Testing Independence for Bivariate Current Status Data

By A. Adam Ding
Email: ding@neu.edu
Department of Math, Northeastern University
Boston, MA 02115, U.S.A.

and Weijing Wang
Email: wjwang@stat.nctu.edu.tw
Institute of Statistics, National Chiao-Tung University
Hsin-Chu, Taiwan, R.O.C.

Summary

This paper develops a nonparametric procedure for testing marginal independence based on bivariate current status data. Asymptotic properties of the proposed tests are derived and their finite sample performance is studied via simulations. The method is applied to analyze data from a community-based study of cardiovascular epidemiology in Taiwan.

Some key words: Epidemiology; Interval Censoring; Lifetime data; Nonparametric Analysis; Cochran-Mantel-Haenszel Test; Two-by-Two Tables.

Aidong Adam Ding is Associate Professor, Department of Math, Northeastern University, Boston, MA 02115 (Email: a.ding@neu.edu). Weijing Wang is Associate Professor, Institute of Statistics, National Chiao-Tung University, Hsin-Chu, Taiwan, R.O.C. (Email: wjwang@stat.nctu.edu.tw). The authors would like to thank the editors and reviewers of this article and Sophia Shing for their helpful comments and suggestions. W. Wang was supported on this research by a grant from the Taiwan National Science Council. This work benefited from the allocation of computer time at the Northeastern University Advanced Scientific Computation Center (NU-ASCC).

1 Introduction

Current status data commonly arises in animal tumorigenicity and epidemiological investigations of the natural history of a disease. Specifically, the researcher only has the information about whether the failure time of interest lies before or after the observed monitoring time. Such data structure is also called interval censoring of the case I (Groeneboom & Wellner, 1992). In this article, we consider the bivariate case. Bivariate analysis is useful when one wants to investigate the dependent relationship between two variables. Our work was motivated by a community-based study of cardiovascular diseases in Taiwan. The purpose of the study was to investigate whether the onset ages of patients with some common cardiovascular diseases, specifically hypertension, diabetes mellitus and hypercholesterolemia, are correlated with each other. Because the natural history of these chronic diseases was difficult to trace precisely, the data only contained information about whether or not a subject under the study had already developed the diseases and, about the subject's current age at the time of the study.

Let (T_1, T_2) be a pair of failure times of interest and C_j be the monitoring time of T_j $(j = 1, 2)$, respectively. Bivariate current status data are of the form, $\{C_1, C_2, \delta_1 = I(T_1 \leq C_1), \delta_2 = I(T_2 \leq C_2)\}$. The observed data are of the form, $(C_{1k}, C_{2k}, \delta_{1k}, \delta_{2k})$ $(k = 1, \ldots, n)$, which are independent identically distributed replications of $(C_1, C_2, \delta_1, \delta_2)$. Note that when the two failure times are measured from the same subject, as in the above example, usually $C_1 = C_2$. A number of statistical methods have been developed for univariate current status data. For example, nonparametric estimation of the marginal distribution function has been considered by Ayer et al. (1955), Peto (1973), Turnbull (1976) and Groeneboom & Wellner (1992). The algorithm for computing the nonparametric maximum likelihood estimator for current status data is introduced in Groeneboom & Wellner (1992, p. 66-67). Asymptotic properties of the nonparametric maximum likelihood estimator were examined in Groeneboom & Wellner (1992). It is shown that this estimator converges pointwise at rate $n^{1/3}$ to a complex limiting distribution related to Brownian motion. Properties of smooth functionals of the nonparametric maximum likelihood estimator were studied in Groeneboom & Wellner (1992) and Huang & Wellner (1995). Semiparametric analysis of regression models for current status data have been studied by Finkelstein (1986), Rabinowitz, Tsiatis & Aragon (1995), Rossini & Tsiatis (1996), and Lin, Oakes & Ying (1998), just to name a few. Although bivariate analysis of current status data has many interesting applications, there has not been much literature in this direction to date. Wang & Ding (2000) considered semi-parametric estimation of the association parameter in a bivariate copula model.

The main objective of this work is to develop a nonparametric inference procedure for testing independence between two failure time variables given only bivariate current status data. It is important to note that the semiparametric procedure proposed by Wang & Ding (2000) can be directly applied to test independence only if the parameter under independence is located at the interior of the parameter space. Some modification is required to handle the case when the true parameter lies on the boundary. Several independence tests for bivariate right censored data have been developed by Oakes (1982), Shih & Louis (1996) and Hsu & Prentice (1996), among others. The test proposed by Oakes (1982) is based on estimating Kendall's tau under the null hypothesis. Shih & Louis (1996) studied several test statistics based on marginal martingale residuals.

Our ideas are similar to those discussed in Hsu & Prentice (1996), both of which can be viewed as a generalization of the Mantel-Haenszel test. Specifically Hsu and Prentice constructed a sequence of 2×2 tables formed at observed failure times and then proposed a test statistic based on the merged table. Later we will see that bivariate current status data can naturally be represented by two-by-two tables formed at observed monitoring times. However, the techniques used for current status data are different from those for right-censored data, which extensively use the martingale theory.

The paper is organized as follows. The main result is presented in §2. Simulation analysis and real data analysis are given in §3 and §4, respectively. Section 5 gives the concluding remarks.

2 The proposed methodology

2.1 Preliminary

Let $H_0 : T_1 \perp T_2$ and $H_0' : T_1 \perp T_2 | (C_1, C_2)$. It is obvious that when H_0 is true, H_0' must be true and if H_0' is false, H_0 must be false. Our original goal is to test the hypothesis $H_0 : T_1 \perp T_2$. However given current status data, we believe that it is impossible to capture any departure from H_0 when H_0' is true. For example when $C_1 = C_2$, it is possible to verify whether $F(t,t) = F_1(t)F_2(t)$, which describes independence along the diagonal $C_1 = C_2 = t$. However no information is available to judge whether off-diagonal independence (i.e. $F(t_1, t_2) = F_1(t_1)F_2(t_2)$ for $t_1 \neq t_2$) also holds. In other words, current status data provides only limited information to identify the dependent relationship between T_1 and T_2. Therefore, we focus on deriving tests for the null hypothesis $H_0' : T_1 \perp T_2 | (C_1, C_2)$ in this paper. It should be mentioned that although we set the null hypothesis to be H_0', under the Neyman-Pearson framework, any valid test for testing H_0' is also a valid test for testing H_0 since it gives the correct type-one error rate.

Given that $(C_1, C_2) = (c_1, c_2)$, one can construct the following two-by-two table.

	$\delta_2 = 1$	$\delta_2 = 0$	
$\delta_1 = 1$	$N_{11}(c_1, c_2)$	$N_{10}(c_1, c_2)$	
$\delta_1 = 0$	$N_{01}(c_1, c_2)$	$N_{00}(c_1, c_2)$	
			$N(c_1, c_2)$

The cell counts are defined as $N(c_1, c_2) = \sum_{k=1}^{n} I(C_{1k} = c_1, C_{2k} = c_2)$ and $N_{ij}(c_1, c_2) = \sum_{k=1}^{n} I(C_{1k} = c_1, C_{2k} = c_2, \delta_{1k} = i, \delta_{2k} = j)$ for $i, j = 0, 1$. The proposed test procedure is constructed by merging the tables according to the distribution of (C_1, C_2), which has a form similar to the Cochran-Mantel-Haenszel test (Agresti 1990, p. 231). Some independence tests based on merging several 2×2 tables have been discussed in Agresti (1990), some of which require large observations in each table. Our proposed method, by contrast, is valid whether C_1 and C_2 are discrete or continuous, and can handle empty cells. In fact, the method is particularly designed for merging sparse 2×2 tables, commonly arising in applications for bivariate current status data.

Specifically for the kth patient with monitoring times, (c_{1k}, c_{2k}), one can construct a two-by-two table which has only one entry and three empty cells. Given that $(C_{1k}, C_{2k}) = (c_{1k}, c_{2k})$, the cell counts $\{I(\delta_{1k} = 1, \delta_{2k} = 1)$, $I(\delta_{1k} = 1, \delta_{2k} = 0)$, $I(\delta_{1k} = 0, \delta_{2k} = 1)$, $I(\delta_{1k} = 0, \delta_{2k} = 0)\}$ jointly follow a multinomial distribution with probabilities equal to $\{P_{11}(c_{1k}, c_{2k})$, $P_{10}(c_{1k}, c_{2k})$, $P_{01}(c_{1k}, c_{2k})$ and $P_{00}(c_{1k}, c_{2k})\}$, where $P_{11}(c_{1k}, c_{2k}) = \mathrm{pr}(T_1 \leq c_{1k}, T_2 \leq c_{2k})$, $P_{10}(c_{1k}, c_{2k}) = \mathrm{pr}(T_1 \leq c_{1k}, T_2 > c_{2k})$, $P_{01}(c_{1k}, c_{2k}) = \mathrm{pr}(T_1 > c_{1k}, T_2 \leq c_{2k})$, $P_{00}(c_{1k}, c_{2k}) = \mathrm{pr}(T_1 > c_{1k}, T_2 > c_{2k})$. Under H_0' (as well as under H_0), it follows that $P_{11}(c_{1k}, c_{2k}) = F_1(c_{1k})F_2(c_{2k})$, $P_{10}(c_{1k}, c_{2k}) = F_1(c_{1k})S_2(c_{2k})$, $P_{01}(c_{1k}, c_{2k}) =$

$S_1(c_{1k})F_2(c_{2k})$ and $P_{00}(c_{1k}, c_{2k}) = S_1(c_{1k})S_2(c_{2k})$, where $F_j(t) = \mathrm{pr}(T_j \leq t)$, $S_j(t) = 1 - F_j(t)$ $(j = 1, 2)$, and $G(c_1, c_2) = \mathrm{pr}(C_1 \leq c_1, C_2 \leq c_2)$.

Our idea for testing independence between T_1 and T_2 is to compare observed cell counts in these two-by-two tables with their expected values under H_0'. Large values of the difference indicate departure from the null hypothesis. Combining all the tables, each with three empty cells and a single entry, the observed cell counts in the merged table become

$$N_{ab} = \sum_{k=1}^{n} I(\delta_{1k} = a, \delta_{2k} = b) = \sum_{c_1, c_2} N_{ab}(c_1, c_2) \ (a, b = 0, 1),$$

where the last sum is over all observed censoring time values. Under the null hypothesis H_0' which is conditional on the observed censoring times, the expected counts in the merged table become

$$E_{ab} = \sum_{k=1}^{n} E_{ab,k} = \sum_{k=1}^{n} F_1(c_{1k})^a S_1(c_{1k})^{1-a} F_2(c_{2k})^b S_2(c_{2k})^{1-b} \ (a, b = 0, 1),$$

which can be estimated by plugging in the corresponding marginal nonparametric maximum likelihood estimators of $F_j(\cdot)$ and $S_j(\cdot)$, denoted as $\hat{F}_j(\cdot)$ and $\hat{S}_j(\cdot) = 1 - \hat{F}_j(\cdot)$ $(j = 1, 2)$, respectively. Therefore E_{ij} can be estimated by

$$\hat{E}_{ab} = \sum_{k=1}^{n} \hat{E}_{ab,k} = \sum_{k=1}^{n} \hat{F}_1(c_{1k})^a \hat{S}_1(c_{1k})^{1-a} \hat{F}_2(c_{2k})^b \hat{S}_2(c_{2k})^{1-b} \ (a, b = 0, 1). \tag{1}$$

The max-min formula for computing $\hat{F}_j$ is given by

$$\hat{F}_j(c_{(ji)}) = \max_{l \leq i} \min_{k \geq i} \frac{\sum_{m=l}^{k} \delta_{(jm)}}{k - l + 1},$$

where $c_{(j1)} < \ldots < c_{(jn)}$ are ordered observed values of $(C_{j1}, \ldots, C_{jn})$ and $\delta_{(ji)}$ $(j = 1, 2)$ are the associated indicators for $C_{(ji)}$. Their properties are discussed in Groeneboom & Wellner (1992) and Huang & Wellner (1997). When H_0' is true, $(N_{00} - \hat{E}_{00})/n$ will be close to zero as n is large. Deviation of this measure from zero indicates that association exists between T_1 and T_2.

2.2 The proposed test

Under the null hypothesis, $(N_{11}, N_{10}, N_{01}, N_{00})$ has a multinomial distribution, which is conditional on censoring times, with cell probabilities $(P_{11}, P_{10}, P_{01}, P_{00})$ where

$$P_{ab} = \int\int F_1(c_1)^a S_1(c_1)^{1-a} F_2(c_2)^b S_2(c_2)^{1-b} G_n(dc_1, dc_2) \ (a, b = 0, 1).$$

where $G_n(c_1, c_2) = \sum_{k=1}^n I(C_{1k} \le c_1, C_{2k} \le c_2)/n$ is the empirical estimator of $G(c_1, c_2)$. It is easy to see that $\sum_{a=0,1}\sum_{b=0,1} P_{ab} = 1$. If the marginal functions were known, one can test H_0' using the Pearson Chi-squared statistic $\sum_{a=0,1}\sum_{b=0,1}\frac{(N_{ab}-E_{ab})^2}{E_{ab}}$ with three degrees of freedom. Since $E_{ab}(= nP_{ab})$ is unknown, it is natural to use its estimate $\hat{E}_{ab}$ in the test. However, the degrees of freedom is no longer three after replacing E_{ab} by $\hat{E}_{ab}$ because

$$\hat{E}_{00} - N_{00} = N_{10} - \hat{E}_{10} = \hat{E}_{11} - N_{11} = N_{01} - \hat{E}_{01}. \tag{2}$$

To see why (2) is true, notice that

$$\begin{aligned}
&\hat{E}_{00} - N_{00} + \hat{E}_{10} - N_{10} \\
= &\sum_{k=1}^n \hat{S}_1(c_{1k})\hat{S}_2(c_{2k}) - \sum_{k=1}^n I(\delta_{1k} = 0, \delta_{2k} = 0) + \sum_{k=1}^n [1 - \hat{S}_1(c_{1k})]\hat{S}_2(c_{2k}) - \sum_{k=1}^n I(\delta_{1k} = 1, \delta_{2k} = 0) \\
= &\sum_{k=1}^n \hat{S}_2(c_{2k}) - \sum_{k=1}^n I(\delta_{2k} = 0) \\
= &\sum_{k=1}^n [\hat{S}_2(c_{2k}) - 1 + \delta_{2k}] \\
= &\sum_{k=1}^n [\delta_{2k} - \hat{F}_2(c_{2k})].
\end{aligned}$$

The last quantity equals zero because of the self-consistency property of the univariate current status NPMLE estimator (Groeneboom and Weller 1992). Therefore, $\hat{E}_{00} - N_{00} = N_{10} - \hat{E}_{10}$. Similarly, we can show that $\hat{E}_{00} - N_{00} = N_{01} - \hat{E}_{01}$ and $\hat{E}_{11} - N_{11} = N_{10} - \hat{E}_{10}$. Since the degree of freedom reduces to one after replacing E_{ab} by $\hat{E}_{ab}$, we only need to concentrate on one of the four terms, say, $\hat{E}_{00} - N_{00}$.

Notice that $\hat{E}_{00} - N_{00} = \hat{E}_{11} - N_{11}$ is the sum of the differences between $\delta_{1k}\delta_{2k}$ and their estimated conditional expectations $\hat{F}_1(c_{1k})\hat{F}_2(c_{2k})$ under the null hypothesis. Therefore, a significant derivation from zero of $(\hat{E}_{00} - N_{00})^2$ indicates violation of the null hypothesis. Thus, we propose using the test statistic

$$Q = \frac{(N_{00} - \hat{E}_{00})^2}{\widehat{avar}(N_{00} - \hat{E}_{00})},$$

where $\widehat{avar}(N_{00} - \hat{E}_{00})$ is a consistent estimator of the asymptotic variance of $(N_{00} - \hat{E}_{00})$. Appendix 1 shows that under the null hypothesis, $n^{-1/2}(N_{00} - \hat{E}_{00})$ converges in distribution to $N(0, \sigma^2)$ and hence Q converges in distribution to a χ^2 distribution with one degree of freedom if $\hat{\sigma}^2 = \widehat{avar}(N_{00} - \hat{E}_{00})/n$ is a consistent estimator of σ^2. Under the alternative hypothesis, $n^{-1/2}(N_{00} - \hat{E}_{00})$ converges to a non-zero mean normal distribution.

Alternatively, one can construct a test based on an estimate of

$$E[Cov(\delta_1, \delta_2)|C_1, C_2] = E[\{\delta_1 - F_1(C_1)\}\{\delta_2 - F_2(C_2)\}],$$

where $Cov(\delta_1, \delta_2)|C_1, C_2$ is interpreted as the conditional covariance between δ_1 and δ_2 given (C_1, C_2), whose expectation equals $E[F(C_1, C_2) - F_1(C_1)F_2(C_2)]$ and can be estimated by $(N_{11} - \hat{E}_{11})/n$. As mentioned earlier, the above covariance measure reduces to zero under H_0'. Furthermore, it is easy to see that $E[Cov(\delta_1, \delta_2)|C_1, C_2]$ equals

$$-E[Cov(\delta_1, 1-\delta_2)|C_1, C_2] = -E[Cov(1-\delta_1, \delta_2)|C_1, C_2] = E[Cov(1-\delta_1, 1-\delta_2)|C_1, C_2].$$

When the marginal distributions are completely known, the above identity implies that we only need to consider any one of the four covariance measures. When the marginal distributions are estimated by their nonparametric MLEs, such an argument is still true which can also be justified by equation (2).

Explicit estimation of σ^2, however, is very technically involved due to the complexity of the nonparametric MLEs. The bootstrap method provides a convenient numerical solution to obtain a variance estimator. Specifically from the original data, $\{(C_{1i}, C_{2i}, \delta_{1i}, \delta_{2i})\ i = 1, \ldots, n\}$, one can generate a pseudo dataset, $\{(C_{1k}, C_{2k}, \delta_{1k}^*, \delta_{2k}^*)\ k = 1, \ldots, n\}$, where δ_{jk}^* is a Bernoulli random variable with probability $\hat{F}_j(C_{jk})$ $(j = 1, 2)$. The procedure is repeated m times. Let $(N_{00,r}^* - \hat{E}_{00,r}^*)$ be the counterpart of $(N_{00} - \hat{E}_{00})$

for the rth bootstrapped sample. Then $avar(N_{00} - \hat{E}_{00})$ can be estimated by the sample variance of $(N^*_{00,r} - \hat{E}^*_{00,r})$ $(r = 1, \ldots, m)$, that is,

$$n\hat{\sigma}_b^2 = \sum_{r=1}^{m}(N^*_{00,r} - \hat{E}^*_{00,r} - \bar{R}^*)^2/(m-1),$$

where

$$\bar{R}^* = \sum_{r=1}^{m}(N^*_{00,r} - \hat{E}^*_{00,r})/m$$

is the sample mean. As long as $n, m \to \infty$, $\hat{\sigma}_b^2 \to \sigma^2$. The resulting test statistic

$$Q_b = \frac{(N_{00} - \hat{E}_{00})^2}{n\hat{\sigma}_b^2} \tag{3}$$

converges to χ^2 distribution with one degree of freedom under the null hypothesis. Although the bootstrap method for variance estimation is straightforward and asymptotically valid, the power of the resulting test is not satisfactory in our simulation analysis. We provide an explanation of this phenomenon in §4.

To improve the power, we derive an analytic formula given in Appendix 2-4 for variance estimation. The proposed variance estimator in general is complicated. However, when $C_1 = C_2 = C$ which occurs when the paired measurements are taken from the same subjects, the formula can be simplified to

$$\begin{aligned}\hat{\sigma}_p^2 = \;& n^{-1}\textstyle\sum_{k=1}^{n}[\hat{S}_1(C_{1k})\hat{S}_2(C_{2k})\{1 + \hat{S}_1(C_{1k}) + \hat{S}_2(C_{2k}) - 3\hat{S}_1(C_{1k})\hat{S}_2(C_{2k})\} \\ & +(1-\delta_{1k})(1-\delta_{2k})(\hat{E}_{00,(-k)} - \hat{E}_{00})],\end{aligned} \tag{4}$$

where $\hat{E}_{00}$ is defined in equation (1) and $\hat{E}_{00,(-k)}$ is the delete-one-jackknife version of $\hat{E}_{00}$. Specifically $\hat{E}_{00,(-k)}$ is calculated by removing the kth patient in the estimation of marginal survival functions before plugging into equation (1). Accordingly, one can construct the following test statistic,

$$Q_p = \frac{(N_{00} - \hat{E}_{00})^2}{n\hat{\sigma}_p^2}, \tag{5}$$

which also converges to χ^2 distribution with one degree of freedom as $n \to \infty$.

2.3 Finite-sample adjustment for bias

Although the proposed test has nice asymptotic behavior with the regular convergence rate, bias adjustment is useful especially when the sample size is not large. The bias comes from replacing the unknown marginal functions by their nonparametric maximum likelihood estimators in equations (1) and (4). Let $n^{-1/2}(N_{00}-\hat{E}_{00}) = n^{-1/2}(N_{00}-E_{00})+B_{1n}$ and $\hat{\sigma}_p^2 = \sigma^2 + B_{2n}$. As $n \to \infty$, B_{1n} and B_{2n} will shrink to zero but, when the sample size is not large, they are not ignorable and will result in inaccurate type I error (usually higher than the nominal level).

To improve the finite-sample performance, we can estimate B_{1n} and B_{2n} using the bootstrap method and then eliminate their effect in the testing procedure. Specifically for a bootstrapped sample, $\{(C_{1k}, C_{2k}, \delta^*_{1k}, \delta^*_{2k}) \;\; k = 1, \ldots, n\}$, one can compute the statistics, $U^* = n^{-1/2}(N_{00}-\hat{E}^*_{00})$ and $(\hat{\sigma}^*_p)^2$. The procedure is then repeated m times and let $\hat{U}_b$ and $\hat{\sigma}^2_{pb}$ be the average of the two estimators based on m bootstrap samples. Let $\hat{B}_{1n} = \hat{U}_b - n^{-1/2}(N_{00} - \hat{E}_{00})$ and $\hat{B}_{2n} = \hat{\sigma}^2_{pb} - \hat{\sigma}^2_p$. For relatively large m, say $m = 500$, $\hat{B}_{1n}$ and $\hat{B}_{2n}$ would provide good approximation of the true bias terms. Subtracting these estimated bias terms from the test statistics (3) and (5), we have the bias adjusted versions of test statistics

$$Q_{b(a)} = \frac{(N_{00} - \hat{E}_{00} - n^{1/2}\hat{B}_{1n})^2}{n\hat{\sigma}^2_b} \tag{6}$$

$$Q_{p(a)} = \frac{(N_{00} - \hat{E}_{00} - n^{1/2}\hat{B}_{1n})^2}{n(\hat{\sigma}^2_p - \hat{B}_{2n})}. \tag{7}$$

We will see in Section 3 that the adjusted tests perform much better in finite samples than the unadjusted versions.

2.4 Weight adjustment

Power and efficiency of the test may be improved by including a weight function in the statistic. Notice that one can write $N_{00}-\hat{E}_{00} = \sum_k (N_{00,k}-\hat{E}_{00,k})$. Therefore the modified test is related to the following statistic, $\sum_k W_k(N_{00,k} - \hat{E}_{00,k})$, where $W_k = w(C_{1k}, C_{2k})$ is the weight assigned to the kth subject according to the observed monitoring times. One can write

$$\begin{aligned} Z_W &= n^{-1/2} \sum_k W_k(N_{00,k} - \hat{E}_{00,k}) \\ &= n^{1/2} \int_{c_1} \int_{c_2} w(c_1, c_2) \left\{ N_{00}(c_1, c_2) - \hat{S}_1(c_1)\hat{S}_2(c_2) \right\} G_n(dc_1, dc_2), \end{aligned}$$

where $W_k = w(C_{1k}, C_{2k})$. In Appendix 5, we show that under independence, Z_W converges to a zero-mean normal random variable with variance σ_W^2 given in (A8) which can also be estimated using the bootstrap method. An analytic variance estimator can be easily obtained by modifying $\hat{\sigma}_p^2$. Denote $\hat{\sigma}_W^2$ as a consistent estimator σ_W^2. A weighted test statistic is of the form, $Q_W = Z_W^2/\hat{\sigma}_W^2$, which converges to χ_1^2 as n goes to infinity. Note that when $w(c_1, c_2) \equiv 1$ for all (c_1, c_2), Q_W reduces to Q_p.

The choice of a good weight function depends on the dependence structure under the alternative hypothesis. Following Anderson, Louis, Holm & Harvald (1992), the bivariate survival function of (T_1, T_2) can be expressed as

$$S(t_1, t_2) = S_1(t_1)S_2(t_2)e^{-A(t_1,t_2)},$$

where $A(t_1, t_2)$ measures the dependence structure. Note that under independence, $A(t_1, t_2) = 0$. For other related association measures, refer to Dabrowska (1988) and Prentice & Cai (1992). Our original objective is to choose a weight function which, under the local alternative H_α: $A(t_1, t_2) = n^{-1/2}a(t_1, t_2) + o_p(n^{-1/2})$, maximizes

$$\frac{\left|E\{\sum_k W_k(N_{00,k} - \hat{E}_{00,k})\}\right|^2}{avar\{\sum_k W_k(N_{00,k} - \hat{E}_{00,k})\}}.$$

Note that we only consider the local optimality condition here because the behavior of the statistic is more important in the region near independence and also the analysis can be simplified. Due to the complexity of the plugged-in nonparametric maximum likelihood estimators, we derive the local optimal weight function, denoted as W^*, by maximizing

$$\frac{|E\{\sum_k W_k(N_{00,k} - E_{00,k})\}|^2}{avar\{\sum_k W_k(N_{00,k} - E_{00,k})\}}.$$

We find that the local optimal weight function $w^*(t_1, t_2)$ is proportional to

$$\frac{|a(t_1, t_2)|}{1 - S_1(t_1)S_2(t_2)}.$$

The proof is given in Appendix 6. Because $w^*(t_1, t_2)$ depends on unknown quantities, we can replace it by its estimator denoted as $\hat{w}^*(t_1, t_2)$. Then the test statistic Q_{W^*} can be modified as $Q_{\hat{W}^*} = Z_{\hat{W}^*}^2/\hat{\sigma}_{\hat{W}^*}^2$. When $\hat{w}^*(t_1, t_2)$ converges uniformly to $w^*(t_1, t_2)$, the asymptotic distribution of $Q_{\hat{W}^*}$ is the same as Q_{W^*}. Note that for finite samples, $Q_{\hat{W}^*}$ may not have the advantage of variance reduction due to extra estimation of the marginal functions.

Now we calculate $w^*(t_1, t_2)$ for the Clayton and Frank models (Clayton, 1978; Genest, 1987). The results will be used in the simulation analysis.

Example 1: Clayton's family. The joint survival function is given by

$$S(t_1, t_2) = \{S_1(t_1)^{1-\alpha} + S_2(t_2)^{1-\alpha} - 1\}^{1/(1-\alpha)} \quad (\alpha > 1),$$

where α is an association parameter, related to Kendall's tau (τ), such that $\tau = (\alpha - 1)/(\alpha + 1)$. Let $\alpha = 1 + \delta$ and it follows that

$$\begin{aligned} A(t_1, t_2) &= \frac{1}{\alpha - 1} log\{S_1(t_1)^{\alpha-1} + S_2(t_2)^{\alpha-1} - S_1(t_1)^{\alpha-1} S_2(t_2)^{\alpha-1}\} \\ &= -2\delta log\{S_1(t_1)\} log\{S_2(t_2)\} + o_p(\delta). \end{aligned}$$

Thus for Clayton's model, $a(t_1, t_2) \propto log\{S_1(t_1)\} log\{S_2(t_2)\}$, and hence

$$w^*(t_1, t_2) = \frac{log\{S_1(t_1)\} log\{S_2(t_2)\}}{1 - S_1(t_1) S_2(t_2)}. \tag{8}$$

Example 2: Frank's family. The joint survival function is given by

$$S(t_1, t_2) = log_\alpha\{1 + \frac{(\alpha^{S_1(t_1)} - 1)(\alpha^{S_2(t_2)} - 1)}{\alpha - 1}\} \quad (\alpha > 0).$$

Note that (T_1, T_2) are positively associated when $\alpha < 1$, negatively associated when $\alpha > 1$, and independent when $\alpha \to 1$. Let $\alpha = 1 + \delta$ and it follows that

$$\begin{aligned} A(t_1, t_2) &= -log \left[\frac{log_\alpha\{1 + \frac{(\alpha^{S_1(t_1)} - 1)(\alpha^{S_2(t_2)} - 1)}{\alpha - 1}\}}{S_1(t_1) S_2(t_2)} \right] \\ &= \delta\{1 - S_1(t_1)\}\{1 - S_2(t_2)\}/2 + o_p(\delta). \end{aligned}$$

Thus $a(t_1, t_2) \propto \{1 - S_1(t_1)\}\{1 - S_2(t_2)\}$, and hence for Frank's family,

$$w^*(t_1, t_2) = \frac{\{1 - S_1(t_1)\}\{1 - S_2(t_2)\}}{1 - S_1(t_1) S_2(t_2)}. \tag{9}$$

3 Simulation analysis

A series of simulations were carried out to examine finite sample performance of the proposed tests. The bivariate failure times (T_1, T_2) were generated from the Clayton family using the algorithm in Prentice & Cai (1992) and the Frank family using the algorithm in Genest (1987). We first generated $C_1 = C_2$ from uniform distributions. Performance of the test was measured by the empirical power, based on 4000 runs, which is the relative frequency that the test rejected the null hypothesis at the 0.05 nominal level. We investigated the power behavior under the combination of 10 dependence levels with τ ranging from 0 to 0.45, three prevalence levels (P.L. $\approx$ 0.2, 0.5, 0.8) and two sample sizes ($n = 200,\ 400$). The prevalence level is defined as the expected proportion of observations that reports failure occurrence, that is $\mathrm{pr}(\delta_{1k} = 1) = \mathrm{pr}(\delta_{2k} = 1)$.

The test statistics (3) and (5) without bias adjustment produced incorrect type I errors, many of which even exceed 0.1. Since they are not reliable tests for small samples (i.e. $n = 200, 400$), the data is not shown here. The results of the bias-adjusted tests $Q_{b(a)}$ and $Q_{p(a)}$ in equations (6) and (7) are shown in Table 1 and 2, respectively.

[**Insert Table 1 and Table 2**]

Table 1 shows the results using the bootstrap variance estimator $Q_{b(a)}$. The type I errors for the unweighted tests are close to the 0.05 nominal level. The weighted tests also have type I errors close to 0.05 in most cases. Thus, the test is valid for the two chosen sample sizes. The power increases as the association, measured by τ, becomes higher. However, it looks like the weight adjustment does not improve the power. Recall that the bootstrap estimation is conducted under the assumption of independence while the two failure times are correlated under the alternative hypothesis. The variance estimator obtained using the bootstrap method may be overestimated and hence it offsets the power gain.

The results for $Q_{p(a)}$ using the analytic variance estimator are summarized in Table 2. Although the power is generally higher than that using $Q_{b(a)}$, the type I error seems less accurate in some cases. The unweighted test has the correct type I error except for the case of $n = 200$ and at an 80% prevalence rate. The type I errors for the weighted tests get closer to the nominal level when the sample size increases or prevalence level increases. On the other hand, the weighted tests perform poorly under small sample sizes and low prevalence rates. Now we explain how the prevalence level affects the finite sample performance of the weighted tests. Let us first examine the optimal weight for the Clayton model. Equation (8) suggests to assign larger weights to the tail region of (T_1, T_2). However when $\mathrm{pr}(T_k \leq C_k)$ $(k = 1, 2)$ are small, most observations tend to have big values of $S_j(c_j)$ $(j = 1, 2)$, which are assigned with lower weights. This would reduce the effective sample size and hence the asymptotics kick in more slowly than in the unweighted test. When assuming the Frank model, equation(9) also suggests to assign larger weights to the tail region of (T_1, T_2). Hence, when the prevalence level is low, the weight adjustment also offsets the effective sample size. Nevertheless since the Frank weights are bounded between 0 and 1, the test using Frank's weight produces more accurate type I error than that using Clayton's weight when the correct model is assumed. Generally, under the same sample size, the power is highest when the prevalence rate is around 50%. This is because the effective sample size is larger when the observed failures and survivals are more balanced.

Now we examine whether weight adjustment does indeed improve the local power near the independence. As intended from the theoretical deduction, the weighted tests in most cases have higher power than the unweighted version when $\tau < 0.10$. However, the theory does not ensure that the weights are optimal when association is stronger. When true level of association is moderate or strong, model mis-specification may have a substantially negative effect. For example when $n = 200$, prevalence rate 80% and the Frank's model is mis-specified as the Clayton model with $\tau = 0.3$, the power of the weighted test is only 49% while that of the unweighted test is over 85%.

The above analysis indicates that the unweighted test based on $Q_{p(a)}$ is an accurate and safe choice. It does not rely on any model assumption, and its performance in all cases is satisfactory in terms of the power behavior. The weighted version of $Q_{p(a)}$ may be a consideration only if the sample size is large and some prior information about the model and weak association is available.

The proposed tests were also evaluated under unequal monitoring times. Specifically we set $C_2 = C_1 + 0.1$. The results are similar to those in Table 1 and 2, and hence are omitted.

The effect of the censoring mechanism on the power of the unweighted test is governed by the magnitude of the conditional covariance measure

$$\left[\int\int \{\mathrm{pr}(T_1 > c_1, T_2 > c_2) - S_1(c_1)S_2(c_2)\}\, G(dc_1, dc_2)\right]^2 .$$

Obviously the power is affected by the underlying dependence structure, the censoring pattern and their joint effect. To improve the power by controlling the censoring scheme, the researcher should sample more observations which give a larger value of $|\mathrm{pr}(T_1 > c_1, T_2 > c_2) - S_1(c_1)S_2(c_2)|$. However the suggestion may not be practical since testing independence is usually the first step of analysis.

4 Data analysis

We applied the proposed methodology to analyze a community-based study of cardiovascular diseases conducted from 1991 to 1993 in Taiwan. There were 6314 participants in the study, including 2904 males and 3410 females. The data consisted of measurements of the participants' current age at the time of study and the prevalence indicators of three diseases, namely diabetes mellitus, hypercholesterolemia and hypertension. Denote (T_1, T_2, T_3) as the onset age of diabetes mellitus, hypercholesterolemia and hypertension respectively and C as the subject's age at the monitoring time. Therefore the data are of the form, $(C, \delta_1, \delta_2, \delta_3)$ where $\delta_j = I(T_j \leq C)$ $(j = 1, 2, 3)$. It should be mentioned that we used this example only for illustrative purposes because the prevalence of the three cardiovascular diseases were determined via participants' interviews, health examinations or previous medical history rather than based on formal medical diagnosis. For more detailed description of the data, please refer to Wang & Ding (2000).

[**Insert Table 3**]

Table 3 summarizes the results for testing pairwise independence of (T_1, T_2), (T_1, T_3) and (T_2, T_3). The associations between the onset ages of diabetes mellitus (T_1) and the other two diseases, namely hypercholesterolemia and hypertension, are both very strong with $p-$value close to zero. The association between T_2 and T_3 is significant at 0.05 level but not at 0.01 level.

It is interesting to note that Wang & Ding (2000) assumed that the pairwise dependence structures of the three diseases all follow Clayton's models and then estimated the association parameters. The estimated values of corresponding Kendall's tau between T_i

and T_j, denoted as $\hat{\tau}_{ij}$, are $\hat{\tau}_{12} = 0.304$, $\hat{\tau}_{13} = 0.128$ and $\hat{\tau}_{23} = 0.082$. The corresponding 95% confidence intervals are (0.210, 0.378), (-0.005, 0.230) and (-0.019, 0.165).

5 Concluding Remarks

We have developed a nonparametric method to test independence between two failure time variables when only current status data is available. Since the true failure times (T_1, T_2) are never observed, this inference problem is harder than it looks at first sight. Specifically it is not easy to test independence between (T_1, T_2) without being affected by the distribution of (C_1, C_2), which are often correlated. The proposed testing procedures use only statistics conditional on the censoring times to avoid making assumptions on the censoring distribution. It is possible that the proposed test correctly accepts H_0' while H_0 is false since current status data does not provide information to identify such a condition. Nevertheless, as long as C_1 and C_2 are continuous and their support is large enough to cover the distribution of (T_1, T_2), the above situation is not likely to happen.

Sometimes practitioners may want to apply nonparametric methods, such as independence, rank and permutation tests, which are popular in cross-sectional analysis, to bivariate current status data. Here we discuss why these methods are not applicable. First of all, we discuss an independence test based on the merged two-by-two table with entries N_{00}, N_{01}, N_{10} and N_{11} without using the information of individual censoring times. Independence between the columns and rows implies that

$$\{\int S_1(c_1)dG_1(c_1)\}\{\int S_2(c_2)dG_2(c_2)\} = \int\int S(c_1, c_2)G(dc_1, dc_2),$$

where $G(c_1, c_2)$ is the distribution function of (C_1, C_2) with marginal distribution functions denoted by $G_1(c_1)$ and $G_2(c_2)$. It is easy to see that the above equation holds when not only $S_1(c_1)S_2(c_2) = S(c_1, c_2)$ but also $G_1(c_1)G_2(c_2) = G(c_1, c_2)$. Therefore this test would be valid for testing independence between T_1 and T_2 only under the unrealistic assumption that C_1 and C_2 are also independent.

Now we discuss the validity of some permutation tests. Let $G_n(c_1, c_2)$, $G_{1n}(c_1)$ and $G_{2n}(c_2)$ denote the corresponding empirical distribution functions. One possible alternative is to perform a permutation test by randomly pairing up (C_{1k}, δ_{1k}) with (C_{2j}, δ_{2j}). Such a procedure that breaks up observations of (C_1, C_2) would make the resampled censoring distribution to be $G_{1n}(c_1)G_{2n}(c_2)$ which converges to $G_1(c_1)G_2(c_2)$ instead of $G(c_1, c_2)$. Again, the resulting test would be valid only if C_1 and C_2 are also independent. A second possible mistake is to run a permutation test by keeping (C_{1k}, C_{2k}) together and only randomly pairing up δ_{1i} with δ_{2j}. Notice that $\Pr(\delta_{1k} = 1|C_{1k}) = F_1(C_{1k})$ while $\Pr(\delta_{1i} = 1|C_{1i}) = F_1(C_{1i})$. Hence they can not be exchanged under the null hypothesis. A third possibility is running a nonparametric bootstrap without replacement. That is, to keep $(C_{1k}, \delta_{1k}, C_{2k}, \delta_{2k})$ together and just permute among the individuals. Although this is a valid permutation, which retains the distribution of the original data, it does not provide any information for the variation.

The main purpose of weight adjustment is to improve the power when the true association is near independence. However the weighted test requires a high prevalence rate and large sample size to assure the asymptotic validity. It is also susceptible to model mis-specification and may not offer any advantages when the association is high. Therefore we would suggest the unweighted test which performs quite well in almost all the simulated cases.

The proposed methodology can be easily extended to adjust for the effects of covariates. The assumption of strict independence between the failure times and the monitoring times may be relaxed if their dependence can be accounted for by observed covariate, say Z, such that $T_j \perp C_j|Z$ $(j = 1, 2)$. Accordingly the two-by-two tables should be constructed based on distinct observed values of (C_1, C_2, Z). If Z also affects the marginal distributions, the nonpamatric MLEs $\hat{S}_j(c_j)$ $(j = 1, 2)$ should be replaced by appropriate estimators of $S_j(c_j|Z)$. A candidate of such an estimator is the one proposed by van der Laan and Robins (1998) under the proportional hazard assumption. Then the adjusted estimates of $S_j(c_j|Z)$ $(j = 1, 2)$ are used in estimating the expected counts $\hat{E}_{ab}$ in $Q_{b(a)}$ or $Q_{p(a)}$.

In this article we concentrate on the bivariate case. The tests can be easily generalized to multivariate data of dimensions higher than two. Using similar ideas, we can construct $2 \times 2 \times ... \times 2$ tables and use the test statistic $Q = (N_{00...0} - \hat{E}_{00...0})^2/n\hat{\sigma}^2$. The extension of the variance estimator $\hat{\sigma}^2$ to multivariate data is straightforward.

APPENDICES

APPENDIX 1: *Asymptotic Normality of* $n^{-1/2}(N_{00} - \hat{E}_{00})$

Note that $N_{00} - \hat{E}_{00} = (N_{00} - E_{00}) + (E_{00} - \hat{E}_{00})$. One can write

$$E_{00} = n \int\int S_1(c_1)S_2(c_2)G_n(dc_1, dc_2)$$

and similar expressions apply to E_{10}, E_{01} and E_{00}. The first term $(N_{00} - E_{00})$ can be written explicitly as

$$\sum_{k=1}^{n} \{I(T_{1k} > C_{1k}, T_{2k} > C_{2k}) - S_1(C_{1k})S_2(C_{2k})\}.$$

Hence by the central limit theorem $n^{-1/2}(N_{00}-E_{00})$ converges in distribution to $N(\mu, \sigma_1^2)$, where $\mu = E[S(C_1, C_2) - S_1(C_1)S_2(C_2)]$ and σ_1^2 is the unconditional variance of $I(T_1 > C_1, T_2 > C_2) - S_1(C_1)S_2(C_2)$. Under the null hypothesis, $\mu = 0$ and $\sigma_1^2 = E[S_1(C_1)S_2(C_2)(1 - S_1(C_1)S_2(C_2))]$.

Now we prove asymptotic normality of the second term $n^{-1/2}(E_{00} - \hat{E}_{00})$. By uniform consistency of the marginal nonparametric maximum likelihood estimators and asymptotic properties of an empirical process, it follows that

$$\begin{aligned}
& n^{-1/2}(\hat{E}_{00} - E_{00}) \\
= \; & n^{-1/2} \sum_{k=1}^{n} \{\hat{S}_1(C_{1k})\hat{S}_2(C_{2k}) - S_1(C_{1k})S_2(C_{2k})\} \\
= \; & n^{1/2} \int\int \{\hat{S}_1(c_1)\hat{S}_2(c_2) - S_1(c_1)S_2(c_2)\}G_n(dc_1, dc_2) \\
= \; & n^{1/2} \int\int S_2(c_2)\{\hat{S}_1(c_1) - S_1(c_1)\}G(dc_1, dc_2) + n^{1/2} \int\int S_1(c_1)\{\hat{S}_2(c_2) - S_2(c_2)\}G(dc_1, dc_2) \\
& + rem_n.
\end{aligned}$$

We shall show that the first two terms converges to normal distribution where the remainder term rem_n is of order $o_p(1)$.

First, the remainder term

$$\begin{aligned} rem_n &= n^{1/2}\int\int S_2(c_2)\{\hat{S}_1(c_1)-S_1(c_1)\}[G_n(dc_1,dc_2)-G(dc_1,dc_2)] \\ &\quad +n^{1/2}\int\int S_1(c_1)\{\hat{S}_2(c_2)-S_2(c_2)\}[G_n(dc_1,dc_2)-G(dc_1,dc_2)] \\ &\quad +n^{1/2}\int\int\{\hat{S}_1(c_1)-S_1(c_1)\}\{\hat{S}_2(c_2)-S_2(c_2)\}G_n(dc_1,dc_2) \\ &= I_{1n}+I_{2n}+I_{3n} \end{aligned}$$

Since $\hat{S}_j(c_j)-S_j(c_j), j=1,2$ are of order $O_p(n^{-1/3})$ (Groeneboom and Wellner 1992), the last term $I_{3n}=O_p(n^{1/2-1/3-1/3})=O_p(n^{-1/6})=o_p(1)$. To show the first two terms are of order $o_p(1)$, notice that each integrand only involves one-dimensional empirical survival function $\hat{S}_j$. It is essentially the same proof as in the univariate current status data case, e.g., we can apply arguments similar to those on page 160-161 of Huang & Wellner (1995). Let $\mathcal{S}=\{$S: S is a one-dimensional survival function$\}$, and consider the class of functions $\mathcal{F}=\{S_1(x)(S(y)-S_2(y)):S\in\mathcal{S}\}$. First, uniform entropy for $\mathcal{S}$ is bounded by $K(1/\epsilon)^\lambda, \lambda>1$ because it is contained in the convex hull of the VC graph class of right half lines (Dudley 1987). Since for any $S_{(1)}, S_{(2)}\in\mathcal{S}$,

$$|S_1(x)(S_{(1)}(y)-S_2(y))-S_1(x)(S_{(2)}(y)-S_2(y))|\le|S_{(1)}(y)-S_{(2)}(y)|,$$

the uniform entropy for $\mathcal{F}$ is also bounded by the bound for $\mathcal{S}$, $K(1/\epsilon)^\lambda, \lambda>1$. Therefore, $\mathcal{F}$ is a G-Donsker class by Pollard's theorem (e.g., Dudley 1987). Then we apply Theorem 1.1 of Sheehy and Wellner (1992) to ensure the uniform asymptotic equicontinuity of the empirical process over $\mathcal{F}$, which then implies that the first term $I_{1n}=o_p(1)$. The second term $I_{2n}=o_p(1)$ is proved the same way by symmetry, and we have $rem_n=o_p(1)$.

Next we will show that the first two terms in (A1) are asymptotically normal. Without loss of generality, assume that $G(c_1,c_2)$ is differentiable with respect to both arguments and $g(c_1,c_2)=\partial^2G(c_1,c_2)/\partial c_1\partial c_2$. Denote

$$A_1(x)=\int_{c_1=0}^{x}\left[\int_{c_2=0}^{\infty}S_2(c_2)g(c_1,c_2)dc_2\right]dc_1=\int_{c_1=0}^{x}a_1(c_1)dc_1,$$

$$A_2(x)=\int_{c_2=0}^{x}\left[\int_{c_1=0}^{\infty}S_1(c_1)g(c_1,c_2)dc_1\right]dc_2=\int_{c_2=0}^{x}a_2(c_2)dc_2.$$

Performing integration by parts, it follows that

$$n^{-1/2}\left\{\hat{E}_{00}-E_{00}\right\} = -n^{1/2}\int_{c_1=0}^{\infty} A_1(c_1)d(\hat{S}_1-S_1)(c_1) - n^{1/2}\int_{c_2=0}^{\infty} A_2(c_2)d(\hat{S}_2-S_2)(c_2)+o_p(1)$$

$$= -n^{1/2}\{\nu_1(\hat{S}_1)-\nu_1(S_1)\} - n^{1/2}\{\nu_2(\hat{S}_2)-\nu_2(S_2)\}+o_p(1),$$

where

$$\nu_1(S_1)=\int A_1(c)dS_1(c), \quad \nu_2(S_2)=\int A_2(c)dS_2(c).$$

Notice that $\nu_j(S_j)$ is a "smooth" functional of S_j satisfying the conditions in Theorem 5.1 (Huang & Wellner, 1995). Therefore the asymptotic normality of $n^{1/2}\{\nu_j(\hat{S}_j)-\nu_j(S_j)\}$ $(j=1,2)$ and hence that of $n^{-1/2}(\hat{E}_{00}-E_{00})$ is established.

APPENDIX 2: *Estimation of* σ^2

We have shown that $n^{-1/2}(N_{00}-\hat{E}_{00})$ converges to $N(0,\sigma^2)$. To estimate the asymptotic variance σ^2, we notice that $\sigma^2=\sigma_1^2+\sigma_2^2+2\sigma_{12}$, where $\sigma_1^2=avar\{n^{-1/2}(N_{00}-E_{00})\}$, $\sigma_2^2=avar\{n^{-1/2}(E_{00}-\hat{E}_{00})\}$ and $\sigma_{12}=acov\{n^{-1/2}(N_{00}-E_{00}),n^{-1/2}(E_{00}-\hat{E}_{00})\}$. Since $\sigma_1^2=E[S_1(C_1)S_2(C_2)\{1-S_1(C_1)S_2(C_2)\}]$, it can be estimated consistently by

$$\hat{\sigma}_1^2=n^{-1}\sum_{k=1}^{n}\hat{S}_1(C_{1k})\hat{S}_2(C_{2k})\{1-\hat{S}_1(C_{1k})\hat{S}_2(C_{2k})\}. \tag{A1}$$

We show (in Appendix 3) that

$$\sigma_2^2=\int\int\{F_1(c_1)\frac{a_1(c_1)}{g_1(c_1)}+F_2(c_2)\frac{a_2(c_2)}{g_2(c_2)}\}S_1(c_1)S_2(c_2)G(dc_1,dc_2). \tag{A2}$$

We may estimate σ_2^2 analytically based on the above expression. But the estimator $\hat{\sigma}_2^2$ in general is very complicated. However, the estimation can be simplified if the relationship between C_1 and C_2 is specified. For example, in the most common case when the measurements are taken from the same subjects, we have $C_{1k}=C_{2k}=C_k$ $(k=1,...,n)$. In such a case, the above expression is simplified to

$$\sigma_2^2=\int\int\{S_1(c_1)+S_2(c_2)-2S_1(C_k)S_2(C_k)\}S_1(c_1)S_2(c_2)G(dc_1,dc_2),$$

which can be estimated consistently by

$$\tilde{\sigma}_2^2=\frac{1}{n}\sum_{k=1}^{n}\hat{S}_1(C_k)\hat{S}_2(C_k)\{\hat{S}_1(C_k)+\hat{S}_2(C_k)-2\hat{S}_1(C_k)\hat{S}_2(C_k)\}. \tag{A3}$$

Denote $\hat{E}_{00,(-k)}$ as the delete-one version of $\hat{E}_{00}$ calculated after deleting subject k from the sample. We show (in Appendix 4) that σ_{12} can be consistently estimated by

$$\hat{\sigma}_{12} = \frac{1}{n}\sum_{k=1}^{n}(1-\delta_{1k})(1-\delta_{2k})(\hat{E}_{00,(-k)} - \hat{E}_{00}). \tag{A4}$$

Hence the asymptotic variance σ^2 is estimated consistently by $\hat{\sigma}^2 = \hat{\sigma}_1^2 + \hat{\sigma}_2^2 + 2\hat{\sigma}_{12}$, which is simplified to equation (4) when $C_1 = C_2$.

In finite samples, the above estimator tends to overestimate the true variance. The bias comes from $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ given in equation (A1) and (A3). To see this, we can write

$$\begin{aligned}
& \sigma_1^2 - \hat{\sigma}_1^2 \\
= \; & -\tfrac{1}{n}\textstyle\sum_{i=1}^{n}\{\hat{S}_1(C_{1i})\hat{S}_2(C_{2i}) - S_1(C_{1i})S_2(C_{2i})\}\{1 - \hat{S}_1(C_{1i})\hat{S}_2(C_{2i}) - S_1(C_{1i})S_2(C_{2i})\} \\
= \; & -\tfrac{1}{n}\textstyle\sum_{i=1}^{n}\{\hat{S}_1(C_{1i})\hat{S}_2(C_{2i}) - S_1(C_{1i})S_2(C_{2i})\}\{1 - 2S_1(C_{1i})S_2(C_{2i})\} \\
& +\tfrac{1}{n}\textstyle\sum_{i=1}^{n}\{\hat{S}_1(C_{1i})\hat{S}_2(C_{2i}) - S_1(C_{1i})S_2(C_{2i})\}^2 \\
= \; & r_{1n} + r_{2n}.
\end{aligned}$$

The first term r_{1n} asymptotically is of mean zero and order $O_p(n^{-1/2})$. The second term r_{2n} is of smaller order $O_p(n^{-2/3})$ but always positive. Similarly, $\hat{\sigma}_2^2$ underestimates σ_2^2. Hence the finite sample adjustments for bias in Section 2.3 is necessary.

To save computing time, the bias adjustment can be applied only to $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ rather than the whole $\hat{\sigma}^2$.

APPENDIX 3: *Estimation of* $\sigma_2^2 = avar(n^{-1/2}\{E_{00} - \hat{E}_{00}\})$

Applying Theorem 5.1 (Huang & Wellner, 1995), we can further write

$$n^{-1/2}\left\{\hat{E}_{00} - E_{00}\right\} = n^{-1/2}\sum_{k=1}^{n}[\{\delta_{1k} - F_1(C_{1k})\}\frac{a_1(C_{1k})}{g_1(C_{1k})} + \{\delta_{2k} - F_2(C_{2k})\}\frac{a_2(C_{2k})}{g_2(C_{2k})}] + o_p(1),$$

where $g_1(\cdot)$ and $g_2(\cdot)$ are the marginal density functions of C_1 and C_2, respectively. It follows that

$$\begin{aligned}
\sigma_2^2 &= var[\{\delta_{11} - F_1(C_{11})\}\frac{a_1(C_{11})}{g_1(C_{11})} + \{\delta_{21} - F_2(C_{21})\}\frac{a_2(C_{21})}{g_2(C_{21})}] \\
&= \int F_1(c_1)S_1(c_1)\frac{[a_1(c_1)]^2}{g_1(c_1)}dc_1 + \int F_2(c_2)S_2(c_2)\frac{[a_2(c_2)]^2}{g_2(c_2)}dc_2
\end{aligned}$$

$$= \int\int F_1(c_1)S_1(c_1)\frac{a_1(c_1)}{g_1(c_1)}S_2(c_2)g(c_1,c_2)dc_1dc_2 + \int\int F_2(c_2)S_2(c_2)\frac{a_2(c_2)}{g_2(c_2)}S_1(c_1)g(c_1,c_2)dc_1dc_2$$

$$= \int\int \{F_1(c_1)\frac{a_1(c_1)}{g_1(c_1)} + F_2(c_2)\frac{a_2(c_2)}{g_2(c_2)}\}S_1(c_1)S_2(c_2)G(dc_1,dc_2).$$

Based on the above expression, one may estimate σ_2^2 analytically: estimate $G(\cdot,\cdot)$ by the empirical distribution $G_n(\cdot,\cdot)$; estimate the marginal functions by the corresponding nonparametric maximum likelihood estimators; and estimate $\frac{a_j(c)}{g_j(c)}$ $(j = 1, 2)$ by some nonparametric methods. The last step involves estimating a ratio of density functions nonparametrically. We suggest applying the kernel method to estimate each component. Specifically

$$\frac{\hat{a}_1(c_1)}{\hat{g}_1(c_1)} = \frac{\sum_{k=1}^n \hat{S}_2(C_{2k})K\{(C_{1k}-c_1)/h_1\}}{\sum_{k=1}^n K\{(C_{1k}-c_1)/h_1\}},$$

where the kernel function, $K(\cdot)$, is a symmetric density function and h is the bandwidth, controlling the size of the local neighborhood. Bandwidth selection is often crucial for the kernel method. However the estimation of $\frac{a_j(c)}{g_j(c)}$ $(j = 1, 2)$ is not the ultimate goal, but only to provide a consistent plugged-in estimator. Hence, we do not have to find the optimal kernel and bandwidth which is a difficult topic by itself since there is no obvious optimal criterion here. For computation simplicity, we can take the linear kernel $K(x) = max(1-|x|, 0)$ and bandwidth $h_1 = n^{-1/5}s_1$ (of the conventional optimal order), where s_1 denotes the sample standard deviation of the censoring times $C_{1k}, k = 1, ..., n$. Similarly, the bandwidth $h_2 = n^{-1/5}s_2$. For a thorough discussion on kernel methods, refer to Wand & Jones (1995).

Now a consistent estimator of σ_2^2 is given by

$$\hat{\sigma}_2^2 = \frac{1}{n}\sum_{k=1}^n \hat{S}_1(C_{1k})\hat{S}_2(C_{2k})[\hat{F}_1(C_{1k})\frac{\sum_{j=1}^n \hat{S}_2(C_{2j})K\{(C_{1j}-C_{1k})/h_1\}}{\sum_{j=1}^n K\{(C_{1j}-C_{1k})/h_1\}}$$

$$+\hat{F}_2(C_{2k})\frac{\sum_{j=1}^n \hat{S}_1(C_{1j})K\{(C_{2j}-C_{2k})/h_2\}}{\sum_{j=1}^n K\{(C_{2j}-C_{2k})/h_2\}}]. \qquad \text{(A5)}$$

The above estimator is still rather complicated, and we recommend using simpler estimators in practice based on knowledge of the censoring times. For example, when

$C_1 = C_2 = C$, $\frac{a_1(c)}{g_1(c)} = S_2(c)$ and $\frac{a_2(c)}{g_2(c)} = S_1(c)$. Therefore σ_2^2 can be estimated consistently by

$$\begin{aligned} \tilde{\sigma}_2^2 &= \frac{1}{n}\sum_{k=1}^{n} \hat{S}_1(C_k)\hat{S}_2(C_k)\{\hat{F}_1(C_k)\hat{S}_2(C_k) + \hat{F}_2(C_k)\hat{S}_1(C_k)\} \\ &= \frac{1}{n}\sum_{k=1}^{n} \hat{S}_1(C_k)\hat{S}_2(C_k)\{\hat{S}_1(C_k) + \hat{S}_2(C_k) - 2\hat{S}_1(C_k)\hat{S}_2(C_k)\}, \qquad \text{(A6)} \end{aligned}$$

which reduces to the form in equation (A3).

In another example, if $C_1 \perp C_2$, we have $\frac{a_1(c)}{g_1(c)} = E_G[S_2]$ and $\frac{a_2(c)}{g_2(c)} = E_G[S_1]$. In such a case σ_2^2 can be estimated consistently by

$$\breve{\sigma}_2^2 = \frac{1}{n}\sum_{k=1}^{n} \hat{S}_1(C_{1k})\hat{S}_2(C_{1k})\{\hat{F}_1(C_{1k})\bar{S}_2 + \hat{F}_2(C_{2k})\bar{S}_1\}, \qquad \text{(A7)}$$

where

$$\bar{S}_i = \frac{1}{n}\sum_{j=1}^{n} \hat{S}_i(C_{ij}).$$

Appendix 4: *Estimation of* $\sigma_{12} = n^{-1} acov(N_{00} - E_{00}, E_{00} - \hat{E}_{00})$

Explicit expression of σ_{12} is difficult to obtain due to the complexity of the plugged-in nonparametric maximum likelihood estimators. In addition to the bootstrap approach, we provide another estimation method using the delete-one jackknife method. Conditioning on the monitoring times, $cov(N_{00} - E_{00}, E_{00}) = 0$ and hence $cov(N_{00} - E_{00}, E_{00} - \hat{E}_{00}) = cov(N_{00} - E_{00}, -\hat{E}_{00})$. Notice that

$$cov(N_{00} - E_{00}, -\hat{E}_{00}) = ncov(\{1 - \delta_{1k}\}\{1 - \delta_{2k}\} - S_1(C_{1k})S_2(C_{2k}), -\hat{E}_{00}).$$

Denote $\hat{E}_{00,(-k)}$ as the delete-one version of $\hat{E}_{00}$ calculated after deleting subject k from the sample. It is obvious that $cov(\{-\delta_{1k}\}\{1 - \delta_{2k}\} - S_1(C_{1k})S_2(C_{2k}), \hat{E}_{00,(-k)}) = 0$. Then it follows that

$$cov(N_{00} - E_{00}, E_{00} - \hat{E}_{00}) = ncov(\{1 - \delta_{1k}\}\{1 - \delta_{2k}\} - S_1(C_{1k})S_2(C_{2k}), \hat{E}_{00,(-k)} - \hat{E}_{00}).$$

Because $cov(S_1(C_{1k})S_2(C_{2k}), \hat{E}_{00,(-k)} - \hat{E}_{00}) = o_p(1)$, it follows that

$$cov(N_{00} - E_{00}, E_{00} - \hat{E}_{00}) = ncov(\{1 - \delta_{1k}\}\{1 - \delta_{2k}\}, \hat{E}_{00,(-k)} - \hat{E}_{00}) + o_p(n).$$

Therefore σ_{12} can be consistently estimated by the estimator in equation (A4).

Appendix 5: Asymptotic properties of Z_W under independence

We can write

$$\begin{aligned} Z_W &= n^{1/2}\int_{c_1}\int_{c_2} w(c_1,c_2)\left\{N_{00}(c_1,c_2)-S_1(c_1)S_2(c_2)\right\}G_n(dc_1,dc_2) \\ &\quad +n^{1/2}\int_{c_1}\int_{c_2} w(c_1,c_2)\left\{S_1(c_1)S_2(c_2)-\hat{S}_1(c_1)\hat{S}_2(c_2)\right\}G(dc_1,dc_2)+o_p(1) \\ &= r_{1n}+r_{2n}+o_p(1). \end{aligned}$$

The arguments in Appendix 1 can be applied to show asymptotic normality of r_{1n} and r_{2n}. Therefore, Z_W converges in distribution to a normal random variable with mean zero and variance

$$\sigma_W^2=\sigma_{W1}^2+2\sigma_{W12}+\sigma_{W2}^2. \tag{A8}$$

Here σ_{W1}^2, σ_{W2}^2 and σ_{W12} are the weighted versions of σ_1^2, σ_2^2 and σ_{12}, respectively, where

$$\sigma_{W1}^2=\int_{c_1}\int_{c_2} w^2(c_1,c_2)S_1(c_1)S_2(c_2)\{1-S_1(c_1)S_2(c_2)\}G(dc_1,dc_2)$$

$$\sigma_{W2}^2=\int\int w^2(c_1,c_2)\{F_1(c_1)\frac{a_1(c_1)}{g_1(c_1)}+F_2(c_2)\frac{a_2(c_2)}{g_2(c_2)}\}S_1(c_1)S_2(c_2)G(dc_1,dc_2).$$

Again, we do not have a simple analytical expression for σ_{W12}. But it can be estimated similarly to $\hat{\sigma}_{12}$ in equation (A4) by

$$\hat{\sigma}_{W12}=\frac{1}{n}\sum_{k=1}^{n}W^2(C_{1k},C_{2k})(1-\delta_{1k})(1-\delta_{2k})(\hat{E}_{00,(-k)}-\hat{E}_{00}).$$

The other two components of σ_{W1}^2 and σ_{W2}^2 can be estimated by similarly modifying the estimators of σ_1^2 and σ_2^2 in equations (A1) and (A3).

APPENDIX 6: *Derivation of the local optimal weight function*

Let $\tilde{Z}_W=n^{-1/2}\sum_k W_k(N_{00,k}-E_{00,k})$. Under the alternative hypotheses H_α: $A(t_1,t_2)=n^{-1/2}a(t_1,t_2)+o_p(n^{-1/2})$, $\tilde{Z}_W$ converges in distribution to a normal distribution with mean,

$$n^{1/2}\int\int w(c_1,c_2)S_1(c_1)S_2(c_2)\{e^{-A(c_1,c_2)}-1\}G(dc_1,dc_2)$$

$$= \quad -\int\int w(c_1,c_2)S_1(c_1)S_2(c_2)a(c_1,c_2)G(dc_1,c_2)+o_p(1),$$

and variance

$$\int\int w^2(c_1,c_2)S_1(c_1)S_2(c_2)e^{-A(c_1,c_2)}\{1-S_1(c_1)S_2(c_2)e^{-A(c_1,c_2)}\}G(dc_1,dc_2)$$

$$= \quad \int\int w^2(dc_1,dc_2)S_1(dc_1)S_2(dc_2)\{1-S_1(c_1)S_2(c_2)\}G(dc_1,dc_2)+o_p(1).$$

Hence the local optimal weight function maximizes

$$\frac{\{\int\int w(c_1,c_2)S_1(c_1)S_2(c_2)a(c_1,c_2)G(dc_1,dc_2)\}^2}{\int\int w^2(c_1,c_2)S_1(c_1)S_2(c_2)\{1-S_1(c_1)S_2(c_2)\}G(dc_1,dc_2)}.$$

By the Cauchy-Schwartz inequality, the optimal weight function $w^*(t_1,t_2)$ is proportional to

$$\frac{|a(t_1,t_2)|}{1-S_1(t_1)S_2(t_2)}.$$

REFERENCES

Ayer, M., Brunk, H. D., Ewing, G. M., Reid, W. T. & Silverman, E. (1955), "An empirical distribution function for sampling with incomplete observations," *Ann. Math. Statist.*, 26, 641-7.

Agresti, A. (1990), *Categorical Data Analysis*, New York: John Wiley & Sons.

Anderson, J. E., Louis, T. A., Holm, N. V., & Harvald, B. (1992), "Time-Dependent Association Measures for Bivariate Survival Distributions," *J. Am. Statist. Assoc.*, 87, 641-50.

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

Dabrowska, D. M. (1988), "Kaplan-Meier Estimates on the Plane," *Ann. Statist.*, 16, 1475-89.

Dudley, R. M. (1987), "Universal Donsker classes and metric entropy," *Ann. Prob.*, 15, 1306-26.

Finkelstein, D. M. (1986), "A proportional hazards model for interval-censored failure time data," *Biometrics*, 42, 845-54.

Genest, C. (1987), "Frank's family of bivariate distributions," *Biometrika*, 74, 549-55.

Groeneboom, P. & Wellner, J. A. (1992), *Information Bounds and Non-Parametric Maximum Likelihood Estimation.* Boston: Birkhäuser.

Hsu, L. & Prentice, R. L. (1996), "A generalization of the Mantel-Haenszel tests to bivariate failure time data," *Biometrika*, 83, 905-11.

Huang, J. & Wellner, J. A. (1995), "Asymptotic normality of the NPMLE of linear functionals for interval censored data, Case 1," *Statist. Neer.*, 49, 153-63.

Lin, D. Y., Oakes, D. & Ying, Z. (1998), "Additive hazards regression with current status data," *Biometrika*, 85, 289-98.

Laan, M.J. van der & Robbins, J. M. (1998), “Locally efficient estimation with current status data and time-dependent covariates,” *J. Am. Statist. Assoc.*, 93, 693-701.

Oakes, D. (1982), “A Concordance Test for Independence in the Presence of Bivariate Censoring,” *Biometrics*, 38, 451-55.

Prentice, R. L. & Cai, J. (1992), “Covariance and Survivor Function Estimation Using Censored Multivariate Failure Time Data,” *Biometrika*, 79, 495-512.

Rabinowitz, D., Tsiatis, A. & Aragon, J. (1995), “Regression with interval censored data,” *Biometrika*, 82, 501-13.

Rossini, A. J. & Tsiatis, A. A. (1996), “A semiparametric proportional odds regression model for the analysis of current status data,” *J. Am. Statist. Assoc.*, 91, 713-21.

Sheehy, A. and Wellner, J. A. (1992), “Uniform Donsker classes of functions,” *Ann. Prob.*, 20, 1983-2030.

Shih, J. H. & Louis, T. A. (1996), “Tests of independence for bivariate survival Data,” *Biometrics*, 52, 1440-9.

Turnbull, B. W. (1976), “The empirical distribution function with arbitrarily grouped, censored and truncated data,” *J. R. Statist. Soc. Ser. B*, 38, 290-5.

Wand, M.P. & Jones, M.C. (1995), *Kernel Smoothing.* London: Chapman and Hall.

Wang, W. & Ding, A. A. (2000), “On Assessing the Association for Bivariate Current Status Data,” *Biometrika*, 87, 879-93.

Wang, W. & Wells, M. T. (2000), “Nonparametric Estimation of Kendell’s Tau under Censoring,” *Statistica Sinica*, 10, 1199-1215.

				τ									
T	n	p.l.	W	0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
C	200	.8	U	0.059	0.118	0.275	0.490	0.685	0.842	0.932	0.975	0.996	0.998
			C	0.037	0.127	0.289	0.490	0.674	0.813	0.906	0.958	0.984	0.995
			F	0.053	0.113	0.287	0.520	0.721	0.871	0.942	0.980	0.997	0.999
		.5	U	0.057	0.128	0.300	0.566	0.797	0.934	0.985	0.998	1.000	1.000
			C	0.061	0.100	0.226	0.459	0.704	0.858	0.955	0.988	0.998	1.000
			F	0.056	0.104	0.256	0.514	0.775	0.913	0.980	0.996	0.999	1.000
		.2	U	0.056	0.078	0.142	0.213	0.338	0.471	0.650	0.775	0.891	0.950
			C	0.050	0.046	0.069	0.104	0.169	0.259	0.399	0.540	0.685	0.808
			F	0.051	0.053	0.084	0.124	0.205	0.315	0.463	0.620	0.762	0.867
	400	.8	U	0.061	0.151	0.421	0.719	0.912	0.979	0.998	1.000	1.000	1.000
			C	0.040	0.180	0.435	0.719	0.905	0.971	0.994	0.999	1.000	1.000
			F	0.054	0.154	0.442	0.753	0.940	0.985	0.999	1.000	1.000	1.000
		.5	U	0.054	0.177	0.519	0.828	0.974	0.999	1.000	1.000	1.000	1.000
			C	0.066	0.139	0.413	0.746	0.934	0.991	0.999	1.000	1.000	1.000
			F	0.063	0.149	0.469	0.804	0.966	0.997	1.000	1.000	1.000	1.000
		.2	U	0.052	0.092	0.197	0.342	0.546	0.739	0.880	0.961	0.992	0.997
			C	0.051	0.053	0.087	0.166	0.328	0.506	0.691	0.840	0.941	0.983
			F	0.048	0.057	0.109	0.203	0.386	0.575	0.759	0.895	0.965	0.993
F	200	.8	U	0.059	0.093	0.168	0.293	0.440	0.599	0.757	0.869	0.929	0.973
			C	0.037	0.058	0.091	0.143	0.193	0.280	0.389	0.489	0.584	0.697
			F	0.053	0.075	0.118	0.216	0.318	0.463	0.610	0.749	0.834	0.923
		.5	U	0.057	0.127	0.296	0.550	0.807	0.935	0.980	0.998	1.000	1.000
			C	0.061	0.089	0.173	0.341	0.542	0.723	0.871	0.946	0.986	0.996
			F	0.056	0.099	0.217	0.429	0.677	0.850	0.951	0.987	0.999	1.000
		.2	U	0.056	0.118	0.248	0.416	0.636	0.807	0.928	0.979	0.994	0.999
			C	0.050	0.055	0.106	0.209	0.360	0.550	0.741	0.850	0.931	0.972
			F	0.051	0.070	0.137	0.259	0.439	0.635	0.819	0.905	0.967	0.989
	400	.8	U	0.061	0.114	0.261	0.457	0.685	0.878	0.954	0.988	0.997	1.000
			C	0.040	0.066	0.119	0.200	0.316	0.461	0.616	0.744	0.850	0.926
			F	0.054	0.089	0.176	0.335	0.541	0.748	0.877	0.951	0.985	0.999
		.5	U	0.054	0.175	0.511	0.825	0.973	0.998	1.000	1.000	1.000	1.000
			C	0.066	0.107	0.296	0.572	0.820	0.953	0.991	0.999	1.000	1.000
			F	0.063	0.122	0.383	0.708	0.921	0.989	0.999	1.000	1.000	1.000
		.2	U	0.052	0.146	0.387	0.660	0.883	0.974	0.997	1.000	1.000	1.000
			C	0.051	0.065	0.188	0.386	0.676	0.852	0.951	0.991	0.998	1.000
			F	0.048	0.076	0.228	0.456	0.741	0.904	0.978	0.997	1.000	1.000

Table 1: *Empirical power of $Q_{b(a)}$ based on 4000 replications. The first column "T" lists the true distribution: "C" for the Clayton model and "F" for the Frank model. The second column "n" gives the sample size and the third column "p.l." gives the prevalence level. The fourth column "W" gives the weights: "U" for the unweighted version, "C" for the optimal weight based on the Clayton model and "F" for the optimal weight based on the Frank model.*

				τ									
T	n	p.l.	W	0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
C	200	.8	U	0.073	0.174	0.356	0.570	0.756	0.880	0.949	0.983	0.997	0.999
			C	0.064	0.197	0.394	0.605	0.766	0.877	0.939	0.978	0.993	0.998
			F	0.072	0.205	0.441	0.667	0.835	0.937	0.975	0.993	0.999	1.000
		.5	U	0.050	0.170	0.364	0.640	0.851	0.953	0.992	0.999	1.000	1.000
			C	0.069	0.191	0.372	0.608	0.810	0.922	0.979	0.994	0.999	1.000
			F	0.064	0.198	0.401	0.677	0.876	0.962	0.995	0.999	1.000	1.000
		.2	U	0.041	0.078	0.142	0.223	0.348	0.482	0.654	0.784	0.893	0.949
			C	0.141	0.196	0.271	0.353	0.480	0.578	0.707	0.795	0.883	0.929
			F	0.110	0.161	0.249	0.340	0.477	0.590	0.732	0.825	0.914	0.958
	400	.8	U	0.057	0.197	0.475	0.759	0.925	0.983	0.998	1.000	1.000	1.000
			C	0.053	0.231	0.519	0.786	0.939	0.983	0.997	1.000	1.000	1.000
			F	0.059	0.246	0.579	0.835	0.967	0.993	1.000	1.000	1.000	1.000
		.5	U	0.048	0.219	0.572	0.858	0.981	1.000	1.000	1.000	1.000	1.000
			C	0.060	0.215	0.533	0.821	0.960	0.996	0.999	1.000	1.000	1.000
			F	0.053	0.227	0.602	0.882	0.982	0.999	1.000	1.000	1.000	1.000
		.2	U	0.039	0.088	0.203	0.346	0.560	0.741	0.885	0.962	0.993	0.998
			C	0.080	0.141	0.254	0.370	0.558	0.719	0.842	0.924	0.969	0.990
			F	0.068	0.131	0.249	0.380	0.584	0.753	0.874	0.957	0.988	0.998
F	200	.8	U	0.073	0.147	0.267	0.408	0.573	0.719	0.848	0.930	0.968	0.987
			C	0.064	0.104	0.156	0.222	0.293	0.392	0.514	0.615	0.701	0.799
			F	0.072	0.138	0.248	0.382	0.514	0.642	0.782	0.873	0.925	0.969
		.5	U	0.050	0.167	0.359	0.631	0.859	0.957	0.988	0.999	1.000	1.000
			C	0.069	0.161	0.285	0.480	0.672	0.820	0.918	0.966	0.992	0.998
			F	0.064	0.175	0.332	0.578	0.797	0.924	0.975	0.996	1.000	1.000
		.2	U	0.041	0.115	0.251	0.426	0.642	0.812	0.930	0.979	0.993	0.999
			C	0.141	0.236	0.364	0.510	0.670	0.797	0.900	0.938	0.969	0.987
			F	0.110	0.206	0.350	0.513	0.691	0.829	0.928	0.966	0.988	0.997
	400	.8	U	0.057	0.153	0.338	0.552	0.758	0.915	0.973	0.994	0.998	1.000
			C	0.053	0.100	0.170	0.274	0.411	0.557	0.706	0.812	0.896	0.951
			F	0.059	0.141	0.287	0.488	0.684	0.859	0.940	0.981	0.994	1.000
		.5	U	0.048	0.212	0.566	0.861	0.978	0.999	1.000	1.000	1.000	1.000
			C	0.060	0.162	0.401	0.674	0.879	0.967	0.994	0.999	1.000	1.000
			F	0.053	0.192	0.503	0.802	0.953	0.995	1.000	1.000	1.000	1.000
		.2	U	0.039	0.145	0.394	0.665	0.887	0.973	0.997	1.000	1.000	1.000
			C	0.080	0.189	0.405	0.610	0.821	0.932	0.977	0.993	0.998	1.000
			F	0.068	0.183	0.417	0.649	0.864	0.956	0.990	0.998	1.000	1.000

Table 2: *Empirical power of* $Q_{p(a)}$ *based on* 4000 *replications. The first column "T" lists the true distribution: "C" for the Clayton model and "F" for the Frank model. The second column "n" gives the sample size and the third column "p.l." gives the prevalence level. The fourth column "W" gives the weights: "U" for the unweighted version, "C" for the optimal weight based on the Clayton model and "F" for the optimal weight based on the Frank model.*

Hypothesis	Weight	Value of $Q_{b(a)}$ (p-value)	Value of $Q_{p(a)}$ (p-value)
$T_1 \perp T_2$	Unweighted	26.3148 (0.000)	34.4500 (0.000)
	C-optimal	11.8061 (0.001)	18.0300 (0.000)
	F-optimal	12.1977 (0.000)	18.5347 (0.000)
$T_1 \perp T_3$	Unweighted	31.4113 (0.000)	55.0494 (0.000)
	C-optimal	18.9685 (0.000)	36.6759 (0.000)
	F-optimal	21.2697 (0.000)	41.0896 (0.000)
$T_2 \perp T_3$	Unweighted	5.8107 (0.016)	6.0074 (0.014)
	C-optimal	3.9863 (0.046)	4.4345 (0.035)
	F-optimal	4.2934 (0.038)	4.7365 (0.030)

Table 3: *Pairwise independent tests for the onset ages of patients with three cardiovascular diseases, diabetes mellitus (T_1), hypercholesterolemia (T_2) an hypertension (T_3). The second column indicates the assigned weights, namely unweighted, the optimal weight based on the Clayton model (C-optimal) and the optimal weight based on Frank's model (F-optimal).*