1, K_{i2} = k_2)$, for $i \in I(K_1 > 0, K_2 > 0)$. For the estimation of g_2 (and similarly for g_3), the estimand at (t_1, t_2) is the 'proportion' of the population of whom first members of pairs are representative who fail due to cause k_1, conditional on failure time $T_1 = t_1$ and having a pair partner who fails later than t_2. To estimate this, suppose we simply apply a smoothing method to the dataset $\{t_{i1}, t_{i2}, \text{prop}(K_1 = k_1)\},\ i \in I(K_1 > 0)$ where 'prop()' indicates the proportion of observations satisfying the condition in the parenthesis on the area $T_1 = t_1$ and $T_2 \geq t_2$. Unfortunately, this strategy very likely will estimate $\Pr(K_1 = k_1 | T_1 = t_1, T_2 = t_2)$ rather than $\Pr(K_1 = k_1 | T_1 = t_1, T_2 \geq t_2)$, if the time variable is continuous, because it is highly likely that no observation satisfies $T_1 = t_1$ and $T_2 > t_2$ in this case. To circumvent this complexity, we use observations from the 'band', $\{t_1 - w/2 < T_1 < t_1 + w/2, T_2 \geq t_2\}$, rather than from the 'line', $\{T_1 = t_1, T_2 \geq t_2\}$, where w is an appropriately chosen positive number. Specifically, we calculate the proportion of the 1st individuals who failed due to cause k_1 among non-censored individuals, $\#I(K_1 = k_1)/\#I(K_1 > 0)$, from the rectangular area, $\{t_{i1} - w/2 < T_1 < t_{i1} + w/2, T_2 \geq t_{i2}\}$, and consider this as a response value in the input data for smoothing.

Any smoothing method that can take continuous or binomial response variables and two explanatory variables can be used. Based on the authors' work described elsewhere, we used generalized additive models with Gaussian family and bivariate smoothing function: $g(E(y)) = f(x_1, x_2)$ where $g(E(y))$ is an identity link function. After obtaining the smoothed estimates of g_1, g_2, and g_3 above, pointwise multiplication and division leads to the estimate of the shape component for any (t_1, t_2). This estimator will be called 'Shape2' hereafter.

4.2.4 Estimation of the size component

The second multiplicand of the CCSHR, the size component, is same as the conditional hazard ratio under the framework defined in Section 2.2. A considerable literature has discussed parametric modeling of the CHR as a function of bivariate failure times (Clayton & Cuzick, 1985; Genest & MacKay, 1986; Oakes, 1989; Shih & Louis, 1995). A much sparser literature had addressed nonparametric estimation. In this paper, we evaluated a method that does not impose any parametric assumption in estimating the CHR, allowing it to be fully time-varying as a function of t_1 and t_2, in comparison to two parametric approaches:

Method 1: Nonparametric Estimator – The first method nonparametrically estimates a local version of Kendall's τ. Consider two realizations from the same bivariate failure time distribution, $T^{(a)} = (T_1^{(a)}, T_2^{(a)})$ and $T^{(b)} = (T_1^{(b)}, T_2^{(b)})$, and denote their corresponding pairwise minimum as $(T_1^{(ab)}, T_2^{(ab)}) = (\min(T_1^{(a)}, T_1^{(b)}), \min(T_2^{(a)}, T_2^{(b)}))$. Recall that $T^{(a)}$ and $T^{(b)}$ are concordant if $(T_1^{(a)} - T_1^{(b)})(T_2^{(a)} - T_2^{(b)}) > 0$ and are discordant if $(T_1^{(a)} - T_1^{(b)})(T_2^{(a)} - T_2^{(b)}) < 0$. As noted by Oakes (1989; Equation (4.2)), the CHR can be written as a ratio of concordance and discordance probabilities and hence relates directly to a local version of Kendall's τ, $\tau(t_1, t_2) = P\{(T_1^{(a)} - T_1^{(b)})(T_2^{(a)} - T_2^{(b)}) > 0 | (T_1^{(ab)}, T_2^{(ab)}) = (t_1, t_2)\} - P\{(T_1^{(a)} - T_1^{(b)})(T_2^{(a)} - T_2^{(b)}) < 0 | (T_1^{(ab)}, T_2^{(ab)}) = (t_1, t_2)\}$. Therefore, we propose to estimate the CHR by first obtaining a smoothed estimator of this local Kendall's τ, and then back-transform as $\theta(t_1, t_2) = \dfrac{1 + \tau(t_1, t_2)}{1 - \tau(t_1, t_2)}$ to obtain the CHR estimator.

Our Kendall's τ estimator takes failure times T_1 and T_2 as two explanatory variables and a concordance-discordance indicator as a response variable. As a first step, then, we need to prepare a dataset of concordances and discordances. Beginning with bivariate failure times data on n pairs of individuals, we create a dataset of all available pairings of them ($n \times (n-1)/2$ pairings). For each of these pairings, for example, $T^{(i)} = (T_1^{(i)}, T_2^{(i)})$ and $T^{(j)} = (T_1^{(j)}, T_2^{(j)})$, we obtain $(\min(T_1^{(i)}, T_1^{(j)}), \min(T_2^{(i)}, T_2^{(j)}))$ and a concordance status indicator. The concordance status indicator is defined as +1 if $(T_1^{(i)} - T_1^{(j)})(T_2^{(i)} - T_2^{(j)}) > 0$, -1 if $(T_1^{(i)} - T_1^{(j)})(T_2^{(i)} - T_2^{(j)}) < 0$, and 0 if $(T_1^{(i)} - T_1^{(j)})(T_2^{(i)} - T_2^{(j)}) = 0$, or we may randomly assign +1 or -1 in this latter case if a binary outcome is required in the selected smoothing method. Then, we smooth concordance status data in terms of the pairwise minimum times. Any smoothing method which can take two explanatory variables can be used. As a result of an extensive simulation study comparing the performance of these smoothing methods for local Kendall's τ estimation, reported elsewhere, we recommend the GAM (Hastie & Tibshirani, 1986), $g(E(y)) = f(x_1, x_2)$ where $g(E(y))$ is identity or logit link function. To address censoring, we used multiple imputation among methods reported in Chapter 2 of this thesis. For logit link, one transforms the (1,-1) data to (1,0) and then back again. This nonparametric estimator of the size component will be called 'Size1' estimator.

Method 2: Shih and Louis Estimator, Method 3: Alternative Parametric Estimator – These estimators are similar in the sense that they estimate a copula parameter, α, for a specific, parametrically specified copula $C(S_1(t_1), S_2(t_2); \alpha) = S(t_1, t_2)$ which links the joint survival function with two marginal survival functions. We propose, as in the preceding thesis paper, to work with random 'standardized' arguments transformed to be uniformly distributed. Then, the relationship between the survival function and the copula function is $S(t_1, t_2) = C(1 - u_1, 1 - u_2; \alpha)$ with $S(t_m) = 1 - u_m$.

The second estimator is the maximum likelihood estimator due to Shih and Louis (1995). The likelihood function for the copula parameter is defined as $\prod_i L(\alpha; u_{1i}, u_{2i}) = \prod_i c(u_{1i}, u_{2i}; \alpha)^{\delta_{1i}\delta_{2i}} \dfrac{\partial c(u_{1i}, u_{2i}; \alpha)^{\delta_{1i}(1-\delta_{2i})}}{\partial u_{1i}} \dfrac{\partial c(u_{1i}, u_{2i}; \alpha)^{\delta_{2i}(1-\delta_{1i})}}{\partial u_{2i}} C(u_{1i}, u_{2i}; \alpha)^{(1-\delta_{1i})(1-\delta_{2i})}$ where i is the pair index from 1 to n and δ_{1i} and δ_{2i} are failure-censoring indicators. The estimator of α maximizes this function.

The third strategy chooses the parameter value which minimizes the mean absolute deviation between implied and estimated versions of the local Kendall's τ across the entire bivariate time domain. This can be accomplished using optimization software such as 'optim' function of R. The implied Kendall's τ is calculated from the one-to-one relationship between the specified copula (hence, bivariate survival) function using the CHR formula, $\theta(t_1, t_2) = \dfrac{f(t_1,t_2)S(t_1,t_2)}{\frac{\partial S(s_1,t_2)}{\partial s_1}\Big|_{s_1=t_1} \cdot \frac{\partial S(t_1,s_2)}{\partial s_2}\Big|_{s_2=t_2}}$, and the one-to-one relationship between CHR and local Kendall's τ, $\tau(t_1,t_2) = \dfrac{\theta(t_1,t_2)-1}{\theta(t_1,t_2)+1}$. The estimated Kendall's τ, $\hat{\tau}(t_i,t_j)$, can be obtained by the smoothing method as described in the first method of this section. The mean absolute deviation (MAD) between them can be approximated by $\frac{1}{99^2}\sum_{i=1}^{99}\sum_{j=1}^{99}\Big|\tau(u_i,u_j) - \tilde{\tau}(u_i,u_j)\Big|$ where $u_k = \dfrac{k}{100}$. As above, here we employed standardized times U_i rather than crude failure times T_i.

A challenge for Methods 2 and 3 is that the correct copula type is not known when analyzing data in practice. In our simulation study to be described shortly, we estimated fits (separately by Methods 2 and 3) for each of three Archimedean copulas: Clayton, Frank, and Gumbel. Then, the copula type achieving the best fit to the data as assessed by the MAD defined just above was chosen as the final fit. Methods 2 and 3 will be called 'Size2' and 'Size3' estimators hereafter.

4.3 Simulation studies

A set of simulation studies was designed to assess the performance of the proposed estimators and compare them with existing methods. We generated simulated datasets of various association structures from copula models. We estimated the shape and size components separately using the methods described in the previous section, then obtained the CCSHR by multiplying various combinations of shape and size component estimators. We assessed the estimators' accuracy and variability at specific time points.

4.3.1 Methods

We need to create a simulated dataset with four variables in which each observation consists of a pair of failure times and a pair of associated failure causes. In this simulation study, we assumed failure times to be uniformly distributed between 0 and 1 (reflecting conversion to zero-one scale by applying one minus survival function transformation). By standardizing the bivariate time domain into $[0, 1] \times [0, 1]$, we can compare the association structures of datasets with different time scales.

The bivariate failure times were created using the 'rCopula' function in the R 'copula' package with a specified copula type, dimension (two in this paper), and copula parameter value. To generate failure causes, we employ the Bandeen-Roche and Liang framework where the proportions of each failure cause for a given pair at time t have a beta distribution with two-dimensional parameter $\delta(t) = (\delta_1(t), \delta_2(t))$ decomposed into the multiplication of a parameter Δ representing the strength of cohesion of failure causes within a pair and the mean function of the proportions of each failure cause, $R(t) = (R_1(t), R_2(t))$. For each pair and each time point t, a beta distributed random number was generated and was compared with two standard uniformly distributed random numbers. If the beta random number was greater than the first uniform random number, the first individual's failure cause was set to 1, otherwise it was 2. The second individual's failure cause was similarly defined.

To create a simulated dataset, we must designate sample size, association structure (a copula type and parameter), and the distribution governing allocation of two failure causes (Δ and $R(t)$). We used two different sample sizes, 500 and 1,000. Among numerous types of copulas, we chose Clayton and Gumbel copulas. The Clayton copula represents an association structure with constant CHR over time, and Gumbel copula represents a structure with CHR that is decreasing over time. We designed six data generation scenarios. Scenarios 1, 2, and 6 employed a Clayton copula with parameter 1, which is equivalent to Pearson correlation coefficient 0.48. Scenario 3 employed a Gumbel copula with parameter 2.5 and scenarios 4 and 5 with parameter 1.125, which are equivalent to correlation coefficients of 0.79 and 0.29, respectively. The parameter Δ was fixed to 1 for all scenarios. The function of the proportion of the first failure cause $R_1(t)$ was assumed to be constant for scenarios $1 \sim 4$ where the values were 0.2, 0.8, 0.5, and 0.5, respectively. We used $R_1(t) = 0.4 \times \dfrac{1}{1 + \exp\left(-(t - 0.5) \times 10\right)} + 0.3$ for scenarios 5 and 6 which represents an S-shaped curve increasing from 0.3 to 0.7 when t changes from 0 to 1. Scenario 7 used the complement of $R(t)$ in scenarios 5 and 6, that is, a decreasing S-shaped curve from 0.7 to 0.3. Scenarios $1 \sim 4$ replicated scenarios evaluated by Bandeen-Roche and Ning (2008), who generated bivariate failure time data from gamma frailty with mean 1 and variance 1 (equivalent to scenarios 1 and 2) and positive stable frailty with $\alpha = 0.4$ and 0.8 (equivalent to scenarios 3 and 4).

We introduced two methods of estimation for the shape component and three methods for the size component in Sections 2.3 and 2.4. The combination of 'Shape1' and 'Size1' will be referred to as 'Method 1', that of 'Shape2' and 'Size1' as 'Method 2', that of 'Shape1' and 'Size2' as 'Method 3', and that of 'Shape1' and 'Size3' as 'Method 4.' Method 2 is a completely nonparametric method where both components are estimated by smoothing, and the other three methods are semiparametric.

Table 4.1: Labels for the shape and size components, and their combinations

	New nonparametric (Size1)	Shih and Louis (Size2)	New parametric (Size3)
Bandeen-Roche and Liang (Shape1)	Method 1		
New method (Shape2)	Method 2	Method 3	Method 4

Bandeen-Roche and Ning (2008) evaluated their estimators at four quadrants which were bisected at the medians of the first and second individuals' failure times. Their estimator was evaluated for specific rectangular regions, but our estimator is continuously varying and point-specific, thus is not directly comparable. Since our estimator supposes that the time variables were standardized to zero-one scale, we bisected each time axis at its median, 0.5, and then chose a representative point from each quadrant at which to evaluate our estimators. For comparison, we chose a representative point from each quadrant for which the true values of the CCSHR, calculated from the Bandeen-Roche and Liang formula, are closest to the area-specific CCSHR for the four quadrants in the 2008 paper. These representative points were $(0.20, 0.20)$, $(0.40, 0.60)$, $(0.60, 0.40)$, and $(0.70, 0.70)$.

Each scenario was repeated 300 times. For each replicate of simulated data, two shape estimates, three size estimates, and four CCSHR estimates were recorded at $(0.20, 0.20)$, $(0.40, 0.60)$, $(0.60, 0.40)$, and $(0.70, 0.70)$. We also evaluated contour plots of differences between true and estimated local Kendall's τ, to visualize variation in accuracy over time. To study the variability of our estimates, we used the bootstrap method. The bootstrapped samples were sampled as pairs from the original bivariate failure time data of size $n = 500$ or $1,000$ with sizes the same as that of the original dataset. For each of the 300 bootstrapped samples, we obtained two shape estimates, three size estimates, and four CCSHR estimates at the same points described above. The mean, standard deviation, 2.5th percentile, and 97.5th percentile at the same points from 300 bootstrapped samples were also collected. We could obtain the coverage probability that the 95% bootstrap confidence intervals (from the 2.5th percentile to the 97.5th percentile) include the true CCSHR value.

4.3.2 Results

First, we examine the biases of the separate estimators of the shape and size components (Table 4.2). In the following discussion, the bias is reported as $\frac{\text{(Estimate)} - \text{(True value)}}{\text{(True value)}}$, and we will mainly discuss the sample size 1,000 unless otherwise specified. In the 1st

Table 4.2: Shape and size components estimates and their biases from simulation studies

Copula type		Shape component			Size component			
Cause allocation	Location	True	Shape1	Shape2	True	Size1	Size2	Size3
	(0.20,0.20)	3.000	3.023 (0.008)	3.024 (0.008)	2.000	2.004 (0.002)	2.003 (0.002)	2.006 (0.003)
Clayton(1)	(0.40,0.60)	3.000	3.018 (0.006)	3.038 (0.013)	2.000	2.006 (0.003)	2.003 (0.002)	2.006 (0.003)
R = 0.2	(0.60,0.40)	3.000	3.018 (0.006)	3.035 (0.012)	2.000	2.013 (0.006)	2.003 (0.002)	2.006 (0.003)
	(0.70,0.70)	3.000	3.038 (0.013)	3.03 (0.010)	2.000	2.026 (0.013)	2.003 (0.002)	2.006 (0.003)
	(0.20,0.20)	1.125	1.122 (-0.003)	1.124 (-0.001)	2.000	2.004 (0.002)	2.003 (0.002)	2.006 (0.003)
Clayton(1)	(0.40,0.60)	1.125	1.123 (-0.002)	1.127 (0.002)	2.000	2.006 (0.003)	2.003 (0.002)	2.006 (0.003)
R = 0.8	(0.60,0.40)	1.125	1.123 (-0.002)	1.128 (0.003)	2.000	2.013 (0.006)	2.003 (0.002)	2.006 (0.003)
	(0.70,0.70)	1.125	1.124 (-0.001)	1.127 (0.002)	2.000	2.026 (0.013)	2.003 (0.002)	2.006 (0.003)
	(0.20,0.20)	1.500	1.502 (0.001)	1.505 (0.003)	6.094	5.944 (-0.025)	6.096 (0.000)	6.158 (0.011)
Gumbel(2.5)	(0.40,0.60)	1.500	1.501 (0.001)	1.494 (-0.004)	2.506	2.531 (0.010)	2.511 (0.002)	2.526 (0.008)
R = 0.5	(0.60,0.40)	1.500	1.501 (0.001)	1.501 (0.001)	2.506	2.518 (0.005)	2.511 (0.002)	2.526 (0.008)
	(0.70,0.70)	1.500	1.502 (0.001)	1.497 (-0.002)	1.944	1.935 (-0.005)	1.946 (0.001)	1.957 (0.006)
	(0.20,0.20)	1.500	1.499 (0.000)	1.507 (0.005)	1.643	1.68 (0.022)	1.656 (0.008)	1.656 (0.008)
Gumbel(1.125)	(0.40,0.60)	1.500	1.496 (-0.003)	1.508 (0.005)	1.199	1.255 (0.047)	1.208 (0.008)	1.208 (0.008)
R = 0.5	(0.60,0.40)	1.500	1.496 (-0.003)	1.508 (0.005)	1.199	1.256 (0.048)	1.208 (0.008)	1.208 (0.008)
	(0.70,0.70)	1.500	1.503 (0.002)	1.506 (0.004)	1.119	1.181 (0.055)	1.122 (0.003)	1.121 (0.002)
	(0.20,0.20)	2.068	1.802 (-0.128)	1.725 (-0.166)	1.643	1.68 (0.022)	1.656 (0.008)	1.656 (0.008)
Gumbel(1.125)	(0.40,0.60)	1.500	1.592 (0.061)	1.405 (-0.063)	1.199	1.255 (0.047)	1.208 (0.008)	1.208 (0.008)
R = 0.3 ↑ 0.7	(0.60,0.40)	1.500	1.592 (0.061)	1.399 (-0.067)	1.199	1.256 (0.048)	1.208 (0.008)	1.208 (0.008)
	(0.70,0.70)	1.266	1.362 (0.076)	1.259 (-0.006)	1.119	1.181 (0.055)	1.122 (0.003)	1.121 (0.002)
	(0.20,0.20)	2.068	1.86 (-0.101)	1.811 (-0.124)	2.000	2.004 (0.002)	2.004 (0.002)	2.008 (0.004)
Clayton(1)	(0.40,0.60)	1.500	1.653 (0.102)	1.422 (-0.052)	2.000	2.008 (0.004)	2.004 (0.002)	2.008 (0.004)
R = 0.3 ↑ 0.7	(0.60,0.40)	1.500	1.653 (0.102)	1.42 (-0.053)	2.000	2.014 (0.007)	2.004 (0.002)	2.008 (0.004)
	(0.70,0.70)	1.266	1.358 (0.073)	1.264 (-0.002)	2.000	2.03 (0.015)	2.004 (0.002)	2.008 (0.004)
	(0.20,0.20)	1.266	1.332 (0.052)	1.337 (0.056)	1.643	1.675 (0.019)	1.654 (0.007)	1.657 (0.009)
Gumbel(1.125)	(0.40,0.60)	1.500	1.453 (-0.031)	1.677 (0.118)	1.199	1.256 (0.048)	1.208 (0.008)	1.208 (0.008)
R = 0.7 ↓ 0.3	(0.60,0.40)	1.500	1.453 (-0.031)	1.676 (0.117)	1.199	1.257 (0.048)	1.208 (0.008)	1.208 (0.008)
	(0.70,0.70)	2.068	1.742 (-0.158)	1.961 (-0.052)	1.119	1.182 (0.056)	1.122 (0.003)	1.121 (0.002)

scenario, both shape component estimators exhibited biases of at most 1.3%, across all the evaluation points. In scenarios $2 \sim 4$, the biases did not exceed 0.5%. Bias differences across quadrants seemed to reflect random variation. In scenarios 5, 6, and 7, where $R(t)$ was time-varying, the shape estimators were biased up to 16.6%. In scenarios 5 and 6, the Shape2 method consistently underestimated the strength of association, whereas the Shape1 estimator underestimated in the 1st quadrant and overestimated in the other quadrants. For both estimators, the magnitude of the bias was largest in the 1st quadrant, where the $R(t)$ was smallest, and smallest in the 4th quadrant. In scenario 7, the biases were smaller than those for scenarios 5 and 6, but the direction of the biases were opposite. Evaluation over the entire time domain (by contour plots; not shown here) showed that Shape1 estimator biases inflated greatly as the difference between T_1 and T_2 grew, likely due to the Shape1 estimator dependence on $\min(T_1, T_2)$, while Shape2 estimator biases did not depend substantially on time.

The biases of the size component for Clayton failure times (scenarios 1, 2, and 6) were at most 1.5%, 0.2%, and 0.4% in the Size1, Size2, and Size3 estimator, respectively. The biases for the Size1 method were time-varying, while it was constant in the Size2 and Size3 method. This is obvious because the algorithms in the Size2 and Size3 methods must have chosen Clayton copula as the simulated datasets failure time association structure which gives constant size component estimates. In Gumbel copula scenarios (scenarios 3, 4, 5, and 7), the Size1 estimator showed biases of at most 5.6% while those for the Size2 and Size3 estimators were less than 1.1%.

Table 4.3: CCSHR estimates, their biases and coverage probabilities from simulation studies

Copula type			Method 1				Method 2				Method 3				Method 4			
Cause allocation	Location	True	Mean	Bias	B.SD	Cov	Mean	Bias	B.SD	Cov	Mean	Bias	B.SD	Cov	Mean	Bias	B.SD	Cov
Clayton(1) R = 0.2	(0.20,0.20)	6.000	6.041	0.007	0.720	0.963	6.044	0.007	0.959	0.977	6.059	0.010	0.825	0.973	6.064	0.011	0.887	0.973
	(0.40,0.60)	6.000	6.051	0.009	0.738	0.963	6.089	0.015	1.034	0.997	6.085	0.014	0.900	0.983	6.092	0.015	0.938	0.987
	(0.60,0.40)	6.000	6.080	0.013	0.740	0.960	6.115	0.019	1.040	0.990	6.080	0.013	0.903	0.987	6.085	0.014	0.940	0.980
	(0.70,0.70)	6.000	6.145	0.024	0.889	0.967	6.123	0.021	1.114	0.987	6.069	0.012	0.946	0.977	6.074	0.012	0.980	0.983
Clayton(1) R = 0.8	(0.20,0.20)	2.250	2.243	-0.003	0.198	0.930	2.247	-0.001	0.203	0.930	2.251	0.000	0.095	0.957	2.254	0.002	0.145	0.957
	(0.40,0.60)	2.250	2.252	0.001	0.195	0.943	2.261	0.005	0.204	0.940	2.258	0.004	0.097	0.950	2.260	0.004	0.133	0.953
	(0.60,0.40)	2.250	2.262	0.005	0.195	0.963	2.273	0.010	0.204	0.970	2.260	0.004	0.097	0.947	2.263	0.006	0.132	0.957
	(0.70,0.70)	2.250	2.274	0.011	0.220	0.960	2.281	0.014	0.228	0.957	2.258	0.004	0.101	0.960	2.261	0.005	0.136	0.950
Gumbel(2.5) R = 0.5	(0.20,0.20)	9.142	8.909	-0.025	0.961	0.927	8.931	-0.023	1.069	0.940	9.175	0.004	0.735	0.970	9.253	0.012	1.011	0.977
	(0.40,0.60)	3.759	3.797	0.010	0.350	0.927	3.781	0.006	0.403	0.937	3.750	-0.002	0.300	0.973	3.772	0.003	0.333	0.967
	(0.60,0.40)	3.759	3.779	0.005	0.348	0.943	3.779	0.005	0.402	0.957	3.767	0.002	0.300	0.980	3.788	0.008	0.334	0.960
	(0.70,0.70)	2.916	2.908	-0.003	0.301	0.970	2.900	-0.005	0.334	0.960	2.913	-0.001	0.214	0.980	2.927	0.004	0.240	0.973
Gumbel(1.125) R = 0.5	(0.20,0.20)	2.465	2.513	0.019	0.245	0.950	2.525	0.024	0.270	0.953	2.498	0.013	0.210	0.947	2.496	0.013	0.231	0.943
	(0.40,0.60)	1.799	1.877	0.043	0.178	0.950	1.891	0.051	0.208	0.963	1.822	0.013	0.139	0.993	1.822	0.013	0.138	0.997
	(0.60,0.40)	1.799	1.876	0.043	0.179	0.933	1.891	0.051	0.207	0.953	1.822	0.013	0.139	0.990	1.822	0.013	0.138	0.973
	(0.70,0.70)	1.679	1.778	0.059	0.164	0.913	1.781	0.061	0.181	0.943	1.690	0.007	0.110	0.980	1.688	0.005	0.110	0.983
Gumbel(1.125) R = 0.3 ↑ 0.7	(0.20,0.20)	3.398	3.028	-0.109	0.325	0.790	2.890	-0.149	0.400	0.863	2.859	-0.159	0.343	0.853	2.856	-0.160	0.363	0.870
	(0.40,0.60)	1.799	1.998	0.111	0.201	0.840	1.763	-0.020	0.196	0.977	1.698	-0.056	0.127	0.977	1.698	-0.056	0.126	0.970
	(0.60,0.40)	1.799	2.000	0.112	0.201	0.823	1.755	-0.024	0.196	0.957	1.691	-0.060	0.127	0.973	1.692	-0.059	0.126	0.963
	(0.70,0.70)	1.418	1.609	0.135	0.146	0.717	1.489	0.050	0.141	0.950	1.411	-0.005	0.073	0.973	1.411	-0.005	0.074	0.967
Clayton(1) R = 0.3 ↑ 0.7	(0.20,0.20)	4.135	3.727	-0.099	0.392	0.810	3.630	-0.122	0.487	0.898	3.629	-0.122	0.373	0.902	3.637	-0.120	0.426	0.915
	(0.40,0.60)	3.000	3.316	0.105	0.331	0.843	2.857	-0.048	0.310	0.922	2.850	-0.050	0.198	0.919	2.856	-0.048	0.232	0.942
	(0.60,0.40)	3.000	3.327	0.109	0.332	0.843	2.861	-0.046	0.310	0.956	2.846	-0.051	0.197	0.956	2.853	-0.049	0.230	0.942
	(0.70,0.70)	2.533	2.753	0.087	0.289	0.907	2.567	0.013	0.282	0.956	2.533	0.000	0.149	0.956	2.538	0.002	0.181	0.959
Gumbel(1.125) R = 0.7 ↓ 0.3	(0.20,0.20)	2.028	2.238	0.104	0.207	0.810	2.241	0.105	0.213	0.924	2.213	0.091	0.155	0.894	2.217	0.093	0.176	0.928
	(0.40,0.60)	1.799	1.823	0.013	0.167	0.950	2.106	0.171	0.223	0.826	2.026	0.126	0.151	0.795	2.026	0.126	0.149	0.788
	(0.60,0.40)	1.799	1.824	0.014	0.167	0.950	2.107	0.171	0.224	0.822	2.025	0.126	0.152	0.799	2.025	0.126	0.150	0.807
	(0.70,0.70)	2.169	2.057	-0.052	0.203	0.930	2.318	0.069	0.293	0.951	2.199	0.014	0.216	0.970	2.198	0.013	0.216	0.970

* B.SD: Bootstrap Standard Deviation, Cov: Probability of 95% CI including the true value

We proceeded to assess the performances of the CCSHR estimators (Table 4.3). In scenarios $1 \sim 4$, the biases were no greater than 6.1% for Methods 1 and 2, and no greater than 1.5% for Methods 3 and 4. The difference between Methods 3 and 4 was negligible, and likewise for Methods 1 and 2. In scenarios 5 and 6, the bias was no greater than 12%, in all quadrants but the first, for Method 1, no greater than 5% for Method 2, and no greater than 6% for Methods 3 and 4. For these methods, the bias in the first quadrant, (0.20,0.20), was larger than those in the other quadrants. Method 1 exhibited considerably higher bias than the other methods for scenarios 5 and 6, except in the first quadrant, but was the most accurate estimator in scenario 7.

Estimator variability, as indicated by the bootstrap SD, for the most part tracked the true value of the estimand. In most scenarios, Method 3 exhibited a substantially lower bootstrap SD than the other three methods: Method 1 was the best performer in scenario 1, and in scenario 5, none of them dominated the others. In scenarios $1 \sim 4$, the probabilities of 95% confidence interval coverage were greater than 0.91 for Methods 1 and 2 and at least 0.94 for Methods 3 and 4. In scenarios $5 \sim 7$, however, the coverage probability was as low as 0.72 for Method 1 and 0.80 for the other methods (in scenario 7). In the scenarios with n=500, the bootstrap SD was about 50% higher than $n = 1,000$ scenarios.

The directions of these biases were mainly decided by the size component in scenarios 3 and 4 and by the shape component in scenarios 5, 6, and 7. Scenario 7 seemed to present a particular challenge for estimation, perhaps due to the combination of increasing shape association together with decreasing size association over time.

Based on the results of the simulation studies, we recommend Method 3. It provides the most precise and accurate size component and CCSHR estimates in various scenarios while allowing flexibility in the shape component estimation. All of the methods, however, proved capable of tracking the general shape of the CCSHR association over time.

4.4 Data analysis

The estimators studied in the previous sections were applied to data from the Cache County Study on Memory Health and Aging (Breitner et al., 1999). This study was conducted to investigate the prevalence of dementia. It is known that the onset of dementia aggregates in families (Hendrie, 1998), and the heritability is higher for early-age onset than late-age (Silverman, 2005). The study collected information on dementia onset from the permanent residents of Cache County, Utah, U.S.A., aged 65 and over (the 'proband') on themselves and all their family members. Thus, the Cache County dataset is appropriate for assessing whether our estimator adequately expresses time-dependent association between failure causes and failure times within a pair.

To simplify the analysis, we included only the participant's mother and the oldest sibling inclusive of self. Pairs without information available for both members were excluded; pairs for which either member died or became demented before age 55 also were excluded. The resulting dataset consisted of 3,635 pairs' times of event occurrence and event indicators: 0 for censoring or living without dementia at the end of the study, 1 for dementia, and 2 for death without dementia. The proportion of data censored was 60.4$\%$ among the eldest siblings and 4.0$\%$ among mothers. Among those experiencing events, $13 \sim 14\%$ of participants experienced dementia before death. Both members of a pair became demented in 40 pairs, both members died without dementia in 1,132 pairs, and the members failed of different causes in 259 pairs. The primary purpose of our analysis was to assess association between dementia onset in families. Because only a small proportion of failures were due to dementia, we also conducted an analysis considering death as the failure cause of interest and dementia as the secondary cause.

To assess the variability of estimators in our analysis, we used bootstrapping. Three hundred bootstrap samples with the same size as the original dataset (3,635 pairs) were generated. Each bootstrap sample was created by random selection of pairs with replacement. We obtained bootstrap standard errors and 95$\%$ confidence intervals from the bootstrapped samples using the percentile method described in the previous section.

Times of event occurrence ranged from 55 to 104. These were transformed to lie between 0 and 1 by computing their Kaplan-Meier functions, separately for mothers and children, and transforming the times to cumulative incidence probabilities. Size and shape estimators studied in the previous section were applied to these transformed times. Below we display values for each at $(T_1', T_2') = (0.25, 0.25), (0.25, 0.75), (0.75, 0.25)$, and $(0.75, 0.75)$, which correspond to children's ages 77 and 91 and mothers' ages 73 and 88, as well as a contour plot of the estimated CCSHR function.

The results considering dementia as the primary cause of interest are summarized in

Tables 4.4 (size and shape components) and 4.5 (CCSHR). The Shape2 estimator showed associations that varied by region but were strong in each. On the other hand, the Shape1 estimator showed modest association regardless of region. The Size1 and Size3 estimators produced similar results of weakly positive association in (early,early) failure times and virtually no association otherwise. The Size2 estimates suggested near-independence of first-failure times. Bootstrap standard errors were considerably smaller for Method 1 than the other methods even after considering the magnitude of the estimates.

Relatively high bootstrap standard errors of the Shape2 estimator implies that this estimator may not be stable if the proportion of the failure cause of primary interest is too low or if there is heavy censoring. The magnitudes of the size component estimators indicate there is modest or little association between failure times, and the association between onset times of dementia mainly comes from the association between failure causes rather than the association between first-failure times. The magnitudes of associations for Methods $2 \sim 4$ are more consistent with prior estimates (Bandeen-Roche & Liang, 2002; Bandeen-Roche & Ning, 2008) than those for Method 1.

Table 4.4: Shape and size components estimates from Cache County data with dementia as cause 1

	Shape1	Shape2	Size1	Size2	Size3
(0.25,0.25)	1.76 (0.22)	6.29 (1.49)	1.19 (0.07)	1.04 (0.02)	1.19 (0.05)
(0.25,0.75)	1.76 (0.22)	5.03 (1.58)	1.24 (0.08)	1.02 (0.02)	1.07 (0.02)
(0.75,0.25)	1.76 (0.22)	3.67 (1.23)	0.85 (0.08)	1.02 (0.02)	1.07 (0.02)
(0.75,0.75)	1.67 (0.20)	5.87 (2.84)	0.95 (0.10)	1.02 (0.02)	1.03 (0.02)

If we consider death as the main failure cause, the association was generally weaker

Table 4.5: CCSHR estimates and their bootstrap standard errors and 95% confidence intervals from Cache County data with dementia as cause 1

		Method 1	Method 2	Method 3	Method 4
(0.25,0.25)	Mean (SD)	2.08 (0.28)	7.48 (1.86)	6.52 (1.54)	7.52 (1.82)
	95% CI	(1.57,2.61)	(4.44,11.93)	(3.98,10.22)	(4.49,11.83)
(0.25,0.75)	Mean (SD)	2.18 (0.32)	6.24 (2.00)	5.13 (1.62)	5.37 (1.69)
	95% CI	(1.57,2.61)	(3.41,11.41)	(2.97,9.10)	(3.10,9.44)
(0.75,0.25)	Mean (SD)	1.50 (0.23)	3.14 (1.12)	3.74 (1.26)	3.91 (1.31)
	95% CI	(1.05,1.97)	(1.58,5.99)	(1.86,7.13)	(1.95,7.32)
(0.75,0.75)	Mean (SD)	1.59 (0.26)	5.60 (2.75)	5.97 (2.89)	6.04 (2.92)
	95% CI	(1.16,2.12)	(2.20,13.00)	(2.36,13.29)	(2.41,13.44)

Table 4.6: Shape and size components estimates from Cache County data with death as cause 1

	Shape1	Shape2	Size1	Size2	Size3
(0.25,0.25)	1.02 (0.01)	3.17 (0.18)	1.19 (0.07)	1.04 (0.02)	1.19 (0.05)
(0.25,0.75)	1.02 (0.01)	3.38 (0.28)	1.24 (0.08)	1.02 (0.02)	1.07 (0.02)
(0.75,0.25)	1.02 (0.01)	1.51 (0.09)	0.85 (0.08)	1.02 (0.02)	1.07 (0.02)
(0.75,0.75)	1.02 (0.01)	1.24 (0.06)	0.95 (0.10)	1.02 (0.02)	1.03 (0.02)

than when dementia was the main failure cause. See Tables 4.6 and 4.7. The Shape2 estimates indicated association that was modest to strong for early failure times, but became weaker as time increased; while the Shape1 estimates indicated virtually zero association. The size estimates were, of course, identical as for the dementia-based analysis. The bootstrap standard errors for the shape estimators were substantially smaller than for dementia outcomes mainly due to the high proportion of death compared to dementia. Here, the standard errors were smallest for Method 1 except for (late,late) failiure times.

We provide the contour plot of the estimated CCSHR from the Method 3 in the dementia-based analysis (Figure 4.1). Consistent with existing knowledge, it indicates cause-specific

Table 4.7: CCSHR estimates and their bootstrap standard errors and 95% confidence i ntervals from Cache County data with death as cause 1

		Method 1	Method 2	Method 3	Method 4
(0.25,0.25)	Mean (SD)	1.21 (0.07)	3.77 (0.33)	3.29 (0.20)	3.79 (0.30)
	95% CI	(1.08,1.34)	(3.11,4.41)	(2.91,3.69)	(3.23,4.42)
(0.25,0.75)	Mean (SD)	1.26 (0.09)	4.19 (0.51)	3.44 (0.29)	3.61 (0.31)
	95% CI	(1.11,1.44)	(3.21,5.16)	(2.94,4.05)	(3.04,4.31)
(0.75,0.25)	Mean (SD)	0.87 (0.09)	1.29 (0.16)	1.54 (0.09)	1.61 (0.10)
	95% CI	(0.71,1.06)	(1.04,1.60)	(1.39,1.75)	(1.45,1.83)
(0.75,0.75)	Mean (SD)	0.98 (0.10)	1.19 (0.14)	1.26 (0.07)	1.28 (0.07)
	95% CI	(0.79,1.21)	(0.95,1.48)	(1.14,1.42)	(1.16,1.46)

association that is very strong in (early,early) failure time region and modestly strong in the other regions. We believe that the very high values in the upper part of the plot reflects instability due to the sparsity of dementia onset. If we use the method of deciding a boundary where the estimates are considered reliable (reported in Chapter 2 of this thesis), the estimates in the area of (early,early) failures are trustworthy.

4.5 Conclusion

This paper addressed the estimation of two multiplicative components of the CCSHR. The shape component, representing the association between failure causes, has been estimated parametrically by assuming a Dirichlet (or beta) distribution of failure cause allocation (Bandeen-Roche & Liang, 2002), and by a composite likelihood method proposed by Shih and Albert (2010). We suggested a nonparametric method which estimates Shih and Albert's alternative representation of the shape component by smoothing. The size com-

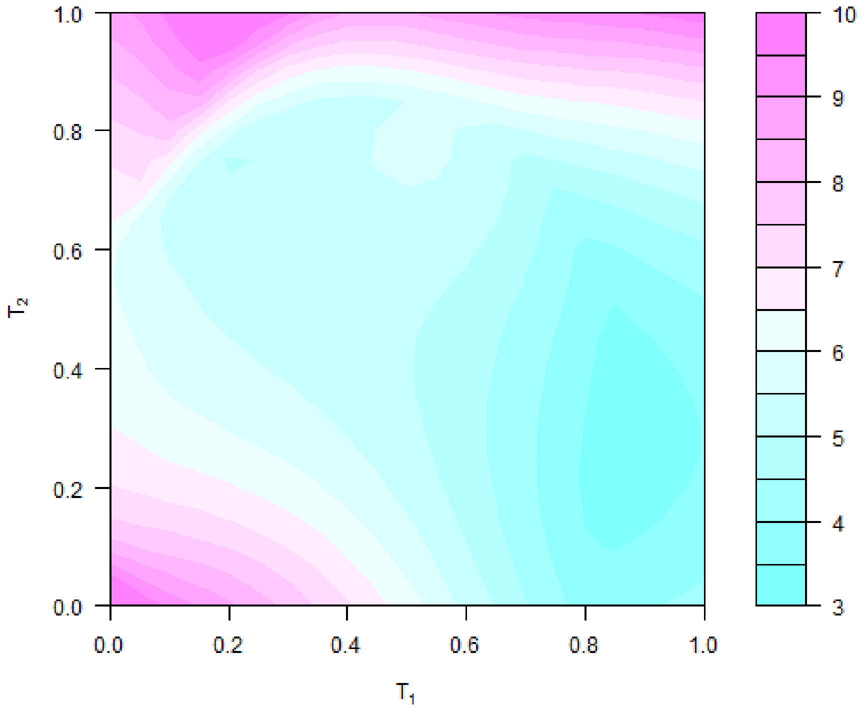


Figure 4.1: Contour plot of CCSHR estimates from Cache County Study data analysis

ponent, representing the association between failure times, has been estimated by various methods including one proposed by Shih and Louis (1995). We studied both parametric and nonparametric alternatives of this method. We conducted a set of simulation studies of these methods, where we varied copula types and strength of association, beta distribution parameters Δ, which controls the aggregation of failure causes within a pair, and $R(t)$, the constant or time-varying allocation of failure causes. The estimators we proposed generally performed well, but the shape component estimators tended to too shallowly represent time variation in $R(t)$. We also conducted an analysis of association in times to dementia onset and in times to death, using data from the Cache County Study on Memory Health and Aging. Our new shape association estimator (Shape2) suggested occurrence of dementia among family members to be more strongly associated than death without dementia, and the size component estimators suggested that the association between event occurrence times is slightly stronger in early ages than in late ages. The different methods of shape component estimation varied substantially in their assessments of the strengths of association.

In the simulation studies, it was observed that the shape component estimator was more biased when $R(t)$ was small. This may reflect the small size of the data subset engaged in the estimation in these cases and the greater value of the estimand which can be seen from Equation (4.12). We also found the shape component estimators to be more biased than the size component estimators, especially when $R(t)$ was time-varying, even after considering the magnitude of $R(t)$. The reason of this bias is unclear. The Shape1 estimator was severely biased for large differences between T_1 and T_2 in the increasing $R(t)$ scenario. The Shape1 estimator assumes that the beta-distributed stochastic process of failure cause allocation divided by its mean function is a martingale. This assumption reduced the temporal dependence of the estimator on T_1 and T_2 into a dependence upon the minimum of T_1 and T_2, but it may not appropriately capture association in many cases in which the true association depends on T_1 and T_2 separately. Consequently, further development of parametric estimation of the shape component in cases of time-dependent $R(t)$ is needed. Moreover, our methods remain subject to the assumption of independence between shape and size frailty which was imposed when the CCSHR was initially proposed. How to accommodate covariation between these quantities is an open question. Considering accuracy and precision, we recommend the combination of Shape2 and Size2 estimators, Method 3.

In analyses restricted to a few time points in quadrants, we failed to find convincingly stronger familial aggregation of early- than late-onset dementia as Silverman (2005) did. However, a contour plot displaying the association fully as a function of time provided indication of this. The stronger association between cause-specific failure times in early ages was primarily tied to the association between failure causes rather than the association between failure times, thus appeared specific to the dementia disease rather than a more general propensity to become sick or die. This finding emphasizes how the separation of the shape and size components of the CCSHR may help interpret the source of association. However, this result should be interpreted with caution because of the small proportion of mother-child pairs in which there was shared dementia onset.

Overall, the methods we introduced in this paper demonstrated usefulness in detecting and describing the strength of association between failure causes and failure times. There are some aspects, however, requiring further work to make these methods more useful. A measure of estimator variability over the whole time domain should be developed. Further investigation and development are needed regarding treatment of censoring and the estimator biases that were observed in regions of data sparsity.

Chapter 5

Conclusion

This thesis studied and developed statistical methodologies to measure the strength of the association among clustered failure times, with an emphasis on the competing risks setting. In Chapter 2, I developed a nonparametric estimator of the local version of Kendall's τ. Based on the simple idea of smoothing the concordance-discordance indicator as a function of bivariate failure times, this method enabled easy visualization and interpretation of the association. Logistic regression and smoothing methods such as Loess, generalized additive models (GAM), and multivariate adaptive regression splines (MARS) were assessed, among which GAM was considered to perform best in terms of RMSDs. We also compared approaches for dealing with censored data: we adapted existing methods to estimate global Kendall's τ to be suited for the localized version of Kendall's τ and also suggested a novel multiple imputation method based on Dabrowska's bivariate density estimator.

Chapter 3 investigated the sensitivity of the estimation of the CCSHR, by the method proposed by Bandeen-Roche and Liang (2002), to that method's underlying statistical assumptions. To assess the assumption of Dirichlet distribution governing association between causes of failure (via a 'shape' frailty), we developed a new estimator based on the logit-normal distribution assumption and compared it with an existing one based on the Dirichlet distribution. There was very little difference in performance between these estimators, even when one was applied to data generated from the other model. Rather, we discovered that misspecification of failure-time association ('size') structure, rather than the Dirichlet shape mechanism, was the major source of poor performance reported in the original Bandeen-Roche and Liang paper. Such mis-specification is easily addressed by more flexible modeling of the CHR. To assess the independence assumption between frailties governing association in times-to-failure ('size' frailty) and the shape frailty, we generated data from dependent size and shape frailty variables and applied the Bandeen-Roche and Liang estimator to these data. Violation of independence assumption, which crucially underlay the development of simplified CCSHR estimator, was shown to have a huge impact on the estimators.

Chapter 4 aimed to develop a completely nonparametric estimator of the CCSHR based on separate estimators of shape and size components of the CCSHR, then multiplied. To estimate the size component, I suggested using a nonparametric estimator of the CHR developed in Chapter 2, and I also proposed a modified parametric estimator. The shape component estimator was motivated from Shih and Albert's alternative representation of the CCSHR, and could be estimated by applying nonparametric regression methods to each of its multiplicands. Various combinations of these methods were assessed. Multiplication of our nonparametric shape component estimator and the Shih and Louis (1995) likelihood-based CHR estimator was identified as the best strategy.

A number of nonparametric approaches to estimating association between failure times have been previously proposed. Prentice and Cai (1992), Hsu and Prentice (1996), Fan, Hsu, and Prentice (2000), Fan, Prentice, and Hsu (2000), Sankaran, Abraham, and Antony (2006), and Nair and Sankaran (2010) all proposed such nonparametric estimators. The Kendall's τ-based method suggested in Chapter 2 adds an alternative to these existing methods which has a clear advantage of easy visualization and intuitive interpretation. That Kendall's τ ranges from -1 to +1 may assist in distinguishing temporal differences in strength of association across applications, as compared to measures which may diverge to infinity. In addition, it is interpreted as a difference of probabilities of concordance and discordance, thus, absolute values of the estimator are easily interpretable, whereas many of the other estimators only allow ready relative comparison between different time points or data sets.

The main appeal of the nonparametric estimators of the shape component proposed in Chapter 4 is smooth description over the entire time domain while maintaining the advantage of the shape and size component decomposition. The measures proposed by Cheng and colleagues (2007, 2009) or Scheike and colleagues (2010, 2012) do not provide this. Shih and Albert (2010) adopted the same decomposition framework as in my work, but their estimate is piecewise constant on a 'binned' time domain.

A common limitation of our approaches in Chapters 2 and 4 is instability of estimates in regions where the data are sparse. We suggested a method to define a boundary of reliable estimation based on the data density, so the analyst can limit estimation to a region in which the estimator is most trustworthy. Our estimators are not yet equipped with convenient inferential procedures by which to judge uncertainty; rather, we relied on bootstrapping. We observed moderate biases even for our preferred estimators in a number of the most challenging scenarios. None of the methods of dealing with censoring in Chapter 2 clearly outperformed the others, and the performance was not satisfactory under heavy censoring. Still, the tools I have developed proved capable of capturing overall shapes of relationships, and should provide useful new tools for visualizing failure time associations.

There are several future directions for this work. Development of readier inferential procedures, including simultaneous confidence bands for the estimators, is a priority. The assumption of independence between size and shape frailty in the estimation of the CC-SHR enables simplified estimation, but may be easily violated in real data. Thus, further study of this phenomenon and work to develop a new estimator which can accommodate potential covariation would be worthwhile. Functions to accomplish estimation are available from the author; however, development of user-friendly software to implement the methods proposed in this thesis, such as an R package, is needed.

A.1 Comparison of smoothing methods for estimating local Kendall's τ

A.1.1 Comparison of true and estimated local Kendall's τ

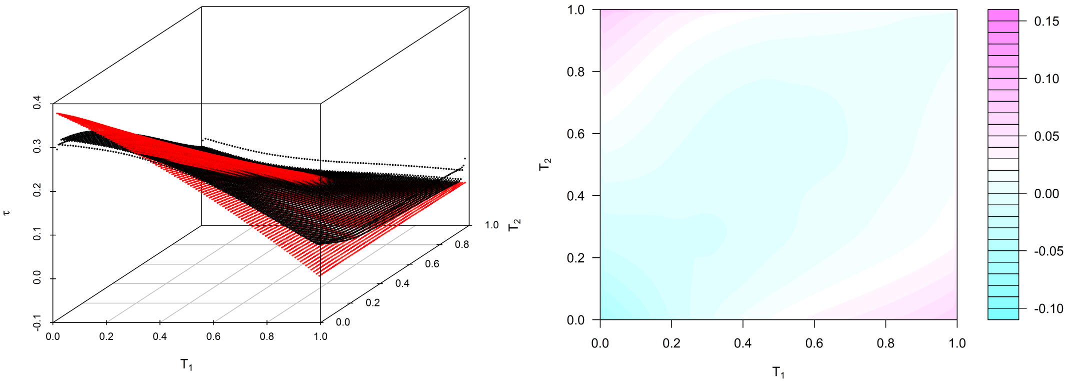
The data were generated from Frank copula with parameter 1.9 or Gumbel copula with parameter 1.26 (correlation 0.3).

On the left panel, the true values are in red and the estimates are in black.

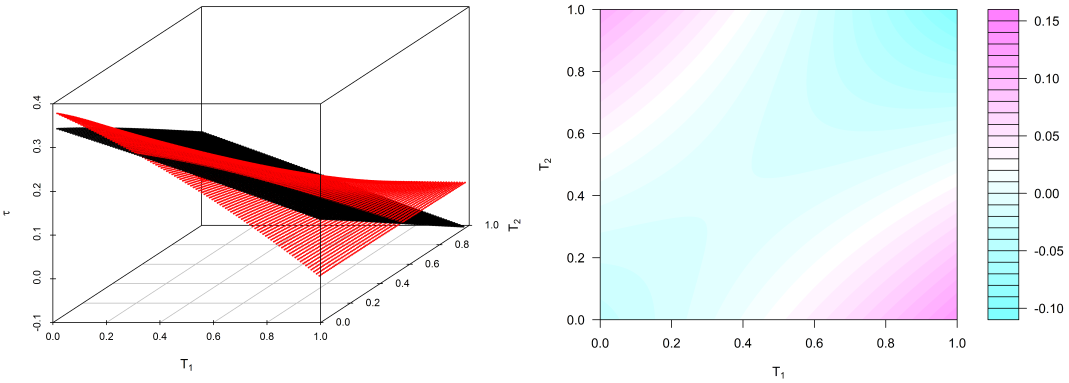
The right panel is the difference between the true values and the estimates.

1. Frank (1.9)

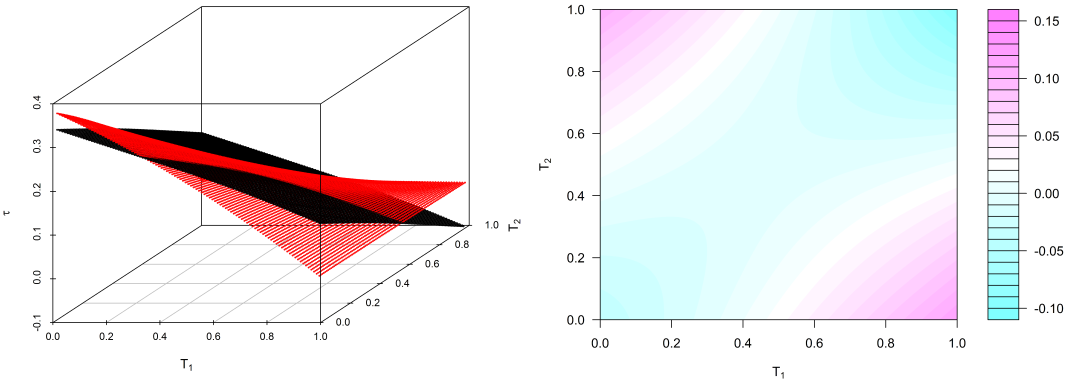
(1) Loess



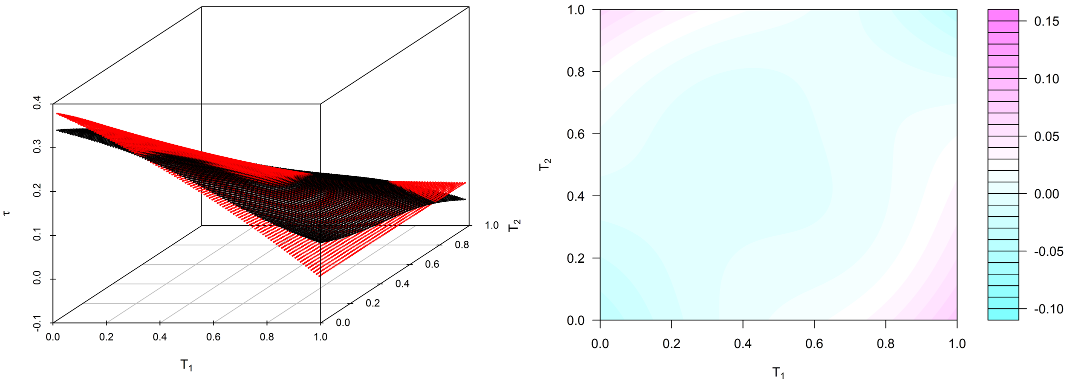
(2) Logistic regression



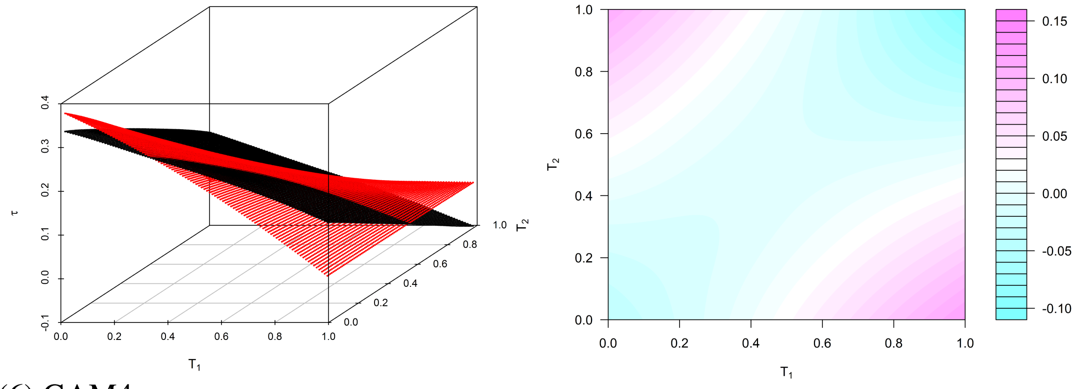
(3) GAM1



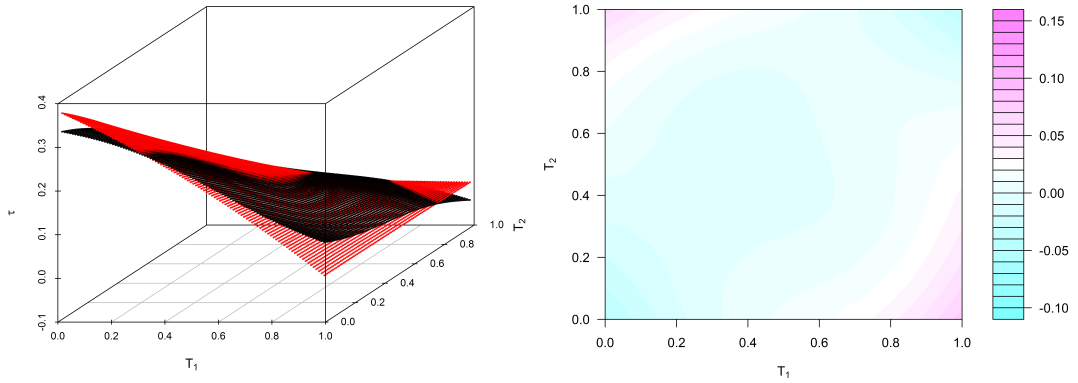
(4) GAM2



(5) GAM3



(6) GAM4



(7) MARS

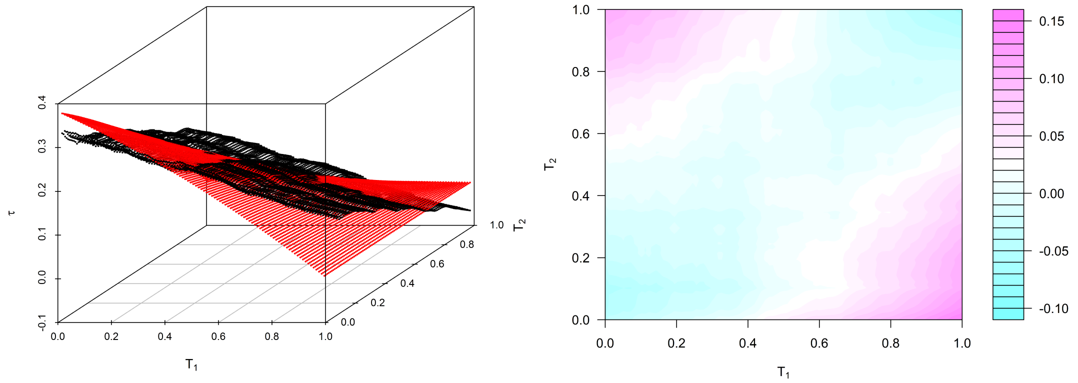
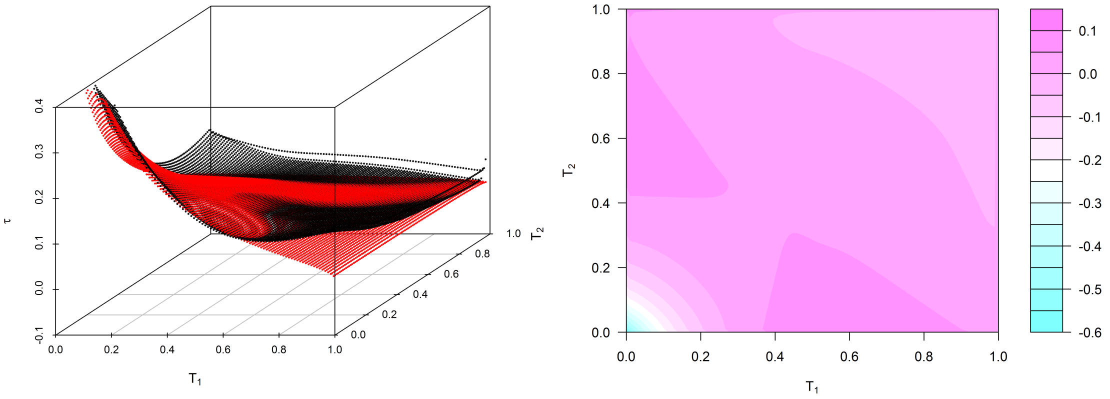


Table A.1: True and estimated values of local Kendall's τ for Frank (1.9) copula

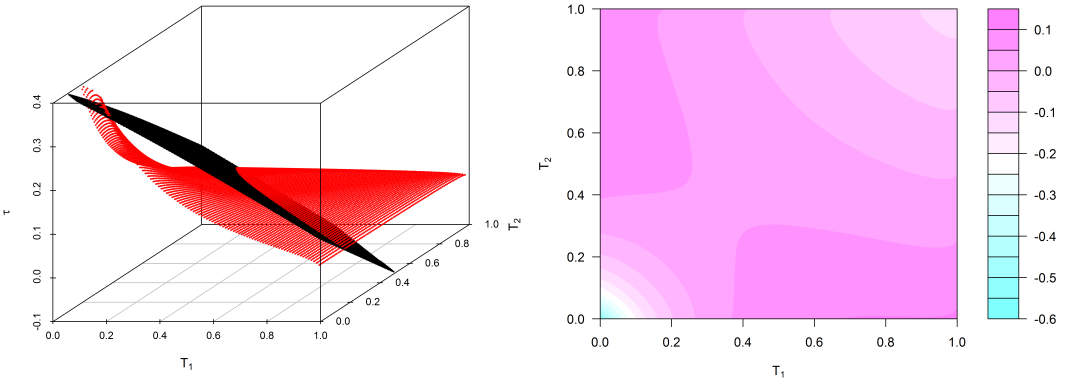
Location	True	Loess	Logistic	GAM1	GAM2	GAM3	GAM4	MARS
(0.2,0.2)	0.276	0.264	0.261	0.260	0.267	0.260	0.268	0.262
(0.2,0.5)	0.191	0.190	0.192	0.193	0.187	0.193	0.187	0.196
(0.2,0.8)	0.083	0.100	0.120	0.121	0.092	0.123	0.092	0.137
(0.5,0.2)	0.191	0.192	0.198	0.199	0.194	0.200	0.194	0.207
(0.5,0.5)	0.138	0.133	0.127	0.132	0.133	0.132	0.133	0.141
(0.5,0.8)	0.063	0.065	0.054	0.061	0.063	0.060	0.064	0.081
(0.8,0.2)	0.083	0.104	0.132	0.132	0.101	0.134	0.101	0.139
(0.8,0.5)	0.063	0.070	0.060	0.065	0.072	0.064	0.072	0.073
(0.8,0.8)	0.031	0.036	-0.013	-0.007	0.034	-0.008	0.033	0.014

2. Gumbel (1.26)

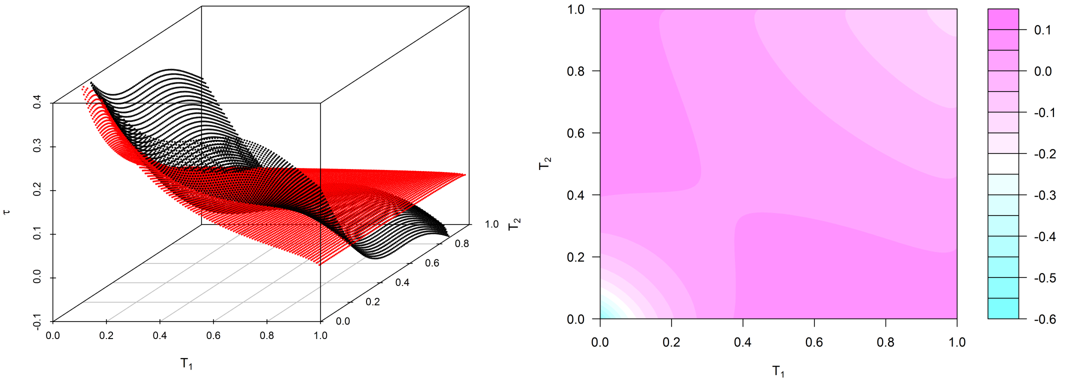
(1) Loess



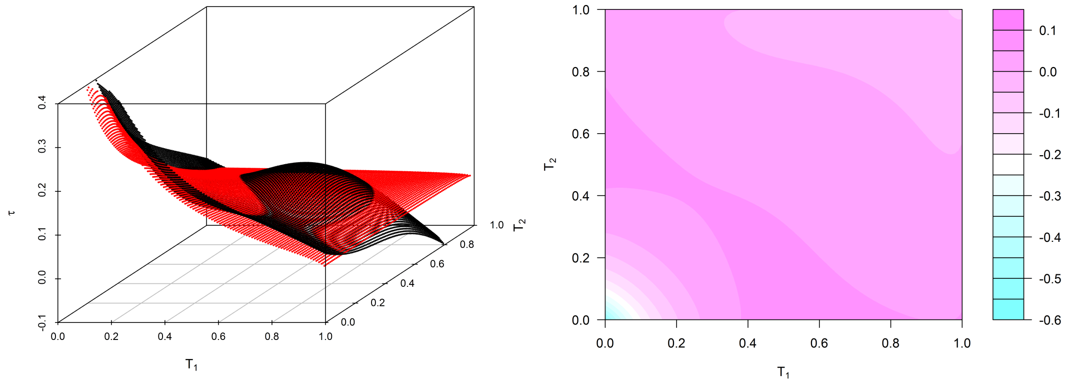
(2) Logistic regression



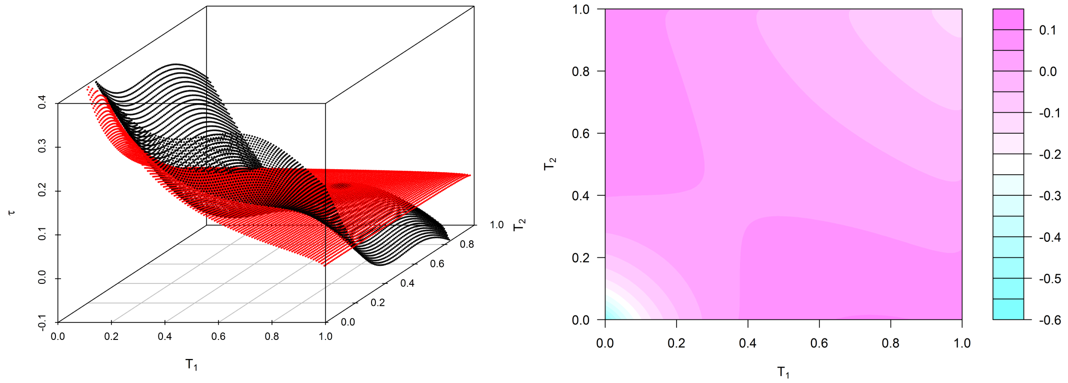
(3) GAM1



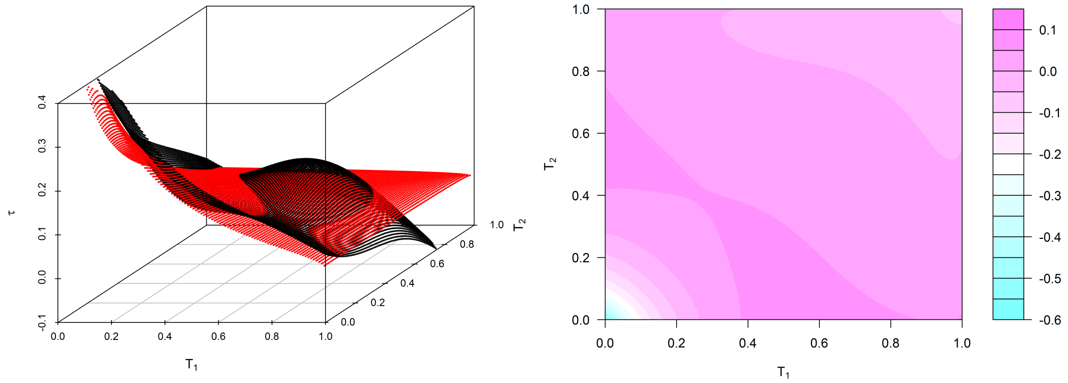
(4) GAM2



(5) GAM3



(6) GAM4



(7) MARS

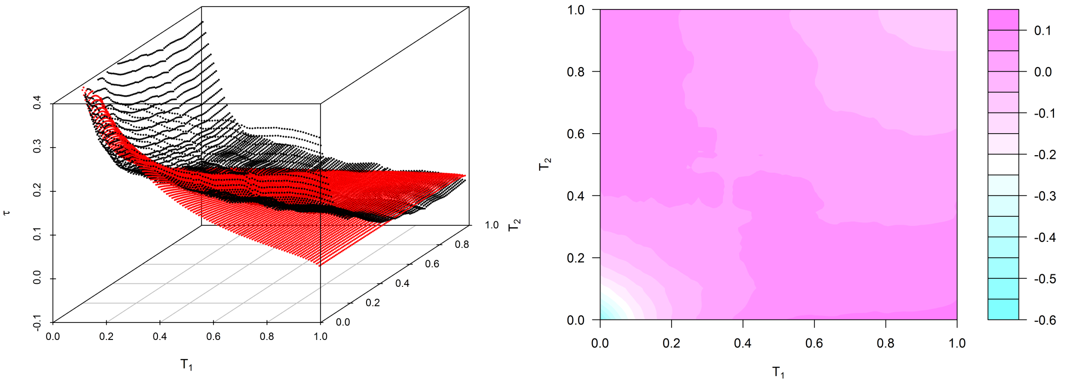


Table A.2: True and estimated values of local Kendall's τ for Gumbel (1.26) copula

Location	True	Loess	Logistic	GAM1	GAM2	GAM3	GAM4	MARS
(0.2,0.2)	0.252	0.237	0.302	0.242	0.248	0.246	0.254	0.195
(0.2,0.5)	0.137	0.121	0.195	0.143	0.112	0.144	0.107	0.130
(0.2,0.8)	0.070	0.077	0.083	0.120	0.062	0.124	0.062	0.106
(0.5,0.2)	0.137	0.128	0.197	0.151	0.113	0.152	0.108	0.141
(0.5,0.5)	0.098	0.095	0.086	0.052	0.120	0.048	0.124	0.076
(0.5,0.8)	0.060	0.062	-0.028	0.029	0.072	0.028	0.074	0.052
(0.8,0.2)	0.070	0.087	0.088	0.132	0.072	0.135	0.071	0.114
(0.8,0.5)	0.060	0.062	-0.025	0.032	0.080	0.030	0.082	0.049
(0.8,0.8)	0.045	0.041	-0.138	0.010	0.002	0.010	-0.008	0.025

A.1.2 Decomposition of RMSD into variance and bias squared

We selected nine points on the bivariate time domain and obtained variances and squared biases from two copula types

1. Frank (1.9)

Table A.3: Variance, Frank (1.9)

Location	Loess	Logistic	GAM1	GAM2	GAM3	GAM4	MARS
(0.2,0.2)	71	26	41	53	40	52	128
(0.2,0.5)	69	23	45	53	45	54	141
(0.2,0.8)	109	60	65	119	65	120	138
(0.5,0.2)	83	25	57	61	56	62	165
(0.5,0.5)	77	17	48	74	50	74	156
(0.5,0.8)	61	49	58	75	59	76	154
(0.8,0.2)	90	60	65	108	66	111	156
(0.8,0.5)	54	47	51	70	53	71	145
(0.8,0.8)	50	74	54	64	55	65	109

Units are 1.0×10^{-3}

Table A.4: Bias^2, Frank (1.9)

Location	Loess	Logistic	GAM1	GAM2	GAM3	GAM4	MARS
(0.2,0.2)	2	2	3	1	2	1	2
(0.2,0.5)	0	0	0	0	0	0	0
(0.2,0.8)	3	14	14	1	16	1	29
(0.5,0.2)	0	1	1	0	1	0	3
(0.5,0.5)	0	1	0	0	0	0	0
(0.5,0.8)	0	1	0	0	0	0	3
(0.8,0.2)	4	24	24	3	26	3	31
(0.8,0.5)	0	0	0	1	0	1	1
(0.8,0.8)	0	20	14	0	15	0	3

Units are 1.0×10^{-3}

2. Gumbel (1.26)

Table A.5: Variance, Gumbel (1.26)

Location	Loess	Logistic	GAM1	GAM2	GAM3	GAM4	MARS
(0.2,0.2)	65	23	39	48	38	48	194
(0.2,0.5)	70	28	54	63	56	67	165
(0.2,0.8)	94	77	64	93	66	95	157
(0.5,0.2)	86	29	63	69	65	73	203
(0.5,0.5)	74	19	42	64	43	63	126
(0.5,0.8)	53	52	45	69	47	71	107
(0.8,0.2)	90	76	65	100	66	102	159
(0.8,0.5)	49	50	48	75	50	78	136
(0.8,0.8)	59	64	55	95	58	100	127

Units are 1.0×10^{-3}

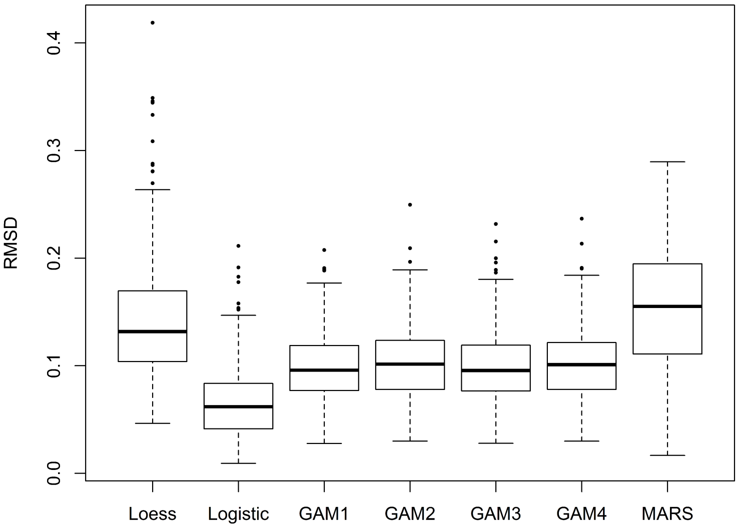
Table A.6: Bias^2, Gumbel (1.26)

Location	Loess	Logistic	GAM1	GAM2	GAM3	GAM4	MARS
(0.2,0.2)	2	25	1	0	0	0	32
(0.2,0.5)	2	34	0	6	1	9	0
(0.2,0.8)	0	2	25	1	28	1	12
(0.5,0.2)	1	37	2	6	3	8	0
(0.5,0.5)	0	1	21	5	25	7	5
(0.5,0.8)	0	77	9	1	11	2	1
(0.8,0.2)	3	3	37	0	42	0	19
(0.8,0.5)	0	73	8	4	9	5	1
(0.8,0.8)	0	332	12	18	12	28	4

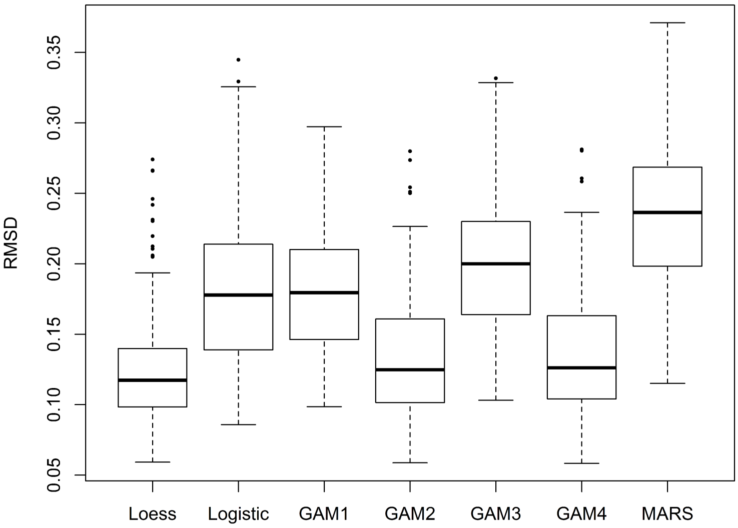
Units are 1.0×10^{-3}

A.1.3 Boxplots of RMSDs for 300 replicates

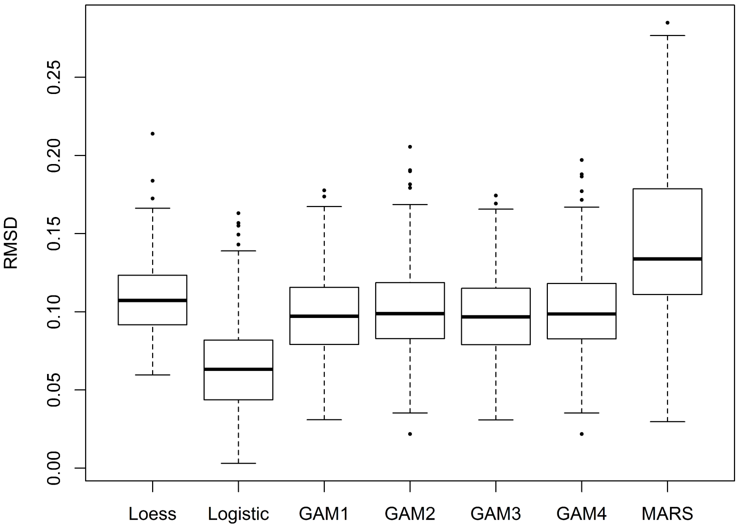
(1) Clayton (-0.53) (Corr = -0.5)



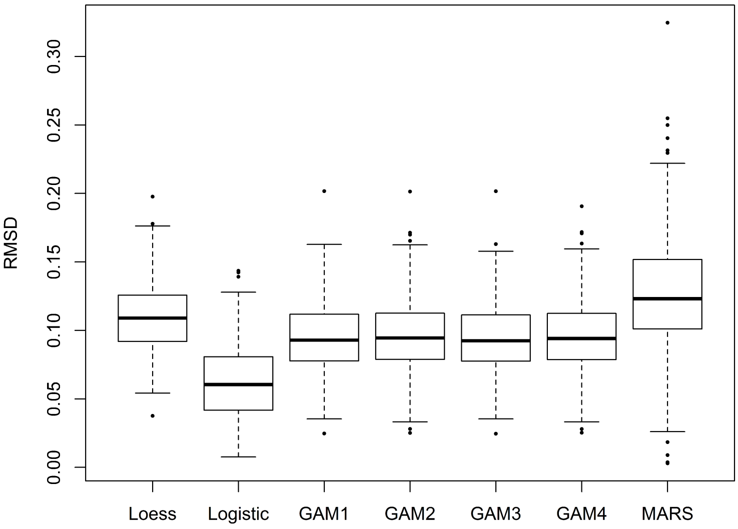
(2) Frank (-3.5) (Corr = -0.5)



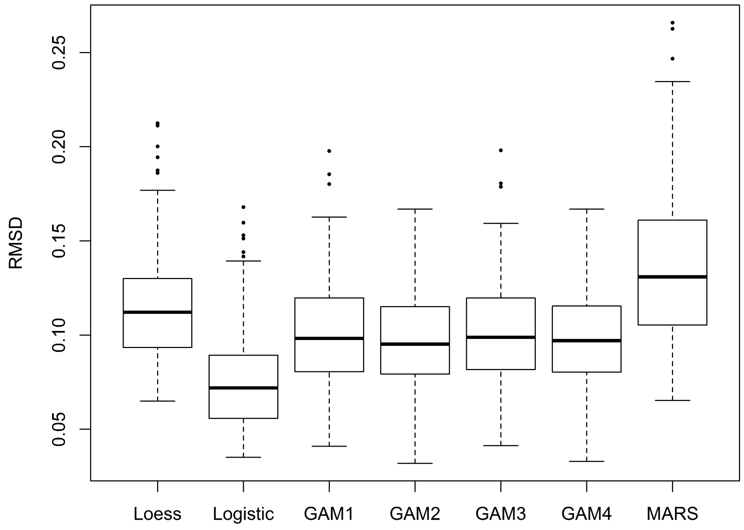
(3) Independent (Corr = 0)



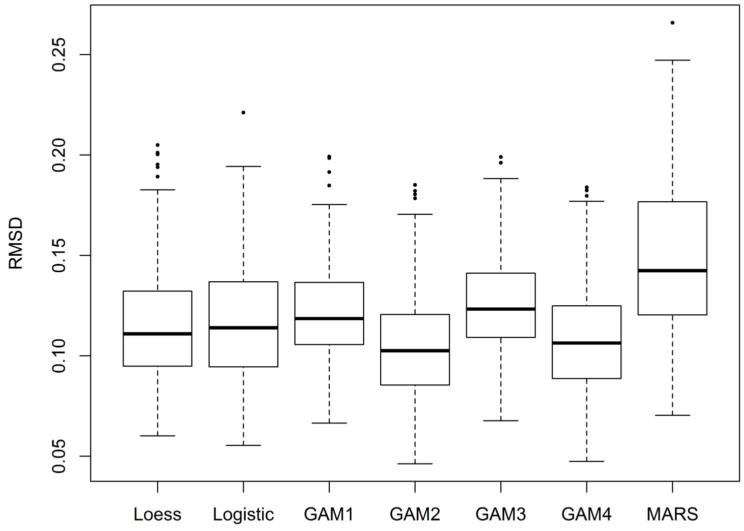
(4) Clayton (0.5) (Corr = 0.3)



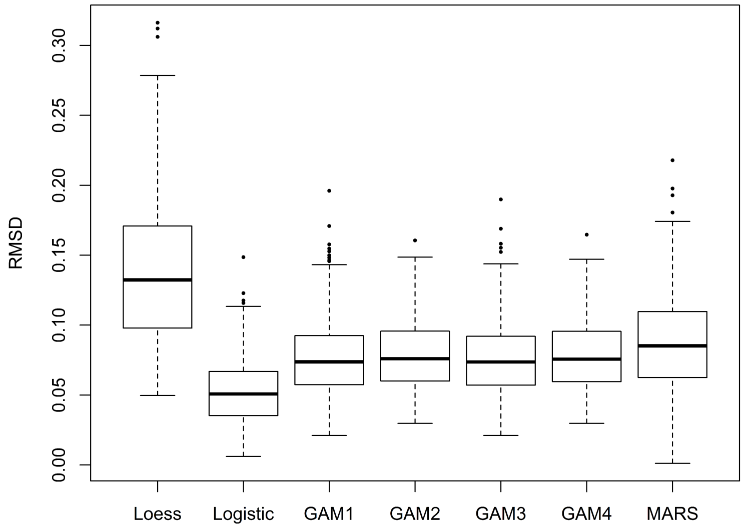
(5) Frank (1.9) (Corr = 0.3)



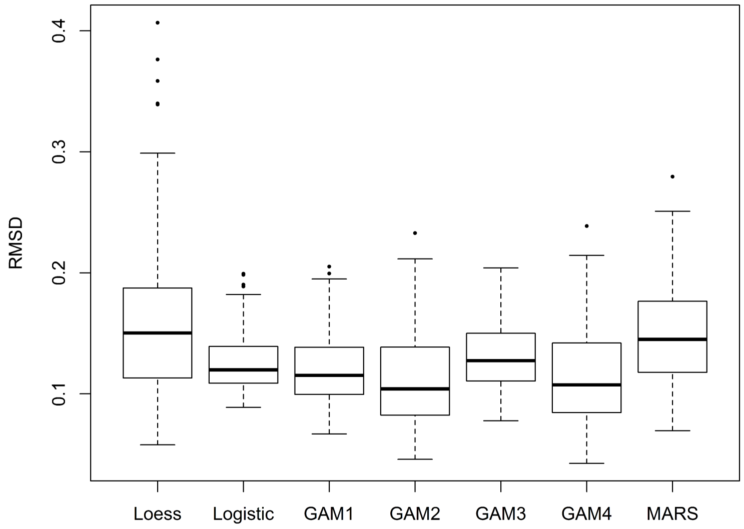
(6) Gumbel (1.26) (Corr = 0.3)



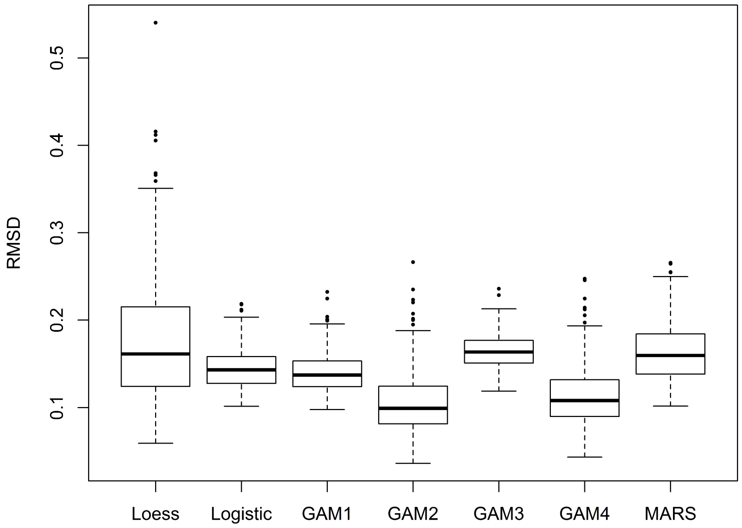
(7) Clayton (2.15) (Corr = 0.7)



(8) Frank (5.8) (Corr = 0.7)



(9) Gumbel (2.07) (Corr = 0.7)



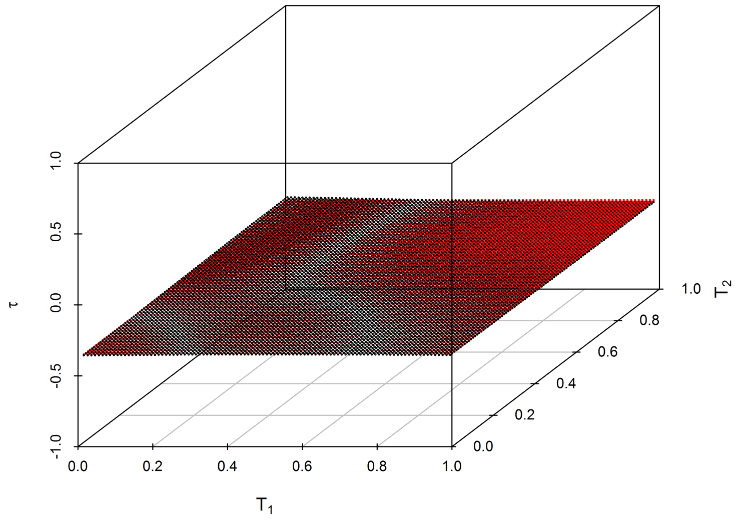
A.2 Comparison of true and estimated local Kendall's τ for various association structure

A.2.1 Mean of 300 replicates

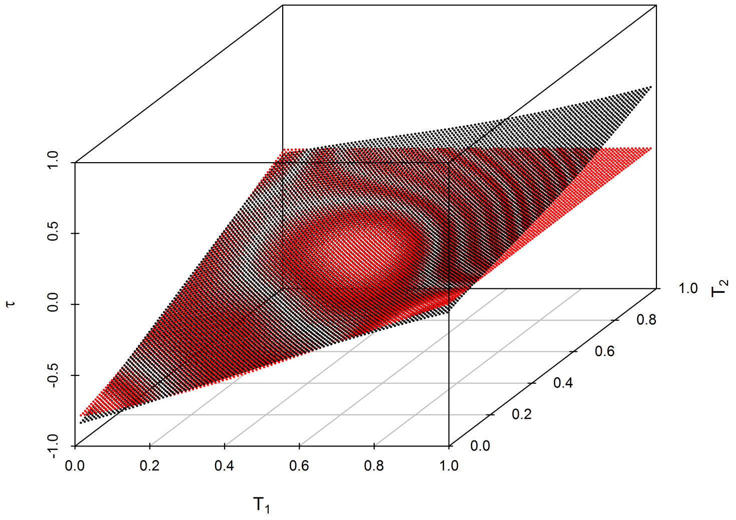
We used GAM with bivariate function and Gaussian family.

True values are in red and the estimates are in black.

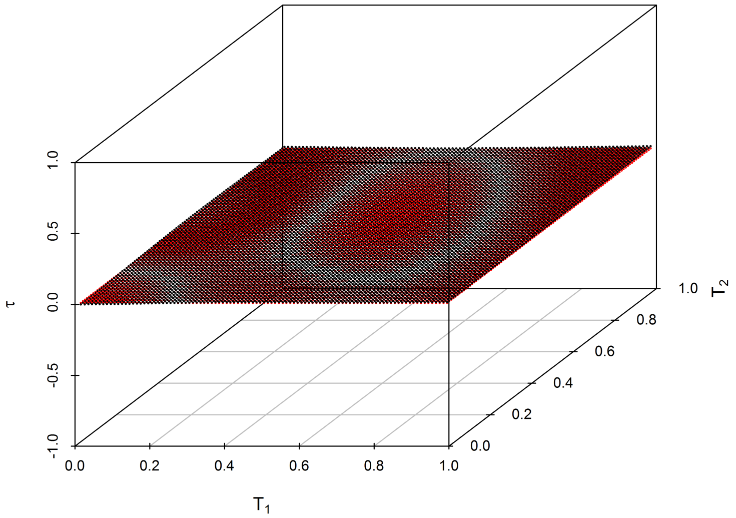
(1) Clayton (-0.53) (Corr = -0.5)



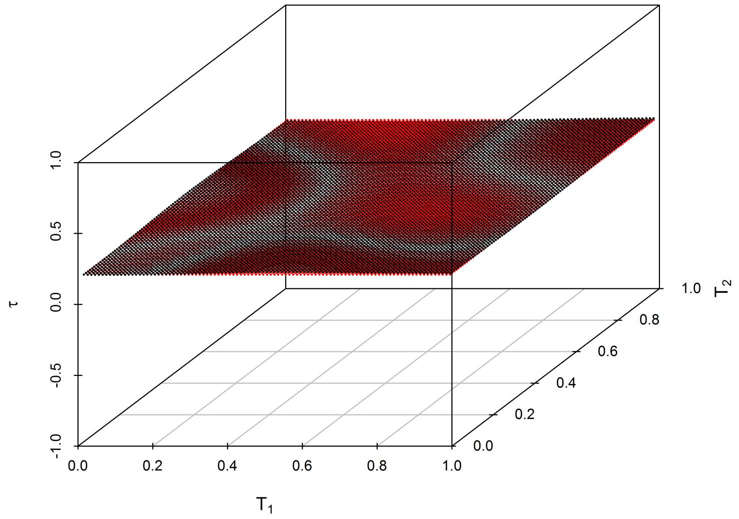
(2) Frank (-3.5) (Corr = -0.5)



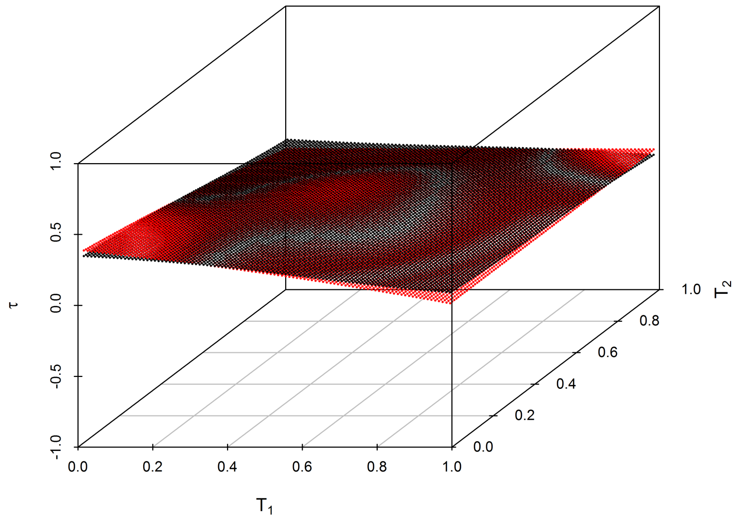
(3) Independent (Corr = 0)



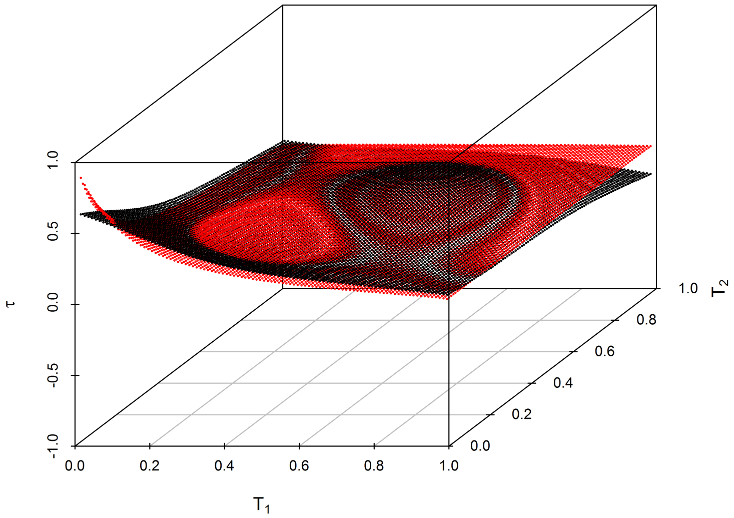
(4) Clayton (0.5) (Corr = 0.3)



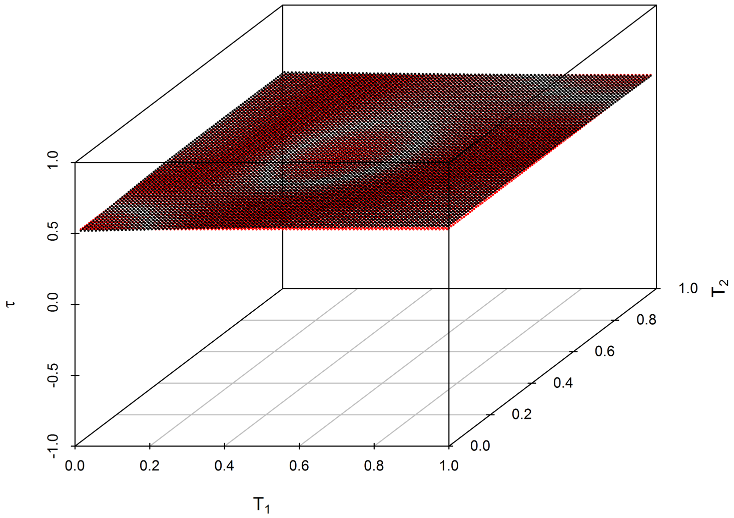
(5) Frank (1.9) (Corr = 0.3)



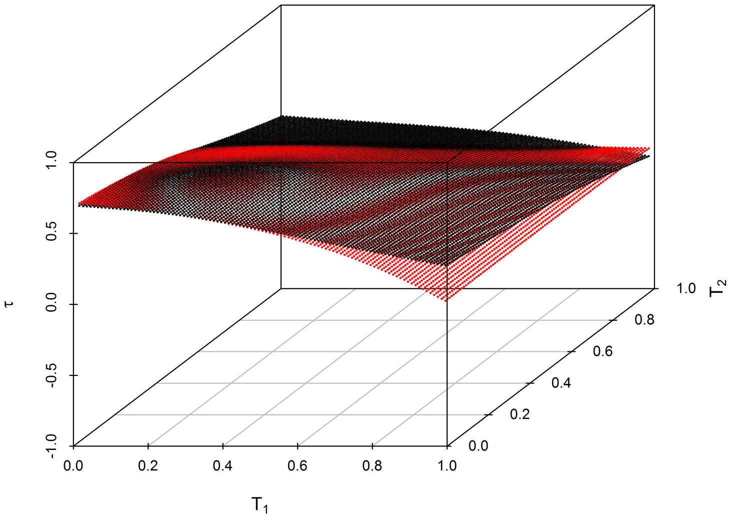
(6) Gumbel (1.26) (Corr = 0.3)



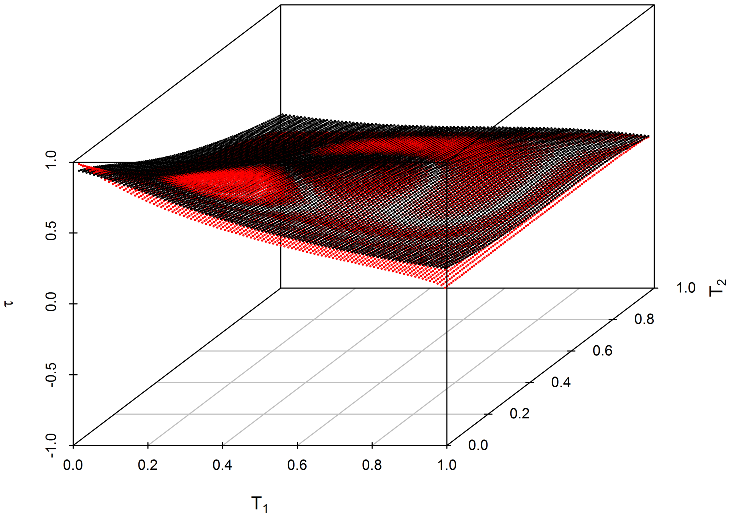
(7) Clayton (2.15) (Corr = 0.7)



(8) Frank (5.8) (Corr = 0.7)



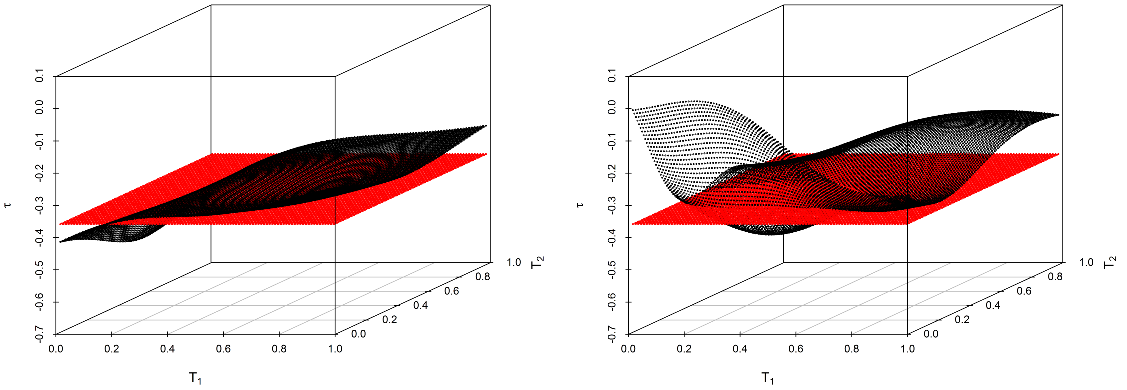
(9) Gumbel (2.07) (Corr = 0.7)



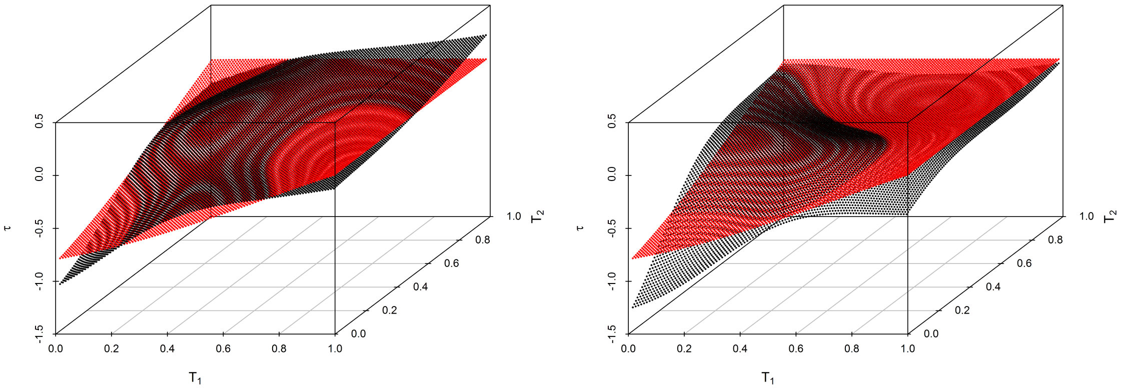
A.2.2 5th and 95th percentiles of RMSD

Two simulated datasets were selected among 300 replicates whose RMSDs are at the 5th and 95th percentiles as examples of ‘good’ and ‘bad’ estimates. The left panel is for 5th percentile and the right panel is for 95th percentile.

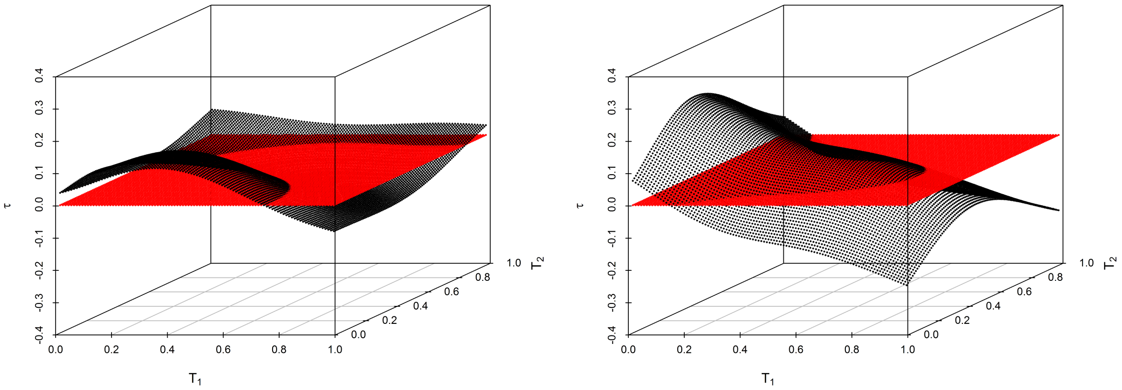
(1) Clayton (-0.53) (Corr = -0.5)



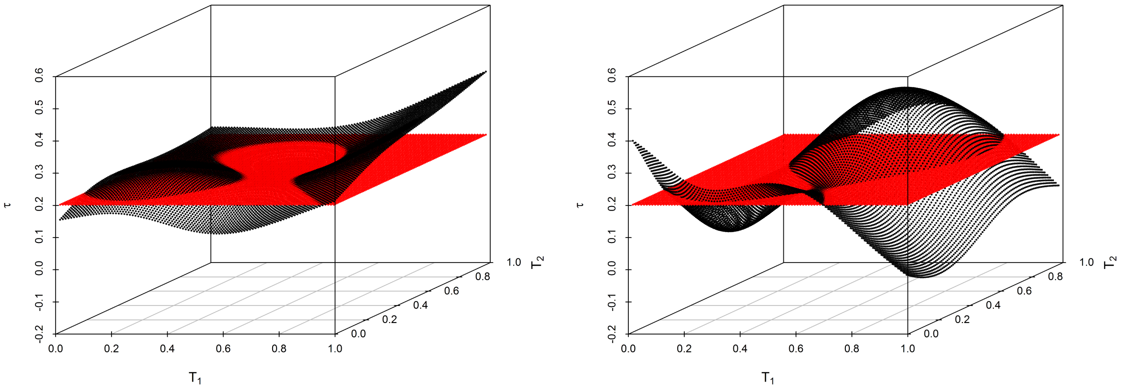
(2) Frank (-3.5) (Corr = -0.5)



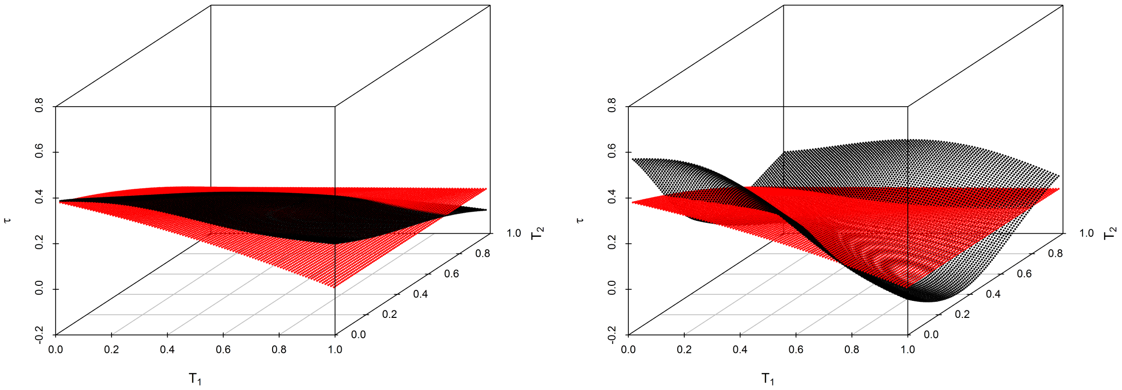
(3) Independent (Corr = 0)



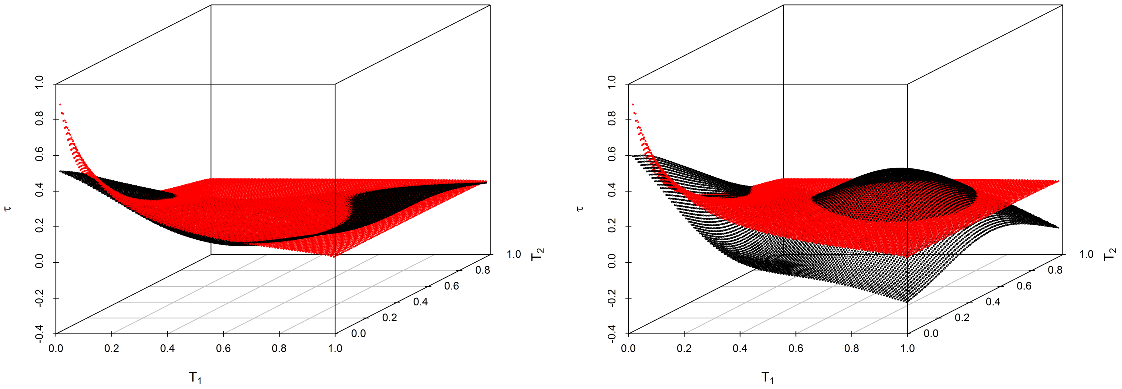
(4) Clayton (0.5) (Corr = 0.3)



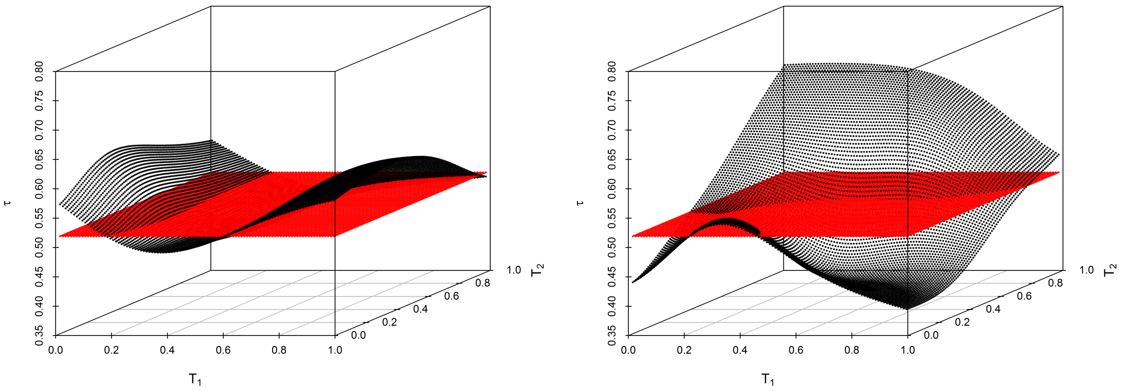
(5) Frank (1.9) (Corr = 0.3)



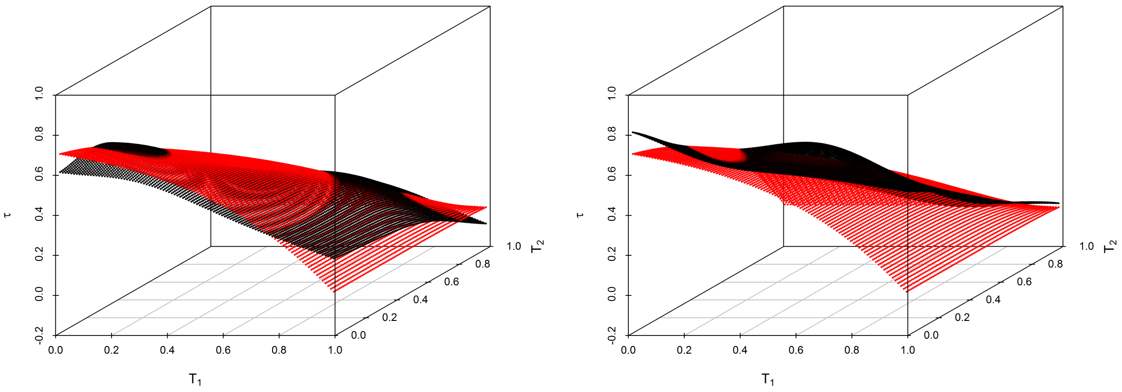
(6) Gumbel (1.26) (Corr = 0.3) (Corr = -0.5)



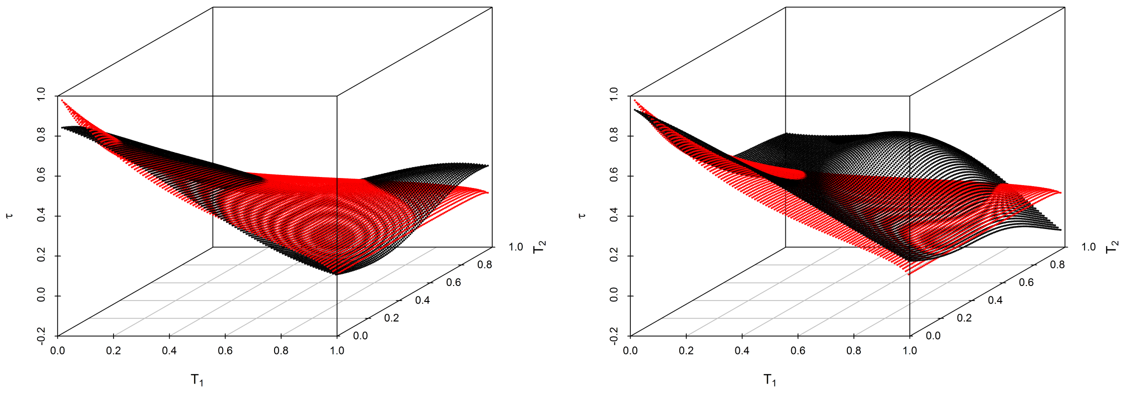
(7) Clayton (2.15) (Corr = 0.7)



(8) Frank (5.8) (Corr = 0.7)



(9) Gumbel (2.07) (Corr = 0.7)



A.3 Censoring-treatment techniques

We compared various censoring-treatment techniques in terms of RMSD for various association structure between failure times and between censoring times. We also presented the decomposition of RMSD into variance and bias squared.

NC: Assuming there is no censoring

CC: Using only complete case pairs

For Brown, MI, and Chen method, we presented the efficacy which is defined as

$$\frac{(\text{RMSD of CC}) - (\text{RMSD of each method})}{(\text{RMSD of CC}) - (\text{RMSD of NC})}.$$

A.3.1 RMSD

(1) 30% censored

Table A.7: RMSD, 30% censored

Failure	Censoring	NC	CC	Brown		MI		Chen	
Clayton	Positive	0.101	0.162	0.189	-44.2%	0.167	-7.6%	0.138	39.3%
Clayton	Negative	0.101	0.263	0.284	-12.6%	0.215	29.5%	0.168	58.7%
Clayton	Indep.	0.101	0.222	0.246	-19.3%	0.190	26.9%	0.152	58.4%
Frank	Positive	0.124	0.187	0.172	24.7%	0.130	90.6%	0.171	25.3%
Frank	Negative	0.124	0.245	0.205	32.9%	0.142	84.9%	0.199	37.9%
Frank	Indep.	0.124	0.211	0.191	22.9%	0.139	83.4%	0.186	29.4%
Gumbel	Positive	0.124	0.197	0.181	21.4%	0.139	80.3%	0.177	27.4%
Gumbel	Negative	0.124	0.265	0.211	38.6%	0.156	77.3%	0.217	34.6%
Gumbel	Indep.	0.124	0.229	0.197	30.2%	0.149	76.7%	0.197	30.4%
Indep.	Positive	0.129	0.172	0.118	125.0%	0.131	94.1%	0.179	-16.1%
Indep.	Negative	0.129	0.202	0.105	132.5%	0.134	93.2%	0.200	3.3%
Indep.	Indep.	0.129	0.187	0.108	136.4%	0.132	93.9%	0.191	-6.1%

(2) 50% censored

Table A.8: RMSD, 50% censored

Failure	Censoring	NC	CC	Brown		MI		Chen	
Clayton	Positive	0.101	0.230	0.243	-10.3%	0.241	-8.5%	0.179	39.3%
Clayton	Negative	0.101	0.465	0.365	27.6%	0.334	36.0%	0.331	37.0%
Clayton	Indep.	0.101	0.370	0.330	14.6%	0.294	28.3%	0.249	44.9%
Frank	Positive	0.124	0.254	0.225	22.3%	0.167	67.3%	0.222	25.2%
Frank	Negative	0.124	0.427	0.264	53.9%	0.200	75.0%	0.341	28.2%
Frank	Indep.	0.124	0.342	0.256	39.5%	0.183	73.1%	0.279	28.8%
Gumbel	Positive	0.124	0.272	0.239	22.1%	0.178	63.4%	0.250	15.0%
Gumbel	Negative	0.124	0.476	0.253	63.3%	0.213	74.9%	0.369	30.4%
Gumbel	Indep.	0.124	0.372	0.258	46.1%	0.191	73.2%	0.306	26.8%
Indep.	Positive	0.129	0.245	0.122	105.6%	0.146	84.9%	0.248	-2.5%
Indep.	Negative	0.129	0.390	0.112	106.6%	0.147	93.2%	0.384	2.4%
Indep.	Indep.	0.129	0.304	0.115	108.1%	0.147	89.5%	0.312	-4.7%

A.3.2 Variance

Definition: Variance of estimates across 300 replicates was averaged over the entire time domain

(1) 30% censored

Table A.9: Variance, 30% censored

Failure	Censoring	NC	CC	Brown	MI	Chen
Clayton	Positive	0.011	0.023	0.011	0.012	0.021
Clayton	Negative	0.011	0.049	0.010	0.012	0.031
Clayton	Indep.	0.011	0.037	0.010	0.011	0.025
Frank	Positive	0.015	0.033	0.014	0.015	0.030
Frank	Negative	0.015	0.053	0.012	0.015	0.042
Frank	Indep.	0.015	0.040	0.013	0.016	0.036
Gumbel	Positive	0.017	0.041	0.015	0.017	0.035
Gumbel	Negative	0.017	0.070	0.013	0.018	0.056
Gumbel	Indep.	0.017	0.054	0.014	0.019	0.045
Indep.	Positive	0.018	0.031	0.015	0.018	0.034
Indep.	Negative	0.018	0.043	0.012	0.019	0.042
Indep.	Indep.	0.018	0.038	0.013	0.019	0.039

(2) 50% censored

Table A.10: Variance, 50% censored

Failure	Censoring	NC	CC	Brown	MI	Chen
Clayton	Positive	0.011	0.046	0.013	0.013	0.034
Clayton	Negative	0.011	0.214	0.013	0.014	0.124
Clayton	Indep.	0.011	0.113	0.014	0.014	0.069
Frank	Positive	0.015	0.063	0.018	0.017	0.051
Frank	Negative	0.015	0.179	0.015	0.015	0.129
Frank	Indep.	0.015	0.110	0.019	0.016	0.085
Gumbel	Positive	0.017	0.079	0.021	0.020	0.073
Gumbel	Negative	0.017	0.231	0.017	0.018	0.161
Gumbel	Indep.	0.017	0.141	0.018	0.018	0.112
Indep.	Positive	0.018	0.063	0.017	0.023	0.068
Indep.	Negative	0.018	0.172	0.015	0.022	0.167
Indep.	Indep.	0.018	0.101	0.015	0.022	0.108

A.3.3 Bias

Definition: Mean of squared bias across 300 replicates was averaged over the entire time domain, then square root of it was taken

(1) 30% censored

Table A.11: Bias, 30% censored

Failure	Censoring	NC	CC	Brown	MI	Chen
Clayton	Positive	0.017	0.070	0.165	0.145	0.013
Clayton	Negative	0.017	0.155	0.278	0.209	0.024
Clayton	Indep.	0.017	0.124	0.236	0.178	0.021
Frank	Positive	0.060	0.083	0.132	0.058	0.079
Frank	Negative	0.060	0.107	0.180	0.080	0.081
Frank	Indep.	0.060	0.092	0.162	0.071	0.076
Gumbel	Positive	0.034	0.056	0.141	0.065	0.049
Gumbel	Negative	0.034	0.094	0.185	0.093	0.038
Gumbel	Indep.	0.034	0.074	0.168	0.077	0.039
Indep.	Positive	0.016	0.033	0.008	0.012	0.016
Indep.	Negative	0.016	0.030	0.012	0.015	0.029
Indep.	Indep.	0.016	0.019	0.011	0.011	0.036

(2) 50% censored

Table A.12: Bias, 50% censored

Failure	Censoring	NC	CC	Brown	MI	Chen
Clayton	Positive	0.017	0.096	0.219	0.233	0.031
Clayton	Negative	0.017	0.169	0.352	0.338	0.048
Clayton	Indep.	0.017	0.188	0.316	0.296	0.039
Frank	Positive	0.060	0.099	0.192	0.113	0.090
Frank	Negative	0.060	0.144	0.247	0.165	0.107
Frank	Indep.	0.060	0.127	0.226	0.136	0.096
Gumbel	Positive	0.034	0.068	0.207	0.119	0.054
Gumbel	Negative	0.034	0.163	0.235	0.170	0.054
Gumbel	Indep.	0.034	0.127	0.233	0.143	0.053
Indep.	Positive	0.016	0.046	0.005	0.013	0.020
Indep.	Negative	0.016	0.034	0.006	0.014	0.038
Indep.	Indep.	0.016	0.028	0.004	0.015	0.019

A.4 Generation of correlated failure times with marginally exponential distribution

To generate 'disease' failure time of the first component of a pair, we generated gamma distributed random numbers as in the third set of simulation studies (see methods). Then we used the fact that $\frac{\log(1-\log(U)/A)}{l_1 \times (t-1)}$ is exponentially distributed where U is uniformly distributed, l_1 is the exponential parameter, and A is gamma distributed with a shape parameter $1/(t-1)$ and a scale parameter 1. The 'disease' failure time for the second component and the 'death' failure times for two components were generate similarly.

To see that this method yields the distributions as claimed, let us consider a univariate frailty model with a random effect denoted by α, with distribution G and Laplace transformation $p(x) = E(e^{-x\alpha})$, where the marginal survival function for individual j in the cluster is $S_j(t) = \int \{S_j^*(t)\}^a dG(a)$. Then, $-\log S_j^*(t) = q[S_j(t)]$, that is, $S_j(t) = p[-\log S_j^*(t)]$ where q is the inverse function of p (See Equation (1) of Bandeen-Roche and Liang (1996)). For exponential distribution, $S_j(t) = e^{-\lambda t}$ and for Clayton copula, $p(u) = (1+u)^{\frac{1}{1-\theta}}$. Thus,

$$\begin{aligned} e^{-\lambda T} &= (1 - \log S_j^*(T))^{\frac{1}{1-\theta}} \\ -\lambda T &= \frac{1}{1-\theta} \log(1 - \log S_j^*(t)) \\ T &= \frac{1}{\lambda(\theta-1)} \log(1 - \log S_j^*(t)). \end{aligned} \tag{A.4.1}$$

For gamma frailty, conditionally on frailty, $S_j^*(T)^A$ is uniformly distributed, thus $\log(S_j^*(T)) =$

$\log(U)/A$. Then,

$$T = \frac{1}{\theta - 1} \log(1 - \log(U)/A). \tag{A.4.2}$$

Bibliography

Aitchison, J. (1982). The statistical-analysis of compositional data. *Journal of the Royal Statistical Society Series B-Methodological*, 44(2):139–177.

Aitchison, J. and Shen, S. M. (1980). Logistic-normal distributions - some properties and uses. *Biometrika*, 67(2):261–272.

Anderson, J. E., Louis, T. A., Holm, N. V., and Harvald, B. (1992). Time-dependent association measures for bivariate survival distributions. *Journal of the American Statistical Association*, 87(419):641–650.

Bandeen-Roche, K. and Liang, K. Y. (2002). Modelling multivariate failure time associations in the presence of a competing risk. *Biometrika*, 89(2):299–314.

Bandeen-Roche, K. and Ning, J. (2008). Nonparametric estimation of bivariate failure time associations in the presence of a competing risk. *Biometrika*, 95(1):221–232.

Bandeen-Roche, K. and Ruppert, D. (1991). Source apportionment with one source unknown. *Chemometrics and Intelligent Laboratory Systems*, 10(1-2):169–184.

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., and Khachaturian, A. (1999). Apoe-epsilon 4 count predicts age when prevalence of ad increases, then declines - the cache county study. *Neurology*, 53(2):321–331.

Brown, Jr., B. W., Hollander, M., and Korwar, R. M. (1974). *Nonparametric tests of independence for censored data, with applications to heart transplant studies*, pages 327–354. Society for Industrial and Applied Mathematics, Philadelphia.

Chen, M. C. and Bandeen-Roche, K. (2005). A diagnostic for association in bivariate survival models. *Lifetime Data Analysis*, 11(2):245–264.

Cheng, Y. and Fine, J. P. (2008). Nonparametric estimation of cause-specific cross hazard ratio with bivariate competing risks data. *Biometrika*, 95(1):233–240.

Cheng, Y. and Fine, J. P. (2012). Cumulative incidence association models for bivariate competing risks data. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 74:183–202.

Cheng, Y., Fine, J. P., and Bandeen-Roche, K. (2010). Association analyses of clustered competing risks data via cross hazard ratio. *Biostatistics*, 11(1):82–92.

Cheng, Y., Fine, J. P., and Kosorok, M. R. (2007). Nonparametric association analysis of bivariate competing-risks data. *Journal of the American Statistical Association*, 102(480):1407–1415.

Cheng, Y., Fine, J. P., and Kosorok, M. R. (2009). Nonparametric association analysis of exchangeable clustered competing risks data. *Biometrics*, 65(2):385–393.

Clayton, D. and Cuzick, J. (1985). Multivariate generalizations of the proportional hazards model. *Journal of the Royal Statistical Society Series A-Statistics in Society*, 148:82–117.

Clayton, D. G. (1978). Model for association in bivariate life tables and its application in epidemiological-studies of familial tendency in chronic disease incidence. *Biometrika*, 65(1):141–151.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829–836.

Cleveland, W. S. and Devlin, S. J. (1988). Locally weighted regression - an approach to regression-analysis by local fitting. *Journal of the American Statistical Association*, 83(403):596–610.

Dabrowska, D. M. (1988). Kaplan-meier estimate on the plane. *Annals of Statistics*, 16(4):1475–1489.

Fan, J., Hsu, L., and Prentice, R. L. (2000a). Dependence estimation over a finite bivariate failure time region. *Lifetime Data Analysis*, 6(4):343–355.

Fan, J. J. and Prentice, R. L. (2002). Covariate-adjusted dependence estimation on a finite bivariate failure time region. *Statistica Sinica*, 12(3):689–705.

Fan, J. J., Prentice, R. L., and Hsu, L. (2000b). A class of weighted dependence measures for bivariate failure time data. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 62:181–190.

Friedman, J. H. (1991). Multivariate adaptive regression splines. *Annals of Statistics*, 19(1):1–67.

Genest, C., Ghoudi, K., and Rivest, L. P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3):543–552.

Genest, C. and MacKay, R. J. (1986). Copules archimdiennes et families de lois bidimensionnelles dont les marges sont donnes. *Canadian Journal of Statistics*, 14(2):145–159.

Glidden, D. V. (2000). A two-stage estimator of the dependence parameter for the clayton-oakes model. *Lifetime Data Analysis*, 6(2):141–156.

Gorfine, M. and Hsu, L. (2011). Frailty-based competing risks model for multivariate survival data. *Biometrics*, 67(2):415–426.

Hastie, T. and Tibshirani, R. (1986). Generalized additive models. *Statistical science*, pages 297–310.

Hendrie, H. C. (1998). Epidemiology of dementia and alzheimer's disease. *American Journal of Geriatric Psychiatry*, 6(2):S3–S18.

Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika*, 73(3):671–678.

Hougaard, P. (2000). *Analysis of multivariate survival data*. Statistics for biology and health. Springer, New York.

Hsu, L. and Prentice, R. L. (1996). On assessing the strength of dependency between failure time variates. *Biometrika*, 83(3):491–506.

Liang, K.-Y., Self, S. G., Bandeen-Roche, K. J., and Zeger, S. L. (1995). Some recent ments for regression analysis of multivariate failure time data. *Lifetime data analysis*, 1(4):403–415.

Nair, N. U. and Sankaran, P. G. (2010). A new measure of association for bivariate survival data. *Journal of Statistical Planning and Inference*, 140(9):2569–2581.

Nelder, J. and Wedderburn, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society Series A-Statistics in Society*, 135:370–384.

Nielsen, G. G., Gill, R. D., Andersen, P. K., and Sorensen, T. I. A. (1992). A counting process approach to maximum-likelihood-estimation in frailty models. *Scandinavian Journal of Statistics*, 19(1):25–43.

Ning, J. and Bandeen-Roche, K. (2014). Estimation of time-dependent association for bivariate failure times in the presence of a competing risk. *Biometrics*, 70(1):10–20.

Oakes, D. (1982). A model for association in bivariate survival-data. *Journal of the Royal Statistical Society Series B-Methodological*, 44(3):414–422.

Oakes, D. (1986). Semiparametric inference in a model for association in bivariate survival-data. *Biometrika*, 73(2):353–361.

Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, 84(406):487–493.

Prentice, R. L. (2014). Self-consistent nonparametric maximum likelihood estimator of the bivariate survivor function. *Biometrika*, 101(3):505–518.

Prentice, R. L. and Cai, J. (1992). Covariance and survivor function estimation using censored multivariate failure time data. *Biometrika*, 79(3):495–512.

Pruitt, R. C. (1991). On negative mass assigned by the bivariate kaplan-meier estimator. *Annals of Statistics*, 19(1):443–453.

R Core Team (2013). R: A language and environment for statistical computing.

Ripatti, S., Larsen, K., and Palmgren, J. (2002). Maximum likelihood inference for multivariate frailty models using an automated monte carlo em algorithm. *Lifetime Data Analysis*, 8(4):349–360.

Ripatti, S. and Palmgren, J. (2000). Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*, 56(4):1016–1022.

Sankaran, P. G., Abraham, B., and Antony, A. A. (2006). A dependence measure for bivariate failure time data. *METRON*, 64(3):327–341.

Scheike, T. H. and Sun, Y. Q. (2012). On cross-odds ratio for multivariate competing risks data. *Biostatistics*, 13(4):680–694.

Scheike, T. H., Sun, Y. Q., Zhang, M. J., and Jensen, T. K. (2010). A semiparametric random effects model for multivariate competing risks data. *Biometrika*, 97(1):133–145.

Shih, J. H. and Albert, P. S. (2010). Modeling familial association of ages at onset of disease in the presence of competing risk. *Biometrics*, 66(4):1012–1023.

Shih, J. H. and Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, 51(4):1384–1399.

Silverman, J. M., Ciresi, G., Smith, C., Marin, D. B., and Schnaider-Beeri, M. (2005). Variability of familial risk of alzheimer disease across the late life span. *Archives of General Psychiatry*, 62(5):565–573.

van der Laan, M. J. (1996). Nonparametric estimation of the bivariate survival function with truncated data. *Journal of Multivariate Analysis*, 58(1):107–131.

Varadhan, R., Xue, Q. L., and Bandeen-Roche, K. (2014). Semicompeting risks in aging research: methods, issues and needs. *Lifetime Data Analysis*, 20(4):538–562.

Vaupel, J. W., Manton, K. G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3):439–54.

CURRICULUM VITAE

JEONGYONG KIM

jkim339@jhu.edu

615 N. Wolfe St. E3034

Baltimore, MD 21205

Date of Birth: Mar 22nd, 1983

Place of Birth: Seoul, Republic of Korea

EDUCATION

2017 **Johns Hopkins Bloomberg School of Public Health**, Baltimore, MD

Ph.D. in Biostatistics

Thesis title: *Statistical Methods for Multivariate Failure-Time Data under Competing Risks*

Advisor: Prof. Karen Bandeen-Roche

2008 **Seoul National University**, Seoul, Republic of Korea

M.S. in Statistics

2006 **Seoul National University**, Seoul, Republic of Korea

B.S. in Industrial Engineering and B.S. in Statistics

CURRICULUM VITAE

PROFESSIONAL EXPERIENCE

2015 **Research Assistant**

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

Advisor: Dr. Anbrasi Edward

2014 - 2015 **Research Assistant**

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

Advisor: Dr. Parul Christian; *Principal Investigator*: Dr. Keith P. West Jr.

2013 - 2014 **Research Assistant**

Department of Population, Family and Reproductive Health, Johns Hopkins Bloomberg School of Public Health

Advisor: Dr. Vladimir Canudas-Romo

2010 - 2013 **Research Assistant**

School of Medicine (CHECK), Johns Hopkins University

Advisor: Dr. Karen-Bandeen-Roche; *Principal Investigator*: Dr. L. Ebony Boulware

HONORS AND AWARDS

2006-07	SNU Development Fund Scholarship
2006	Best Alumni Award (SNU College of Engineering)
2001-06	Scholarship for Academic Excellence

CURRICULUM VITAE

Publications

PUBLISHED/SUBMITTED

Kim J, Bandeen-Roche K (2017). Parametric Estimation of Association in Bivariate Failure-time Data Subject to Competing Risks: Sensitivity to Underlying Assumptions. *Revision Invited for Lifetime Data Analysis*.

Christian P, **Kim J**, Mehra S, Shaikh S, Ali H, Shamim AA, Wu L, Klemm Rolf, Labrique A, West Jr KP (2016). Effects of prenatal multiple micronutrient supplementation on growth and cognition through 2 years of age in rural Bangladesh: the JiVitA-3 Trial. *The American Journal of Clinical Nutrition, 104*(4) 1175-1182.

Michels WM, Jaar BG, Ephraim PL, Liu Y, Miskulin DC, Tangri N, Crews DC, Scialla JJ, Shafi T, Sozio SM, Bandeen-Roche K, Cook CJ, Meyer KB, Boulware LE, DEcIDE Network Patient Outcomes in End-Stage Renal Disease Study Investigators (2016). Intravenous iron administration strategies and anemia management in hemodialysis patients.*Nephrology Dialysis Transplantation*. *(As a member of DEcIDE Study Investigators)*

Tangri N, Miskulin DC, Zhou J, Bandeen-Roche K, Michels WM, Ephraim PL, McDermott A, Crews DC, Scialla JJ, Sozio SM, Shafi T, Jaar BG, Meyer K, Boulware LE, DEcIDE Network Patient Outcomes in End-Stage Renal Disease Study Investigators (2015). Effect of intravenous iron use on hospitalizations in patients undergoing hemodialysis: a comparative effectiveness analysis from the DEcIDE-ESRD study. *Nephrology Dialysis Transplantation, 30*(4), 667-675. *(As a member of DEcIDE Study Investigators)*

Miskulin DC, Tangri N, Bandeen-Roche K, Zhou J, McDermott A, Meyer KB, Ephraim PL, Michels WM, Jaar BG, Crews DC, Scialla JJ, Sozio SM, Shafi T, Wu AW, Cook C, Boulware LE, DEcIDE Network Patient Outcomes in End-Stage Renal Disease Study Investigators (2015). Intravenous iron exposure and mortality in patients on hemodialysis. *Clinical Journal of the American Society of Nephrology*, CJN-03370414. *(As a member of DEcIDE Study Investigators)*

Scialla JJ, Liu J, Crews DC, Guo H, Bandeen-Roche K, Ephraim PL, Tangri N, Sozio SM, Shafi T, Miskulin DC, Michels WM, Jaar BG, Wu AW1, Powe NR, Boulware LE, DEcIDE Network Patient Outcomes in End-Stage Renal Disease Study Investigators (2014). An instrumental variable approach finds no associated harm or benefit with early dialysis initiation in the United States. *Kidney International, 86*(4), 798-809. *(As a member of DEcIDE Study Investigators)*

Crews DC, Scialla JJ, Boulware LE, Navaneethan SD, Nally JV, Liu X, Arrigain S, Schold JD, Ephraim PL, Jolly SE, Sozio SM, Michels WM, Miskulin DC, Tangri N, Shafi T, Wu AW, Bandeen-Roche K,DEcIDE Network Patient Outcomes in End-Stage Renal Disease Study Investigators (2014). Comparative Effectiveness of Early Versus Conventional Timing of Dialysis Initiation in Advanced CKD. *American Journal of Kidney Diseases, 63*(5), 806-815. *(As a member of DEcIDE Study Investigators)*

Shafi T, Sozio SM, Bandeen-Roche K, Ephraim PL, Luly JR, St. Peter WL, McDermott A, Scialla JJ, Crews DC, Tangri N, Miskulin DC, Michels WM, Jaar BG, Herzog CA, Zager PG, Meyer KB, Wu AW, Boulware LE, DEcIDE Network Patient Outcomes in End-Stage

Renal Disease Study Investigators (2014). Predialysis systolic BP variability and outcomes in hemodialysis patients. *Journal of the American Society of Nephrology, 25*(4), 799-809. *(As a member of DEcIDE Study Investigators)*

Crews DC, Scialla JJ, Liu J, Guo H, Bandeen-Roche K, Ephraim PL, Jaar BG, Sozio SM, Miskulin DC, Tangri N, Shafi T, Meyer KB, Wu AW, Powe NR, Boulware LE, DEcIDE Network Patient Outcomes in End-Stage Renal Disease Study Investigators (2014). Predialysis Health, Dialysis Timing, and Outcomes among Older United States Adults. *Journal of the American Society of Nephrology, 25*(2), 370-379. *(As a member of DEcIDE Study Investigators)*

Miskulin DC, Zhou J, Tangri N, Bandeen-Roche K, Cook C, Ephraim PL, Crews DC, Scialla JJ, Sozio SM, Shafi T, Jaar BG, Boulware LE, DEcIDE Network Patient Outcomes in End-Stage Renal Disease Study Investigators. (2013). Trends in anemia management in US hemodialysis patients 20042010. *BMC Nephrology, 14*, 265. *(As a member of DEcIDE Study Investigators)*

St Peter WL, Sozio SM, Shafi T, Ephraim PL, Luly J, McDermott A, Bandeen-Roche K, Meyer KB, Crews DC, Scialla JJ, Miskulin DC, Tangri N, Jaar BG, Michels WM, Wu AW, Boulware LE, DEcIDE Network Patient Outcomes in End-Stage Renal Disease Study Investigators (2013), Patterns in blood pressure medication use in US incident dialysis patients over the first 6 months. *BMC Nephrology, 14*, 249. *(As a member of DEcIDE Study Investigators)*

WORKING PAPERS

Kim J, Bandeen-Roche K. Nonparametric Estimation of Association in Bivariate Failure-time Data.

Kim J, Bandeen-Roche K. Nonparametric and Semiparametric Estimation of Association in Bivariate Failure-time Data under Competing Risks.

TEACHING

INSTRUCTOR

2007	Statistics Lab

TEACHING ASSISTANT

2013-15	Statistics for Laboratory Scientists
2012	Multilevel Statistical Models in Public Health
2012	Analysis of Longitudinal Data
2011	Statistics for Psychosocial Research
2011-14	Statistical Methods in Public Health
2010-12	Statistical Reasoning in Public Health
2007	Multivariate Analysis
2006	Statistics

Professional Activities

Volunteer	ENAR Spring Meeting, Baltimore, Maryland, 2014
Member	Survival, Longitudinal, and Multivariate (SLAM) Data Working Group