Published in final edited form as:

Stat Methods Med Res. 2016 December ; 25(6): 2488–2505. doi:10.1177/0962280214526193.

Analysis of Clustered Competing Risks Data using Subdistribution Hazard Models with Multivariate Frailties

Il Do Ha,

Department of Asset Management, Daegu Haany University, Gyeongsan, South Korea

Nicholos J. Christian,

Department of Epidemiology, University of Pittsburgh, Pittsburgh, USA

Jong-Hyeon Jeong,

Department of Biostatistics, University of Pittsburgh, Pittsburgh, USA

Junwoo Park, and

Department of Statistics, Seoul National University, Seoul, South Korea

Youngjo Lee

Department of Statistics, Seoul National University, Seoul, South Korea

Abstract

Competing risks data often exist within a center in multicenter randomized clinical trials where the treatment effects or baseline risks may vary among centers. In this paper we propose a subdistribution hazard regression model with multivariate frailty to investigate heterogeneity in treatment effects among centers from multicenter clinical trials. For inference, we develop a hierarchical likelihood (or h-likelihood) method, which obviates the need for an intractable integration over the frailty terms. We show that the profile likelihood function derived from the h-likelihood is identical to the partial likelihood, and hence it can be extended to the weighted partial likelihood for the subdistribution hazard frailty models. The proposed method is illustrated with a dataset from a multicenter clinical trial on breast cancer as well as with a simulation study. We also demonstrate how to present heterogeneity in treatment effects among centers by using a confidence interval for the frailty for each individual center and how to perform a statistical test for such heterogeneity using a restricted h-likelihood.

Keywords

Competing risks; Hierarchical likelihood; Multivariate frailty; Random treatment-by-center interaction; Subdistribution hazard

Conflict of interest statement

The Authors declare that there is no conflict of interest.

1 Introduction

Competing risks (CR) data arise when an occurrence of an event precludes other type of events from being observed.¹ Two broad classes of models for analyzing the CR data have been developed based on Cox's proportional hazards (PH) models; one is to model the cause-specific hazard of the different event types² and the other is to model the subhazard (i.e. the hazard function of a subdistribution) for the event of interest³. In particular, the subhazard model by Fine and Gray³, often referred to as the Fine-Gray model, directly associates covariate effects with the cumulative probability of a specific cause of events over time, i.e. the cumulative incidence function (CIF), whereas the cause-specific hazard model associates the covariate effects with the cause-specific hazard function. Therefore, if one is interested in direct statistical inference on the cumulative probability of cause-specific events, Fine-Gray model would be more appropriate.

In this paper we model the subhazard for the CR data from multi-center randomized clinical trials, which are often observed within a cluster (e.g. center). In many applications involving CR data, individual events within a cluster may be correlated due to unobserved shared factors across individuals. The Fine-Gray model, however, takes no account for such correlation, which can be modelled by the frailty (or random effect)^{4,5}. Thus Katsahian et al.⁶ and Christian⁷ have extended the Fine-Gray model to a subhazard frailty model with a random center effect only. It would be practically more useful to model heterogeneity in treatment effect among centers as well as the random center effect, where the two random effects might be correlated.^{8,9} Here the heterogeneity in treatment effect among centers can be modelled as an additive frailty term to a regression coefficient for the baseline treatment effect without heterogeneity.¹⁰ In this paper, we extend the standard correlated frailty modelling approach^{5,10-12} to a subhazard frailty modelling approach to handling the potential heterogeneity in treatment effect in CR data from multi-center clinical trials.

For inference, we develop a hierarchical likelihood (or h-likelihood; Lee and Nelder¹³) method; it obviates integration itself over the frailty distributions and also gives a statistically efficient estimation procedure for various random-effect models^{10,13,14}. In particular, we show that the profile likelihood function derived from the h-likelihood is identical to the partial likelihood, and hence it can be extended to the weighted partial likelihood for the subdistribution hazard frailty models. The proposed method is illustrated with time-to-event data from a phase III breast cancer trials (B-14) conducted by the National Surgical Adjuvant Breast and Bowel Project (NSABP), which consist of 2,817 patients from 167 centers^{15,16}, as well as a simulation study. We also demonstrate via the practical data set how to present heterogeneity in treatment effect among centers by using a confidence interval^{17,18} for the frailty for each individual center, not the parameters for the frailty distribution, and how to perform a statistical test for such heterogeneity using a restricted h-likelihood¹⁴.

The paper is organized as follows. In Section 2, we propose a formulation of sub-hazard frailty models. In Section 3, we show how the h-likelihood can be extended to the subhazard frailty models, and develop the h-likelihood estimation procedure for fitting the models. A simulation study is conducted to evaluate the performance of the proposed method in

Section 4. The new method is illustrated using the breast cancer data set in Section 5. Finally, we discuss our method in Section 6. The technical derivations and additional simulation results are presented in Appendix and Supplementary Material, respectively.

2 Formulation for subhazard frailty models

Suppose that the data consist of censored time-to-event observations collected from q centers (or clusters). We also assume that there are L distinct event types in each center. For subject j of center i , let T_{ij} be the time to the first event and let $\varepsilon_{ij} \in \{1, 2, \dots, L\}$ be the corresponding cause of event ($i = 1, \dots, q, j = 1, \dots, n_i, n = \sum_i n_i$). Then observable random variables become $Y_{ij} = \min(T_{ij}, C_{ij})$ and $\xi_{ij} = I(T_{ij} < C_{ij})\varepsilon_{ij}$, where C_{ij} is the independent censoring time, $\xi_{ij} \in \{0, 1, 2, \dots, L\}$ and $I(\cdot)$ is the indicator function. The CIF of events from cause 1 (i.e. $\varepsilon_{ij} = 1$) is defined by

$$F_1(t) = \Pr(T_{ij} \leq t, \varepsilon_{ij} = 1),$$

which represents the probability that an individual will experience an event of Type 1 by time t . The corresponding hazard function of subdistribution (subhazard function) is defined by

$$\lambda_1^s(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \Pr\{t \leq T_{ij} \leq t + \Delta t, \varepsilon_{ij} = 1 | T_{ij} \geq t \cup (T_{ij} < t \cap \varepsilon_{ij} \neq 1)\} = -d \log \{1 - F_1(t)\} / dt.$$

For simplicity, in this paper we consider the two event types ($L = 1, 2$). Thus, $\xi_{ij} \in \{0, 1, 2\}$ it is 1 for an event of interest, 2 for a competing event and 0 for censoring.

Fine and Gray³ first introduced a way to directly associate the effects of covariates with CIF, which models the subhazard for the event of interest, $L = 1$. Furthermore, Katsahian et al.⁶ and Christian⁷ have extended Fine-Gray model to a subhazard frailty model with only one random component (i.e. random center effect) to analyze multi-center competing risks data.

In this paper we show that, for the purpose of more systematic analysis, the model above needs to be extended to a general subhazard frailty model allowing multiple random components (e.g. random center and random treatment effect) and their correlation, as in Ha et al.¹⁰. Here, random treatment effect means random treatment-by-center interaction. In particular, a model allowing for the correlation between random center and random treatment effects can properly account for the heterogeneities from the treatment effect across centers as well as between-center variation.¹⁰ Denote by $v_i = (v_{i0}, v_{i1}, \dots, v_{i,r-1})^T$ an r -dimensional vector of unobserved log-frailties (random effects) associated with the i th ($i = 1, \dots, q$) center. Note that in (1), $u_i = \exp(v_i)$ (i.e. $v_i = \log u_i$) are referred to frailties^{5,8,10}. As described in Fine and Gray³ and Ha et al.¹⁹, we assume that given v_i , $(T_{ij}, \varepsilon_{ij})$ and C_{ij} , $j = 1, \dots, n_i$, are conditionally independent, and that given v_i , C_{ij} , $j = 1, \dots, n_i$, are non-informative about v_i . Suppose that we are interested in assessing the effects of covariates on the conditional CIF for cause 1 given the log-frailties v_i , defined by $F_1(t|v_i) = \Pr(T_{ij} \leq t, \varepsilon_{ij} = 1 | v_i)$. The conditional subhazard function for cause 1 given v_i ,

$$\lambda_{ij1}^s(t|v_i) = -d \log \{1 - F_1(t|v_i)\} / dt, \text{ is modeled as}$$

$$\lambda_{ij1}^s(t|v_i) = \lambda_{01}^s(t) \exp(\eta_{ij}), \quad (1)$$

where $\lambda_{01}^s(\cdot)$ is the unknown baseline subhazard function,

$$\eta_{ij} = x_{ij}^T \beta + z_{ij}^T v_i$$

is the linear predictor for the log-hazard, and $x_{ij} = (x_{ij1}, \dots, x_{ijp})^T$ and $z_{ij} = (z_{ij1}, \dots, z_{ijr})^T$ are $p \times 1$ and $r \times 1$ covariate vectors corresponding to fixed effects $\beta = (\beta_1, \dots, \beta_p)^T$ and log-frailties v_i , respectively. We assume that the log-frailties v_i are independent and follow a multivariate normal distribution, $v_i \sim N(0, \Sigma_i(\theta))$, where the covariance matrix $\Sigma_i(\theta)$ depends on a vector of unknown parameters θ . The normal distribution has been used for modelling multi-component²⁰ and correlated frailties⁹.

Model (1) includes some well-known models as special cases. In a multicenter medical study, let v_{i0} be a random intercept or random center effect that modifies the baseline risk for center i , and let v_{i1} be associated with the treatment effect, i.e., a random treatment effect (or random treatment-by-center interaction). In (1), if we consider $z_{ij} = 1$ and $v_i = v_{i0}$ for all i, j , this becomes the random center or shared frailty model^{6,7} with

$$\eta_{ij} = x_{ij}^T \beta + v_{i0}, \quad (2)$$

where $v_{i0} \sim N(0, \sigma_0^2)$ for all i . Model (2) can be extended as follows. Let β_1 be the main treatment effect associated with the treatment indicator x_{ij1} and let β_m ($m = 2, \dots, p$) be the fixed effects corresponding to the covariates x_{ijm} . Our two random components leads to a bivariate model^{9,10} with

$$\eta_{ij} = v_{i0} + (\beta_1 + v_{i1})x_{ij1} + \sum_{m=2}^p \beta_m x_{ijm}, \quad (3)$$

which is easily derived by taking $z_{ij} = (1, x_{ij1})^T$ and $v_i = (v_{i0}, v_{i1})^T$ in (1). Here, to maintain the invariance of model to parametrization of the treatment effect we allow a general covariance matrix^{9,14} between v_{i0} and v_{i1} within a cluster:

$$\Sigma_i \equiv \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix}, \quad (4)$$

where the correlation is denoted by $\rho = \sigma_{01}/(\sigma_0 \sigma_1)$. The bivariate normal model (3) with (4) is very useful for investigating heterogeneity in the baseline risk and the treatment effect across centers.

3 H-likelihood estimation

In this section we show how to construct systematically the h-likelihood estimation procedure for fitting the semiparametric subhazard frailty model (1). For this, we first show how to construct the h-likelihood, and then propose the estimation procedure. The general case for incomplete (censoring) data is presented here as in Pintilie¹ because the proposed method can be directly applied to complete (no censoring) data.

3.1 H-likelihood construction

Let $y_{(k)}$ denotes the k th ($k = 1, \dots, D$) smallest distinct event time of Type 1 among the y_{ij} 's, where y_{ij} is the observed value of Y_{ij} and D is the total number of distinct Type 1 events. Let $R_{(k)}$ denotes the risk set at $y_{(k)}$:

$$R_{(k)} = R(y_{(k)}) = \left\{ (i, j) : y_{ij} \geq y_{(k)} \text{ or } (y_{ij} \leq y_{(k)} \text{ and } \varepsilon_{ij} \neq 1) \right\}.$$

Note that as compared to the classical Cox model, the risk set $R_{(k)}$ comprises individuals who have not failed from any cause by $y_{(k)}$ but also those who have previously failed from competing causes.^{3,6} Since the functional form of baseline subhazard function $\lambda_{01}^s(t)$ is unknown, following Breslow's²¹ idea and the reformulation (equation (3.4), page 80) used by Fan and Li²² without competing risks, at each y_{ij} , the baseline cumulative subhazard function $\Lambda_{01}^s(t)$ can be written as

$$\Lambda_{01}^s(y_{ij}) = \sum_k \lambda_{01k}^s I \left\{ (i, j) \in R_{(k)} \right\},$$

where $\lambda_{01k}^s = \lambda_{01}^s(y_{(k)})$ is the subhazard function for cause 1 events at $y_{(k)}$. Let $\delta_{ij} = I(\xi_{ij} = 1)$ be an event indicator representing whether subject j of center i experiences a Type 1 event. Along the lines of Lee and Nelder¹³ and Ha et al.¹⁹, the hierarchical loglikelihood (h-likelihood) for subhazard frailty models (1) is defined by

$$h = h(\beta, v, \lambda_{01}^s, \theta) = \sum_{ij} \ell_{1ij} + \sum_i \ell_{2i}, \quad (5)$$

where

$$\begin{aligned} \sum_{ij} \ell_{1ij} = & \sum_{ij} \delta_{ij} \{ \log \lambda_{01}^s(y_{ij}) + \eta_{ij} \} - \sum_{ij} \{ \Lambda_{01}^s(y_{ij}) \exp(\eta_{ij}) \} = \sum_k d_{(k)} \log \lambda_{01k}^s \\ & + \sum_{ij} \delta_{ij} \eta_{ij} - \sum_{ij} \left[\sum_k \lambda_{01k}^s I \left\{ (i, j) \in R_{(k)} \right\} \exp(\eta_{ij}) \right] \end{aligned}$$

is the sum of the logarithm of the conditional density function for Y_{ij} and δ_{ij} given v_i , i.e. the ordinary log-likelihood for censored survival data given v_i , $\ell_{1ij} = \ell_{1ij}(\beta, \lambda_{01}^s; y_{ij}, \delta_{ij} | v_i)$, and

$$\ell_{2i} = \ell_{2i}(\theta; v_i) = -\frac{1}{2} [\log \det \{2\pi \sum_i(\theta)\}] - \frac{1}{2} v_i^T \sum_i(\theta)^{-1} v_i$$

is the logarithm of the density function for v_i with parameters $\theta = (\sigma_0^2, \sigma_1^2, \sigma_{01})^T$, i.e. the log-likelihood for v_i , and $\eta_{ij} = x_{ij}^T \beta + z_{ij}^T v_i$. Here, $v = (v_1^T, \dots, v_q^T)^T$, $\lambda_{01}^s = (\lambda_{011}^s, \dots, \lambda_{01D}^s)^T$, and $d_{(k)}$ is the number of the events of interest at $y_{(k)}$. As the number of λ_{01k}^s 's can increase with the number of distinct event times, the function $\lambda_{01}^s(t)$ is potentially of high dimension. Accordingly, for estimation of (β, v) Ha et al.¹⁹ proposed the use of the profiled h-likelihood h^* from which λ_{01}^s in (5) is eliminated:

$$h^* = h|_{\lambda_{01}^s = \hat{\lambda}_{01}^s} = \sum_{ij} \ell_{1ij}^* + \sum_i \ell_{2i}, \quad (6)$$

where

$$\hat{\lambda}_{01k}^s(\beta, v) = \frac{d_{(k)}}{\sum_{(i,j) \in R_{(k)}} \exp(\eta_{ij})},$$

are solutions of the estimating equations, $\partial h / \partial \lambda_{01k}^s = 0$, for $k = 1, \dots, D$. Note here that

$$\sum_{ij} \ell_{1ij}^* = \sum_{ij} \ell_{1ij}|_{\lambda_{01}^s = \hat{\lambda}_{01}^s} = \sum_k d_{(k)} \log \hat{\lambda}_{01k}^s + \sum_{ij} \delta_{ij} \eta_{ij} - \sum_k d_{(k)} = \sum_{ij} \delta_{ij} \eta_{ij} - \sum_k d_{(k)} \log \left\{ \sum_{(i,j) \in R_{(k)}} \exp(\eta_{ij}) \right\}$$

with a constant term $\sum_k d_{(k)} \{\log d_{(k)} - 1\}$ eliminated, so that h^* becomes the penalized partial likelihood (PPL)^{11,23}: see also Appendix C of Ha et al.¹⁰. In particular, the first term, $\sum_{ij} \ell_{1ij}^*$, of h^* in (6) can be viewed as the log-partial likelihood for the Fine-Gray model given v_i , by treating the observed event times y_{ij} 's as complete outcomes.

In the case of right censoring under competing risks, Fine and Gray³ developed a weighted score function based on the complete-data partial likelihood. Thus the inverse probability of censoring weighting (IPCW) by Fine and Gray³ can be equally applied to the first term of h^* as in Pintilie^{1,24} and Katsahian et al.⁶ Accordingly, a weighted partial h-likelihood h_w^* based on the IPCW is defined by

$$h_w^* = \ell_{1w}^* + \sum_i \ell_{2i}. \quad (7)$$

Here

$$l_{1w}^* = \sum_{ij} \delta_{ij} \eta_{ij} - \sum_k d_{(k)} \log \left\{ \sum_{(i,j) \in R_{(k)}} w_{ij} \exp(\eta_{ij}) \right\},$$

where

$$w_{ij} = w_{ij}(y_{(k)}) = \frac{\hat{G}(y_{(k)})}{\hat{G}(y_{ij} \wedge y_{(k)})}$$

is the weight of subject j of center i at $y_{(k)}$, and $\hat{G}(\cdot)$ is the Kaplan-Meier estimate of the survival function for the censoring times. Here, $w_{ij} = 1$ as long as individuals have not failed (i.e. $y_{ij} > y_{(k)}$; the first condition of $R_{(k)}$), whereas $w_{ij} < 1$ and decreasing over time if they failed from Type 2 (i.e. $y_{ij} \leq y_{(k)}$) and $\varepsilon_{ij} < 1$; the second condition of $R_{(k)}$ ^{1,24}. Note that h_w^* in (7) is an extension of the weighted log-partial likelihood^{1,24,25} for the Fine-Gray model to the subhazard frailty models (1). We can show that, under the subhazard shared model (2), h_w^* is also equivalent to the PPL of Katsahian et al.⁶, by combining $R_{(k)}$ and w_{ij} as in (17) of Appendix.

3.2 Estimation procedure

Now, Ha et al.'s¹⁰ procedures for standard correlated frailty models without competing risks can be extended to the subhazard model (1) by using h_w^* in (7). That is, given frailty parameters θ , the maximum h-likelihood (MHL) estimators of $\tau = (\beta^T, v^T)^T$ are obtained by solving the joint estimating equations, $\partial h_w^* / \partial \tau = 0$. In Appendix we show that given θ , the joint equations lead to Ha and Lee's²⁶ MHL estimator score equations for τ :

$$\begin{pmatrix} X^T W^* X & X^T W^* Z \\ Z^T W^* X & Z^T W^* Z + U \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{v} \end{pmatrix} = \begin{pmatrix} X^T w^* \\ Z^T w^* \end{pmatrix}, \quad (8)$$

where X and Z are $n \times p$ and $n \times q^*$ model matrices for β and v whose ij th row vectors are x_{ij}^T and z_{ij}^T , respectively, W^* is the symmetric weight matrix given in (16) of Appendix, and $U = -\partial^2 l_2 / \partial v^2 = \text{BD}(\sum_1^{-1}, \dots, \sum_q^{-1})$ is a $q^* \times q^*$ matrix; $q^* = q \times r$ and $\text{BD}(\cdot)$ denotes a block diagonal matrix. Here $w^* = W^* \eta + (\delta - \mu)$ with $\eta = X\beta + Zv$ and

$\mu = \exp(\log w + \log \Lambda_{01}^s + \eta)$. Note the λ_{01k}^s terms in both W^* and w^* are evaluated at $\hat{\lambda}_{01k}^w$ given in (16) of Appendix. For the Fine-Gray model³ without frailty, they reduce to a simple form:

$$(X^T W^* X) \hat{\beta} = X^T w^*.$$

We thus see that the estimating equations (8) provide new generalized iterative least squares equations for the Fine-Gray model: see also Ha and Lee²⁶.

For estimation of θ we use the adjusted partial h-likelihood (i.e. restricted h-likelihood¹⁴) $p_{\tau}(h_w^*)$, given by

$$p_{\tau}(h_w^*) = \left[h_w^* - \frac{1}{2} \log \det \{ H_w / (2\pi) \} \right] \Big|_{\tau=\hat{\tau}}, \quad (9)$$

where $\hat{\tau} = \hat{\tau}(\theta) = (\hat{\beta}^T(\theta), \hat{v}^T(\theta))^T$ and $H_w = H(h_w^*; \tau) = -\partial^2 h_w^* / \partial \tau^2$ is an information matrix for τ . Note that $p_{\tau}(h_w^*)$ is a function of θ only because it has already eliminated τ from h_w^* and that the additional term in (9) is an adjusted form for such elimination^{13,14}, leading to restricted maximum likelihood (REML) estimator for θ . The REML estimator for θ are obtained by solving iteratively

$$\frac{\partial p_{\tau}(h_w^*)}{\partial \theta} = 0. \quad (10)$$

Note here that

$$\frac{\partial p_{\tau}(h_w^*)}{\partial \theta} = -\frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta} \right) - \frac{1}{2} \hat{v}^T \left(\frac{\partial \Sigma^{-1}}{\partial \theta} \right) \hat{v} - \frac{1}{2} \text{tr} \left(\hat{H}_w^{-1} \frac{\partial \hat{H}_w}{\partial \theta} \right),$$

where $\Sigma = \text{BD}(\Sigma_1, \dots, \Sigma_q)$ is the $q^* \times q^*$ block diagonal matrix and

$\hat{H}_w = \hat{H}_w(\theta) = H(h_w^*; \tau)|_{\tau=\hat{\tau}(\theta)}$. Here, the equations (10) are solved using the Newton-Raphson method with the Hessian matrix, $-\partial^2 p_{\tau}(h_w^*) / \partial \theta^2$. Note that in implementing (10) we allow the $\partial \hat{v} / \partial \theta$ term^{10,14,26} and that the computations of $\partial \hat{H}_w / \partial \theta$ and $-\partial^2 p_{\tau}(h_w^*) / \partial \theta^2$ follow those of Ha et al.¹⁰ using $p_{\tau}(h^*)$.

The approximated standard-error (SE) estimates for $\hat{\tau} - \tau$ and $\hat{\theta}$ are obtained from the inverses of the corresponding Hessian matrices, $H_w = -\partial^2 h_w^* / \partial \tau^2$ and $-\partial^2 p_{\tau}(h_w^*) / \partial \theta^2$, respectively.^{10,14} In particular, Fine and Gray³ proposed a robust/sandwich variance estimator to estimate $\text{var}(\hat{\beta})$ using an empirical process theory because the martingale properties break down under the Fine-Gray model due to the use of IPCW and thus the standard asymptotic theories are no longer valid. Furthermore, in the subhazard frailty model (2) with one frailty term Katsahian and Boudreau²⁷ presented a robust variance estimator of $\hat{\beta}$ using Gray's²⁸ method, estimated from

$$V(\tau) = H_w^{-1} H_1 H_w^{-1}, \quad (11)$$

where $\tau = (\beta^T, v^T)^T$ and $H_1 = H(\ell_{1w}^*; \tau) = -\partial^2 \ell_{1w}^* / \partial \tau^2$. However, the proposed method, H_w^{-1} , has been also used as a variance estimator in the context of the PPL: see Verweij and Van Houwelingen²⁹ and Therneau et al.¹² We investigate the performance of the two variances of

$\hat{\beta}$ by simulation studies in next section. Accordingly, the current estimation procedure can be implemented by replacing the risk indicator matrix in Ha et al.'s¹⁰ procedure with a weighted risk indicator matrix (i.e. M in (17) of Appendix) which contains both the weights w and the risk set R used for modelling the subhazard function: see also Ruan and Gray.³⁰ Further quantities (e.g. confidence intervals of frailties) are also directly applied.

In summary, the estimates of τ and θ are obtained by alternating between the two estimating equations (8) and (10) until convergence is achieved^{10,26}. At convergence, we compute the SEs of $\hat{\tau} - \tau$ and $\hat{\theta}$. Note that the h-likelihood procedure performs well under any restrictions for cluster size n_i such as $n_i = 1$ and unbalanced cases.^{10,20,31} The equations (8), which estimates all random effects simultaneously from the weighted h-likelihood in (7), may influence the consistency of fixed parameters (β , θ), particularly for a small cluster size n_i . However, here the resulting biases decrease quickly as n rather than q increases^{26,31}; see also the simulation results of Section 4. Furthermore, when n_i is very small, the biases can be further reduced using the Laplace approximation based on the h-likelihood: see Lee et al.¹⁴

For a subhazard shared frailty model (2), the procedures proposed by Katsahian et al.⁶ and Katsahian and Boudreau²⁷ are based on the PPL^{11,23}. Given frailty parameters θ , the PPL and h-likelihood methods provide the same estimates for β and v . However, the two methods do not yield the same final results for β and v because they give different estimators for θ .^{10,26} That is, the PPL method ignores the $\partial \hat{v} / \partial \theta$ term in solving the estimating equations of θ , given in (10); this leads to an underestimation of the parameters θ , particularly when the cluster size n_i is small: see also simulation results by Ha and Lee²⁶ and Christian⁷.

4 Simulation study

Simulation study is conducted to evaluate the performance of the proposed method under the subhazard frailty models (3) with a general correlation structure (4) by using 1000 replications of simulated data.

The simulated data are generated using a method similar to that of Fine and Gray³ and Katsahian and Boudreau²⁷. We consider two covariates $x_{ij} = (x_{ij1}, x_{ij2})^T$ and a bivariate normal (BN) random effect $v_i = (v_{i0}, v_{i1})^T$ with mean 0 and covariance matrix having σ_0^2 , σ_1^2 and σ_{01} . The conditional subdistribution for Type 1 events given x_{ij} and v_i is given by

$$F_1(t|x_{ij}, v_i) = P(T_{ij} \leq t, \varepsilon_{ij} = 1 | x_{ij}, v_i) = 1 - [1 - p(1 - e^{-t})]^{exp(\eta_{ij}^{(1)})}$$

where $p = P(\varepsilon_{ij} = 1 | x_{ij} = (0, 0), v_i = (0, 0))$ is the proportion of Type 1 events and $\eta_{ij}^{(1)} = v_{i0} + (\beta_{11} + v_{i1})x_{ij1} + \beta_{12}x_{ij2}$. Here β_{11} and β_{12} are regression parameters for Type 1 events. Thus the conditional distribution function of T_{ij} given a Type 1 event as well as x_{ij} and v_i is given by

$$F(t|x_{ij}, v_i, \varepsilon_{ij}=1) = \frac{1 - [1 - p(1 - e^{-t})]^{exp(\eta_{ij}^{(1)})}}{1 - (1 - p)^{exp(\eta_{ij}^{(1)})}}. \quad (12)$$

Times to Type 1 event of interest are then generated from the distribution function above (12) using the probability integral transformation, conditional on x_{ij} and v_i . The conditional subdistribution for Type 2 events is simply obtained by taking $P(\varepsilon_{ij}=2|x_{ij}, v_i) = 1 - P(\varepsilon_{ij}=1|x_{ij}, v_i)$ and using an exponential distribution with rate $exp(\eta_{ij}^{(2)})$ for $P(T_{ij} \leq t|\varepsilon_{ij}=2, x_{ij}, v_i)$, where $\eta_{ij}^{(2)} = v_{i0} + (\beta_{21} + v_{i1})x_{ij1} + \beta_{22}x_{ij2}$, and β_{21} and β_{22} are regression parameters for Type 2 events. Thus the conditional distribution function of T_{ij} given a Type 2 event as well as x_{ij} and v_i is given by

$$F(t|x_{ij}, v_i, \varepsilon_{ij}=2) = 1 - exp\{-exp(\eta_{ij}^{(2)})t\}. \quad (13)$$

As before, Type 2 event times (times-to-CR event) are generated from the distribution function (13) using the probability integral transformation.

Following Fine and Gray³ and Katsahian and Boudreau²⁷, we consider the two cases of the true parameter values for p and β :

Case A: $(p, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}) = (0.3, 0.5, 0.5, -0.5, 0.5)$,

Case B: $(p, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}) = (0.6, 1, -1, 1, 1)$.

Following Katsahian and Boudreau²⁷, we also consider the three sample sizes: $n = \sum_{i=1}^q n_i$ with $n = 200, 400$ and 1000 , and $(q, n) = (20, 10), (20, 20)$ and $(50, 20)$. The covariates x_{ij1} are generated from a Bernoulli random variable with probability 0.5 in order to mimic the binary treatment covariate of the multi-center study, and x_{ij2} are from a standard normal distribution. The covariance parameters of the random effects are $\sigma_0^2 = \sigma_1^2 = 0.5$ and $\sigma_{01} = -0.25$, leading to $\rho = -0.5$. Though not reported here, we found similar results for $\sigma_{01} = 0.25$. Censoring times are generated from a Uniform(a, b) distribution where the values of a and b were empirically selected to achieve the approximate right censoring rate, low (around 25%) and high (around 50%). That is, in Case A we used Uniform(1, 2.6) and Uniform(0.45, 1) for the censoring rates 25% and 50%, respectively, and in Case B we used Uniform(0.2, 1.7) and Uniform(0, 0.7) for the censoring rates 25% and 50%, respectively.

For the 1000 replications we computed the mean, standard deviation (SD), and the mean of the estimated standard errors (denoted by SE1) for $\hat{\beta}_1 = (\hat{\beta}_{11}, \hat{\beta}_{12})^T$ and $\hat{\theta} = (\hat{\sigma}_0^2, \hat{\sigma}_1^2, \hat{\sigma}_{01})^T$, respectively. The SE1s for $\hat{\beta}_1$ and $\hat{\theta}$ are, respectively, obtained from

$H_w^{-1} = \{-\partial^2 h_w^* / \partial(\beta_1, v)^2\}^{-1}$ and $\{-\partial^2 p_{\beta_1, v}(h_w^*) / \partial\theta^2\}^{-1}$. For comparison of standard errors of $\hat{\beta}_1$, we calculated the mean of estimated standard errors (denoted by SE2) using a

robust variance formula in (11). All computations were done using SAS/IML. The simulation results are summarized in Table 1.

Our method overall performs well even when the censoring rate is as high as 50%. In particular, the increase of n or n_i , rather than q , reduces bias more effectively: see also Ha and Lee²⁶. In Table 1, the SD is the empirical estimates of $\{var(\hat{\beta}_{1j})\}^{1/2}$ ($j=1, 2$), and SE1 and SE2 are the averages of the proposed and robust standard-error estimates for $\hat{\beta}_{1j}$, respectively. Our SE1($\hat{\beta}_{1j}$) work well as judged by the very good agreement between SE1 and SD. Similarly, our SE1s for $\hat{\theta}$ also perform well. On the other hand, the SE2($\hat{\beta}_{11}$) are seriously underestimated even if n increases, but the SE2($\hat{\beta}_{12}$) work well. A possible reason is that the SE2 may be sensitive to the random-effect structures because β_{11} depends on the random treatment-by-center interaction (v_{i1}) via the same covariates x_{i1} , but β_{12} does not. The trends in Table 2 are similar to those evident in Table 1. In particular, with a smaller sample as in $n = 200$ the biases of frailty-parameter estimators $\hat{\theta}$ are largely reduced, as compared to those in Table 1.

In addition, we carried out the simulation studies of Cases A and B above under a subhazard shared frailty model (2) with $\eta_{ij} = v_{i0} + \beta_{11}x_{i1} + \beta_{12}x_{i2}$ and σ_0^2 . Here, for the censoring and covariate patterns, we follow the schemes of Katsahian and Boudreau²⁷. Furthermore, the performances of likelihood ratio tests (LRT) based on the h-likelihood are evaluated for testing $H_0: \sigma_0^2 = 0$ versus $H_1: \sigma_0^2 > 0$, which lies on the boundary of the parameter space. The LRT statistics are calculated as $-2\{p_\beta(h_w^*) - p_{\beta,v}(h_w^*)\}$, where $p_\beta(h_w^*)$ is the likelihood value under $H_0: \sigma_0^2 = 0$ (i.e. $v_{i0} = 0$ for all i). For the purpose, we investigate the nominal level and power of the LRT at the 5% level based on the asymptotic chi-square mixture distribution, $(\chi_0^2 + \chi_1^2)/2$ which gives critical value 2.71 at 5% level.³² Here σ_0^2 ranges from 0, 0.1 to 1. The conclusions obtained are as follows (see also Tables S1 and S2 in Supplementary Material): (i) The trends of the estimates of the fixed parameters (β_1, θ) are quite similar to those presented in Table 1, except for the standard errors for $\hat{\beta}_1$. We have found that the results of SE1 and SE2 for $\hat{\beta}_{1j}$ ($j=1, 2$) are about the same; here x_{i1} depends on β_{11} , but not the random center effect v_{i0} . (ii) The LRT statistic under $\sigma_0^2 = 0$ is somewhat conservative because the observed size (i.e. the observed rate of rejecting H_0 at the 5% level under $\sigma_0^2 = 0$) is less than the nominal level 5% as shown in Katsahian and Boudreau²⁷, but it becomes closer to the nominal level 5% with sample size n_i or n , especially for the Case B. (iii) The power of all tests increases as σ_0^2 and/or n increase. These results indicate that the LRT statistic overall performs well under the shared model (2).

5 A practical example

5.1 The data

We re-examine the data from the B-14 randomized multicenter breast cancer trial conducted by the NSABP^{15,16}. The 2,817 eligible patients from 167 distinct centers were followed up

for about 20 years since randomization. The number of patients per center varied from 1 to 241, with a mean of 16.9 and median of 8. The patients were randomized to one of two treatment arms, tamoxifen (1413 patients) or placebo (1404 patients). The average age of patients was 55 and the average tumor size was about 2 centimeters. The aim of this analysis is to investigate the effect of treatment on local or regional recurrence. Here we consider two event types. The first type is local or regional recurrence (Type 1) and the second type is a new primary cancer, distance recurrence or death (Type 2); only the event that occurs firstly is of interest in this analysis, so that the repeated event times are not considered. Table 3 gives the number of first observed event types in this data set; Type 1 is an event of interest (314 patients; 11.15%), Type 2 is an event of competing risk (1303 patients; 46.25%), and no-events until the last follow-up are censoring (1200 patients; 42.60%).

Table 3 also shows the number of first observed event types by two treatment arms. Figure 1 presents the estimated CIFs¹ for the two treatment arms. The tamoxifen group has lower CIFs compared to placebo group for both Type 1 and Type 2. For Type 1 the difference of CIFs of two arms seems to be large, whereas for Type 2 it does not. In particular, the estimated probability that a patient of tamoxifen group will experience Type 1 event within ten years after surgery is 5%, while for a patient of placebo group it is 10%.

5.2 Analyses using subhazard models

For the data analysis we consider the three covariates of interest: treatment (x_{ij1} is 1 for tamoxifen and 0 for placebo), age (x_{ij2}) and tumor size (x_{ij3}) as continuous covariates. Let v_{j0} and v_{j1} be random center effects and random treatment effects (i.e. random treatment-by-center interaction), respectively. Following Ha et al.¹⁰, we consider the three submodels of (1) for the time to Type 1 event, which include the proportional subhazards model without random effects (i.e. Fine-Gray model) and two subhazard frailty models. In other words, we consider the following three models, $\lambda_{1ij}^s(t|v) = \lambda_{01}^s(t) \exp(\eta_{ij})$ with η_{ij} allowing two frailty structures (M2 and M3): Here $(v_{j0}, v_{j1}) \sim BN$ means that $v_{j0} \sim N(0, \sigma_0^2)$, $v_{j1} \sim N(0, \sigma_1^2)$ and $\rho = \text{Corr}(v_{j0}, v_{j1})$.

$$\text{M1 (F-G): } \eta_{ij} = \beta_1 x_{ij1} + \beta_2 x_{ij2} + \beta_3 x_{ij3};$$

$$\text{M2 (Center): } \eta_{ij} = v_{j0} + \beta_1 x_{ij1} + \beta_2 x_{ij2} + \beta_3 x_{ij3}, \text{ with } v_{j0} \sim N(0, \sigma_0^2);$$

$$\text{M3 (Corr): } \eta_{ij} = v_{j0} + (\beta_1 + v_{j1})x_{ij1} + \beta_2 x_{ij2} + \beta_3 x_{ij3}, \text{ with } (v_{j0}, v_{j1}) \sim BN,$$

where ‘F-G’, ‘Center’ and ‘Corr’ indicate Fine-Gray model without frailties, subhazard frailty model with the random center effect v_{j0} and subhazard correlated frailty model with $\rho = \text{Corr}(v_{j0}, v_{j1})$, respectively. Here M3 ($\sigma_0^2 > 0, \sigma_1^2 > 0, \rho \neq 0$) is our full model and the others are various simplifications of it by assuming null components, i.e. M1 ($v_{j0} = 0, v_{j1} = 0; \sigma_0^2 = 0, \sigma_1^2 = 0$) and M2 ($v_{j1} = 0; \sigma_0^2 > 0, \sigma_1^2 = 0$). The estimation results are listed in Table 4.

In all the three subhazard models the two fixed effects ($\beta_j, j = 1, 2$) are significant, except for β_3 . In particular, the use of tamoxifen (tamoxifen = 1) significantly reduces the risk of local or regional recurrence (Type 1 event) as compared to patients who receive placebo (placebo = 0). We also observe that overall, there are no substantial changes in the fixed-effects estimates, although the effect of main treatment (β_1) becomes slightly weaker due to the

increased standard error when the two random components and their correlation are included as in M3. In M2 and M3, the variance components (σ_0^2 and σ_1^2) indicate the amount of variation between centers in baseline risk (i.e. center effect) and in the treatment effect, respectively. Here, the estimate and SE of σ_1^2 are relatively larger than those of σ_0^2 , which is also confirmed in Figure 2. Furthermore, the correlated model M3 explains the degree of dependency between the two random components (i.e. the random center effect v_0 and the random treatment-by-center interaction v_1). The estimate of ρ ($\hat{\rho} = -0.721$) gives a negative value, indicating that the two predicted random components ($\hat{v}_0$ and $\hat{v}_1$) have a negative correlation. In particular, the estimate of β_1 in M3 is negative; we see that a decreasing value of v_{1i} corresponds to a larger treatment effect. Thus, the negative correlation leads to the conclusion that treatment confers more benefit in centers with a higher baseline risk. This is consistent with the findings by Rondeau et al.⁹ in the context of meta-analysis and by Ha et al.¹⁰ in that of multi-center trials.

5.3 Investigating and testing for heterogeneity

We demonstrate how to investigate heterogeneity related to treatment effect over centers using confidence intervals^{17,18} for frailties of the individual centers. Note that the standard intervals using $p_\tau(h_w^*)$ in (9) can be null due to zero estimation of the variance components, especially for small sample sizes or small variance components.^{18,33,34} Thus we follow Ha et al.'s¹⁸ modification of (9) to deal with such shrinkage issue¹² for frailty. That is, for estimation of frailty parameters θ we use a further adjusted likelihood, defined by

$$p_{\text{adj}} = p_\tau(h_w^*) + \log \det(\sum_i), \quad (14)$$

which leads to non-negative variance-component estimators. Following Ha et al.¹⁸ and (14), the individual $(1 - \alpha)$ -level h-likelihood confidence intervals for the unidimensional components v_k of random effects v are of the form

$$\hat{v}_k \pm z_{\alpha/2} \cdot \text{SE}(\hat{v}_k - v_k), \quad (15)$$

where $\hat{v}$ maximizes the profile h-likelihood h_w^* in (7), $z_{\alpha/2}$ is the normal quantile with probability $\alpha/2$ in the right tail, and $\text{SE}(\hat{v}_k - v_k)$ are obtained from $H(h_w^*; \hat{\beta}, \hat{v})^{-1}$. In particular, Ha et al.¹⁸ have shown via numerical studies that in a general class of frailty models without competing risks, the adjusted h-likelihood interval (15) preserves well the nominal interval. Figure 2 shows the estimates and 95% confidence intervals^{17,18} for the random effects in the 167 centers using the subhazard correlated model M3. Here, centers are ordered by the number of patients entered. Figures 2(a) and 2(b) give the confidence intervals for the random center effect (v_0) and the random treatment-by-center interaction (v_1), respectively. Overall, the lengths of the intervals are seen to decrease as the number of patients per center increases: see also Vaida and Xu⁸ and Ha et al.¹⁰.

Figure 2(a) indicates overall homogeneity in the baseline risk across 167 centers (i.e. no variation in random center effect). Figure 2(b) also shows there is no substantial variation in the effect of treatment across centers although three centers (148, 164 and 165) among 167 centers noticeably stand out. Note here that the centers (148, 165) and 164 provide the lowest and the highest treatment-by-center interactions, respectively, but that the corresponding three intervals include zero; this indicates there is little treatment-by-center interaction in this data set. Thus, in this multicenter trial there is little variation in the treatment effects across centers and the treatment is shown to be effective. These results suggest that the treatment effect may be generalized to a broader patient population as in the findings by Yamaguchi and Ohashi³⁵ and Ha et al.¹⁰

Now, we show this heterogeneity can also be tested via the restricted h-likelihood in (9). Recently, Katsahian and Boudreau²⁷ proposed how to test such heterogeneity (i.e. variation in random center effect) using the PPL method under the subhazard model (M2) with random center effect only. However, the heterogeneities from random treatment-by center interaction as well as random center effect should be simultaneously tested. For this purpose we again consider the three models in Section 5.2. Note that although we report the SEs of the σ^2 s in Table 4, one should not use them for testing $H_0: \sigma_0^2 = 0$.^{8, 10} Firstly, the null hypothesis $H_0: \sigma_0^2 = 0$ (i.e. no center effect) lies on the boundary of the parameter space. Consider the difference of deviance $-2p_{\beta,v}(h_w^*)$ between two models (M1 and M2). As shown in the simulation study of Section 4, under the null hypothesis this LR test statistic asymptotically follows a chi-square mixture $(\chi_0^2 + \chi_1^2)/2$.^{10,32,36} From Table 4, we obtain the deviance difference between M1 and M2 to be 1.1 (p-value=0.147), indicating that the random center effect is not significant (i.e. $\sigma_0^2 = 0$). Furthermore, the difference in deviance between M2 and M3 is 3.7 (p-value=0.106) with an asymptotic $(\chi_1^2 + \chi_2^2)/2$ statistic³⁷, leading to $\sigma_1^2 = 0$. Accordingly, we find that there are no substantial variations for the baseline risk and treatment effect over centers, which confirms the homogeneity evident in Figure 2.

6 Discussion

We have shown that the proposed correlated frailty modelling approach based on the h-likelihood provides systematically more informative results for multi-center competing risk data. We have also demonstrated via a practical example how to investigate the heterogeneity related to treatment effect over centers and how to test such heterogeneity.

Our h-likelihood procedure can be also applied for fitting the cause-specific PH frailty models^{7,38}, by using the risk set in the classical Cox model. It would be an interesting comprehensive analysis to compare the inference results from both the sub-hazard and cause-specific frailty models for the Type 1 event in the breast cancer dataset presented in Section 5. We have shown via a simulation study that the LR tests based on the h-likelihood perform well under the subhazard shared frailty model (2) with one frailty parameter only. However, a simulation study under the subhazard frailty model (3) with a general correlation structure (4) might be challenging due to the number of parameters to be dealt with, and hence it would be an interesting future work. Another further work is to investigate the

performance of the proposed method via a simulation study when the cluster size n_i is random or data-directed unbalanced case¹⁰ because in current simulation settings, n_i is always fixed and at least 10.

The subhazard frailty models (1) implicitly assume that the frailty effects for the event of interest are independent of those for the other types of events (i.e. competing events). Developing an extended frailty modelling approach to allow a correlation between both events would be an interesting topic for future work.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. 2010-0021165). This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MEST) (No. 2011-0030810). Dr. Jeong's research was supported in part by National Institute of Health (NIH) grants 5-U10-CA69974-09 and 5-U10-CA69651-11.

Appendix: Proof of estimating equations in (8)

The weighted partial h-likelihood h_w^* in (7) can be expressed as

$$h_w^* = h_w |_{\lambda_{01}^s = \hat{\lambda}_{01}^w}, \quad (16)$$

where $h_w = \ell_w + \sum_i \ell_i$ and $\ell_w = \sum_k d_{(k)} \log \lambda_{01k}^s + \sum_{ij} \delta_{ij} \eta_{ij} - \sum_{ij} \mu_{ij}$. Here $\mu_{ij} = \Lambda_{01}^s(y_{ij}) w_{ij} \exp(\eta_{ij})$ with $\Lambda_{01}^s(y_{ij}) = \sum_k \lambda_{01k}^s I \{ (i, j) \in R_{(k)} \}$, and

$$\hat{\lambda}_{01k}^w(\beta, v) = \frac{d_{(k)}}{\sum_{(i,j) \in R_{(k)}} w_{ij} \exp(\eta_{ij})},$$

are solutions of the estimating equations, $\partial h_w / \partial \lambda_{01k}^s = 0$, for $k = 1, \dots, D$. Thus, Ha et al.'s¹⁰ procedures for standard correlated frailty models without competing risks can be extended to the subhazard model (1) by using h_w^* in (16). That is, given frailty parameters θ , the MHL estimators of $\tau = (\beta^T, v^T)^T$ are obtained by solving the joint estimating equations, $\partial h_w^* / \partial \tau = 0$. Here, the calculations in Ha and Lee²⁶ and Ha et al.¹⁰ show that

$$\partial h_w^* / \partial \tau = \partial h_w / \partial \tau |_{\lambda_{01}^s = \hat{\lambda}_{01}^w} = \left\{ E^T(\delta - \mu) - F^T \tau \right\} |_{\lambda_{01}^s = \hat{\lambda}_{01}^w}$$

since $h_w/\tau = (\eta/\tau)(h_w/\eta)$ with $\eta = X\beta + Zv = E\tau$. Here $E = (X, Z)$, $F = BD(0, \sum_1^{-1}, \dots, \sum_q^{-1})$, and δ and μ are the $n \times 1$ vectors of δ_{ij} 's, and μ_{ij} 's, respectively. Note that the vector μ can be written as a simple form by using a weighted risk indicator matrix M which contains the weight w_{ij} as well as the risk set $R_{(k)}$. Let L be the $n \times 1$ vector of L_{ij} 's with $L_{ij} = \Lambda_{01}^s(y_{ij})w_{ij}$. Since $\Lambda_{01}^s(y_{ij}) = \sum_k \lambda_{01k}^s I\{(i, j) \in R_{(k)}\}$ and $w_{ij} = \hat{G}(y_{(k)})/\hat{G}(y_{ij} \wedge y_{(k)})$, we have $L = MAJ$, where M is the $n \times D$ weighted-risk indicator matrix whose (ij, k) th element is $m_{ij,k}$, $A = \text{diag}(\lambda_{01k}^s)$ is the $D \times D$ diagonal matrix and J is the $D \times 1$ vector with one. This gives $\mu = W_0(MAJ)$ with $W_0 = \text{diag}\{\exp(\eta_{ij})\}$. Note here that $m_{ij,k}$ are constructed by combining $R_{(k)}$ and w_{ij} as in Ruan and Gray³⁰:

$$m_{ij,k} = I\{y_{ij} \geq y_{(k)} \text{ or } (y_{ij} \leq y_{(k)} \text{ and } \varepsilon_{ij} \neq 1)\} \left\{ \hat{G}(y_{(k)})/\hat{G}(y_{ij} \wedge y_{(k)}) \right\} = I\{y_{ij} \geq y_{(k)}\} + I\{y_{ij} \leq y_{(k)} \text{ and } \varepsilon_{ij} \neq 1\} \left\{ \hat{G}(y_{(k)})/\hat{G}(y_{ij}) \right\}.$$

(17)

This is also equivalent to the weights by Katsahian et al.⁶ and Katsahian and Boudreau²⁷ because $m_{ij,k}$ are equal to one as long as individuals have not failed by time $y_{(k)}$ (i.e. $y_{ij} \geq y_{(k)}$), and below 1 and decreasing over time if they failed from another type (Type 2) before $y_{(k)}$ (i.e. $y_{ij} < y_{(k)}$ and $\varepsilon_{ij} = 1$), and zero otherwise (e.g. they failed from Type 1 or have been right censored).

Furthermore, using the computation of Ha and Lee²⁶, we have

$$-\partial^2 h_w^*/\partial \tau^2 = E^T W^* E + F, \quad (18)$$

where $W^* = W_1 - W_2$, $W_1 = \text{diag}(\mu)$, $W_2 = (W_0 M) C^{-1} (W_0 M)^T$, $C = \text{diag}\{d_{(k)}/(\lambda_{01k}^s)^2\}$ is the $D \times D$ diagonal matrix, and $F = BD(0, U)$. Following Ha and Lee²⁶ and (18), we can show that given θ , the MHL estimators of $\tau = (\beta^T, v^T)^T$ are obtained from the following score equations:

$$(E^T W^* E + F)\hat{\tau} = E^T w^*,$$

leading to (8). Here $w^* = W^* \eta + (\delta - \mu)$. Note here that the λ_{01k}^s terms in W^* and w^* are evaluated at their estimates $\hat{\lambda}_{01k}^w = d_{(k)}/M_k^T \psi$, where M_k is the k th component vector of $M = (M_1, \dots, M_D)$ and ψ is the vector of $\exp(\eta_{ij})$'s. This completes the proof.

References

1. Pintilie, M. Competing risks: A practical perspective. Wiley; 2006.

2. Prentice R, Kalbfleisch JD, Peterson AV, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. *Biometrics*. 1978; 34:541–554. [PubMed: 373811]
3. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*. 1999; 94:496–509.
4. Hougaard, P. *Analysis of Multivariate Survival Data*. New York: Springer; 2000.
5. Duchateau, L., Janssen, P. *The Frailty Models*. New York: Springer; 2008.
6. Katsahian S, Resche-Rigon M, Chevret S, Porcher R. Analysing multicentre competing risk data with a mixed proportional hazards model for the subdistribution. *Statistics in Medicine*. 2006; 25:4267–4278. [PubMed: 16960919]
7. Christian, NJ. PhD thesis. Department of Biostatistics, University of Pittsburgh; 2011. Hierarchical likelihood inference on clustered competing risk data.
8. Vaida F, Xu R. Proportional hazards model with random effects. *Statistics in Medicine*. 2000; 19:3309–3324. [PubMed: 11122497]
9. Rondeau V, Michiels S, Liqueur B, Pignon JP. Investigating trial and treatment heterogeneity in an individual patient data meta-analysis of survival data by means of the penalized maximum likelihood approach. *Statistics in Medicine*. 2008; 27:1894–1910. [PubMed: 18069745]
10. Ha ID, Sylvester R, Legrand C, MacKenzie G. Frailty modelling for survival data from multi-centre clinical trials. *Statistics in Medicine*. 2011; 30:2144–2159. [PubMed: 21563206]
11. Ripatti S, Palmgren J. Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*. 2000; 56:1016–1022. [PubMed: 11129456]
12. Therneau TM, Grambsch PM, Pankratz VS. Penalized survival models and frailty. *Journal of Computational and Graphical Statistics*. 2003; 12:156–175.
13. Lee Y, Nelder JA. Hierarchical generalized linear models (with discussion). *Journal of the Royal Statistical Society, Series B*. 1996; 58:619–678.
14. Lee, Y., Nelder, JA., Pawitan, Y. *Generalised Linear Models with Random Effects: Unified Analysis via h-Likelihood*. Chapman and Hall; 2006.
15. Fisher B, Costantino J, Redmond C, et al. A randomized clinical trial evaluating tamoxifen in the treatment of patients with node-negative breast cancer who have estrogen receptor-positive tumors. *New England Journal of Medicine*. 1989; 320:479–484. [PubMed: 2644532]
16. Fisher B, Dignam J, Bryant J, et al. Five versus more than five years of tamoxifen therapy for breast cancer patients with negative lymph nodes and estrogen receptor- positive tumors. *Journal of the National Cancer Institute*. 1996; 88:1529–1542. [PubMed: 8901851]
17. Lee Y, Nelder JA. Likelihood inference for models with unobservables: another view (with discussion). *Statistical Science*. 2009; 24:255–293.
18. Ha ID, Vaida F, Lee Y. Interval estimation of random effects in proportional hazards models with frailties. *Statistical Methods in Medical Research*. 2013 Published online: 29/January/2013.
19. Ha ID, Lee Y, Song JK. Hierarchical likelihood approach for frailty models. *Biometrika*. 2001; 88:233–243.
20. Ha ID, Lee Y, MacKenzie G. Model selection for multi-component frailty models. *Statistics in Medicine*. 2007; 26:4790–4807. [PubMed: 17476647]
21. Breslow NE. Discussion of Professor Cox's paper. *Journal of the Royal Statistical Society, Series B*. 1972; 34:216–217.
22. Fan J, Li R. Variable selection for Cox's proportional hazards model and frailty model. *The Annals of Statistics*. 2002; 30:74–99.
23. McGilchrist C. REML estimation for survival models with frailty. *Biometrics*. 1993; 49:221–225. [PubMed: 8513103]
24. Pintilie M. Analysing and interpreting competing risk data. *Statistics in Medicine*. 2007; 26:1360–1367. [PubMed: 16900575]
25. Kuk D, Varadhan R. Model selection in competing risks regression. *Statistics in Medicine*. 2013; 32:3077–3088. [PubMed: 23436643]
26. Ha ID, Lee Y. Estimating frailty models via Poisson hierarchical generalized linear models. *Journal of Computational and Graphical Statistics*. 2003; 12:663–681.

27. Katsahian S, Boudreau C. Estimating and testing for center effects in competing risks. *Statistics in Medicine*. 2011; 30:1608–1617. [PubMed: 21341296]
28. Gray RJ. Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *Journal of the American Statistical Association*. 1992; 87:942–951.
29. Verweij JM, Van Houwelingen HC. Penalized likelihood in Cox regression. *Statistics in Medicine*. 1994; 13:2427–2436. [PubMed: 7701144]
30. Ruan PK, Gray RJ. Analyses of cumulative incidence functions via nonparametric multiple imputation. *Statistics in Medicine*. 2008; 27:5709–5724. [PubMed: 18712779]
31. Ha ID, Noh M, Lee Y. Bias reduction of likelihood estimators in semiparametric frailty models. *Scandinavian Journal of Statistics*. 2010; 37:307–320.
32. Self SG, Liang KY. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*. 1987; 82:605–610.
33. Morris CN. Mixed model prediction and small area estimation. *Test*. 2006; 15:72–76.
34. Li H, Lahiri P. An adjusted maximum likelihood method for solving small area estimation problems. *Journal of Multivariate Analysis*. 2010; 101:882–892. [PubMed: 20161652]
35. Yamaguchi T, Ohashi Y. Investigating centre effects in a multi-centre clinical trial of superficial bladder cancer. *Statistics in Medicine*. 1999; 18:1961–1971. [PubMed: 10440879]
36. Ha ID, Lee Y. Multilevel mixed linear models for survival data. *Lifetime Data Analysis*. 2005; 11:131–142. [PubMed: 15747594]
37. Verbeke G, Molenberghs G. The use of score test for inference on variance components. *Biometrics*. 2003; 59:254–262. [PubMed: 12926710]
38. Gorfine M, Hsu L. Frailty-based competing risks model for multivariate survival data. *Biometrics*. 2011; 67:415–426. [PubMed: 20707868]

Figure 1. Estimated CIFs for tamoxifen vs placebo for the two types of events in the breast cancer data.

Figure 2. Random effects of 167 centers in the breast cancer data (event of interest is Type 1) and their 95% confidence intervals, under subhazard correlated frailty model (M3); (a) random center effects (v_{j0}); (b) random treatment-by-center interaction (i.e. random treatment effects) (v_{j1}); Centers are sorted in increasing order of number of patients.

Table 1

Case A: $(p, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}) = (0.3, 0.5, 0.5, -0.5, 0.5)$; Simulation results for the estimation of parameters over 1000 replications under the subhazard correlated frailty model

Censoring	Sample Size	Parameter	True	Mean	SD	SE1	SE2
25%	$n = 200$ ($q = 20, n_i = 10$)	β_{11}	0.5	0.502	0.307	0.306	0.254
		β_{12}	0.5	0.511	0.137	0.134	0.130
		σ_0^2	0.5	0.534	0.451	0.439	–
		σ_1^2	0.5	0.583	0.654	0.652	–
		σ_0	–0.25	–0.285	0.471	0.438	–
		β_{11}	0.5	0.487	0.245	0.238	0.179
	$n = 400$ ($q = 20, n_i = 20$)	β_{12}	0.5	0.505	0.090	0.092	0.091
		σ_0^2	0.5	0.497	0.300	0.303	–
		σ_1^2	0.5	0.505	0.377	0.405	–
		σ_0	–0.25	–0.245	0.286	0.284	–
		β_{11}	0.5	0.490	0.157	0.150	0.112
		β_{12}	0.5	0.497	0.056	0.058	0.057
50%	$n = 1000$ ($q = 50, n_i = 20$)	σ_0^2	0.5	0.492	0.191	0.182	–
		σ_1^2	0.5	0.486	0.234	0.233	–
		σ_0	–0.25	–0.241	0.174	0.168	–
		β_{11}	0.5	0.512	0.362	0.357	0.309
		β_{12}	0.5	0.515	0.164	0.159	0.156
		σ_0^2	0.5	0.550	0.539	0.558	–
	$n = 200$ ($q = 20, n_i = 10$)	β_{11}	0.5	0.512	0.362	0.357	0.309
		β_{12}	0.5	0.515	0.164	0.159	0.156
		σ_0^2	0.5	0.550	0.539	0.558	–
		σ_1^2	0.5	0.583	0.654	0.652	–
		σ_0	–0.25	–0.285	0.471	0.438	–
		β_{11}	0.5	0.487	0.245	0.238	0.179

Censoring	Sample Size	Parameter	True	Mean	SD	SE1	SE2
		σ_1^2	0.5	0.603	0.758	0.821	-
		σ_0	-0.25	-0.294	0.561	0.559	-
	$n = 400$ $(q = 20, n_i = 20)$	β_{11}	0.5	0.487	0.281	0.270	0.217
		β_{12}	0.5	0.503	0.109	0.110	0.108
		σ_0^2	0.5	0.487	0.342	0.352	-
		σ_1^2	0.5	0.511	0.469	0.503	-
		σ_0	-0.25	-0.237	0.327	0.341	-
	$n = 1000$ $(q = 50, n_i = 20)$	β_{11}	0.5	0.491	0.178	0.168	0.136
		β_{12}	0.5	0.497	0.066	0.069	0.068
		σ_0^2	0.5	0.471	0.214	0.210	-
		σ_1^2	0.5	0.466	0.287	0.286	-
		σ_0	-0.25	-0.222	0.209	0.201	-

SE1, mean of estimated standard errors using H_w^{-1} and $(-\partial^2 p_{\tau}(h_w^*) / \partial \theta^2)^{-1}$ for $\beta = (\beta_{11}, \beta_{12})^T$ and $\theta = (\sigma_0^2, \sigma_1^2, \sigma_{01})^T$, respectively, over 1000 simulations;
SE2, mean of estimated standard errors using $H_w^{-1} H_1 H_w^{-1}$ for $\beta = (\beta_{11}, \beta_{12})^T$ over 1000 simulations; SD, standard deviation of estimates over 1000 simulations, is defined by $\{\sum_i (\hat{\psi}^{(i)} - \overline{\hat{\psi}})^2 / 999\}^{1/2}$, where $\hat{\psi}^{(i)}$ is the estimate of ψ in the i th replication and $\overline{\hat{\psi}} = \sum_i \hat{\psi}^{(i)} / 1000$ is the mean of $\hat{\psi}^{(i)}$, $\sigma_0^2, \sigma_1^2, \sigma_{01}$.

Table 2

Case B: $(p, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}) = (0.6, 1, -1, 1, 1)$; Simulation results for the estimation of parameters over 1000 replications under the subhazard correlated frailty model

Censoring	Sample Size	Parameter	True	Mean	SD	SE1	SE2
25%	$n = 200$ ($q = 20, n_i = 10$)	β_{11}	1	1.003	0.274	0.272	0.217
		β_{12}	-1	-1.006	0.140	0.129	0.124
		σ_0^2	0.5	0.505	0.372	0.357	-
		σ_1^2	0.5	0.520	0.472	0.493	-
		σ_0	-0.25	-0.254	0.353	0.347	-
		β_{11}	1	0.991	0.215	0.217	0.152
	$n = 400$ ($q = 20, n_i = 20$)	β_{12}	-1	-1.004	0.093	0.088	0.087
		σ_0^2	0.5	0.490	0.265	0.259	-
		σ_1^2	0.5	0.493	0.321	0.315	-
		σ_0	-0.25	-0.245	0.244	0.234	-
		β_{11}	1	0.983	0.145	0.137	0.095
		β_{12}	-1	-0.999	0.057	0.056	0.054
50%	$n = 1000$ ($q = 50, n_i = 20$)	σ_0^2	0.5	0.486	0.166	0.160	-
		σ_1^2	0.5	0.477	0.192	0.191	-
		σ_0	-0.25	-0.237	0.150	0.143	-
		β_{11}	1	1.011	0.335	0.330	0.278
		β_{12}	-1	-1.028	0.172	0.159	0.153
		σ_0^2	0.5	0.527	0.458	0.490	-

Censoring	Sample Size	Parameter	True	Mean	SD	SE1	SE2
	$n = 400$ ($q = 20, n_i = 20$)	σ_1^2	0.5	0.578	0.666	0.699	-
		σ_{01}	-0.25	-0.279	0.477	0.486	-
		β_{11}	1	0.986	0.257	0.251	0.193
		β_{12}	-1	-1.006	0.118	0.108	0.106
		σ_0^2	0.5	0.489	0.327	0.329	-
		σ_1^2	0.5	0.518	0.411	0.449	-
		σ_{01}	-0.25	-0.249	0.314	0.316	-
		β_{11}	1	0.983	0.164	0.157	0.121
		β_{12}	-1	-1.002	0.068	0.068	0.066
		σ_0^2	0.5	0.487	0.199	0.198	-
	$n = 1000$ ($q = 50, n_i = 20$)	σ_1^2	0.5	0.478	0.248	0.249	-
		σ_{01}	-0.25	-0.238	0.187	0.185	-

SE1, mean of estimated standard errors using H_w^{-1} and $(-\partial^2 p_{\tau}(h_w^*) / \partial \theta^2)^{-1}$ for $\beta = (\beta_{11}, \beta_{12})^T$ and $\theta = (\sigma_0^2, \sigma_1^2, \sigma_{01})^T$, respectively, over 1000 simulations;
SE2, mean of estimated standard errors using $H_w^{-1} H_1 H_w^{-1}$ for $\beta = (\beta_{11}, \beta_{12})^T$ over 1000 simulations; SD, standard deviation of estimates over 1000 simulations, is defined by $\{\sum_i (\hat{\psi}^{(i)} - \overline{\hat{\psi}})^2 / 999\}^{1/2}$, where $\hat{\psi}^{(i)}$ is the estimate of ψ in the i th replication and $\overline{\hat{\psi}} = \sum_i \hat{\psi}^{(i)} / 1000$ is the mean of $\hat{\psi}^{(i)}$, $\sigma_0^2, \sigma_1^2, \sigma_{01}$.

Table 3First observed event type by two treatment arms ($n = 2817$ patients)

Types of Event	Placebo	Tamoxifen	Total
Type 1: Local or regional recurrence	205	109	314 (11.15%)
Type 2: Distance recurrence, second primary or death	671	632	1303 (46.25%)
No event (Censoring)	537	663	1200 (42.60%)

Table 4

Results for fitting the three subhazard models to Type 1 event of the breast cancer data

Model	Treatment β_1 (SE)	Age β_2 (SE)	Tumor size β_3 (SE)	$\hat{\sigma}_0^2$ (SE)	$\hat{\sigma}_1^2$ (SE)	$\hat{\sigma}_{01}[\hat{\rho}]$ (SE)	$-2p_{\beta,v}(h_w^*)$
M1 (F-G)	-0.667 (0.119)	-0.026 (0.005)	0.082 (0.042)	-	-	-	4870.5
M2 (Center)	-0.672 (0.119)	-0.026 (0.005)	0.081 (0.042)	0.043 (0.051)	-	-	4869.4
M3 (Corr)	-0.658 (0.137)	-0.026 (0.005)	0.079 (0.043)	0.091 (0.026)	0.249 (0.073)	-0.108 [-0.721] (0.037)	4865.7

M1, proportional subhazard model (Fine-Gray model) without frailties;

M2, subhazard shared frailty model with random center effect only;

M3, subhazard correlated frailty model with ρ ;

$\hat{\sigma}_0^2$ and $\hat{\sigma}_1^2$, the variances of random center effect and random treatment-by-center interaction, respectively;

σ_{01} and ρ , the corresponding covariance and correlation with $\rho = \sigma_{01}/(\sigma_0\sigma_1)$;

SE, the estimated standard error for regression and frailty parameters;

$p_{\beta,v}(h_w^*)$, restricted h-likelihood in (9).