



Published in final edited form as:

Biometrics. 2011 June ; 67(2): 415–426. doi:10.1111/j.1541-0420.2010.01470.x.

Frailty-Based Competing Risks Model for Multivariate Survival Data

Malka Gorfine and

Faculty of Industrial Engineering and Management, Technion—Israel Institute of Technology
Technion City, Haifa 32000, Israel, gorfim@ie.technion.ac.il

Li Hsu

Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle,
Washington 98109-1024, U.S.A. lih@fhcrc.org

SUMMARY

In this work, we provide a new class of frailty-based competing risks models for clustered failure times data. This class is based on expanding the competing risks model of Prentice et al. (1978, *Biometrics* **34**, 541–554) to incorporate frailty variates, with the use of cause-specific proportional hazards frailty models for all the causes. Parametric and nonparametric maximum likelihood estimators are proposed. The main advantages of the proposed class of models, in contrast to the existing models, are: (1) the inclusion of covariates; (2) the flexible structure of the dependency among the various types of failure times within a cluster; and (3) the unspecified within-subject dependency structure. The proposed estimation procedures produce the most efficient parametric and semiparametric estimators and are easy to implement. Simulation studies show that the proposed methods perform very well in practical situations.

Keywords

Competing risks; Frailty model; Multivariate survival analysis; Nonparametric maximum likelihood estimator

1. Introduction

In many experimental settings, survival data arise in a clustered form. For example, in multicenter clinical trials, where failure times are observed for multiple patients in each center, and in family studies, where ages at onset of a disease are recorded for multiple members of the same family. Frequently, factors that vary by clusters may exert a powerful influence on study outcomes. Such cluster effects lead to dependencies among outcomes within each cluster. It is well known that inferences that ignore the clustering effect can be seriously misleading (Hougaard, 2000, and references therein). Frailty models (Clayton and Cuzick, 1985) are becoming increasingly popular for analyzing clustered survival data. In this approach, the frailties (the random effects) often enter the model multiplicatively to the hazard function in order to model the correlation among outcomes within a cluster.

Most of the published works dealing with clustered survival data concentrate on estimating the parameters of one type of event at a time, under the assumption that the censoring mechanism is independent of the event type of interest conditional on the covariates.

However, this assumption is violated when multiple types of events occur and follow-up of

some of the events is censored by the onset of the other events. For example, consider a breast cancer family study with BRCA1 mutation as a covariate. The BRCA1 mutation is known to markedly increase the risk of multiple cancers: breast, ovarian, and testis cancers (Risch et al., 2006). Thus, in estimation of breast cancer risk, for example, a subject may be censored due to death from ovarian cancer or other related causes that occur before the onset of breast cancer that might otherwise occur. It is well known (Prentice et al., 1978; Kalbfleisch and Prentice, 2002, Chapter 8, and references therein) that violation of the independent censoring mechanism as described above may produce biased estimates of cause-specific marginal hazard function. A possible way to address this problem is to use a competing risks framework (Kalbfleisch and Prentice, 2002) where the observable information for a subject is the time to *first* event among the possible competing events or censoring for noncompeting reasons.

Limited work has been done on clustered failure time data with competing risks. Bandeen-Roche and Liang (2002) presented statistical methods for analyzing multivariate failure times in the presence of competing risks with no covariates and under a multiplicative frailty effect. Their idea is to model the time to the earliest failure from *any* cause so that the overall hazard, which is the summation over the type-specific hazards, equals the standard frailty model with no competing risks. As a result, the type-specific hazards are being interpreted as the partition of the overall hazard to the various causes. The work of Chatterjee, Hartge, and Wacholder (2003) deals with competing risks kin-cohort data where the covariate of the relatives is unobserved. A piecewise constant nonparametric model of the cause-specific hazard functions with dichotomous covariate is used. The cause-specific hazard function is of Prentice et al. (1978). The estimation technique is based on the “composite-likelihood” approach, where it is assumed that the relatives are independent given the proband’s covariate. By using the composite-likelihood, the problem is being simplified to a competing risks problem with no clustered data, and that can be solved by the method presented in Prentice et al. (1978). Hence, the cause-specific hazards are consistently estimated, however, the dependence of various competing risks among cluster members, which is of interest by itself, is not estimated. Katki, Chen, and Parmigiani (2007) considered a competing risks model for family data with the use of positive-stable copula model for modeling the dependence among the possible outcomes. However, this specific choice of copula function induces identical dependence parameters for any pair of outcomes, an assumption that would likely be violated. Chen et al. (2008) developed a nonparametric estimator of the cumulative incidence functions under a cause-specific hazard model of Prentice et al. (1978) with no covariates. The estimators are based on the nonparametric Nelson–Aalen (Aalen, 1978) estimator of each cause-specific hazard function. It is shown that the estimators are consistent to the true cumulative incidence functions under the competing risks model even if event times within cluster are correlated. As in Chatterjee et al. (2003), the dependence of various competing risks among cluster members is not estimated.

In this work, we propose a new class of flexible frailty models for competing risks analysis of clustered survival data with covariates, and proportional hazards frailty model for each failure type. Flexibility is in the manner of a flexible correlation structure being allowed among failure types within a cluster. It will be shown that this class of models includes the model of Bandeen-Roche and Liang (2002) as a special case. Section 2 describes the model and Section 3 provides examples for the frailty process. The likelihood function and the parametric and nonparametric maximum likelihood estimators (parametric maximum likelihood estimators (PMLEs) and nonparametric maximum likelihood estimators (NPMLEs)) are presented in Section 4. Simulation results are presented in Section 5, where the proposed PMLEs and NPMLEs are being contrasted with a “naive” approach in which

the competing risks are mistakenly considered as independent censoring. A data example is given in Section 6 and a short discussion is given in Section 7.

2. The Proposed Model

We adopt the competing risks model of Prentice et al. (1978) and Kalbfleisch and Prentice (2002) with the required modifications under a frailty model setting. By this, we avoid determining the joint distribution of all failure types for a subject, as required by the approach that uses latent or potential failure times. As a result, our current approach will not address the question, for example, whether the failure times of different failure types within a subject are associated, as we are mainly concerned about the cluster-level associations of failure times of the same or different failure types between subjects. For details regarding completely modeling the probability laws for competing failure times, readers are referred to Crowder (2001) and references therein.

Consider n independent clusters, with cluster i , $i = 1, \dots, n$, having m_i members. For subject k of cluster i , let T_{ik}^o and C_{ik} be the time to the *first* failure and the censoring time, respectively, and \mathbf{Z}_{ik} be a vector of time-independent covariates. For simplicity, we assume time-independent covariates, but all the developments presented in this article can be extended readily to the case of time-dependent covariates. Let $J_{ik} \in \{1, \dots, L\}$ be the type of the first observed failure and $J_{ik} = 0$ if censored. Define

$T_{ik} = \min(T_{ik}^o, C_{ik})$ and $\delta_{ik} = I(T_{ik}^o \leq C_{ik})$. For cluster i , let $\boldsymbol{\varepsilon}_i(t) = \{\varepsilon_{i1}(t), \dots, \varepsilon_{iL}(t)\}$ be a vector-valued random process, $t \in [0, \tau]$ where τ is the maximal follow-up time. Define \mathcal{F}_{it} to be the σ -algebra generated by $\boldsymbol{\varepsilon}_i(t)$ through its history up to and including time t : $\mathcal{F}_{it} = \{\boldsymbol{\varepsilon}_i(u), 0 \leq u \leq t\}$.

The overall hazard function of the k th member of cluster i at time t , given the regression vector \mathbf{Z}_{ik} and the cluster-level σ -algebra \mathcal{F}_{it} , is given by

$$\lambda_{ik}(t | \mathbf{Z}_{ik}, \mathcal{F}_{it}) = \lim_{h \downarrow 0} \frac{1}{h} P_r \left(t \leq T_{ik}^o < t+h \mid T_{ik}^o \geq t, \mathbf{Z}_{ik}, \mathcal{F}_{it} \right).$$

The cause-specific hazard functions are defined by

$$\begin{aligned} \lambda_{jik}(t | \mathbf{Z}_{ik}, \mathcal{F}_{it}) &= \lim_{h \downarrow 0} \frac{1}{h} P_r \left(t \leq T_{ik}^o < t+h, \times J_{ik} = j \mid T_{ik}^o \geq t, \mathbf{Z}_{ik}, \mathcal{F}_{it} \right) \\ &= \lambda_{0j}(t) \exp \left\{ \boldsymbol{\beta}_j^T \mathbf{Z}_{ik} + \varepsilon_{ij}(t) \right\} \end{aligned}$$

for $j = 1, \dots, L$, where $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_L$ and $\lambda_{01}(\cdot), \dots, \lambda_{0L}(\cdot)$ are cause-specific regression coefficients' vectors and cause-specific baseline hazard functions, respectively. Although the vector of covariates \mathbf{Z}_{ik} is the same for all the failure types, the regression coefficients' vector can vary among the failure types. Thus, a coefficient can be set to 0 in case its respective covariate is absent for a particular event. For simplicity of presentation, we assume that the regression coefficients' vectors and the cause-specific baseline hazard functions are identical among a cluster's members.

If only one of the failure types can occur, then $\lambda_{ik}(t | \mathbf{Z}_{ik}, \mathcal{F}_{it}) = \sum_{j=1}^L \lambda_{jik}(t | \mathbf{Z}_{ik}, \mathcal{F}_{it})$. It follows that the cause-specific densities equal $f_{jik}(t | \mathbf{Z}_{ik}, \mathcal{F}_{it}) = \lambda_{jik}(t | \mathbf{Z}_{ik}, \mathcal{F}_{it}) S_{ik}(t | \mathbf{Z}_{ik}, \mathcal{F}_{it})$, $j = 1, \dots, L$, where $S_{ik}(t | \mathbf{Z}_{ik}, \mathcal{F}_{it}) = \exp \left\{ - \int_0^t \sum_{j=1}^L \lambda_{jik}(u | \mathbf{Z}_{ik}, \mathcal{F}_{it}) du \right\}$. Hence, the

likelihood function can be written in terms of the cause-specific hazards. For the cause-specific survival function, we have $S_{jik}(t | \mathbf{Z}_{ik}, \mathcal{F}_{iu}) = 1 - \int_0^t f_{jik}(u | \mathbf{Z}_{ik}, \mathcal{F}_{iu}) du$.

The associations between outcomes of members of cluster i are induced by the latent cluster-specific frailty process history up to the cluster's maximum observed time, denoted by \mathcal{F}_{iu} . Conditional on the cluster's frailty process history and the observed covariates, the survival times within cluster i are assumed to be independent. At any given time $\mathbf{t}_i = (t_{i1}, \dots, t_{im_i})$, $\boldsymbol{\varepsilon}_i(\mathbf{t}_i) = \{\boldsymbol{\varepsilon}_i(t_{i1})^T, \dots, \boldsymbol{\varepsilon}_i(t_{im_i})^T\}^T$, $i = 1, \dots, n$ are assumed to be independent with density denoted by $f\{\boldsymbol{\varepsilon}_i(\mathbf{t}_i) | \theta(\mathbf{t}_i)\}$ where $\theta(\mathbf{t})$ is a vector of unknown parameters.

A multivariate competing risks model involves three types of dependence: (i) Dependence of failure times of the same failure type between cluster members (e.g., mother and her daughter with breast cancer). (ii) Dependence of failure times of different failure types between cluster members (e.g., mother with breast cancer and her daughter with ovarian cancer). (iii) Dependence of failure times of different failure types within a subject (e.g., a mother with breast cancer and ovarian cancer). With the above proposed formulation, associations of types (i) and (ii) are determined by the random process $\boldsymbol{\varepsilon}_i(t)$. The dependency of type (iii) could come from different sources. Each source can be classified into one of two types: (a) risks that are being shared by all the family members (e.g., exposure to air pollution), and (b) risks that are subject-specific (e.g., average number of cigarettes smoked per day). The type (a) risks are at the cluster level, and those of type (b) are at the subject level. The cluster-level dependency is defined by $\boldsymbol{\varepsilon}_i(t)$, however, the subject-level dependency structure is left unspecified.

3. Examples

Various models can be considered for the distribution of $\boldsymbol{\varepsilon}_i(t)$, $i = 1, \dots, n$. Each model imposes different dependency structure within a cluster. Four examples are presented below.

Example 1: In the special case where $\boldsymbol{\beta}_j = \mathbf{0}$ for all $j = 1, \dots, L$, $\lambda_{0j}(t) = \lambda_0(t)$ and $\exp\{\boldsymbol{\varepsilon}_{ij}(t)\} = \omega_j v_{ij}(t)$ for all $j = 1, \dots, L$ and $t \in [0, \tau]$ such that ω_j is a single nonnegative random variable and independent of the nonnegative random vector $\{v_{i1}(t), \dots, v_{iL}(t)\}$, the above cause-specific hazard model becomes the model provided by Bandeen-Roche and Liang (2002, equation (6)).

Example 2: If we impose $\boldsymbol{\varepsilon}_i(t) = \boldsymbol{\varepsilon}_i$, $\boldsymbol{\varepsilon}_{ij} = \boldsymbol{\varepsilon}_i$ for all $t \in [0, \tau]$ and $j = 1, \dots, L$ and let $\boldsymbol{\varepsilon}_i$ be a positive-stable random variable, then the model reduces to the one considered by Katki et al. (2007) with identical dependence parameters for any pair of outcomes.

Example 3: As a practical motivating example, consider the case where for each $j = 1, \dots, L$, $\boldsymbol{\varepsilon}_{ij}(t) = \boldsymbol{\varepsilon}_{ij}^{(1)}$ if the observed time is smaller than or equal to a pre-specified constant, $t_o^{(j)}$, and $\boldsymbol{\varepsilon}_{ij}(t) = \boldsymbol{\varepsilon}_{ij}^{(2)}$ otherwise. Let $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1^{(1)}, \boldsymbol{\varepsilon}_1^{(2)}, \dots, \boldsymbol{\varepsilon}_L^{(1)}, \boldsymbol{\varepsilon}_L^{(2)})^T$ with density function $f(\boldsymbol{\varepsilon} | \theta)$ and unknown vector of parameters θ . Then, for each family i , $\boldsymbol{\varepsilon}_i$ is being determined by the observed data of the family and its distribution, $f(\boldsymbol{\varepsilon}_i | \theta_i)$ —the marginal distribution of $f(\boldsymbol{\varepsilon} | \theta)$, since $\boldsymbol{\varepsilon}_i$ and θ_i are subvectors of $\boldsymbol{\varepsilon}$ and θ , respectively (including $\boldsymbol{\varepsilon}$ and θ). In family studies, this corresponds to a situation that early onset cases (e.g., breast or ovarian cancers) may have a different etiology from cases with late onset.

Example 4: Let $\boldsymbol{\varepsilon}_i(t) \equiv \boldsymbol{\varepsilon}_i$ for all $t \in [0, \tau]$, and consider a multivariate normal distribution for $\boldsymbol{\varepsilon}_i$, which provides a flexible and unrestricted covariance matrix. Since this case is considered in the simulation study and the real data analysis, we explore our proposed model under the multivariate normal frailty distribution, by using the cause-specific cross-ratio of Bandeen-Roche and Liang (2002). This provides us a simple tool in understanding the dependency imposed by this specific frailty model.

The cross-ratio function of Bandeen-Roche and Liang (2002) can be interpreted as a local dependency measure of two failure times of different types. Specifically, let (T_{i1}^o, T_{i2}^o) and (J_{i1}, J_{i2}) be survival times and failure types of two cluster members, respectively. The cause-specific cross-ratio of T_{i1}^o and T_{i2}^o of types $J_{i1} = j_1$ and $J_{i2} = j_2$ is defined as

$$CR_{j_1, j_2}(t_1, t_2) = \frac{S(t_1, t_2) f(t_1, t_2, j_1, j_2)}{\int_{t_2}^{\infty} \sum_{j=1}^L f(t_1, u, j_1, j) du \int_{t_1}^{\infty} \sum_{j=1}^L f(u, t_2, j, j_2) du}$$

where $S(\cdot, \cdot)$ denotes the survival function of (T_{i1}^o, T_{i2}^o) and $f(\cdot, \cdot, j_1, j_2)$ denotes the density function of (T_{i1}^o, T_{i2}^o) of types (j_1, j_2) . Hence, $CR_{j_1, j_2}(t_1, t_2)$ is the ratio of the risk of one member to fail at t_1 due to cause j_1 given the other member failed at t_2 due to cause j_2 to the risk given the other member not yet failed at all by t_2 . $CR_{j_1, j_2}(t_1, t_2) = 1$ indicates independence of (T_{i1}^o, T_{i2}^o) of j_1 and j_2 failure types at (t_1, t_2) , and $CR_{j_1, j_2}(t_1, t_2) > 1 (< 1)$ indicates positive (negative) correlated (T_{i1}^o, T_{i2}^o) of j_1 and j_2 failure types at (t_1, t_2) . Straightforward algebra shows that under the multivariate normal frailty model, if $corr(\epsilon_{j_1}, \epsilon_{j_2}) = 0$, $CR_{j_1, j_2}(t_1, t_2) = 1$ for all $t_1, t_2 \in [0, \tau]$. Figure 1 depicts the above cross-ratio function of two failure types, $j_1 = 1$ and $j_2 = 2$, under the multivariate normal frailty model with mean zero and various covariance matrices. For simplicity, there are no covariates in the model and each conditional hazard function at time t , given the frailty variates, equals $1/(1-t)0 \leq t \leq 1$. It is evident that the cross ratio is a nonlinear decreasing function of t_1 and t_2 as often the situation in family studies (Fan, Prentice, and Hsu, 2000). The cross-ratio function is also a nonlinear increasing function of the variances/covariance components of the covariance matrix. The above conclusions hold for any multivariate normal frailty model, and the hazard functions and the regression coefficients affect the slopes of the CR curves. It should also be noted that the covariance components of the covariance matrix are allowed to be negative, yielding $CR_{j_1, j_2}(t_1, t_2) < 1$.

4. The Likelihood Function, PMLEs, and NPMLEs

Denote the observed data of cluster $i, i = 1, \dots, n$, by $(\mathbf{T}_i, \mathbf{Z}_i, \mathbf{J}_i, \boldsymbol{\delta}_i)$ where $\mathbf{T}_i = (T_{i1}, \dots, T_{im_i})$, $\mathbf{Z}_i = (\mathbf{Z}_{i1}^T, \dots, \mathbf{Z}_{im_i}^T)^T$, $\mathbf{J}_i = (J_{i1}, \dots, J_{im_i})$, and $\boldsymbol{\delta}_i = (\delta_{i1}, \dots, \delta_{im_i})$. For the construction of the likelihood function, we assume the following: (1) Conditional on \mathbf{Z}_i and \mathcal{F}_{it} , the censoring times are independent of the failure times and noninformative for the frailty process and $\{\beta_j, \lambda_{0j}, j = 1, \dots, L\}$. (2) The frailty process is independent of \mathbf{Z}_i . (3) The covariates' effect is subject specific, namely $f_{T_{ik}, J_{ik}}(\mathbf{Z}_i, \boldsymbol{\epsilon}_i, t, j | \mathbf{Z}_i, \mathcal{F}_{it}) = f_{T_{ik}, J_{ik}}(\mathbf{Z}_{ik}, \boldsymbol{\epsilon}_i, t, j | \mathbf{Z}_{ik}, \mathcal{F}_{it})$. Hence, the likelihood function, under any choice of distributions $f\{\boldsymbol{\epsilon}_i(\mathbf{t}_i) | \theta(\mathbf{t}_i)\}$, is proportional to

$$L = \prod_{i=1}^n \prod_{\mathbf{t}_i > \mathbf{0}} \int \prod_{k=1}^{m_i} \times [\lambda_{0j_{ik}}(t_{ik}) \exp\{\beta_{j_{ik}}^T \mathbf{Z}_{ik} + \epsilon_{j_{ik}}(t_{ik})\}]^{dN_{ik}(t_{ik})} \exp\left[-\sum_{j=1}^L \exp\{\beta_j^T \mathbf{Z}_{ik} + \epsilon_{ij}(t_{ik})\} \lambda_{0j}(t_{ik})\right] \times \{1 - dN_{ik}(t_{ik})\} Y_{ik}(t_{ik}) f\{\boldsymbol{\epsilon}_i(\mathbf{t}_i) | \theta(\mathbf{t}_i)\} d\boldsymbol{\epsilon}_i(\mathbf{t}_i), \tag{1}$$

where $N_{ik}(t) = \delta_{ik} I(T_{ik} \leq t)$ and $Y_{ik}(t) = I(T_{ik} \geq t)$ are the failure counting process and the at-risk process, respectively. Alternatively, let the counting processes $\{N_{ik}^{(j)}(t), t \geq 0\}$ be defined at time t by $N_{ik}^{(j)}(t) = N_{ik}(t) I(J_{ik} = j), j = 1, \dots, L$. Then, the above likelihood can also be written as

$$L = \prod_{i=1}^n \prod_{t_i > 0} \int \prod_{k=1}^{m_i} \prod_{j=1}^L \left[\lambda_{0j}(t_{ik}) \exp \left\{ \beta_j^T \mathbf{Z}_{ik} + \varepsilon_{ij}(t_{ik}) \right\} \right]^{dN_{ik}^{(j)}(t_{ik})} \exp \left[- \exp \left\{ \beta_j^T \mathbf{Z}_{ik} + \varepsilon_{ij}(t_{ik}) \right\} \lambda_{0j}(t_{ik}) dt_{ik} \times \left\{ 1 - dN_{ik}^{(j)}(t_{ik}) \right\} Y_{ik}(t_{ik}) \right] f \{ \boldsymbol{\varepsilon}_i(\mathbf{t}_i) \mid \theta(\mathbf{t}_i) \} \tag{2}$$

It is easy to see that if the $\varepsilon_{ij}(\cdot)$'s, $j = 1, \dots, L$, are independent, the above likelihood function factors into a separate components for each cause-specific hazard function $\lambda_j \{ t \mid \mathbf{Z}, \varepsilon_j(t) \}$. The j th factor is precisely the likelihood that would be obtained if failures of types other than j were regarded as being censored and the time-dependent frailty variate $\varepsilon_{ij}(\cdot)$ accounts for the unobservable cluster-level effect related to failure type j . This result is in the spirit of Prentice et al. (1978, Section 2). However, the case of independent $\varepsilon_{ij}(\cdot)$'s does not imply that the L risks are not competing with each other. It only means that the failure time of type j of one cluster member is independent of the failure time of type $k \neq j$ of any other cluster member. However, different failure types can still compete with each other at subject level.

Based on the above likelihood function, various estimation procedures can be proposed. In what follows, we provide two estimation techniques, one that yields PMLEs for the case where $\lambda_{01}(\cdot), \dots, \lambda_{0L}(\cdot)$ are of a parametric form, and another that provides NPMLEs when the baseline hazard functions are left unspecified.

4.1 PMLEs

In parametric regression models, where the baseline hazard functions follow a parametric form such as the generalized gamma, log-normal, or log-logistic, one may maximize the likelihood function (2) directly with respect to $\beta_j, j = 1, \dots, J$, the parameters of the frailty process and the parameters involved with $\Lambda_{0j}, j = 1, \dots, J$, by using numerical integration. Alternatively, we provide a simple EM-algorithm for computing the PMLEs: (1) Estimate the conditional expectation of $\exp\{\varepsilon_{ij}(t)\}$ given the observed data and the current parameters' values, denoted by $\hat{\mu}_{ij}(t)$. This conditional expectation can be written explicitly for each specific density $f\{\boldsymbol{\varepsilon}_i(\cdot) \mid \theta(\cdot)\}$ by using Bayes' theorem. (2) Estimate $\beta_j, j = 1, \dots, J$ and the parameters involved with $\lambda_{0j}(\cdot), j = 1, \dots, J$ by maximizing

$$\prod_{i=1}^n \prod_{t_i > 0} \prod_{k=1}^{m_i} \prod_{j=1}^L \left[\lambda_{0j}(t_{ik}) \exp \left\{ \beta_j^T \mathbf{Z}_{ik} + \log \hat{\mu}_{ij}(t_{ik}) \right\} \right]^{dN_{ik}^{(j)}(t_{ik})} \exp \left[- \exp \left\{ \beta_j^T \mathbf{Z}_{ik} + \log \hat{\mu}_{ij}(t_{ik}) \right\} \lambda_{0j}(t_{ik}) dt_{ik} \left\{ 1 - dN_{ik}^{(j)}(t_{ik}) \right\} Y_{ik}(t_{ik}) \right].$$

(3) Estimate $\theta(\cdot)$ by maximizing $\sum_{i=1}^n \hat{E}[\log f\{\boldsymbol{\varepsilon}_i(\cdot) \mid \theta(\cdot)\}]$, where \hat{E} is defined as the conditional expectation given the observed data and the current parameters' estimated values. Again, this conditional expectation can be written explicitly for each specific density by using Bayes' theorem. (4) Repeat Steps (1)–(3) until convergence is reached with respect to all the parameters' estimates.

Continuing Example 3 above, for each $i = 1 \dots, n, k = 1, \dots, m_i$, and $j = 1, \dots, L$, let

$$\tilde{\varepsilon}_{ikj} = \varepsilon_j^{(1)} I(T_{ik} \leq t_o^{(j)}) + \varepsilon_j^{(2)} I(T_{ik} > t_o^{(j)}), \delta_{ik}^{(j)} = \delta_{ik} I(J_{ik} = j) \text{ and}$$

$$V_i = \exp \left(\sum_{k=1}^{m_i} \sum_{j=1}^L \left[\tilde{\varepsilon}_{ikj} \delta_{ik}^{(j)} - \exp \left\{ \beta_j^T \mathbf{Z}_{ik} + \tilde{\varepsilon}_{ikj} \right\} \Lambda_{0j}(T_{ik}) \right] \right).$$

Then, the conditional expectation of $\exp\{\varepsilon_{ij}(t)\}$ becomes

$$\frac{\int \exp(\varepsilon_j^{(h)}) V_i f(\varepsilon_j^{(h)}, \varepsilon_i | \tilde{\theta}) d\varepsilon_j^{(h)} d\varepsilon_i}{\int V_i f(\varepsilon_i | \theta_i) d\varepsilon_i}$$

with $h = 1$ if $t \leq t_o^{(j)}$, and $h = 2$ otherwise, where $\tilde{\theta}$ is a subvector of θ and is being determined by the joint distribution of $(\varepsilon_j^{(h)}, \varepsilon_j)$. In addition, let $\theta^{(r-1)}$ be the current estimated values of θ . Then, the r th estimate of θ is the solution of the following equations:

$$\frac{1}{\int V_i f(\varepsilon_i | \theta_i^{(r-1)}) d\varepsilon_i} \sum_{i=1}^n \int \frac{\partial \log f(\varepsilon_i | \theta_i)}{\partial \theta_i} V_i f(\varepsilon_i | \theta_i^{(r-1)}) d\varepsilon_i = \mathbf{0}.$$

Usually, the integrals in the above calculations require numerical integration procedures such as Gaussian quadrature approximations.

It is well known (Kalbfleisch and Prentice, 2002, among many others) that parametric regression models often involve stronger distribution assumptions regarding the baseline hazard function than it is suitable to make. If the parametric models of the baseline hazard functions are incorrect, the above PMLEs are inconsistent. Therefore, in the following section, we provide an estimation technique for unspecified baseline hazard functions that is expected to suffer from some efficiency loss in comparison to the PMLEs, as long as the baseline hazards are correctly specified.

4.2 NPMLs

The following NPMLs is based on the seminal paper of Zeng and Lin (2007). That is, we regard the maximum partial likelihood estimators and the associated Breslow (1972) estimators of the cumulative baseline hazard functions as NPMLs in that they maximize the nonparametric likelihood where the cumulative baseline hazard functions are regarded as infinite dimensional parameters.

Specifically, the NPMLs can be obtained by the following EM algorithm. For $j = 1, \dots, L$ define

$$S_j^{(0)}(\beta_j, t) = \sum_{i=1}^n \sum_{k=1}^{m_i} Y_{ik}(t) \exp\{\beta_j^T \mathbf{Z}_{ik} + \varepsilon_{ij}(t)\}$$

and

$$S_j^{(1)}(\beta_j, t) = \sum_{i=1}^n \sum_{k=1}^{m_i} Y_{ik}(t) \mathbf{Z}_{ik} \exp\{\beta_j^T \mathbf{Z}_{ik} + \varepsilon_{ij}(t)\}.$$

In the M-step, we solve the complete-data score equation conditional on the observed data and the current parameter estimates. Hence for the estimation of $\mathcal{B} = (\beta_1^T, \dots, \beta_L^T)$ we solve

$$\sum_{i=1}^n \sum_{k=1}^{m_i} \int_0^{\infty} \left[\mathbf{Z}_{ik} - \frac{\widehat{S}_j^{(1)}(\beta_j, t)}{\widehat{S}_j^{(0)}(\beta_j, t)} \right] dN_{ik}^{(j)}(t) = \mathbf{0}, j=1, \dots, L, \quad (3)$$

where $\widehat{S}_j^{(m)}(\beta_j, t), m=0, 1$ are defined analogously to $S_j^{(m)}(\beta_j, t)$ while replacing the unknown $\exp\{\varepsilon_{ij}(t)\}$ by its conditional expectation given the observed data and the current parameters' values, $\hat{\mu}_{ij}(t)$ of Step (1) of the PMLEs algorithm.

For the estimation of the cumulative baseline hazard functions,

$\Lambda_{0j}(t) = \int_0^t \lambda_{0j}(u) du, j=1, \dots, L$, we define the estimator of the j th cumulative baseline hazard by a step function with jumps at the observed failure times of type j , and is given by

$$\widehat{\Lambda}_{0j}(t) = \int_0^t \frac{\sum_{i=1}^n \sum_{k=1}^{m_i} dN_{ik}^{(j)}(s)}{\widehat{S}_j^{(0)}(\beta_j, s)} j=1, \dots, L. \quad (4)$$

Finally, we estimate $\theta(\cdot)$ as in Step (3) of the PMLEs algorithm.

To summarize, the following is our proposed NPMLE estimation procedure: (i) Given the values of θ and $\Lambda_{0j}, j=1, \dots, L$, estimate $\beta_j, j=1, \dots, L$, by solving (3). (ii) Given the values of θ and $\beta_j, j=1, \dots, L$, estimate $\Lambda_{0j}, j=1, \dots, L$, by using (4). (iii) Given the values of β_j and $\Lambda_{0j}, j=1, \dots, L$, estimate θ by maximizing $\sum_{i=1}^n \widehat{E}[\log f\{\varepsilon_i(\cdot) | \theta(\cdot)\}]$. (iv) Repeat Steps (i)–(iii) until convergence is reached with respect to all the parameters' estimates.

Denote the true values of β, θ , and $\mathcal{A} = (\Lambda_{01}, \dots, \Lambda_{0L})$ by β^o, θ^o , and \mathcal{A}^o and their NPMLEs by $\hat{\beta}, \hat{\theta}$, and $\hat{\mathcal{A}}$. Based on Zeng and Lin (2007), it can be shown that for bounded cluster size m_i as the total number of clusters n goes to infinity: (i) $\hat{\beta}$ is strongly consistent for β^o . (ii) $\hat{\theta}$ is strongly consistent for θ^o . (iii) $\hat{\mathcal{A}}(\cdot)$ uniformly converges to $\mathcal{A}^o(\cdot)$. (iv) $n^{1/2}\{\hat{\beta} - \beta^o, \hat{\theta} - \theta^o, \hat{\mathcal{A}} - \mathcal{A}^o\}$ converges weakly to a zero-mean Gaussian process, and the limiting covariance matrix of $\{\hat{\beta}, \hat{\theta}\}$ achieves the semiparametric efficiency bound. A detailed list of the technical conditions required for the above asymptotic results can be found in Zeng and Lin (2007, Appendix B, Conditions D1–D4 and D6–D8).

For the variance and covariance estimates, we treat the likelihood as a parametric likelihood of the regression coefficients and the jump sizes of the baseline hazard functions. Hence, the inverted observed information matrix is the variance–covariance estimated matrix of all the parameters' estimates $\{\hat{\beta}, \hat{\theta}, \hat{\mathcal{A}}\}$. However, in practice, this procedure often does not work because of the large dimension of the observed information matrix. When the interest lies only in the regression coefficients and the frailty process parameters, the profile likelihood approach of Murphy and van der Vaart (2000) is recommended. For estimating the variance of the cumulative baseline hazard estimator as well, the bootstrap approach is recommended (e.g., Kosorok, Lee, and Fine, 2004), with cluster as the sampling unit.

5. Simulation Study

The following simulation studies are aimed to evaluate the finite-sample properties of the proposed model and estimation procedure. We considered $L = 2$ competing risks, 500 and 1000 clusters and two members within each cluster. The cluster-level dependency was constructed through a zero-mean multivariate normal distribution of the frailty variate $\varepsilon(t) \equiv$

$\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2)$ with $\theta = (\sigma_1^2, \sigma_2^2, \rho_{12}) = (1, 1.5, 0.5)$. In addition, we considered a single $U[0, 1]$ distributed covariate with regression coefficients $\beta_1 = 0.5$ and $\beta_2 = 2.5$ for the two failure types, respectively. Given $(\beta_1, \beta_2, \varepsilon_1, \varepsilon_2)$ and the covariate, we generated the earliest failure time, T^o , of each cluster member, from an exponential distribution with parameter $\gamma = \sum_j \exp(\beta_j Z + \varepsilon_j)$. Note that $Pr(J = j | T^o = t) = \gamma_j / \gamma$ where $\gamma_j = \exp(\beta_j Z + \varepsilon_j)$. Hence, for generating the failure type of each cluster member, a Bernoulli random value, x , was sampled with probability of success γ_1 / γ , so that $x = 1 (x = 0)$ implies $J = 1 (J = 2)$. In addition, independent censoring times were generated from $U[0, 0.3]$. The integrals were approximated by a two-dimensional Gauss–Hermite quadrature with 20×20 function evaluations from a one-dimensional Gauss–Hermite quadrature. Results are based on 200 simulated data sets. Table 1 summarizes the censoring rates and the average number of observed events under each scenario.

We contrasted the proposed PMLEs and NPMLEs with those of the naive analysis that treats competing risks as independent censoring (see Appendix for details). It should be noted that the naive estimation procedure coincides with the proposed PMLE and NPMLE in case the frailty variates $\varepsilon_j, j = 1, \dots, J$ are independent, which in the current context, $\rho_{12} = 0$.

The results are presented in Table 2 which include the empirical means and standard deviations of the estimates. It is evident that both PMLEs and NPMLEs perform very well in terms of bias and PMLEs are generally more efficient than NPMLEs. In contrast, the naive approach suffers from substantially biased estimate of the regression coefficient β_1 and some bias in estimation of σ_1^2 and Λ_{01} . Almost no bias is presented in the estimates of β_2, σ_2^2 , and Λ_{02} . This can be explained by the sampling design in which we set $\beta_1 = 0.5$ and $\beta_2 = 2.5$, so that event type $j = 2$ is rarely censored by event type $j = 1$, while event type $j = 1$ is often censored by event type $j = 2$.

In order to evaluate the efficiency loss due to the unnecessary modeling of the dependency between the frailty variates, we provide, in Table 3, simulation results of the above setting but with $\rho_{12} = 0$ and $(\beta_1, \beta_2) = (1, 1)$ or $(\beta_1, \beta_2) = (0.5, 2.5)$. Under $(\beta_1, \beta_2) = (1, 1)$, we observed negligible efficiency loss, if any, with respect to all the parameters for both PMLEs and NPMLEs. With $(\beta_1, \beta_2) = (0.5, 2.5)$, where event of type $j = 2$ tends to censor the other event type but not vice versa, the NPMLE of β_1 corresponds to $j = 1$ suffers from some efficiency loss (relative efficiency is about 80%) in comparison to the other regression coefficient β_2 (relative efficiency is about 109%). In all the other cases, the efficiency loss, if any, is very small.

Next, we consider the robustness of the proposed model and the estimation technique when the frailty distribution of $\boldsymbol{\varepsilon}$ is misspecified. Specifically, we assumed a multivariate normal frailty model but the true frailty variates were generated from a multivariate skewed-normal distribution (Azzalini and Dalla Valle, 1996). An L -dimensional random variable \mathbf{X} is said to have a multivariate skewed-normal distribution if it is continuous with density function $2\phi_L(\mathbf{x}; \Omega)\Phi(\boldsymbol{\alpha}^T \mathbf{x})$ where $\mathbf{x} \in \mathcal{R}^L$, $\phi_L(\mathbf{x}; \Omega)$ is the L -dimensional normal density with zero mean and correlation matrix Ω , $\Phi(\cdot)$ is the standard normal distribution function, and $\boldsymbol{\alpha}$ is a L -dimensional vector. When the shape parameter $\boldsymbol{\alpha}$, which regulates skewness, equals zero, the above density reduces to the multivariate normal density with zero mean and correlation matrix Ω . Figures 2 and 3 provide a graphical demonstration of the discrepancy between the bivariate normal distribution and the skewed-normal distribution with $\boldsymbol{\alpha} = (50, 50)^T$ and correlation parameter of 0.8, used in the following simulation results. Table 4 presents simulation results under the skewed-normal distribution with shape parameter of $\boldsymbol{\alpha} = (50, 50)^T$ and the correlation parameter value of Ω equals 0.8. The rest of the sampling design is the same as of Table 2. The results indicate that even under misspecification of the frailty

distribution the PMLEs and NPMLEs of \mathcal{B} and \mathcal{A} based on the multivariate normal distribution, introduce only a small bias in contrast to the naive approach treating the other event as censored.

Table 5 aimed to study the effect of misspecification of the baseline hazard functions. The true baseline hazard functions are $\Lambda_{0j}(t) = \psi_j t^{2/3}$ with $\psi_j = 1, j = 1, 2$, and the analysis was performed as if $\Lambda_{0j}(t) = \psi_j t, j = 1, 2$. As expected, the PMLEs are severely biased with respect to all the parameters, while the NPMLEs are not.

In the Web Appendix we present additional simulation results for the following cases: (1) small number of clusters, such as 100; (2) two independent covariates; (3) two dependent covariates; and (4) skewed covariate. In short, it is evident that our methods performs reasonably well under these scenarios as well.

6. Real Data Analysis: The Washington Ashkenazi Kin-Cohort Study

To illustrate the utility of our proposed model and estimation procedure, we analyze data from the Washington Ashkenazi Kin-Cohort Study (WAS; Struwing et al., 1997). In this study, blood samples and questionnaires were collected from Ashkenazi Jewish men and women volunteers living in the Washington, DC area. Based on blood samples, volunteers were tested for specific mutations in BRCA1 and BRCA2 genes. The questionnaire included information on cancer and mortality history of the first-degree relatives of the volunteers.

Currently, we consider a subset of the data consisting of female first-degree relatives of volunteers (mother, sisters and daughters); two competing events: breast ($j = 1$) and ovarian ($j = 2$) cancers; and a binary covariate that takes a value of 1 or 0 according to presence or absence of any BRCA1/2 mutations in the volunteer's blood sample. The data consists of 4153 families with 2–8 relatives (12206 subjects). A total of 896 and 99 subjects were observed with breast cancer and ovarian cancer, respectively, as their first event. Only 10 families were observed with the two types of events among the relatives (exactly one event of each type). For the analysis of these data we used only the nonparametric approach as it is not obvious which parametric form would fit best for the misspecified baseline hazard functions, if any. Given our simulation results for the misspecified baseline hazard function, the NPMLEs may be more appropriate here.

Table 6 presents the NPMLEs based on the two-dimensional normal frailty model for the association of outcomes of family members. Namely, $(\varepsilon_{i1}, \varepsilon_{i2})^T$ are bivariate time-independent normally distributed with mean zero and the covariance matrix consists of the parameters $(\sigma_1^2, \sigma_2^2, \rho_{12})$. We used a time-independent distribution for the frailty process due to the small number of families with two types of events. The weighted bootstrap approach of Kosorok et al. (2004, Section 4) was used for calculating the standard errors and a total of 50 bootstrap data sets were generated. It is evident that the presence of any BRCA1/2 mutation in the volunteer's blood sample increases the risk of breast or ovarian cancers among her first-degree relatives, and the cumulative baseline hazard of breast cancer is much higher than that of ovarian cancer. In addition, the familial dependence of relatives' ovarian cancer ages at onset ($\widehat{\sigma}_2^2=1.7240$) is much stronger than that of breast cancer ($\widehat{\sigma}_1^2=0.8718$). The familial dependence between age at breast cancer onset of one relative and age at ovarian cancer onset of another relative was observed to be relatively weak ($\widehat{\rho}_{12} = 0.1247$). Without adjusting for the volunteer's BRCA1/2 mutations' status, the familial

Supplementary Materials

The Web Appendix, referred to in Section 5, is available under the Paper Information link at the *Biometrics* website <http://www.biometrics.tibs.org>.

dependencies for breast cancer, ovarian cancer, and the correlation between ovarian and breast cancer are 0.9404, 1.7527, and 0.2099, respectively, indicating that volunteer's BRCA1/2 mutations' status explained partly the familial dependence of breast and ovarian cancers and particularly the correlation between breast and ovarian cancers. Two possible naive analyses exist, where in each analysis the risk's parameters of one type of cancer are being estimated and the other type of cancer is being considered as independent censoring. In these data, the estimates of the naive models are fairly close to those based on the proposed competing risks model. For example, $\hat{\beta}_1$ and $\hat{\beta}_2$ were 1.4055 (SD 0.2239) and 1.5118 (SD 0.4695), respectively, and $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ were 0.9416 (SD 0.5285) and 1.6436 (SD 0.7454), respectively. This is probably because $\hat{\rho}_{12}$ is close to zero. However, it can be seen that the competing risks modeling provides a more comprehensive analysis of multiple failure types in cluster settings than the analysis that deals with one failure type at a time. It can also be seen that the efficiency loss due to estimation of ρ_{12} , when this parameter is close to zero, is very small. This result strengthens our confidence in the general usage of competing risks modeling for this type of data.

7. Discussion

The proposed class of models can be easily extended to the case in which the regression coefficients' vectors or the cause-specific hazard functions also vary among individuals within a cluster. For example, in a family study, mothers and daughters may not have the same regression coefficients or cause-specific baseline hazard functions.

The proposed PMLEs and NPMLEs can be easily implemented. For example, for NPMLEs, within each maximization step of the EM algorithm, the estimates of the regression coefficients and the cumulative baseline hazard functions can be computed by using built-in statistical packages such as `coxph` of R with an offset term.

Our proposed approaches require numerical integration for which many numerical routines have been developed. We used the Gauss–Hermite quadrature. For this method, it is generally recommended that the nodes are symmetric about zero and 15^L function evaluations from one-dimensional Gauss–Hermite quadrature usually provide a very good approximation in most applications, as long as $L \leq 5$. Most statistical and mathematical software packages provide one-dimensional Gauss–Hermite quadrature calculations, which can be easily adopted for the multivariate case.

Fine, Jiang, and Chappell (2001) proposed a semi-competing risks model, in which one event may censor another event, but not vice versa. In continuation of the BRCA1 example in the Introduction, the semi-competing risks model describes more precisely the situation where the onset of breast cancer does not prevent the death from ovarian cancer, but the death from ovarian cancer or any other related cause precludes the onset of breast cancer. The development of our multivariate frailty model to semi-competing risks situation is worth pursuing and could be a topic of future work.

Frailty models require the specification of the distribution of the random effects. Hougaard (2000) provides a comprehensive discussion of the theoretical properties and the fit of various frailty distributions for a noncompeting risks setting. Hsu, Gorfine, and Malone (2007) show by simulation that the bias in the marginal parameters, under a noncompeting risks setting, are generally 10% or lower under the assumed Gamma model and misspecification of the frailty distribution. We also showed here that our model is robust under the assumed multivariate normal frailty model when the true distribution is skewed. However, it is worthwhile to develop diagnostic approaches, tests, and graphical procedures

for model checking and model selection to ensure the fitted model is as close to the true model as possible.

Acknowledgments

This work is supported in part by grants from the National Institutes of Health (P01 CA53996 and R01 AG14358). We thank Nilanjan Chatterjee for facilitating the access of the Washington Ashkenazi Jewish data.

REFERENCES

- Aalen OO. Nonparametric inference for a family of counting processes. *Annals of Statistics*. 1978; 6:701–726.
- Azzalini A, Dalla Valle A. The multivariate skew-normal distribution. *Biometrika*. 1996; 83:715–726.
- Bandeem-Roche K, Liang KY. Modelling multivariate failure times associations in the presence of competing risk. *Biometrika*. 2002; 89:299–313.
- Breslow NE. Discussion of the paper by D. R. Cox. *Journal of the Royal Statistical Society, Series B*. 1972; 34:216–217.
- Chatterjee N, Hartge P, Wacholder S. Adjustment for competing risk in kin-cohort estimation. *Genetic Epidemiology*. 2003; 25:303–313. [PubMed: 14639700]
- Chen BE, Kramer JL, Greene MH, Rosenberg PS. Competing risks analysis of correlated failure time data. *Biometrics*. 2008; 64:172–179. [PubMed: 17680835]
- Clayton DG, Cuzick J. Multivariate generalizations of the proportional hazards model. *Journal of Royal Statistical Society, Series B*. 1985; 34:187–220.
- Crowder, M. *Classical Competing Risks*. London: Chapman & Hall/CRC; 2001.
- Fan J, Prentice RL, Hsu L. A class of weighted dependence measures for bivariate failure time data. *Journal of the Royal Statistical Society, Series B*. 2000; 62:181–190.
- Fine JP, Jiang H, Chappell R. On semi-competing risks data. *Biometrika*. 2001; 88:907–919.
- Hougaard, P. *Analysis of Multivariate Survival Data*. New York: Springer; 2000.
- Hsu L, Gorfine M, Malone K. On robustness of marginal regression coefficient estimates and hazard functions in multivariate survival analysis of family data when the frailty distribution is misspecified. *Statistics in Medicine*. 2007; 26:4657–4678. [PubMed: 17348081]
- Kalbfleisch, JD.; Prentice, RL. *The Statistical Analysis of Failure Time Data*. 2nd edition. New York: Wiley; 2002.
- Katki HA, Chen S, Parmigiani G. Multiple diseases in carrier probability estimation: Accounting for surviving all cancers other than breast and ovary in BRCA1/2. Working Paper 110. 2007
- Kosorok MR, Lee BL, Fine JP. Robust inference for univariate proportional hazards frailty regression models. *Annals of Statistics*. 2004; 32:1448–1491.
- Murphy SA, van der Vaart AW. On profile likelihood. *Journal of the American Statistical Association*. 2000; 95:449–465.
- Prentice RL, Kalbfleisch JD, Peterson AV Jr, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. *Biometrics*. 1978; 34:541–554. [PubMed: 373811]
- Risch HA, McLaughlin JR, Cole DEC, Rosen B, Bradley L, Fan I, Tang J, Li S, Zhang S, Shaw PA, Narod SA. Population BRCA1 and BRCA2 mutation frequencies and cancer penetrances: A kin-cohort study in Ontario, Canada. *Journal of the National Cancer Institute*. 2006; 98:1694–1706. [PubMed: 17148771]
- Struwing JP, Hartge P, Wacholder S, Baker SM, Berlin M, McAdams M, Timmerman MM, Lawrence BC, Tucker MA. The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. *New England Journal of Medicine*. 1997; 336:1401–1408. [PubMed: 9145676]
- Zeng D, Lin DY. Maximum likelihood estimation in semiparametric regression models with censored data. *Journal of the Royal Statistical Society, Series B*. 2007; 69:507–564.

APPENDIX

Estimation Procedures for the Naive PMLEs and NPMLEs

Consider the situation where the j th failure type is of interest and all other failure types are considered as independent censoring events, under a time-independent multivariate normally distributed frailty variate. The naive PMLEs and NPMLEs are obtained by the following steps: (1) Estimate $\exp\{\varepsilon_j\}$ by $\hat{\mu}_{ij}$ which are defined as

$$\frac{\int \exp(\varepsilon_j) \exp \left[\sum_{k=1}^{m_i} \{ \varepsilon_j \delta_{ik}^{(j)} - \exp(\beta_j^T \mathbf{Z}_{ik} + \varepsilon_j) \Lambda_{0j}(T_{ik}) \} \right] f(\varepsilon_j | \theta) d\varepsilon_j}{\int \exp \left[\sum_{k=1}^{m_i} \{ \varepsilon_j \delta_{ik}^{(j)} - \exp(\beta_j^T \mathbf{Z}_{ik} + \varepsilon_j) \Lambda_{0j}(T_{ik}) \} \right] f(\varepsilon_j | \theta) d\varepsilon_j} \quad (\text{A.1})$$

while replacing the unknown parameters with their current estimated values. (2) Estimate β_j and the parameters involved with $\lambda_{0j}(\cdot)$ by maximizing

$$\prod_{i=1}^n \prod_{k=1}^{m_i} \left[\lambda_{0j}(T_{ik}) \exp \{ \beta_j^T \mathbf{Z}_{ik} + \log \widehat{\mu}_{ij} \} \right]^{\delta_{ik}^{(j)}} \exp \left[-\exp \{ \beta_j^T \mathbf{Z}_{ik} + \log \widehat{\mu}_{ij} \} \Lambda_{0j}(T_{ik}) \right]$$

for PMLE, and estimate β_j and $\lambda_{0j}(\cdot)$ by maximizing

$$\prod_{i=1}^n \prod_{k=1}^{m_i} \int \{ \lambda_{0j}(T_{ik}) \exp(\beta_j^T \mathbf{Z}_{ik} + \varepsilon_j) \}^{\delta_{ik}^{(j)}} \exp \{ -\exp(\beta_j^T \mathbf{Z}_{ik} + \varepsilon_j) \Lambda_{0j}(T_{ik}) \} f(\varepsilon_j | \theta) d\varepsilon_j$$

for NPMLE. (3) Estimate $\theta = \sigma_j^2$ by

$$\frac{\int \varepsilon_j^2 \exp \left[\sum_{k=1}^{m_i} \{ \varepsilon_j \delta_{ik}^{(j)} - \exp(\beta_j^T \mathbf{Z}_{ik} + \varepsilon_j) \Lambda_{0j}(T_{ik}) \} \right] f(\varepsilon_j | \theta) d\varepsilon_j}{\int \exp \left[\sum_{k=1}^{m_i} \{ \varepsilon_j \delta_{ik}^{(j)} - \exp(\beta_j^T \mathbf{Z}_{ik} + \varepsilon_j) \Lambda_{0j}(T_{ik}) \} \right] f(\varepsilon_j | \theta) d\varepsilon_j} \quad (\text{A.2})$$

while replacing the unknown parameters with their current estimated values. (4) Repeat steps (1)–(3) until convergence.

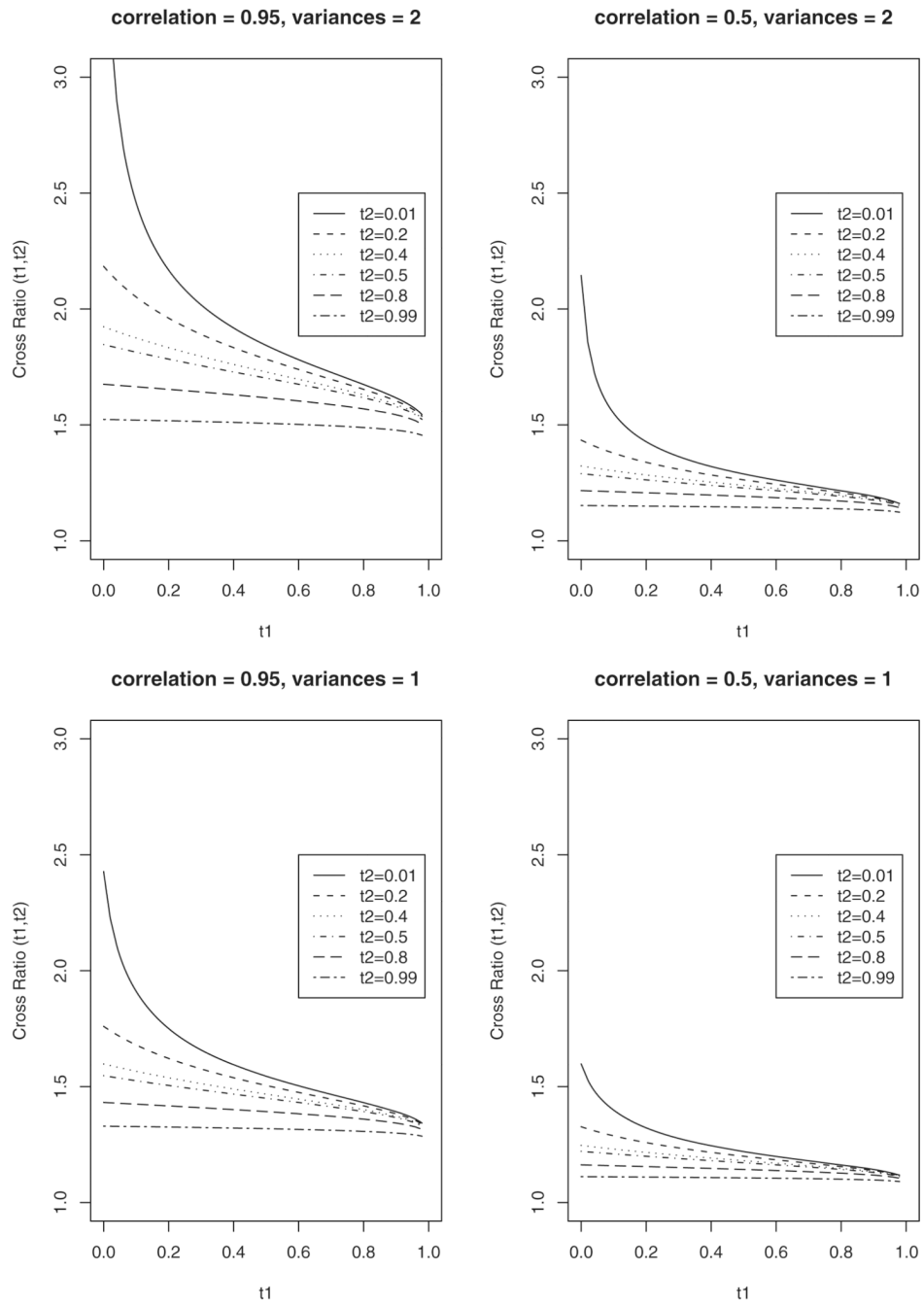


Figure 1. The cross-ratio function under the multivariate normal frailty model with various covariance matrices.

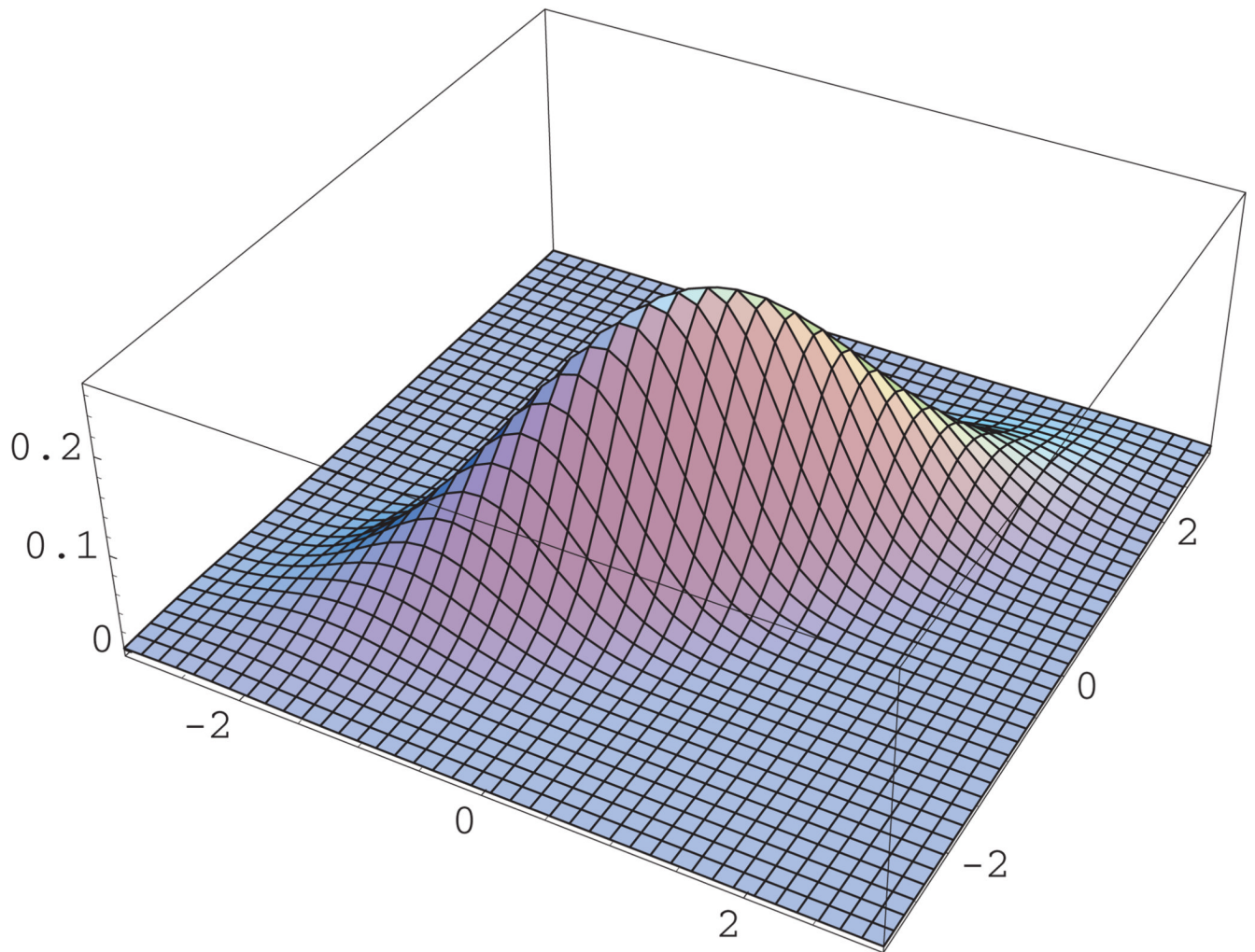


Figure 2. The bivariate normal density function with mean zero and correlation parameter of 0.8. This figure appears in color in the electronic version of this article.

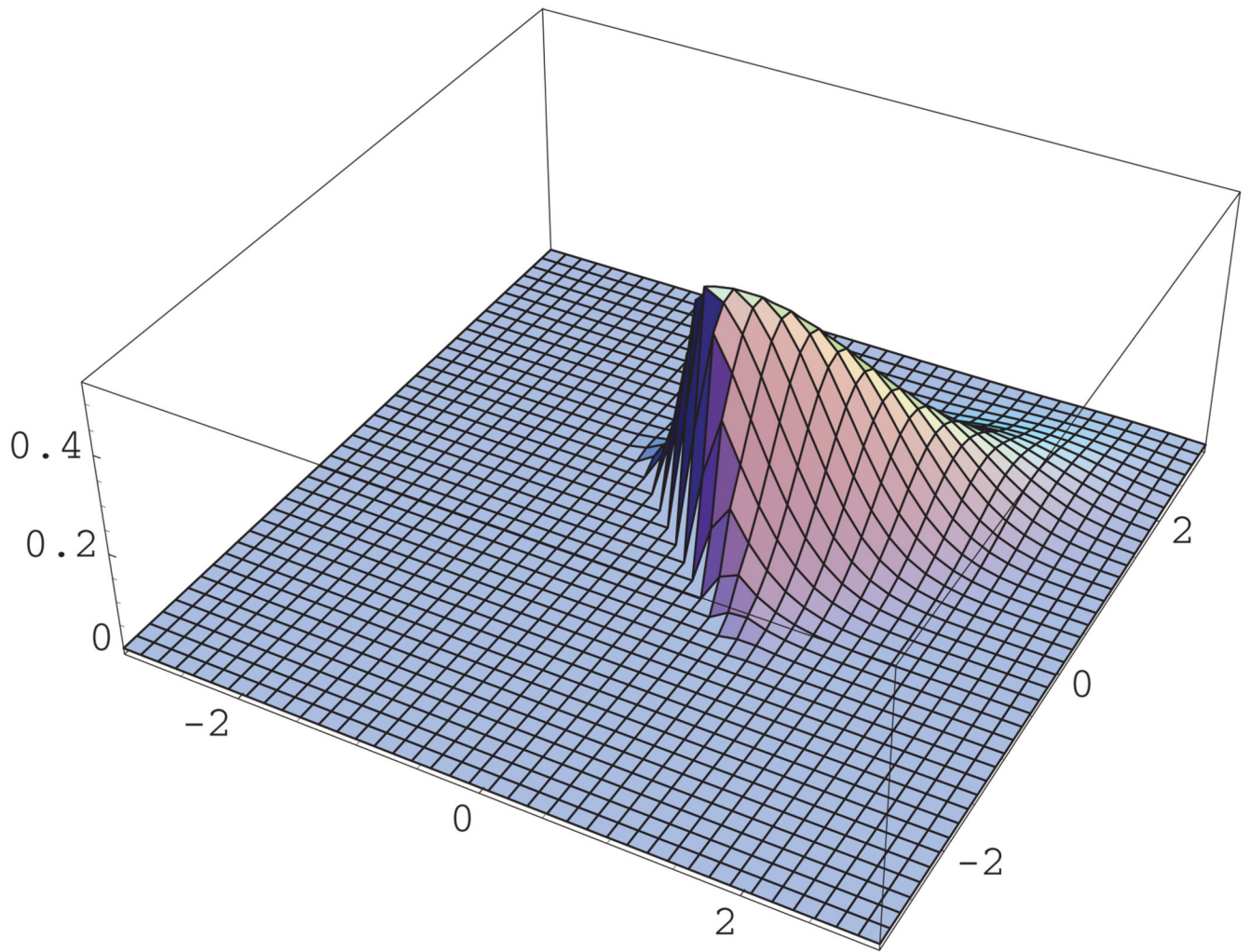


Figure 3. The bivariate skewed-normal density function with correlation parameter of 0.8 and $\boldsymbol{\alpha} = (50, 50)^T$. This figure appears in color in the electronic version of this article.

Table 1

Summary of simulation settings: censoring rates and average (standard deviation, SD) number of observed events under the $\rho_{12} = 0.5$. Each cluster consists of two members.

	500 clusters	1000 clusters
Censoring rate	48%	48%
Average (SD) number of individuals with observed event type $j = 1$	135.0 (12.30)	268.6 (16.26)
Average (SD) number of individuals with observed event type $j = 2$	376.5 (17.9)	754.9 (22.88)
Average (SD) number of families with two different types of observed events	47.7 (6.60)	94.8 (9.25)
Average (SD) number of families with two identical observed events of type $j = 1$	14.7 (3.80)	28.2 (5.35)
Average (SD) number of families with two identical observed events of type $j = 2$	93.7 (8.85)	188.4 (12.19)

Summary of simulation results: the proposed NPMLEs and PMLEs for competing risks versus the naive approach that ignores competing risks under $\rho = 0.5$

Table 2

True value	Proposed NPMLE						Proposed PMLE						$j = 2$ as Censored						$j = 1$ as Censored							
	Mean		SD		Mean		SD		Mean		SD		Mean		SD		Mean		SD		Mean		SD			
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
500 clusters																										
β_1	0.5	0.4790	0.2875	0.4881	0.3674	0.4881	0.3674	0.2857	0.3420	0.3293	0.3652	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
β_2	2.5	2.4720	0.2792	2.4850	0.2323	2.4850	0.2323	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
σ_1^2	1.0	1.0000	0.3724	1.0490	0.3491	1.0490	0.3491	0.9384	0.3878	1.1420	0.4038	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
σ_2^2	1.5	1.4590	0.3369	1.4850	0.2362	1.4850	0.2362	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
ρ_{12}	0.5	0.4833	0.2132	0.4878	0.2059	0.4878	0.2059	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
ψ_1	1.0	-	-	1.0110	0.2252	1.0110	0.2252	-	-	0.9352	0.2240	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
w_2	1.0	-	-	1.0280	0.1797	1.0280	0.1797	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.7330	0.1758	-	
$\Lambda_{01}(0.10)$	0.10	0.1002	0.0243	0.1011	0.0225	0.1011	0.0225	0.1013	0.0257	0.0935	0.0224	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
$\Lambda_{01}(0.15)$	0.15	0.1497	0.0347	0.1516	0.0338	0.1516	0.0338	0.1466	0.0358	0.1403	0.0336	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
$\Lambda_{01}(0.20)$	0.20	0.1969	0.0447	0.2022	0.0450	0.2022	0.0450	0.1883	0.0446	0.1870	0.0448	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
$\Lambda_{01}(0.25)$	0.25	0.2397	0.0559	0.2527	0.0563	0.2527	0.0563	0.2253	0.0541	0.2338	0.0560	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
$\Lambda_{02}(0.10)$	0.10	0.1036	0.0212	0.1028	0.0179	0.1028	0.0179	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.1013	0.0210	0.0973	0.0176	
$\Lambda_{02}(0.15)$	0.15	0.1529	0.0289	0.1542	0.0269	0.1542	0.0269	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.11479	0.0284	0.1460	0.0264	
$\Lambda_{02}(0.20)$	0.20	0.2025	0.0379	0.2056	0.0359	0.2056	0.0359	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.1939	0.0368	0.1947	0.0352	
$\Lambda_{02}(0.25)$	0.25	0.2493	0.0464	0.2570	0.0449	0.2570	0.0449	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.2367	0.0442	0.2433	0.0439	
1000 clusters																										
β_1	0.5	0.5008	0.2636	0.4971	0.2567	0.4971	0.2567	0.3080	0.2563	0.3423	0.2619	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
β_2	2.5	2.4890	0.1817	2.4960	0.1751	2.4960	0.1751	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
σ_1^2	1.0	0.9558	0.2765	0.9656	0.2508	0.9656	0.2508	0.8838	0.2892	1.0640	0.2856	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

True value	Proposed						$j = 2$ as Censored						$j = 1$ as Censored						
	NPMLE			PMLE			NPMLE			PMLE			NPMLE			PMLE			
	Mean	SD		Mean	SD		Mean	SD		Mean	SD		Mean	SD		Mean	SD		
σ_2^2	1.4770	0.2002	1.4890	0.1495	-	-	-	-	-	-	-	-	-	-	-	1.4510	0.2017	1.5360	0.1542
ρ_{12}	0.5014	0.1499	0.4943	0.1412	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ψ_1	-	-	1.0150	0.1637	-	-	0.9330	0.1555	-	-	-	-	-	-	-	-	-	-	-
ψ_2	-	-	1.0070	0.1232	-	-	-	-	-	-	-	-	-	-	-	-	-	0.9633	0.1171
$\Lambda_{01}(0.10)$	0.1011	0.0179	0.1018	0.0164	0.1007	0.0183	0.0933	0.0155	-	-	-	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.15)$	0.1512	0.0258	0.1527	0.0246	0.1461	0.0254	0.1399	0.0233	-	-	-	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.20)$	0.2006	0.0333	0.2036	0.0328	0.1891	0.0316	0.1866	0.0311	-	-	-	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.25)$	0.2465	0.0423	0.2545	0.0410	0.2282	0.0386	0.2333	0.0389	-	-	-	-	-	-	-	-	-	-	-
$\Lambda_{02}(0.10)$	0.1014	0.0129	0.1017	0.0122	-	-	-	-	-	-	-	-	-	-	-	0.0990	0.0126	0.0963	0.0117
$\Lambda_{02}(0.15)$	0.1524	0.0185	0.1525	0.0183	-	-	-	-	-	-	-	-	-	-	-	0.1471	0.0179	0.1444	0.0176
$\Lambda_{02}(0.20)$	0.2019	0.0249	0.2034	0.0244	-	-	-	-	-	-	-	-	-	-	-	0.1929	0.0239	0.1927	0.0234
$\Lambda_{02}(0.25)$	0.2513	0.0315	0.2543	0.0305	-	-	-	-	-	-	-	-	-	-	-	0.2380	0.0298	0.2408	0.0293

Summary of simulation results: the proposed NPMLEs and PMLEs for competing risks versus the naive approach that ignores competing risks under $\rho = 0$

Table 3

True value	Proposed NPMLE						Proposed PMLE						$j = 2$ as Censored						$j = 1$ as Censored								
	Mean		SD		Mean		SD		Mean		SD		Mean		SD		Mean		SD		Mean		SD				
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD			
1000 clusters																											
β_1	1.0	1.0020	0.1990	1.0080	0.1963	0.9946	0.1983	1.0080	0.1959	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
β_2	1.0	0.9944	0.1863	0.9846	0.1894	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
σ_1^2	1.0	0.9523	0.2142	0.9797	0.1842	0.9998	0.2308	0.9737	0.1839	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
σ_2^2	1.5	1.4890	0.2546	1.4920	0.2186	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
ρ_{12}	0.0	-0.0106	0.1537	-0.0125	0.1429	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
ψ_1	1.0	-	-	1.0040	0.1363	-	-	1.0070	0.1342	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
ψ_2	1.0	-	-	1.0240	0.1434	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
$\Lambda_{01}(0.10)$	0.10	0.1012	0.0153	0.1004	0.0142	0.1001	0.0255	0.1007	0.0134	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
$\Lambda_{01}(0.15)$	0.15	0.1507	0.0212	0.1506	0.0213	0.1501	0.0230	0.1510	0.0201	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
$\Lambda_{01}(0.20)$	0.20	0.1999	0.0280	0.2008	0.0273	0.2004	0.0301	0.2014	0.0268	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
$\Lambda_{01}(0.25)$	0.25	0.2464	0.0349	0.2510	0.0341	0.2483	0.0384	0.2517	0.0336	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
$\Lambda_{02}(0.10)$	0.10	0.1018	0.0150	0.1024	0.0143	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
$\Lambda_{02}(0.15)$	0.15	0.1528	0.0216	0.1536	0.0215	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
$\Lambda_{02}(0.20)$	0.20	0.2039	0.0299	0.2048	0.0287	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
$\Lambda_{02}(0.25)$	0.25	0.2523	0.0368	0.2560	0.0359	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
1000 clusters																											
β_1	0.5	0.5156	0.2422	0.5143	0.2357	0.4890	0.2151	0.5181	0.2349	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
β_2	2.5	2.4830	0.1751	2.4890	0.1696	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
σ_1^2	1.0	0.9880	0.2632	0.9434	0.2372	0.9970	0.2702	0.9316	0.2393	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

True value	Proposed						$j = 2$ as Censored						$j = 1$ as Censored					
	NPMLE			PMLE			NPMLE			PMLE			NPMLE			PMLE		
	Mean	SD		Mean	SD		Mean	SD		Mean	SD		Mean	SD		Mean	SD	
σ_2^2	1.4790	0.2048	1.4920	0.1565	-	-	1.4860	0.2097	1.4880	0.1536	-	-	1.4860	0.2097	1.4880	0.1536	-	-
ρ_{12}	-0.0099	0.1726	-0.0130	0.1604	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ψ_1	-	-	1.0170	0.1556	-	-	1.0210	0.1503	-	-	-	-	-	-	-	-	-	-
ψ_2	-	-	1.0200	0.1231	-	-	-	-	1.0220	0.1212	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.10)$	0.1022	0.0173	0.1017	0.0155	0.1010	0.0168	0.1021	0.0150	-	-	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.15)$	0.1525	0.0246	0.1525	0.0233	0.1520	0.0245	0.1531	0.0226	-	-	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.20)$	0.2012	0.0316	0.2034	0.0311	0.2015	0.0323	0.2042	0.0300	-	-	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.25)$	0.2477	0.0391	0.2542	0.0389	0.2462	0.0381	0.2553	0.0376	-	-	-	-	-	-	-	-	-	-
$\Lambda_{02}(0.10)$	0.1020	0.0132	0.1020	0.0123	-	-	-	-	0.1028	0.0139	0.1022	0.0121	-	-	-	-	-	-
$\Lambda_{02}(0.15)$	0.1527	0.0190	0.1530	0.0185	-	-	-	-	0.1539	0.0197	0.1533	0.0182	-	-	-	-	-	-
$\Lambda_{02}(0.20)$	0.2036	0.0240	0.2040	0.0246	-	-	-	-	0.2032	0.0257	0.2044	0.0242	-	-	-	-	-	-
$\Lambda_{02}(0.25)$	0.2514	0.0313	0.2550	0.0308	-	-	-	-	0.2512	0.0335	0.2555	0.0303	-	-	-	-	-	-

Table 4

Summary of simulation results: misspecification of the frailty distribution. The true frailty distribution is a multivariate skewed-normal and the assumed distribution is a multivariate normal.

True value	$j = 2$ as Censored												$j = 1$ as Censored				
	Proposed NPMLE		Proposed PMLE		NPMLE		PMLE		NPMLE		PMLE		NPMLE		PMLE		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
500 clusters																	
β_1	0.5	0.4823	0.3246	0.4858	2.5140	0.2587	0.2983	0.3026	0.3070	-	-	-	-	-	-	-	-
β_2	2.5	2.4880	0.2488	2.5140	0.2360	-	-	-	-	2.4670	0.2444	2.5450	0.2379	-	-	-	-
ψ_1	1.0	-	-	0.9554	0.2012	-	-	0.8868	0.2013	-	-	-	-	-	-	-	-
ψ_2	1.0	-	-	0.9670	0.1533	-	-	-	-	-	-	0.9057	0.1437	-	-	-	-
$\Lambda_{01}(0.10)$	0.10	0.0939	0.0213	0.0955	0.0201	0.0966	0.0231	0.0887	0.0201	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.15)$	0.15	0.1405	0.0309	0.1433	0.0302	0.1392	0.0318	0.1330	0.0302	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.20)$	0.20	0.1864	0.0408	0.1911	0.0402	0.1796	0.0409	0.1774	0.0403	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.25)$	0.25	0.2285	0.0519	0.2389	0.0503	0.2153	0.0498	0.2217	0.0503	-	-	-	-	-	-	-	-
$\Lambda_{02}(0.10)$	0.10	0.0972	0.0171	0.0967	0.0153	-	-	-	-	0.0951	0.0167	0.0906	0.0144	-	-	-	-
$\Lambda_{02}(0.15)$	0.15	0.1447	0.0247	0.1451	0.0229	-	-	-	-	0.1396	0.0237	0.1359	0.0216	-	-	-	-
$\Lambda_{02}(0.20)$	0.20	0.1933	0.0310	0.1934	0.0307	-	-	-	-	0.1841	0.0293	0.1811	0.0287	-	-	-	-
$\Lambda_{02}(0.25)$	0.25	0.2390	0.0389	0.2418	0.0383	-	-	-	-	0.2253	0.0359	0.2264	0.0359	-	-	-	-
1000 clusters																	
β_1	0.5	0.4723	0.2462	0.4841	0.2333	0.2476	0.2277	0.2948	0.2341	-	-	-	-	-	-	-	-
β_2	2.5	2.1550	0.2005	2.5300	0.1813	-	-	-	-	2.4990	0.1999	2.5650	0.1833	-	-	-	-
ψ_1	1.0	-	-	0.9521	0.1432	-	-	0.8742	0.1425	-	-	-	-	-	-	-	-
ψ_2	1.0	-	-	0.9514	0.1120	-	-	-	-	-	-	0.8868	0.1063	-	-	-	-
$\Lambda_{01}(0.10)$	0.10	0.0950	0.0159	0.0951	0.0143	0.0967	0.0170	0.0874	0.0143	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.15)$	0.15	0.1413	0.0216	0.1428	0.0215	0.1387	0.0224	0.1311	0.0214	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.20)$	0.20	0.1888	0.0282	0.1904	0.0286	0.1799	0.0282	0.1748	0.0285	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.25)$	0.25	0.2334	0.0376	0.2380	0.0356	0.2175	0.0364	0.2186	0.0356	-	-	-	-	-	-	-	-
$\Lambda_{02}(0.10)$	0.10	0.0947	0.0130	0.0951	0.0111	-	-	-	-	0.0924	0.0128	0.0887	0.0106	-	-	-	-

True value	Proposed NPMLE						Proposed PMLE						$j = 2$ as Censored						$j = 1$ as Censored						
	Proposed NPMLE			Proposed PMLE			Proposed NPMLE			Proposed PMLE			NPMLE			PMLE			NPMLE			PMLE			
	Mean	SD		Mean	SD		Mean	SD		Mean	SD		Mean	SD		Mean	SD		Mean	SD		Mean	SD		
$\Lambda_{02}(0.15)$	0.15	0.1422	0.0182	0.1427	0.0168	0.0168	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
$\Lambda_{02}(0.20)$	0.20	0.1914	0.0238	0.1903	0.0224	0.0224	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
$\Lambda_{02}(0.25)$	0.25	0.2394	0.0306	0.2379	0.0280	0.0280	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 5

Summary of simulation results: misspecification of the parametric baseline hazard functions. The true baseline hazard functions are $\Lambda_{0j}(t) = \psi_j t^{2/3}, j = 1, 2$ and the assumed functions (under PLMEs) are $\Lambda_{0j}(t) = \psi_j t^j, j = 1, 2$.

True value	Proposed						$j = 2$ as Censored						$j = 1$ as Censored						
	NPMLE			PMLE			NPMLE			PMLE			NPMLE			PMLE			
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
1000 clusters																			
β_1	0.5	0.5065	0.2303	1.4070	0.3159	0.3813	0.2666	0.6406	0.3209	-	-	-	-	-	-	-	-	-	-
β_2	2.5	2.4880	0.1729	3.2670	0.2252	-	-	-	-	2.4740	0.1731	3.5426	0.2387	-	-	-	-	-	-
σ_1^2	1.0	0.9606	0.2383	2.7130	0.3267	0.8576	0.2073	2.8230	0.3065	-	-	-	-	-	-	-	-	-	-
σ_2^2	1.5	1.4740	0.1699	3.1090	0.2115	-	-	-	-	1.4410	0.1714	3.5870	0.2102	-	-	-	-	-	-
ρ_{12}	0.5	0.4984	0.1303	0.8997	0.1065	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ψ_1	1.0	-	-	1.8870	0.2790	-	-	1.4090	0.2491	-	-	-	-	-	-	-	-	-	-
ψ_2	1.0	-	-	1.9630	0.2866	-	-	-	-	-	-	-	-	-	-	-	1.4200	0.2041	-
$\Lambda_{01}(0.10)$	0.2154	0.2165	0.0303	0.1887	0.0279	0.0988	0.0181	0.1409	0.0249	-	-	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.15)$	0.2823	0.2818	0.0378	0.2831	0.0418	0.1443	0.0236	0.2113	0.0374	-	-	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.20)$	0.3420	0.3374	0.0463	0.3740	0.0558	0.1879	0.0306	0.2818	0.0498	-	-	-	-	-	-	-	-	-	-
$\Lambda_{01}(0.25)$	0.3969	0.3833	0.0549	0.4718	0.0697	0.2274	0.0377	0.3523	0.0623	-	-	-	-	-	-	-	-	-	-
$\Lambda_{02}(0.10)$	0.2154	0.2181	0.0235	0.1963	0.0287	-	-	-	-	0.2086	0.0227	0.1420	0.0204	-	-	-	-	-	-
$\Lambda_{02}(0.15)$	0.2823	0.2854	0.0307	0.2945	0.0429	-	-	-	-	0.2700	0.0290	0.2130	0.0306	-	-	-	-	-	-
$\Lambda_{02}(0.20)$	0.3420	0.3433	0.0379	0.3926	0.0573	-	-	-	-	0.3220	0.0352	0.2840	0.0408	-	-	-	-	-	-
$\Lambda_{02}(0.25)$	0.3969	0.3918	0.0441	0.4908	0.0716	-	-	-	-	0.3652	0.0407	0.3550	0.0510	-	-	-	-	-	-

Table 6

Analysis of the Washington Ashkenazi Kin-Cohort Data using the proposed NPMLE approach

	Estimator	Bootstrap SD
Breast cancer		
β_1	1.3910	0.2167
σ_1^2	0.8718	0.5362
$\Lambda_{01}(50)$	0.0298	0.0061
$\Lambda_{01}(60)$	0.0578	0.0108
$\Lambda_{01}(70)$	0.0966	0.0165
$\Lambda_{01}(80)$	0.1395	0.0215
Ovarian cancer		
β_2	1.5420	0.4361
σ_2^2	1.7240	0.7811
$\Lambda_{02}(50)$	0.0014	0.0006
$\Lambda_{02}(60)$	0.0035	0.0014
$\Lambda_{02}(70)$	0.0067	0.0028
$\Lambda_{02}(80)$	0.0096	0.0038
Familial correlation between breast and ovarian cancers		
ρ_{12}	0.1247	0.0981