A semiparametric random effects model for multivariate competing risks data

BY THOMAS H. SCHEIKE

*Department of Biostatistics, University of Copenhagen, Øster Farimagsgade 5,
Copenhagen DK-1014, Denmark*
ts@biostat.ku.dk

YANQING SUN

*Department of Mathematics and Statistics, The University of North Carolina at Charlotte,
9201 University City Boulevard, Charlotte, North Carolina 28223, U.S.A.*
yasun@uncc.edu

MEI-JIE ZHANG

*Division of Biostatistics, Medical College of Wisconsin, 8701 Watertown Plank Road,
Milwaukee, Wisconsin 53226, U.S.A.*
meijie@mcw.edu

AND TINA KOLD JENSEN

*Institute of Public Health, University of Southern Denmark, Winslowparken 17,
Odense DK-5000, Denmark*
tkjensen@health.sdu.dk

SUMMARY

We propose a semiparametric random effects model for multivariate competing risks data when the failures of a particular type are of interest. Under this model, the marginal cumulative incidence functions follow a generalized semiparametric additive model. The associations between the cause-specific failure times can be studied through dependence parameters of copula functions that are allowed to depend on cluster-level covariates. A cross-odds ratio-type measure is proposed to describe the associations between cause-specific failure times, and its relationship to the dependence parameters is explored. We develop a two-stage estimation procedure where the marginal models are estimated in the first stage and the dependence parameters are estimated in the second stage. The large sample properties of the proposed estimators are derived. The proposed procedures are applied to Danish twin data to model the cumulative incidence for the age of natural menopause and to investigate the association in the onset of natural menopause between monozygotic and dizygotic twins.

Some key words: Binomial modelling; Copula function; Cross-odds ratio; Cumulative incidence function; Danish twin data; Estimating equation; Inverse-censoring probability weighting; Two-stage estimation.

1. INTRODUCTION

This work is motivated by a dataset from the Danish twin registry ([Skytthe et al., 2003](#)) on the timing of natural menopause among female Danish twins, where natural menopause and surgical menopause are two competing events and time to menopause may be correlated within each pair of twins.

Much of the early work for analyzing competing risks data has been focused on estimation and modelling the cause-specific hazard function ([Kalbfleisch & Prentice, 2002](#)). The cumulative incidence function by contrast measures the absolute cause-specific risk, that is, the probability of having experienced failure from a particular cause over time. The effect of a covariate on the cause-specific hazard function and cumulative incidence function can be very different. The two measures provide different perspectives for cause-specific failure times; see [Fine & Gray \(1999\)](#), [Katsahian et al. \(2006\)](#), [Latouche et al. \(2007\)](#) and [Beyersmann & Schumacher \(2007\)](#).

[Gray \(1988\)](#) and [Fine & Gray \(1999\)](#) proposed a proportional hazards model for the subdistribution hazard to model the cumulative incidence function directly. A general semiparametric model for the cumulative incidence function is studied by [Scheike et al. \(2008\)](#), where some covariates have time-varying effects and others have constant effects. [Scheike & Zhang \(2008\)](#) proposed some goodness-of-fit procedures for model checking for these semiparametric models. Here we extend this work to deal with clusters by providing robust standard errors, and thereby also extend the recent work by [Chen et al. \(2008\)](#) that considered the k -sample case.

Multivariate failure time data analysis has received much attention in the literature. [Dabrowska \(1988\)](#) proposed a nonparametric estimator for the multivariate survival function. [Gill \(1993\)](#) proposed the concept of the iterated odds ratio to model dependence between failure times. The marginal model approach of [Wei et al. \(1989\)](#) fits univariate models to the marginal distributions without modelling the within-cluster dependence. The efficiency of the marginal approach can be improved by introducing weights that account for failure time dependencies ([Cai & Prentice, 1995](#)). A frailty model approach, on the other hand, attempts to model this dependence by postulating a cluster-specific frailty to account for each cluster's particular susceptibility to failures; see [Clayton \(1978\)](#), [Oakes \(1982, 1989\)](#), [Clayton & Cuzick \(1985\)](#) and [Glidden \(2000\)](#), among others.

There has been little work on developing statistical methods for correlated cause-specific failure times in the competing risks setting. Recently, [Bandein-Roche & Ning \(2008\)](#) and [Bandein-Roche & Liang \(2002\)](#) introduced a nonparametric cause-specific cross hazard ratio association measure for multivariate cause-specific failure times; see also [Cheng & Fine \(2008\)](#). [Cheng et al. \(2007\)](#) studied nonparametric estimation of bivariate cause-specific hazard functions and bivariate cumulative incidence functions. They also proposed two nonparametric association measures in terms of the bivariate cause-specific hazard functions and bivariate cumulative incidence functions. These methods do not deal with covariates and require specification of joint cause-specific intensities for all cause combinations. It may be difficult to apply these measures to cause-specific failure time regressions.

To analyze clustered competing risks data, [Katsahian et al. \(2006\)](#) proposed a frailty model for the subdistribution hazard by multiplying lognormal random cluster effects to the proportional subdistribution hazard model. This model allows subject-specific regression effects. One drawback of this approach is that the frailty parameter can be identified solely from the marginal models and that the marginal models depend on the distribution of the cluster-specific effects.

We suggest an alternative approach to avoid this problem by working on a new semiparametric random effects model. Under this model, the marginal cumulative incidence functions follow a generalized semiparametric additive model that is independent of the frailty parameters and frailties account only for correlations between competing risks failure times within clusters. The

frailty variance will thus reflect solely the amount of variation due to clusters. The concept of cross-odds ratio is introduced to measure the degree of association between two cause-specific failure times within the same cluster.

2. ESTIMATION PROCEDURES WITH MULTIVARIATE COMPETING RISKS DATA

2.1. Model specification

Let K be the number of clusters and n_k the number of subjects within the k th cluster. The failure time and cause of failure for the i th subject in the k th cluster are denoted by T_{ki} and ϵ_{ki} , respectively. Assume that cause $\epsilon_{ki} = 1$ is the primary cause of interest and $\epsilon_{ki} = 2$ for other competing causes. Let $X_{ki} = (1, X_{ki,1}, \dots, X_{ki,p})^T$ and $Z_{ki} = (Z_{ki,1}, \dots, Z_{ki,q})^T$ be the associated covariates. Let C_{ki} be the censoring time. The observed competing risks data can be represented by $\{(\tilde{T}_{ki}, \tilde{\epsilon}_{ki}, X_{ki}, Z_{ki})\}$ for $i = 1, \dots, n_k$ and $k = 1, \dots, K$, where $\tilde{T}_{ki} = \min(T_{ki}, C_{ki})$, $\Delta_{ki} = \mathcal{I}(T_{ki} \leq C_{ki})$ is an indicator for uncensored failure time and $\tilde{\epsilon}_{ki} = \epsilon_{ki} \Delta_{ki}$. The clusters are assumed to be independent but the competing risks failure times within clusters are generally correlated. Let $[0, \tau]$ be the time period from which data are collected. We assume that (T_{ki}, ϵ_{ki}) are independent of C_{ki} given the covariates (X_{ki}, Z_{ki}) .

Let θ_k be independently distributed random variables. Assume that the subjects within clusters are independent conditional on θ_k and (X_{ki}, Z_{ki}) . Let $P_1^*(t | x_{ki}, z_{ki}, \theta_k) = \text{pr}(T_{ki} \leq t, \epsilon_{ki} = 1 | \theta_k, x_{ki}, z_{ki})$ be the cluster-specific conditional cumulative incidence function given θ_k and $(X_{ki}, Z_{ki}) = (x_{ki}, z_{ki})$. We consider the following random effects model:

$$P_1^*(t | x_{ki}, z_{ki}, \theta_k) = 1 - \exp(-\theta_k \Psi_{v_k}^{-1}[\exp\{-\eta(t)^T x_{ki} - (\gamma^T z_{ki})t\}]), \quad (1)$$

where $\eta(t)$ is a $(p+1)$ -dimensional vector of regression functions, γ is a q -dimensional vector of parameters, $\Psi_{v_k}(t) = E_{\theta_k}\{\exp(-\theta_k t) | x_{ki}, z_{ki}\}$ is the Laplace transform of the random effects θ_k conditional on (x_{ki}, z_{ki}) and v_k are the dependence parameters. For example, here we take $\Psi_{v_k}(t)$ as the Laplace transform of a gamma distribution with mean 1 and variance v_k , where v_k can depend on some cluster-level covariates Q_k that may be a subset of (x_{ki}, z_{ki}) . This generalization allows for different degrees of association among different types of clusters. For example, the associations between the times to natural menopause for monozygotic twins may be different from that of dizygotic twins. Here, we consider the following simple regression model:

$$v_k = \alpha^T Q_k, \quad (2)$$

where α is an r -dimensional vector of parameters and Q_k is a regression design vector.

Under (1), the marginal cumulative incidence function, $P_1(t | x_{ki}, z_{ki}) = \text{pr}(T_{ki} \leq t, \epsilon_{ki} = 1 | x_{ki}, z_{ki})$, follows the generalized semiparametric additive model:

$$-\log\{1 - P_1(t | x_{ki}, z_{ki})\} = \eta(t)^T x_{ki} + (\gamma^T z_{ki})t. \quad (3)$$

The joint survival function of the cause-specific failure times has an Archimedean copula function representation,

$$\begin{aligned} E[\{1 - I(T_{ki} \leq t, \epsilon_{ki} = 1)\}\{1 - I(T_{kj} \leq s, \epsilon_{kj} = 1)\} | x_{ki}, z_{ki}, x_{kj}, z_{kj}] \\ = \Psi_{v_k}[\Psi_{v_k}^{-1}\{1 - P_1(t | x_{ki}, z_{ki})\} + \Psi_{v_k}^{-1}\{1 - P_1(s | x_{kj}, z_{kj})\}] \\ = C_k\{1 - P_1(t | x_{ki}, z_{ki}), 1 - P_1(s | x_{kj}, z_{kj})\}, \end{aligned}$$

where $C_k(u, v) = \Psi_{v_k}\{\Psi_{v_k}^{-1}(u) + \Psi_{v_k}^{-1}(v)\}$. Various frailty distributions for θ_k can be used to describe correlations among members of each cluster. When the conditional distribution of θ_k

given (x_{ki}, z_{ki}) follows the gamma distribution with mean 1 and variance of v_k , $C_k(u, v) = (u^{-v_k} + v^{-v_k} - 1)^{-1/v_k}$ is the Clayton copula; see [Clayton \(1978\)](#). Other Archimedean copula families include those of [Frank \(1979\)](#) and [Gumbel \(1960\)](#), which are generated by the Laplace transforms of the log series distribution and positive stable distribution; see also [Hougaard \(2000\)](#).

2.2. Estimation of marginal models

The random effects model (1) implies the semiparametric additive model (3) for the marginal cumulative incidence functions. This suggests that model (1) can be estimated through a two-stage procedure, where the parameters of the marginal models are estimated in the first stage and the dependence parameter is estimated in the second stage. Ignoring the dependence between the failure times, the parameters of the marginal models can be estimated by modifying the procedure of [Scheike et al. \(2008\)](#) for univariate competing risks data. This resembles the marginal approach of [Wei et al. \(1989\)](#).

Let $G_c(t | X_{ki}, Z_{ki}) = \text{pr}(C_{ki} > t | X_{ki}, Z_{ki})$ and $N_{ki}(t) = \mathcal{I}(T_{ki} \leq t, \epsilon_{ki} = 1)$ be the counting process associated with cause 1. Under conditional independence between C_{ki} and (T_{ki}, ϵ_{ki}) given the covariates, we have

$$E \left\{ \frac{\Delta_{ki} N_{ki}(t)}{G_c(T_{ki} | X_{ki}, Z_{ki})} \middle| X_{ki}, Z_{ki} \right\} = P_1(t | X_{ki}, Z_{ki}).$$

We consider estimating equations based on $\Delta_{ki} N_{ki}(t) / G_c(T_{ki} | X_{ki}, Z_{ki})$. Although in practice the censoring distribution $G_c(t | x_{ki}, z_{ki})$ is unknown, it can be estimated by the Kaplan–Meier estimator or by using a regression model to improve the efficiency. For simplicity, we use the Kaplan–Meier estimator $\hat{G}_c(t)$. In the regression case, $G_c(t | x_{ki}, z_{ki})$ will be estimated based on the data $\{(\hat{T}_{ki}, \Delta_{ki}, X_{ki}, Z_{ki}), i = 1, \dots, n_k, k = 1, \dots, K\}$.

Let $P_1^{(k)}(t, \eta, \gamma)$ be the $n_k \times 1$ vector of $P_1(t | x_{ki}, z_{ki})$ for $i = 1, \dots, n_k$, $R^{(k)}(t)$ be the $n_k \times 1$ vector of adjusted responses $\Delta_{ki} N_{ki}(t) / \hat{G}_c(T_{ki})$, $D_\eta^{(k)}(t, \eta, \gamma)$ and $D_\gamma^{(k)}(t, \eta, \gamma)$ be the $n_k \times (p + 1)$ and $n_k \times q$ matrices with the i th rows equal to $D_{\eta, ki}(t, \eta, \gamma) = \partial P_1(t | x_{ki}, z_{ki}) / \partial \eta(t)$ and $D_{\gamma, ki}(t, \eta, \gamma) = \partial P_1(t | x_{ki}, z_{ki}) / \partial \gamma$, respectively. Let $W^{(k)}(t) = \text{diag}\{w_{ki}(t)\}$ where $w_{ki}(t)$ are the weights.

Following [Scheike et al. \(2008\)](#), the marginal regression functions $\eta(t)$ and regression parameters γ can be estimated based on the following estimating equations:

$$U_\eta(t, \eta, \gamma) = D_\eta^\top(t, \eta, \gamma) W(t) \{R(t) - P_1(t, \eta, \gamma)\} = 0, \quad (4)$$

$$U_\gamma(\tau, \eta, \gamma) = \int_0^\tau D_\gamma^\top(t, \eta, \gamma) W(t) \{R(t) - P_1(t, \eta, \gamma)\} dt = 0, \quad (5)$$

where $W(t) = \text{diag}\{W^{(k)}(t)\}$, $R(t) = [\{R^{(1)}(t)\}^\top, \dots, \{R^{(K)}(t)\}^\top]^\top$, and

$$P_1(t, \eta, \gamma) = \left[\{P_1^{(1)}(t, \eta, \gamma)\}^\top, \dots, \{P_1^{(K)}(t, \eta, \gamma)\}^\top \right]^\top,$$

$$D_\eta(t, \eta, \gamma) = \left[\{D_\eta^{(1)}(t, \eta, \gamma)\}^\top, \dots, \{D_\eta^{(K)}(t, \eta, \gamma)\}^\top \right]^\top,$$

$$D_\gamma(t, \eta, \gamma) = \left[\{D_\gamma^{(1)}(t, \eta, \gamma)\}^\top, \dots, \{D_\gamma^{(K)}(t, \eta, \gamma)\}^\top \right]^\top.$$

The estimating equations (4) and (5) can be solved by an iterative procedure similar to that given in [Scheike et al. \(2008\)](#); see the Appendix. The procedure can be implemented using standard software for generalized linear models. However, the standard errors obtained from such software are conservative. Correct estimation of the standard errors is described in the following.

Suppose that there exists a constant N_0 such that $n_k < N_0$ for all k . Under regularity conditions similar to [Scheike et al. \(2008\)](#) we show in the Appendix that as $K \rightarrow \infty$, $K^{\frac{1}{2}}(\hat{\gamma} - \gamma)$ and $K^{\frac{1}{2}}\{\hat{\eta}(t) - \eta(t)\}$ are asymptotically equivalent to the following decompositions

$$K^{\frac{1}{2}} I_{\gamma}^{-1} \sum_{k=1}^K \sum_{i=1}^{n_k} W_{\gamma,ki}(\tau), \quad K^{\frac{1}{2}} \{I_{\eta}(t)\}^{-1} \sum_{k=1}^K \sum_{i=1}^{n_k} W_{\eta,ki}(t),$$

respectively, for which the formula for I_{γ} , $I_{\eta}(t)$, $W_{\gamma,ki}(\tau)$ and $W_{\eta,ki}(t)$ are given in the Appendix. The sums $\sum_{i=1}^{n_k} W_{\gamma,ki}(\tau)$ and $\sum_{i=1}^{n_k} W_{\eta,ki}(t)$ are independent and identically distributed with mean zero for $k = 1, \dots, K$. Thus $K^{\frac{1}{2}}(\hat{\gamma} - \gamma)$ converges in distribution to a mean zero normal random vector by a simple application of the central limit theorem. By an application of Theorem 19.5 of [van der Vaart \(1998\)](#), $K^{\frac{1}{2}}\{\hat{\eta}(t) - \eta(t)\}$ converges weakly to a mean zero Gaussian process on $t \in [0, \tau]$.

Let $\hat{I}_{\gamma}$, $\hat{W}_{\gamma,ki}(\tau)$, $\hat{I}_{\eta}(t)$ and $\hat{W}_{\eta,ki}(t)$ be the estimators of their population counterparts with the plug-in estimates of all needed quantities. The asymptotic variance of $K^{\frac{1}{2}}(\hat{\gamma} - \gamma)$ and $K^{\frac{1}{2}}\{\hat{\eta}(t) - \eta(t)\}$ can be estimated by

$$\begin{aligned} \hat{\Sigma}_{\gamma} &= K \hat{I}_{\gamma}^{-1} \left[\sum_{k=1}^K \left\{ \sum_{i=1}^{n_k} \hat{W}_{\gamma,ki}(\tau) \right\}^{\otimes 2} \right] \hat{I}_{\gamma}^{-1}, \\ \hat{\Sigma}_{\eta}(t) &= K \{\hat{I}_{\eta}(t)\}^{-1} \left[\sum_{k=1}^K \left\{ \sum_{i=1}^{n_k} \hat{W}_{\eta,ki}(t) \right\}^{\otimes 2} \right] \{\hat{I}_{\eta}(t)\}^{-1}, \end{aligned}$$

where $a^{\otimes 2} = aa^T$ for a column vector a . These estimators can be used to construct confidence intervals for γ and $\eta(t)$.

The Gaussian multiplier resampling method can be used to construct simultaneous confidence intervals for $\eta(t)$ over a fixed time interval and make inferences. By application of empirical process theory ([van der Vaart, 1998](#)), $K^{\frac{1}{2}}\{\hat{\eta}(t) - \eta(t)\}$ converges weakly to the same mean zero Gaussian limit processes as $K^{\frac{1}{2}}\{\hat{I}_{\eta}(t)\}^{-1} \sum_{k=1}^K \sum_{i=1}^{n_k} G_k \hat{W}_{\eta,ki}(t)$, given the observed data, where $G_1, \dots, G_K$ are independent standard normal random variables. We refer to Lemma 1 of [Sun & Wu \(2005\)](#) for a result that applies directly to this setting. Hence a $(1 - \alpha) \times 100\%$ asymptotic confidence band for $\hat{\eta}(t)$ can be constructed by repeated resampling of $G_1, \dots, G_K$.

2.3. Estimation of dependence parameters

Let

$$V_{kij}(t) = \{G_c(T_{ki})\}^{-1} \Delta_{ki} N_{ki}(t) \{G_c(T_{kj})\}^{-1} \Delta_{kj} N_{kj}(t), \quad (6)$$

where i, j are two subjects within cluster k . It follows that, for $i \neq j$,

$$\begin{aligned} v_{kij}(t, \alpha, \eta, \gamma) &= E\{V_{kij}(t) \mid X_{ki}, Z_{ki}, X_{kj}, Z_{kj}\} \\ &= 1 - \exp\{-\Lambda_{ki}(t)\} - \exp\{-\Lambda_{kj}(t)\} + \Psi_{v_k}(\Psi_{v_k}^{-1}[\exp\{-\Lambda_{ki}(t)\}] + \Psi_{v_k}^{-1}[\exp\{-\Lambda_{kj}(t)\}]), \quad (7) \end{aligned}$$

where $v_k = \alpha^T Q_k$ under (2) and $\Lambda_{ki}(t) = \eta(t)^T X_{ki} + (\gamma^T Z_{ki})t$.

Let H_k denote the index set for the k th cluster. Replacing $G_c(t)$ by $\hat{G}_c(t)$ in (6) to obtain $\hat{V}_{kij}(t)$, we consider the following estimating equation for α :

$$U_{\alpha}(\alpha, \hat{\eta}, \hat{\gamma}, \hat{G}_c) = \int_0^{\tau} \sum_{k=1}^K \sum_{i,j \in H_k, i \neq j} D_{\alpha, kij}(t, \alpha, \hat{\eta}, \hat{\gamma}) \{\hat{V}_{kij}(t) - v_{kij}(t, \alpha, \hat{\eta}, \hat{\gamma})\} dt = 0, \quad (8)$$

where $D_{\alpha, kij}(t, \alpha, \hat{\eta}, \hat{\gamma}) = \partial v_{kij}(t, \alpha, \hat{\eta}, \hat{\gamma}) / \partial \alpha$, and $\hat{\eta}(t)$, $\hat{\gamma}$ are the estimators of $\eta(t)$, γ from the marginal model. Let $\hat{\alpha}$ be the estimator of α solving (8). Then the cluster-specific dependence parameter is given by $\hat{v}_k = \hat{\alpha}^\top Q_k$. By Taylor approximation,

$$K^{\frac{1}{2}}(\hat{\alpha} - \alpha) = - \left\{ K^{-1} \frac{\partial U_\alpha(\alpha, \eta, \gamma, G_c)}{\partial \alpha} \right\}^{-1} K^{-\frac{1}{2}} U_\alpha(\alpha, \hat{\eta}, \hat{\gamma}, \hat{G}_c) + o_p(1).$$

We show in the Appendix that $K^{\frac{1}{2}}(\hat{\alpha} - \alpha)$ has the following approximation:

$$K^{\frac{1}{2}}(\hat{\alpha} - \alpha) = K^{-\frac{1}{2}} \{ K^{-1} \mathcal{I}_\alpha(\alpha, \eta, \gamma) \}^{-1} \left\{ \sum_{k=1}^K W_{\alpha, k}(\alpha, \eta, \gamma, G_c) \right\} + o_p(1),$$

for which the formulae for $W_{\alpha, k}(\alpha, \eta, \gamma, G_c)$ and $\mathcal{I}_\alpha(\alpha, \eta, \gamma)$ are given in the Appendix. It follows that $K^{\frac{1}{2}}(\hat{\alpha} - \alpha)$ is asymptotically normal with mean zero and its variance can be estimated by

$$\hat{\Sigma}_\alpha = K \hat{\mathcal{I}}_\alpha^{-1} \left\{ \sum_{k=1}^K (\hat{W}_{\alpha, k})^{\otimes 2} \right\} \hat{\mathcal{I}}_\alpha^{-1},$$

where $\hat{\mathcal{I}}_\alpha$ and $\hat{W}_{\alpha, k}$ are the empirical counterparts of $\mathcal{I}_\alpha(\alpha, \eta, \gamma)$ and $W_{\alpha, k}(\alpha, \eta, \gamma, G_c)$, respectively.

2.4. Cross-odds ratio

There seems to be no simple direct extension of the association measures for multivariate failure times to the competing risks data. The popular Kendall's tau-like interpretation of dependence parameter is no longer applicable to multivariate cause-specific failure times since cumulative incidence functions are in general not distribution functions.

We introduce a cross-odds ratio to describe the dependence between cause-specific failure times. Consider the case of bivariate failure times (T_{ki}, T_{kj}) with causes $(\epsilon_{ki}, \epsilon_{kj})$. We define the following association measure for the failure times due to cause 1 between individual i and j in the same cluster k :

$$\pi_{i|j}(t) = \frac{\text{pr}(T_{ki} \leq t, \epsilon_{ki} = 1 \mid T_{kj} \leq t, \epsilon_{kj} = 1) / \text{pr}\{(T_{ki} \leq t, \epsilon_{ki} = 1)^c \mid T_{kj} \leq t, \epsilon_{kj} = 1\}}{\text{pr}(T_{ki} \leq t, \epsilon_{ki} = 1) / \text{pr}\{(T_{ki} \leq t, \epsilon_{ki} = 1)^c\}}, \quad (9)$$

where $(A)^c$ means the event complementary to A . This measure of association resembles the odds ratio for the event probabilities and it may vary with time. If the events $(T_{ki} \leq t, \epsilon_{ki} = 1)$ and $(T_{kj} \leq t, \epsilon_{kj} = 1)$ are independent, then $\pi_{i|j}(t) = 1$. If they are positively associated with $\text{pr}(T_{ki} \leq t, \epsilon_{ki} = 1 \mid T_{kj} \leq t, \epsilon_{kj} = 1) > \text{pr}(T_{ki} \leq t, \epsilon_{ki} = 1)$, then $\pi_{i|j}(t) > 1$; otherwise $\pi_{i|j}(t) < 1$. If $\text{pr}(T_{ki} \leq t, \epsilon_{ki} = 1 \mid T_{kj} \leq t, \epsilon_{kj} = 1) = 1$, thus perfectly positive correlation, then $\pi_{i|j}(t) = \infty$. If $\text{pr}(T_{ki} \leq t, \epsilon_{ki} = 1 \mid T_{kj} \leq t, \epsilon_{kj} = 1) = 0$, then $\pi_{i|j}(t) = 0$. Further, this measure is symmetric about the individuals i and j provided that the two have the same marginal subprobabilities, i.e., $\text{pr}(T_{ki} \leq t, \epsilon_{ki} = 1) = \text{pr}(T_{kj} \leq t, \epsilon_{kj} = 1)$.

In the presence of covariates, the conditional cross-odds ratio given the covariates, $\pi_{i|j}(t \mid x_{ki}, z_{ki}, x_{kj}, z_{kj})$, can be similarly defined as in (9) by replacing the probabilities with the conditional probabilities given the covariates. An important advantage of our definition of the cross-odds ratio is that it is readily available once model (1) is estimated.

When the random effects θ_k follow a gamma distribution with mean 1 and variance v_k , $\Psi_{v_k}(x) = (1 + v_k x)^{-1/v_k}$, for $x > -1/v_k$ and its inverse function is $\Psi_{v_k}^{-1}(y) = (y^{-v_k} - 1)/v_k$ for $y > 0$. Once the marginal cumulative incidence model and the frailty variance are estimated the cross-odds ratio $\pi_{i|j}$ is simple to estimate. The cross-odds ratio is closely related to the dependence parameter v_k of the Clayton Archimedean copula family. Indeed, when all the covariates are

cluster level covariates, i.e. $(x_{ki}, z_{ki}) = (x_{kj}, z_{kj})$, $\pi_{i|j}(t | x_{ki}, z_{ki}, x_{kj}, z_{kj}) = 2$ for $v_k = 1$, and $\pi_{i|j}(t | x_{ki}, z_{ki}, x_{kj}, z_{kj}) \approx v_k + 1$ for moderate values of v_k in the range of $[0, 2]$. The association modelled by the Clayton copula is always positive. A larger value of v_k indicates stronger positive association in terms of times to cause-specific failure times within clusters. Other copula functions may be used to model negative as well as positive associations.

3. DANISH TWIN DATA

We identified 503 pairs of female twins born between 1931 and 1952 through the Danish Twin Registry. There are 269 monozygotic twin pairs and 234 dizygotic twin pairs identified by [Skytthe et al. \(2003\)](#). Before they reach natural menopause, some women may experience surgical menopause as a result of removal of the uterus, cervix or ovaries and others may be treated with hormones that affect the timing of menopause. Surgical menopause and hormonal treatment are therefore competing risks for natural menopause. Of 1006 twins, 416 twins experienced natural menopause, 57 twins had surgical related menopause, 183 twins underwent hormonal treatments before natural menopause, 246 twins had censored event times. There were 104 twins with unknown failure types or censoring status and they were excluded from the analysis. These summaries do not reflect that the probabilities of different events changed considerably over the time cohorts. We partition the data into four cohorts of almost equal size according to birthdate. Cohort 1 from 1931–1936 has 120 twin pairs, cohort 2 from 1937–1942 has 123 twin pairs, cohort 3 from 1943–1947 has 121 twin pairs, and cohort 4 from 1948–1952 has 139 twin pairs. In cohort 1, there was zero censoring, 154 natural menopause, 22 surgical menopause, 40 hormonal treatments, and 24 cases with unknown failure types or censoring status. These numbers were (3, 137, 13, 64, 29) for cohort 2, (65, 89, 11, 54, 23) for cohort 3, and (178, 36, 11, 25, 28) for cohort 4. In the analysis, we allow the censoring distribution to depend on cohorts.

We first estimate the marginal cumulative incidence of natural menopause excluding the cohort structure. Let $x_{ki} = \{1, I_{ki}(\text{dizygotic twin})\}^T$, where $I_{ki}(\text{dizygotic twin})$ is 1 if the i th twin from the k th family is dizygotic for $i = 1, 2$ and $k = 1, \dots, 503$. Under model (1) without covariate z_{ki} , the marginal cumulative incidence is $-\log\{1 - P_1(t | x_{ki})\} = \eta(t)^T x_{ki}$, which leads to the nonparametric estimates for the distributions of the ages to natural menopause for the monozygotic and dizygotic twins. The estimates, 95% confidence intervals and 95% confidence bands based on the Gaussian multiplier resampling technique are shown in Fig. 1(a) for the monozygotic twins. The plots for the dizygotic twins are similar and omitted to save space.

To examine whether the cumulative incidence of natural menopause changes over the different cohorts, we estimate the cumulative incidence function for natural menopause by adjusting for time cohort effects. Let $I_{ki}(\text{cohort } 2)$ be the cohort indicator of whether the twins in the k th family are born between 1937 and 1942. In addition to x_{ki} defined above, we introduce cohort covariates z_{ki} to the random effects model (1), where $z_{ki} = \{I_{ki}(\text{cohort } 2), I_{ki}(\text{cohort } 3), I_{ki}(\text{cohort } 4)\}^T$. The marginal cumulative incidence functions then follow the semiparametric additive model $-\log\{1 - P_1(t | x_{ki}, z_{ki})\} = \eta(t)^T x_{ki} + (\gamma^T z_{ki})t$. This model assumes separate nonparametric baseline functions for monozygotic and dizygotic twins and constant time cohort effects. The goodness-of-fit procedures of [Scheike & Zhang \(2008\)](#) indicate that the cohorts can well be described as having constant effects. Figure 1 (b) shows the estimated cumulative incidence functions of natural menopause for cohort 1 and 4 for monozygotic twins. There is a clear trend over time in that the later cohort experience their natural menopause later. For cohort 4 we only estimate the cumulative incidence curve for the ages between 40 and 55 due to censoring. The estimated cohort effects $\hat{\gamma}$ are $(-1.03, -3.88, -5.72) \times 10^{-2}$ with the standard errors of $(0.65, 0.58, 0.53) \times 10^{-2}$, respectively.

Fig. 1. Estimated probability of natural menopause in Danish twins data. The estimated marginal cumulative incidence functions for monozygotic twins are shown in panel (a) and the upper curves in panel (b) are for cohort 1 and the lower curves are for cohort 4. The dashed lines are 95% confidence intervals and dotted lines are 95% confidence bands. The estimated probabilities of monozygotic twins in cohort 1 having both experienced natural menopause are shown in panel (c) under model (1) (dark dashed line), under independence (solid thin line), for $\theta_k = \theta_{k,0.95}$ (top dashed line) and for $\theta_k = \theta_{k,0.05}$ (bottom dashed line). Panel (d) shows similar plots for dizygotic twins.

We now examine whether there is an association in the occurrence of natural menopause between twins of the same family under the random effects model (1). To allow for different degrees of association for different types of twins, we let $\nu_k = \alpha_0 + \alpha_1 x_k$ under (2), where x_k is 0 for monozygotic twins and 1 for dizygotic twins. Let $\rho_1 = \alpha_0$ and $\rho_2 = \alpha_0 + \alpha_1$. Using the estimation procedure presented in § 2.3, we get $\hat{\rho}_1 = 1.04$ with the standard error of 0.42, and $\hat{\rho}_2 = 0.39$ with the standard error of 0.29. Hence, a significant association exists in the monozygotic twins but not for the dizygotic twins.

The joint probability of both twins experiencing natural menopause by age t depends on the degree of association between the twins. This probability is $\nu_{kij}(t)$ given in (7) under the random effects model (1). For monozygotic twins in cohort 1, this joint probability is plotted in Fig. 1 (c) against age. The random effect θ_k in model (1) represents family-to-family variability and induces dependence between twins. Let $\theta_{k,0.95}$ and $\theta_{k,0.05}$ be the 95th percentile and the 5th percentile of the gamma distribution with mean 1 and variance ν_k , respectively. The joint probabilities of both twins experiencing natural menopause by age t in a twin pair with $\theta_k = \theta_{k,0.95}$ and $\theta_k = \theta_{k,0.05}$ can be expressed as $P_1^*(t | x_{ki}, z_{ki}, \theta_{k,0.95})P_1^*(t | x_{kj}, z_{kj}, \theta_{k,0.95})$ and $P_1^*(t | x_{ki}, z_{ki}, \theta_{k,0.05})P_1^*(t | x_{kj}, z_{kj}, \theta_{k,0.05})$, respectively. For monozygotic twins in cohort 1, these joint probabilities are given in Fig. 1 (c). The plot of the joint probability of both twins experiencing natural menopause by age t , $P_1(t | x_{ki}, z_{ki})P_1(t | x_{kj}, z_{kj})$, under the marginal model (3) and the presumed independence between twins is also given in Fig. 1 (c). Similar plots for dizygotic twins are given in Fig. 1 (d). Other quantities such as the conditional probabilities of

Table 1. Summaries of the simulation results based on 1000 realizations for K twin pairs

K	ν	Estimation of $\gamma (= 0.5)$			Estimation of ν		
		Mean	MSE	SSE	Mean	MSE	SSE
100	2	0.51	0.29	0.31	2.12	0.83	0.86
	1	0.51	0.29	0.30	1.08	0.54	0.54
	0.5	0.52	0.29	0.30	0.61	0.42	0.37
200	2	0.51	0.21	0.21	2.06	0.56	0.56
	1	0.50	0.21	0.20	1.02	0.36	0.36
	0.5	0.51	0.21	0.21	0.54	0.28	0.27

Mean, average of estimates; MSE, mean of estimated standard errors; SSE, sample standard error of estimates.

the second twin experiencing natural menopause given that the first twin has already experienced her natural menopause can also be calculated and plotted.

Finally, we calculate the cross-odds ratio to show how the onset of natural menopause for one twin affects the probability of the onset for the other twin. The cross-odds ratios for monozygotic twins are around 2. This indicates that, if one twin has experienced natural menopause, the odds of the second twin experiencing natural menopause by the same age essentially doubles the unconditional odds of natural menopause for the second twin. For the dizygotic twins, we find a cross-odds ratio that lies around 1.4.

4. SIMULATION STUDY

We conducted a small simulation study for model (1) to evaluate small sample performance of the proposed method. We take $K = 100$ and 200 pairs and let the random effects follow a Gamma distribution with mean 1 and variance ν . We consider a situation with one covariate that is uniformly distributed on $[0, 1]$. The baseline function is set as $\eta(t) = 0.5t$ and the regression coefficient $\gamma = 0.5$. The censoring times are generated from the uniform distribution on $[0, 2]$. The probability of observing the cause of interest is 0.42, which is comparable to the Danish twin dataset where natural menopause occurred in about 40% of the twins.

Table 1 summarizes the simulation results for $\hat{\gamma}$ and $\hat{\nu}$ based on 1000 realizations for $\gamma = 0.5$ and ν equal to 2, 1 and 0.5. The estimator $\hat{\gamma}$ performs satisfactorily at all levels of the random effect variance ν and for both sample sizes. The estimator $\hat{\gamma}$ is essentially unbiased and its variance is well estimated by the robust variance estimator. The biases of the estimator for ν are reasonably small, and the estimator for the variance of $\hat{\nu}$ perform well. The simulations also find the baseline function well approximated by its estimator, which are not reported here.

5. DISCUSSION

The random effects model (1) is estimated through a two-stage approach where the marginal parameters are estimated separately from the dependence parameters. Additional efficiency can be gained by jointly estimating both the marginal parameters and the dependence parameters and by using all the joint observations of the inverse probability of censoring weighted responses for the dependence parameters. This issue deserves further attention and investigation.

The asymptotic distribution of $\hat{\pi}_{i|j}(t | x_{ki}, z_{ki}, x_{kj}, z_{kj})$ can be derived using the delta method. Confidence intervals and hypothesis testing procedures about the cross-odds ratio can be derived accordingly. The cross-odds ratio may also be extended in different ways to show how the association persists across the time-scale.

ACKNOWLEDGMENT

The research was partially supported by grants from the National Cancer Institute to Scheike and Zhang, and by grants from the National Science Foundation and the National Institutes of Health to Sun. Thomas Scheike was also supported by a grant from the Danish Research Council. We appreciate the valuable comments of the reviewers and the editor, which have improved the content and presentation of this paper.

APPENDIX

An iterative procedure for the estimators $\hat{\gamma}$ and $\hat{\eta}(t)$

The estimating equations (4) and (5) can be solved simultaneously in a way similar to [Scheike et al. \(2008\)](#). By the Taylor expansion of $P_1(t, \eta, \gamma)$ around the current value of the estimates $(\hat{\eta}^{(m)}(t), \hat{\gamma}^{(m)})$, we have

$$P_1(t, \eta, \gamma) = P_1(t, \hat{\eta}^{(m)}, \hat{\gamma}^{(m)}) + D_\gamma(t, \hat{\eta}^{(m)}, \hat{\gamma}^{(m)})\{\gamma - \hat{\gamma}^{(m)}\} \\ + D_\eta(t, \hat{\eta}^{(m)}, \hat{\gamma}^{(m)})\{\eta(t) - \hat{\eta}^{(m)}(t)\} + o_p(K^{-\frac{1}{2}}).$$

Let $D_\eta^{(m)}(t) = D_\eta(t, \hat{\eta}^{(m)}, \hat{\gamma}^{(m)})$, $D_\gamma^{(m)}(t) = D_\gamma(t, \hat{\eta}^{(m)}, \hat{\gamma}^{(m)})$ and $P_1^{(m)}(t) = P_1(t, \hat{\eta}^{(m)}, \hat{\gamma}^{(m)})$. We have the following approximated estimating equations:

$$D_\eta^{(m)}(t)^\top W(t)[R(t) - P_1^{(m)}(t) - D_\eta^{(m)}(t)\{\eta(t) - \hat{\eta}^{(m)}(t)\} - D_\gamma^{(m)}(t)\{\gamma - \hat{\gamma}^{(m)}\}] = 0, \quad (\text{A1})$$

$$\int_0^\tau D_\gamma^{(m)}(t)^\top W(t)[R(t) - P_1^{(m)}(t) - D_\eta^{(m)}(t)\{\eta(t) - \hat{\eta}^{(m)}(t)\} - D_\gamma^{(m)}(t)\{\gamma - \hat{\gamma}^{(m)}\}] dt = 0. \quad (\text{A2})$$

Solving the equations (A1) and (A2) for γ and $\eta(t)$, the $(m+1)$ th iterative estimators are given by $\hat{\gamma}^{(m+1)} = \hat{\gamma}^{(m)} + \{I_\gamma^{(m)}\}^{-1} B_\gamma^{(m)}$ and $\hat{\eta}^{(m+1)}(t) = \hat{\eta}^{(m)}(t) + \{I_\eta^{(m)}(t)\}^{-1} \{D_\eta^{(m)}(t)\}^\top W(t)\{R(t) - P_1^{(m)}(t) - D_\gamma^{(m)}(t)\{I_\gamma^{(m)}\}^{-1} B_\gamma^{(m)}\}$, where

$$I_\gamma^{(m)} = \int_0^\tau D_\gamma^{(m)}(t)^\top W(t) H^{(m)}(t) D_\gamma^{(m)}(t) dt, \\ B_\gamma^{(m)} = \int_0^\tau D_\gamma^{(m)}(t)^\top W(t) H^{(m)}(t) \{R(t) - P_1^{(m)}(t)\} dt, \\ H^{(m)}(t) = I - D_\eta^{(m)}(t) \{I_\eta^{(m)}(t)\}^{-1} D_\eta^{(m)}(t)^\top W(t), \\ I_\eta^{(m)}(t) = D_\eta^{(m)}(t)^\top W(t) D_\eta^{(m)}(t). \quad (\text{A3})$$

On the asymptotic distributions of $K^{\frac{1}{2}}(\hat{\gamma} - \gamma)$ and $K^{\frac{1}{2}}\{\hat{\eta}(t) - \eta(t)\}$

Let $D_\eta(t)$, $D_\gamma(t)$ and $P_1(t, \eta, \gamma)$ be the corresponding terms of $D_\eta^{(m)}(t)$, $D_\gamma^{(m)}(t)$ and $P_1^{(m)}(t)$ evaluated at the true values of $\{\eta(t), \gamma\}$. Similarly, I_γ , $I_\eta(t)$, B_γ and $H(t)$ are the corresponding terms for (A3). By applying the Taylor approximation,

$$K^{\frac{1}{2}}(\hat{\gamma} - \gamma) = K^{\frac{1}{2}} I_\gamma^{-1} B_\gamma + o_p(1).$$

Let $Q(t) = D_\gamma^\top(t) W(t) D_\eta(t) \{I_\eta(t)\}^{-1}$, $\zeta_{\gamma,ki}(t) = \{D_{\gamma,ki}(t) - Q(t) D_{\eta,ki}(t)\} w_{ki}(t)$, $\tilde{B}_\gamma = \sum_{k,i} \int_0^\tau \zeta_{\gamma,ki}(t) dt$, $\{G_c(T_{ki})^{-1}\} \Delta_{ki} N_{ki}(t) - P_1(t, x_{ki}, z_{ki})\}$, and $\Delta_\gamma = \sum_{k,i} \int_0^\tau \{D_{\gamma,ki}(t) - Q(t) D_{\eta,ki}(t)\} w_{ki}(t) \Delta_{ki} N_{ki}(t) \{\hat{G}_c(T_{ki}) G_c(T_{ki})\}^{-1} \{G_c(T_{ki}) - \hat{G}_c(T_{ki})\} dt$. Then $B_\gamma = \tilde{B}_\gamma + \Delta_\gamma$. Note that

$$\Delta_{ki} N_{ki}(t) \frac{\hat{G}_c(T_{ki}) - G_c(T_{ki})}{\hat{G}_c(T_{ki}) G_c(T_{ki})} = - \frac{\Delta_{ki} N_{ki}(t)}{G_c(T_{ki})} \sum_{l=1}^K \sum_{j=1}^{n_l} \int_0^\tau \mathcal{I}(s \leq T_{ki}) \frac{dM_{lj}^C(s)}{Y(s)} + o_p(K^{-\frac{1}{2}}),$$

where $Y(s) = \sum_{k,i} Y_{ki}(s) = \sum_{k,i} \mathcal{I}(\tilde{T}_{ki} \geq s)$ and $M_{lj}^C(s) = \mathcal{I}(\tilde{T}_{lj} \leq s, \Delta_{lj} = 0) - \int_0^s Y_{lj}(u) d\{-\log G_c(u)\}$. We have $\Delta_\gamma = \sum_{k,i} \int_0^\tau \psi_{\gamma,ki}(t) dt + o_p(K^{\frac{1}{2}})$, where $\psi_{\gamma,ki}(t) = \int_0^\tau \{Y(s)\}^{-1} q_\gamma(s, t) dM_{ki}^C(s)$ and $q_\gamma(s, t) = \sum_{l,j} \{G_c(T_{lj})\}^{-1} \{D_{\gamma,lj}(t) - Q(t)D_{\eta,lj}\} w_{lj}(t) \Delta_{lj} N_{lj}(t) \mathcal{I}(s \leq T_{lj} \leq t)$. Thus,

$$K^{\frac{1}{2}}(\hat{\gamma} - \gamma) = K^{\frac{1}{2}} I_\gamma^{-1} \sum_{k,i} W_{\gamma,ki}(\tau) + o_p(1), \quad (\text{A4})$$

where $W_{\gamma,ki}(\tau) = \sum_{k,i} \int_0^\tau \{\zeta_{\gamma,ki}(t) + \psi_{\gamma,ki}(t)\} dt$. Let $\hat{I}_\gamma$ and $\hat{W}_{\gamma,ki}(\tau)$ be estimators of I_γ and $W_{\gamma,ki}(\tau)$ obtained by inserting the estimators $\hat{\eta}(t)$, $\hat{\gamma}$ and $\hat{M}_{ki}^C(t)$. The asymptotic variance of $K^{\frac{1}{2}}(\hat{\gamma} - \gamma)$ can be consistently estimated by $\hat{\Sigma}_\gamma = K \hat{I}_\gamma^{-1} [\sum_{k=1}^K \{\sum_{i=1}^{n_k} \hat{W}_{\gamma,ki}(\tau)\}^{\otimes 2}] \hat{I}_\gamma^{-1}$.

Similarly, by using a Taylor expansion, we have

$$K^{\frac{1}{2}}\{\hat{\eta}(t) - \eta(t)\} = K^{\frac{1}{2}}\{I_\eta(t)\}^{-1} \sum_{k,i} W_{\eta,ki}(t) + o_p(1), \quad (\text{A5})$$

where

$$\begin{aligned} W_{\eta,ki}(t) &= \{\zeta_{\eta,ki}(t) + \psi_{\eta,ki}(t) - D_\eta^\tau(t)W(t)D_\gamma(t)I_\gamma^{-1}W_{\gamma,ki}(\tau)\}, \\ \zeta_{\eta,ki}(t) &= D_{\eta,ki}(t)w_{ki}(t) \left\{ \frac{\Delta_{ki}N_{ki}(t)}{G_c(T_{ki})} - P_1(t, x_{ki}, z_{ki}) \right\}, \\ \psi_{\eta,i}(t) &= \int_0^\tau \left\{ \sum_{l,j} \frac{D_{\eta,lj}(t)w_{lj}(t)\Delta_{lj}N_{lj}(t)}{G_c(T_{lj})} \mathcal{I}(s \leq T_{lj} \leq t) \right\} \frac{dM_{ki}^C(s)}{Y(s)}. \end{aligned}$$

Both $I_\eta(t)$ and $W_{\eta,ki}(t)$ can be consistently estimated by the plug-in estimators, denoted as $\hat{I}_\eta(t)$ and $\hat{W}_{\eta,ki}(t)$, and the variance of $K^{\frac{1}{2}}\{\hat{\eta}(t) - \eta(t)\}$ can be estimated by $\hat{\Sigma}_\eta(t) = K\{\hat{I}_\eta(t)\}^{-1} [\sum_{k=1}^K \{\sum_{i=1}^{n_k} \hat{W}_{\eta,ki}(t)\}^{\otimes 2}] \{\hat{I}_\eta(t)\}^{-1}$.

On the asymptotic distribution of $K^{\frac{1}{2}}(\hat{\alpha} - \alpha)$

To establish the asymptotic distribution for $K^{\frac{1}{2}}(\hat{\alpha} - \alpha)$, we consider the decomposition

$$\begin{aligned} U_\alpha(\alpha, \hat{\eta}, \hat{\gamma}, \hat{G}_c) &= \int_0^\tau \sum_{k=1}^K \sum_{i,j \in H_k, i \neq j} D_{\alpha,ki}(t, \alpha, \hat{\eta}, \hat{\gamma}) \{V_{ki}(t) - v_{ki}(t, \alpha, \eta, \gamma)\} dt \\ &\quad + \int_0^\tau \sum_{k=1}^K \sum_{i,j \in H_k, i \neq j} D_{\alpha,ki}(t, \alpha, \hat{\eta}, \hat{\gamma}) \{\hat{V}_{ki}(t) - V_{ki}(t)\} dt \\ &\quad - \int_0^\tau \sum_{k=1}^K \sum_{i,j \in H_k, i \neq j} D_{\alpha,ki}(t, v_k, \hat{\eta}, \hat{\gamma}) \{v_{ki}(t, \alpha, \hat{\eta}, \hat{\gamma}) - v_{ki}(t, \alpha, \eta, \gamma)\} dt. \end{aligned} \quad (\text{A6})$$

Since $\{\hat{G}_c(T_{ki})\}^{-1} \{\hat{G}_c(T_{ki}) - G_c(T_{ki})\} = -\sum_{l=1}^K \sum_{m=1}^{n_l} \int_0^\tau \mathcal{I}(s \leq T_{ki}) \{Y(s)\}^{-1} dM_{lm}^C(s) + o_p(K^{-\frac{1}{2}})$, the second term of $U_\alpha(\alpha, \hat{\eta}, \hat{\gamma}, \hat{G}_c)$ in (A6) equals

$$\begin{aligned} &2 \int_0^\tau \sum_{k=1}^K \sum_{i,j \in H_k, i \neq j} D_{\alpha,ki}(t, \alpha, \eta, \gamma) V_{ki}(t) \sum_{l=1}^K \sum_{m=1}^{n_l} \int_0^\tau \mathcal{I}(s \leq T_{ki}) \frac{dM_{lm}^C(s)}{Y(s)} dt + o_p(K^{\frac{1}{2}}) \\ &= \sum_{l=1}^K \sum_{m=1}^{n_l} \int_0^\tau \int_0^\tau q_\alpha(s, t) dM_{lm}^C(s) dt + o_p(K^{\frac{1}{2}}), \end{aligned} \quad (\text{A7})$$

where $q_\alpha(s, t) = 2\{\sum_{k=1}^K \sum_{i,j \in H_k, i \neq j} D_{\alpha,ki}(t, \alpha, \eta, \gamma) V_{ki}(t) \mathcal{I}(s \leq T_{ki})\} \{Y(s)\}^{-1}$.

By a Taylor expansion at the true values of $\{\eta(t), \gamma\}$, the third term of (A6) equals

$$K \int_0^\tau q_x(t, \alpha, \eta, \gamma) \{\hat{\eta}(t) - \eta(t)\} dt + K \int_0^\tau q_z(t, \alpha, \eta, \gamma) (\hat{\gamma} - \gamma) dt + o_p(K^{\frac{1}{2}}), \quad (\text{A8})$$

where

$$q_x(t, \alpha, \eta, \gamma) = K^{-1} \sum_{k=1}^K \sum_{i,j \in H_k, i \neq j} D_{\alpha, kij}(t, \alpha, \eta, \gamma) \left\{ \frac{\partial v_{kij}}{\partial \Lambda_{ki}(t)} x_{ki}^\top + \frac{\partial v_{kij}}{\partial \Lambda_{kj}(t)} x_{kj}^\top \right\},$$

$$q_z(t, \alpha, \eta, \gamma) = K^{-1} \sum_{k=1}^K \sum_{i,j \in H_k, i \neq j} D_{\alpha, kij}(t, \alpha, \eta, \gamma) \left\{ \frac{\partial v_{kij}}{\partial \Lambda_{ki}(t)} z_{ki}^\top + \frac{\partial v_{kij}}{\partial \Lambda_{kj}(t)} z_{kj}^\top \right\} t.$$

By a Taylor approximation,

$$K^{\frac{1}{2}}(\hat{\alpha} - \alpha) = - \left\{ K^{-1} \frac{\partial U_\alpha(\alpha, \eta, \gamma, G_c)}{\partial \alpha} \right\}^{-1} K^{-1/2} U_\alpha(\alpha, \hat{\eta}, \hat{\gamma}, \hat{G}_c) + o_p(1).$$

Let $\mathcal{I}_\alpha(\alpha, \eta, \gamma) = \sum_{k=1}^K \sum_{i,j \in H_k, i \neq j} \int_0^\tau \{\partial v_{kij}(t, \alpha, \eta, \gamma) / \partial \alpha\}^{\otimes 2} dt$. It follows from equations (A4), (A5), (A6), (A7) and (A8) that

$$K^{\frac{1}{2}}(\hat{\alpha} - \alpha) = K^{-1/2} \left\{ K^{-1} \mathcal{I}_\alpha(\alpha, \eta, \gamma) \right\}^{-1} \sum_{k=1}^K W_{\alpha, k}(\alpha, \eta, \gamma, G_c) + o_p(1),$$

where

$$\begin{aligned} W_{\alpha, k}(\alpha, \eta, \gamma, G_c) = & \sum_{i,j \in H_k, i \neq j} \int_0^\tau D_{\alpha, kij}(t, v_k, \eta, \gamma) \{V_{kij}(t) - v_{kij}(t, \alpha, \eta, \gamma)\} dt \\ & + \sum_{i=1}^{n_k} \int_0^\tau \int_0^\tau q_\alpha(s, t) dM_{ik}^C(s) dt \\ & - \sum_{i=1}^{n_k} \int_0^\tau q_x(t, \alpha, \eta, \gamma) n_k \{I_\eta(t)\}^{-1} W_{\eta, ki}(t) dt \\ & - \sum_{i=1}^{n_k} \int_0^\tau q_z(t, \alpha, \eta, \gamma) n_k I_\gamma^{-1} W_{\gamma, ki}(\tau) dt, \end{aligned}$$

and $I_\eta(t)$, $W_{\eta, ki}(t)$, I_γ and $W_{\gamma, ki}(\tau)$ are given following (A3), (A4) and (A5). By the central limit theorem $K^{\frac{1}{2}}(\hat{\alpha} - \alpha)$ is asymptotically normal with mean zero and a variance that can be estimated by $\hat{\Sigma}_\alpha = K \hat{\mathcal{I}}_\alpha^{-1} \sum_{k=1}^K \{\hat{W}_{\alpha, k}\}^{\otimes 2} \hat{\mathcal{I}}_\alpha^{-1}$, where $\hat{\mathcal{I}}_\alpha$ and $\hat{W}_{\alpha, k}$ are the empirical counterparts of $\mathcal{I}_\alpha(\alpha, \eta, \gamma)$ and $W_{\alpha, k}(\alpha, \eta, \gamma, G_c)$, respectively, obtained by inserting the estimators of α , $\eta(t)$, γ and by replacing $\hat{M}_{ki}^C(t)$ for $M_{ki}^C(t)$.

REFERENCES

- BANDEEN-ROCHE, K. & LIANG, K.-Y. (2002). Modelling multivariate failure time associations in the presence of a competing risk. *Biometrika* **89**, 299–314.
- BANDEEN-ROCHE, K. & NING, J. (2008). Nonparametric estimation of bivariate failure time associations in the presence of a competing risk. *Biometrika* **95**, 221–32.
- BEYERSMANN, J. & SCHUMACHER, M. (2007). Letter to the editor: misspecified regression model for the subdistribution hazard of a competing risk. *Statist. Med.* **26**, 1649–52.
- CAL, J. & PRENTICE, R. L. (1995). Estimating equations for hazard ratio parameters based on correlated failure time data. *Biometrika* **82**, 151–64.

- CHEN, B. E., KRAMER, J. L., GREENE, M. H. & ROSENBERG, P. S. (2008). Competing risks analysis of correlated failure time data. *Biometrics* **64**, 172–79.
- CHENG, Y. & FINE, J. (2008). Nonparametric estimation of cause-specific cross hazard ratio with bivariate competing risks data. *Biometrika* **95**, 233–40.
- CHENG, Y., FINE, J. P. & KOSOROK, M. R. (2007). Nonparametric association analysis of bivariate competing-risks data. *J. Am. Statist. Assoc.* **102**, 1407–15.
- CLAYTON, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141–51.
- CLAYTON, D. G. & CUZICK, J. (1985). Multivariate generalisations of the proportional hazards model. *J. R. Statist. Soc. A* **148**, 82–117.
- DABROWSKA, D. M. (1988). Kaplan–Meier estimate on the plane. *Ann. Statist.* **16**, 1475–89.
- FINE, J. P. & GRAY, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *J. Am. Statist. Assoc.* **94**, 496–509.
- FRANK, M. J. (1979). On the simultaneous associativity of $f(x, y)$ and $x + y - f(x, y)$. *Aequationes Mathematicae* **19**, 194–226.
- GILL, R. D. (1993). Multivariate survival analysis. *Theory Prob. Appl.* **37**, 18–31.
- GLIDDEN, D. (2000). A two-stage estimator of the dependence parameter for the Clayton-Oakes model. *Lifetime Data Anal.* **6**, 141–56.
- GRAY, R. J. (1988). A class of k -sample tests for comparing the cumulative incidence of a competing risk. *Ann. Statist.* **16**, 1141–54.
- GUMBEL, E. J. (1960). Bivariate exponential distributions. *J. Am. Statist. Assoc.* **55**, 698–707.
- HOUGAARD, P. (2000). *Analysis of Multivariate Survival Data*. New York: Springer.
- KALBFLEISCH, J. D. & PRENTICE, R. L. (2002). *The Statistical Analysis of Failure Time Data*, 2nd ed. New York: Wiley.
- KATSAHIAN, S., RESCHE-RIGON, M., CHEVRET, S. & PORCHER, R. (2006). Analysing multicentre competing risks data with a mixed proportional hazards model for the subdistribution. *Statist. Med.* **25**, 4267–78.
- LATOUCHE, A., BOISSON, V., CHEVRET, S. & PORCHER, R. (2007). Misspecified regression model for the subdistribution hazard of a competing risk. *Statist. Med.* **26**, 965–74.
- OAKES, D. (1982). A model for association in bivariate survival data. *J. R. Statist. Soc. B* **44**, 414–22.
- OAKES, D. (1989). Bivariate survival models induced by frailties. *J. Am. Statist. Assoc.* **84**, 487–93.
- SCHEIKE, T. H. & ZHANG, M.-J. (2008). Flexible competing risks regression modelling and goodness-of-fit. *Lifetime Data Anal.* **14**, 464–83.
- SCHEIKE, T. H., ZHANG, M.-J. & GERDS, T. (2008). Predicting cumulative incidence probability by direct binomial regression. *Biometrika* **95**, 205–20.
- SKYTTH, A., PEDERSEN, N., KAPRIO, J., STAZI, M., HJELMBORG, J., IACHINE, I., VAUPEL, J. & CHRISTENSEN, K. (2003). Longevity studies in genomeutwin. *Twin. Res.* **6**, 448–54.
- SUN, Y. & WU, H. (2005). Semiparametric time-varying coefficients regression model for longitudinal data. *Scand. J. Statist.* **32**, 21–47.
- VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. Cambridge: Cambridge University Press.
- WEI, L. J., LIN, D. Y. & WEISSFELD, L. (1989). Regression analysis of multivariate incomplete failure time data by modelling marginal distributions. *J. Am. Statist. Assoc.* **84**, 1065–73.

[Received October 2008. Revised August 2009]