Published in final edited form as:

Stat Methods Med Res. 2016 April ; 25(2): 936–953. doi:10.1177/0962280212474059.

Interval Estimation of Random Effects in Proportional Hazards Models with Frailties

Il Do Ha,

Department of Asset Management, Daegu Haany University, Gyeongsan 712-715, South Korea
idha@dhu.ac.kr

Florin Vaida, and

Division of Biostatistics and Bioinformatics, Department of Family and Preventive Medicine,
University of California, San Diego, USA fvaida@ucsd.edu

Youngjo Lee

Department of Statistics, Seoul National University, Seoul 151-742, South Korea
youngjo@snu.ac.kr

Abstract

Semi-parametric frailty models are widely used to analyze clustered survival data. In this paper, we propose the use of the hierarchical likelihood interval for individual frailties of the clusters, not the parameters of the frailty distribution. We study the relationship between hierarchical likelihood, empirical Bayesian, and fully Bayesian intervals for frailties. We show that our proposed interval can be interpreted as a frequentist confidence interval and Bayesian credible interval under a uniform prior. We also propose an adjustment of the proposed interval to avoid null intervals. Simulation studies show that the proposed interval preserves the nominal confidence level. The procedure is illustrated using data from a multicenter lung cancer clinical trial.

Keywords

Empirical Bayes; Hierarchical likelihood; Interval estimator; Random effects; Survival analysis

1 Introduction

Frailty models, or proportional hazards models with random effects, are widely used to analyze clustered survival data.¹ It is important to investigate the potential heterogeneity in survival among clusters in order to understand and interpret the variability in the data.² Multivariate semi-parametric frailty models offer a flexible framework for modeling this heterogeneity. Such heterogeneity can be accounted for by random cluster effects on the baseline hazard. Furthermore, the effects of a treatment can vary substantially across participating centers in a multicenter clinical trial³, suggesting the inclusion of random treatment effects.

Conflict of interest statement

The Authors declare that there is no conflict of interest.

In addition to the estimation of random effects, a measure of uncertainty for these point estimates is useful and necessary. In this paper we focus on the interval estimation of the individual random effects, not the parameters of the random effects distribution. The standard methods in use are empirical Bayes (EB) confidence intervals, based on the conditional posterior distribution of random effects given the observed data and the estimated parameter values.² However, EB interval estimators have been criticized for not maintaining the nominal level.⁴ Gray³, Legrand et al.⁵ and Komarek et al.⁶ developed fully Bayesian methods. Lee and Ha⁷ proposed hierarchical likelihood (HL) methods to estimate random effects and their confidence intervals for hierarchical generalized linear models (HGLMs⁸). Recently, HL inferences of random effects are of special interest^{9,10}. In this paper, we extend these methods to semi-parametric frailty models.

One particular difficulty is that in random-effect models such as frailty models, likelihood methods may lead to zero estimates for strictly positive variance components.¹² In turn, this leads to null confidence intervals in the EB and HL approaches, clearly underestimating the variability of the random effects. These intervals will have a coverage probability below the nominal level, especially for small dispersion parameters or for small sample sizes. In the context of Fay-Herriot's¹³ small-area linear mixed models, Morris¹⁴ addressed this problem and proposed an adjustment to the restricted maximum likelihood (REML) estimation. Li and Lahiri¹⁵ advocated the use of this approach for confidence intervals of random effects. Here, we extend the adjustment proposed by Morris¹⁴ to general random-effect models, including HGLMs and frailty models. Li and Lahiri¹⁵ used the bootstrap method¹⁶; however, here we show that for frailty models, the likelihood method is sufficient and bootstrapping is not required. The proposed HL interval can be interpreted as a frequentist confidence interval and fully Bayesian credible interval under a uniform prior.⁹ Through numerical studies, we show that the proposed interval improves the EB interval by maintaining the stated nominal level.

The remainder of this paper is organized as follows. In Section 2, we discuss three interval estimators based on HL, EB, and fully Bayesian approaches. In Section 3, we review frailty models. In Section 4, we study how these intervals can be extended to frailty models, and investigate the relationships among them. In addition, we propose an adjustment to avoid null intervals for frailty models. Simulation studies are presented in Section 5. In Section 6, we conduct a case study using a multicenter lung cancer trial^{2,3} conducted by the Eastern Cooperative Oncology Group (ECOG). Finally, the discussion is presented in Section 7. The technical details are described in Appendix.

2 Interval estimators for random effects

In this section, we review the HL interval and study its relationship with EB and fully Bayesian credible intervals.

2.1 Hierarchical-likelihood confidence intervals

Consider a statistical model, described by the joint density

$$\log f_{\theta}(y, v) = \log f_{\theta}(y|v) + \log f_{\theta}(v), \quad (1)$$

where $f_{\theta}(v)$ and $f_{\theta}(y|v)$ are respectively density functions for v and $y|v$, so that they have different functional forms even though we use the same notation $f_{\theta}(\cdot)$. Here three objects (y , v , θ) denote respectively the response variables, unobservable random variables, and fixed unknown parameters. Fixed unknown parameters θ consist of regression coefficients β and dispersion parameters. When covariates x are available $f_{\theta}(y|v) = f_{\theta}(y|x, v)$. This model allows a two-stage interpretation: first, the unobservable random variables are *realized* as v^* from the distribution $f_{\theta}(v)$; second, the observed data y_o is sampled from $f_{\theta}(y|v^*)$. Once the data y_o are observed, y_o are known; however, v^* remain unknown. For the likelihood inference of unknowns (θ , v^*), Lee and Nelder⁸ proposed the use of the HL defined by

$$h = h(\theta, v^*) = \log f_{\theta}(y_o, v^*) = m + \log f_{\theta}(v^*|y_o), \quad (2)$$

where $m = \log f_{\theta}(y_o)$ is the marginal log-likelihood, with $f_{\theta}(y) = \int f_{\theta}(y|v)f_{\theta}(v)dv$. In this paper, we omit the star in v^* unless it is necessary to highlight the fact that v^* is a fixed unknown.

The maximization of the HL (1) yields the EB-mode estimator for v , without computing $f_{\theta}(v|y_o)$ in (2). Given θ , let $\hat{v}(\theta)$ be a random-effect estimator solving $h' / v = 0$. In HGLMs, (v, β) and dispersion parameters are asymptotically orthogonal⁸; therefore, in estimating (v, β) , we need not consider the information loss caused by estimating the dispersion parameters. The negative Hessian matrix of β and v based on h is given by

$$H(h; \beta, v) \equiv -\partial^2 h / \partial(\beta, v)^2 = \begin{pmatrix} -\partial^2 h / \partial \beta \partial \beta^T & -\partial^2 h / \partial \beta \partial v^T \\ -\partial^2 h / \partial v \partial \beta^T & -\partial^2 h / \partial v \partial v^T \end{pmatrix}. \quad (3)$$

For linear mixed models, Henderson¹⁷ showed that the lower-right-hand corner of $H(h; \beta, v)^{-1}$ gives an estimate of the unconditional mean squared error (UMSE):

$$\text{UMSE} \equiv E_{\theta}[\{\hat{v}(\hat{\theta}) - v\}\{\hat{v}(\hat{\theta}) - v\}^T] = E_{\theta}[\{\hat{v}(\theta) - v\}\{\hat{v}(\theta) - v\}^T] + E_{\theta}[\{\hat{v}(\hat{\theta}) - \hat{v}(\theta)\}\{\hat{v}(\hat{\theta}) - \hat{v}(\theta)\}^T], \quad (4)$$

where $\hat{v}(\hat{\theta}) \equiv \hat{v}(\hat{\theta})|_{\hat{\theta}=\hat{\theta}}$ and $\hat{\theta}$ is either the maximum likelihood (ML) or REML estimator of θ . The second term above is the inflation caused in the UMSE because of the estimation of θ by $\hat{\theta}$. This holds generally in HGLMs.^{8,11} The UMSE above could be used to construct HL confidence intervals for v ($= v^*$) using an asymptotic normal approximation. For example, let A be the lower-right-hand corner of $H(h; \beta, \hat{v})^{-1}$ corresponding to v , with the k th diagonal element a_{kk} . Then, a $(1 - \alpha)$ HL interval for v_k is $\hat{v}_k \pm z_{\alpha/2} \cdot \sqrt{a_{kk}}$, where $z_{\alpha/2}$ is the normal quantile with probability $\alpha/2$ in the right tail. Lee and Nelder⁹ argue that this HL interval can be interpreted as a frequentist confidence interval for v^* . Asymptotically, the HL interval will include the true realized value $v_k = v_k^*$, in the long run, in $(1 - \alpha)$ proportion of cases.

Booth and Hobert¹⁸ showed that in generalized linear mixed models (GLMMs), $H(h; \beta, v)^{-1}$ in (3) gives an estimate of the conditional MSE (CMSE), defined by $\text{CMSE} \equiv E_{\theta}[\{\hat{v}(\hat{\theta}) - v\}\{\hat{v}(\hat{\theta}) - v\}^T | y_o]$, because the UMSE is the first-order approximation to the CMSE. In

the CMSE, the probability statement of the interval is based on the conditional distribution of $\{\hat{v}(\hat{\theta}) - v\} | y_0$. Therefore, the probability statement is confined to the current data y_o , whereas that of the confidence interval depends upon the repeated sampling of y .

2.2 Empirical Bayes intervals

In the Bayesian framework, fixed parameters θ are treated as random variables with a prior distribution $\pi(\theta)$. The Bayesian statistical model is

$$B = \pi(\theta) f_{\theta}(v) f_{\theta}(y|v) = \pi(\theta) f(v|\theta) f(y|v, \theta),$$

where $f(v|\theta) = f_{\theta}(v)$ and $f(y|v, \theta) = f_{\theta}(y|v)$. The maximum likelihood estimator for fixed unknown parameter θ can be interpreted as the maximum posterior estimator under a uniform prior $\pi(\theta) = 1$. Similarly, the HL, i.e. $h = \log B$ can be interpreted as the Bayesian probability under the uniform prior.

The EB interval for v is based on the (conditional) posterior distribution $f_{\hat{\theta}}(v|y_o) = f(v|y_o, \hat{\theta})$, where $\hat{\theta}$ is the ML estimator.¹⁹ The asymptotic variance of v given y_0 is estimated by $\{-d^2 \log f_{\hat{\theta}}(v|y_o)/dv^2\}^{-1}$. This can also be obtained via h in (2) without computing $f_{\hat{\theta}}(v|y_o)$ because $d \log f_{\hat{\theta}}(v|y_o)/dv = dh/dv$ and $\{-d^2 \log f_{\hat{\theta}}(v|y_o)/dv^2\}^{-1} = (-d^2 h/dv^2)^{-1}$. A uni-dimensional EB interval can be derived as follows: for v_k , a scalar component of v , the EB $(1-\alpha)$ -level interval corresponds to the $(1-\alpha)$ credible region of $f_{\hat{\theta}}(v_k|y_o) = f(v_k|y_o, \hat{\theta})$. A normal approximation of the type $\hat{v}_k \pm z_{\alpha/2} \cdot \sigma_k$ is often used, where $\sigma_k^2 = \text{var}(v_k|y_o, \theta)|_{\theta=\hat{\theta}}$ and $z_{\alpha/2}$ is the $\alpha/2$ normal quantile.

However, the downside of the EB interval is that it ignores the uncertainty in $\hat{\theta}$. That is, the EB variance estimator, the inverse of the negative Hessian matrix from $\log f_{\hat{\theta}}(v|y_o)$, is obtained from $(-d^2 h/dv^2)^{-1}$ which is the inverse of bottom right-hand corner of H in (3). To contrast the EB and HL estimators, note that the HL estimator takes into account the covariance between $\hat{\beta}$ and $\hat{v} - v$, corresponding to the off-diagonal term in the Hessian H in (3), whereas this term is ignored in the EB. This covariance term carries information about the uncertainty in estimating β and its effect on $\hat{v} - v$.^{7,9} This implies that the EB method is not satisfactory because it does not reflect the uncertainty caused by estimating θ . Bjørnstad²⁰ also showed that the EB intervals underestimate the uncertainty in v , occasionally quite severely. See also the discussion in Section 3.5 of Carlin and Louis⁴.

2.3 Bayesian credible intervals

The *Bayesian credible interval* for v is based on the marginal posterior

$$\pi(v|y_o) = \int \pi(v, \theta|y_o) d\theta,$$

where θ is integrated out. The marginal posterior variance of v is

$$\text{var}(v|y_o) = E_{\theta|y_o} \{ \text{var}(v|y_o, \theta) \} + \text{var}_{\theta|y_o} \{ E(v|y_o, \theta) \}. \quad (5)$$

Carlin and Gelfand²¹ noted that the EB variance estimate only approximates the first term in (5), and ignores the second. Laird and Louis²² and Carlin and Gelfand²¹ proposed the use of a bootstrap method to estimate the second term. In general, Kass and Steffey²³ used a Laplace approximation to show that under the improper prior $\pi(\theta) = 1$, the estimator for $\text{var}(v|y_o)$ can be obtained from $H(h; \theta, v)^{-1} = \{-h^2/(\theta, v)^2\}^{-1}$. This implies that the marginal posterior distribution of $v|y_o$ is approximately multivariate normal with mean $v\{\theta(y_o)\}$ and the variance obtained from $H(h; \theta, v)^{-1}$; the resulting HL interval can be viewed as an approximate fully Bayesian credible interval under $\pi(\theta) = 1$.

In summary, credible and confidence intervals for the random effects are defined via different considerations. Empirical and credible intervals are based on the observed data y_o , whereas the confidence interval from UMSE is based on the statistical model of y . The HL interval allows both interpretations.

3 Formulation for frailty models

Suppose that the data consist of censored time-to-event observations collected from q clusters. Let T_{ij} be the survival time for the j th observation in the i th cluster, $i = 1, \dots, q, j = 1, \dots, n_i$, $n = \sum_i n_i$. Denote by v_i an s -dimensional vector of unobserved log-frailties (random effects) associated with the i th cluster. Given v_i , the conditional hazard function of T_{ij} is of the form

$$\lambda_{ij}(t|v_i) = \lambda_0(t) \exp(\eta_{ij}), \quad (6)$$

where $\lambda_0(\cdot)$ is the unknown baseline hazard 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 $z_{ij} = (z_{ij1}, \dots, z_{ijs})^T$ are $p \times 1$ and $s \times 1$ 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, $v_i \sim N_s(0, \Sigma)$. Here, the covariance matrix $\Sigma = \Sigma(\phi)$ depends on a vector of unknown parameters ϕ . The normal distribution has been used for modelling multi-component²⁴ and correlated frailties²⁵.

Model (6) includes some well-known models as particular 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 (6), if we consider $z_{ij} = 1$ and $v_i = v_{i0}$ for all i, j , this is the random center or shared frailty model¹ with

$$\eta_{ij} = x_{ij}^T \beta + v_{i0}, \quad (7)$$

where $v_{i0} \sim N(0, \sigma_0^2)$ for all i . Model (7) 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} . By taking $z_{ij} = (x_{ij0}, x_{ij1})^T$ and $v_i = (v_{i0}, v_{i1})^T$, we have a two-random-interactions model² with

$$\eta_{ij} = (\beta_0 + v_{i0})x_{ij0} + (\beta_1 + v_{i1})x_{ij1} + \sum_{m=2}^p \beta_m x_{ijm}, \quad (8)$$

where $v_i = (v_{i0}, v_{i1})^T \sim N_2(0, \Sigma)$. In this model, by taking $x_{ij0} = 1$ for all i, j and $\beta_0 = 0$, we obtain the random coefficient model³ with

$$\eta_{ij} = v_{i0} + (\beta_1 + v_{i1})x_{ij1} + \sum_{m=2}^p \beta_m x_{ijm}. \quad (9)$$

To maintain the invariance of model to parameterization of the treatment effect, we allow a general covariance matrix^{11,25} in (8) and (9):

$$\Sigma \equiv \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix}, \quad (10)$$

where $\rho = \sigma_{01}/(\sigma_0\sigma_1)$ is the correlation between v_{i0} and v_{i1} within a cluster.

4 Interval estimators for random effects in frailty models

We now show how the HL confidence intervals for random effects can be extended to frailty models. A particular problem occurs when one or more frailty variance parameters are estimated to be zero, leading to null HL intervals. We propose a general modification to deal with this issue.

4.1 Hierarchical likelihood for frailty models

For observations j of cluster i , let T_{ij} and C_{ij} be the event and censoring times, respectively, and response variable $y_{ij} = \min\{T_{ij}, C_{ij}\}$ with event indicator $\delta_{ij} = I(T_{ij} < C_{ij})$. Ha et al.²⁶ made the following two assumptions:

Assumption 1. Given v_i , the pairs $\{(T_{ij}, C_{ij}), j = 1, \dots, n_i\}$ are conditionally independent and both T_{ij} and C_{ij} are also conditionally independent for $j = 1, \dots, n_i$.

Assumption 2. Given v_i , the set $\{C_{ij}, j = 1, \dots, n_i\}$ are noninformative about v_i .

In the semi-parametric frailty model (6), the functional form of $\lambda_0(t)$ is unknown. The non-parametric estimator²⁷ of the baseline cumulative hazard function $\Lambda_0(t) = \int_0^t \lambda_0(k)dk$ is a step function with jumps at the observed event times. Restricting ourselves to hazard functions of the above form, we have $\Lambda_0(t) = \sum_{k: Y(k) \leq t} \lambda_{0k}$, where $Y_{(1)} < \dots < Y_{(r)}$ are the ordered distinct event times and $\lambda_{0k} = \lambda_0(Y_{(k)})$. Let y^* be the vector of response variables y_{ij}^* , with $y_{ij}^* = (y_{ij}, \delta_{ij})$.

Once the data are observed, $y_{oij}^* = (y_{oij}, \delta_{oij})$. Under assumptions 1 and 2, the HL for the semi-parametric frailty model (6) is given by

$$h = h\{(\beta, \lambda_0, \phi), v\} = \sum_{ij} \ell_{1ij} + \sum_i \ell_{2i}, \quad (11)$$

where

$$\sum_{ij} \ell_{1ij} = \sum_{ij} \delta_{oij} \{ \log \lambda_0(y_{oij}) + \eta_{ij} \} - \sum_{ij} \{ \Lambda_0(y_{oij}) \exp(\eta_{ij}) \} = \sum_{k=1}^r d_{(k)} \log \lambda_{0k} + \sum_{ij} \delta_{oij} \eta_{ij} - \sum_{k=1}^r \lambda_{0k} \left\{ \sum_{(i,j) \in R_{(k)}} \exp(\eta_{ij}) \right\},$$

$\ell_{1ij} = \ell_{1ij}(\beta, \lambda_0; y_{oij}^* | v_i)$ is the logarithm of the conditional density function for $y_{ij}^* = y_{oij}^*$ given v_i

$$\ell_{2i} = \ell_{2i}(\phi; v_i) = -\frac{1}{2} \{ \log \det(2\pi \Sigma(\phi)) \} - \frac{1}{2} v_i^T \Sigma(\phi)^{-1} v_i$$

is the logarithm of the density function for v_i with parameters ϕ , and $\eta_{ij} = x_{ij}^T \beta + z_{ij}^T v_i$. Here, $v = (v_1^T, \dots, v_q^T)^T$ are vectors of v_i 's, $\lambda_0 = (\lambda_{01}, \dots, \lambda_{0r})^T$, $d_{(k)}$ is the number of events at $y_{(k)}$, and $R_{(k)} = R(y_{(k)}) = \{(i, j) : y_{oij} = y_{(k)}\}$ is the risk set at $y_{(k)}$. Since (β, λ_0, v) and ϕ are asymptotically orthogonal as in HGLMs²⁸, we only need to consider the Hessian matrix of v and $\psi = (\beta^T, \lambda_0^T)^T$. Let y_o^* be the vector of observed data points y_{oij}^* . The definitions of the Hessian matrix (3) and UMSE (4) are respectively extended to

$$H(h; \psi, v) \equiv \partial^2 h / \partial(\psi, v) \partial(\psi, v)^T \text{ and } \text{UMSE} \equiv E_\psi [\{ \hat{v}(\hat{\psi}) - v \} \{ \hat{v}(\hat{\psi}) - v \}^T]. \quad (12)$$

Here, $\hat{v}(\hat{\psi}) \equiv \hat{v}(\hat{\psi})|_{\psi=\hat{\psi}}$, where $\hat{v}(\hat{\psi})$ is the solution to $h' / v = 0$ for a given ψ . Note that $\hat{v}(\psi) = E_\psi(v | y_o^*)$ asymptotically as $N = \min_i \sum_j n_{ij} \rightarrow \infty$. Following Lee and Nelder (1996) and Lee and Ha (2010), we can be shown that the $H(h; \psi, \hat{v})^{-1}$ gives the first-order approximation to the UMSE in (12), leading to a standard error for $\hat{v} - v$ and a HL confidence interval for v .

In semi-parametric frailty models (6), the number of λ_{0k} in ψ increases with sample size n . Thus, $H(h; \psi, \hat{v})^{-1}$ requires inversion of a high-dimensional $(p + q + r)$ matrix. Following Ha et al.²⁶, we propose the use of the profiled HL, h^* , that eliminates λ_0 :

$$h^* = h|_{\lambda_0 = \hat{\lambda}_0} = \sum_k d_{(k)} \log \hat{\lambda}_{0k} + \sum_{ij} \delta_{oij} \eta_{ij} - \sum_k d_{(k)} + \sum_i \ell_{2i}, \quad (13)$$

where

$$\hat{\lambda}_{0k}(\beta, v) = \frac{d_{(k)}}{\sum_{(i,j) \in R_{(k)}} \exp(\eta_{ij})}$$

are solutions of the estimating equations $h' / \lambda_{0k} = 0$ for $k = 1, \dots, r$. Note that h^* is proportional to the penalized partial likelihood of Ripatti and Palmgren²⁹. Again, the covariance estimates for $\hat{v} - v$ are obtained from $H(h^*; \beta, v)^{-1}$, given by

$$H(h^*; \beta, v) = \begin{pmatrix} -\partial^2 h^* / \partial \beta \partial \beta^T & -\partial^2 h^* / \partial \beta \partial v^T \\ -\partial^2 h^* / \partial v \partial \beta^T & -\partial^2 h^* / \partial v \partial v^T \end{pmatrix} = \begin{pmatrix} X^T W^* X & X^T W^* Z \\ Z^T W^* X & Z^T W^* Z + U \end{pmatrix}, \quad (14)$$

where X and Z respectively denote the model matrices for β and v , $U = -2\ell_2 / v^2 = BD(\Sigma^{-1}, \dots, \Sigma^{-1})$ is a $q^* \times q^*$ block diagonal matrix with $q^* = q \times s$, and the weight matrix $W^* = W^*(\beta, v)$ is given in Appendix 2 of Ha and Lee³⁶. With the use of h^* , we need to invert the $(p+q)$ matrix $H(h^*; \beta, v)$, leading to an efficient computation of the confidence interval for v .

As in the previous section, the individual $(1 - \alpha)$ -level HL and EB confidence intervals for the uni-dimensional components v_k of 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 HL h^* in (13). Let A^* be the lower-right-hand corner matrix of dimension s from $H(h^*; \beta, \hat{v})^{-1}$, corresponding to $-2h^* / v v^T$. Then, for HL confidence intervals, $\text{SE}(\hat{v}_k - v_k)$ is $\sqrt{a_{kk}^*}$, where a_{kk}^* is the k th diagonal element of A^* corresponding to v_k . For EB confidence intervals, $\text{SE}(\hat{v}_k - v_k)$ is simply $\sqrt{h_{kk}^*}$, where h_{kk}^* is the k th diagonal element of the matrix $(-2h^* / v v^T)^{-1}$. Therefore, $\text{SE}(\hat{v}_k - v_k)$ using HL is always greater than that using EB method, so that HL interval is larger than the EB interval. In other words, as compared to the EB interval, the HL confidence interval takes into account the correlation between $\hat{v} - v$ and $\hat{\beta}$ reflected in the off-diagonal block of the Hessian matrix $H(h^*; \beta, v)$. For the frailty model (6), Vaida and Xu² used the EB interval based on the conditional

distribution $f_{\hat{\kappa}}(v|y_o^*)$, where $\hat{\kappa} = (\hat{\beta}^T, \hat{\lambda}_0^T, \hat{\phi}^T)^T$ are the ML estimators. As in the case of GLMM, this HL interval enjoys a confidence interval interpretation, based on the UMSE, and is also an approximate Bayesian credible interval under the uniform prior.

Lee and Nelder⁹, Lee and Ha⁷, Lee et al.³⁰ and Ha et al.³¹ have used the HL intervals in (15) for various random-effect models. However, it gives null intervals when variance-component estimates are zero. This can lead to liberal intervals when variance components are small. In this paper, we propose a general modification in order to overcome this shortcoming.

4.2 Non-negative variance-component estimation

As discussed above, for the frailty model (6), the maximum HL estimators of $\tau = (\beta^T, v^T)^T$ are obtained by solving, for given ϕ , the equation

$$\partial h^* / \partial \tau = (\partial h / \partial \tau)|_{\lambda_0 = \hat{\lambda}_0} = 0. \quad (16)$$

The estimation of ϕ can be carried out by extending Lee et al.'s¹¹ restricted likelihood:

$$p_\tau(h^*) = [h^* - \frac{1}{2} \log \det \{H(h^*; \tau) / (2\pi)\}]|_{\tau = \hat{\tau}}, \quad (17)$$

where $\hat{\tau} = \hat{\tau}(\phi) = (\hat{\beta}^T(\phi), \hat{v}^T(\phi))^T$ and, as before, $H(h^*; \tau) = -2h^* / \tau \tau^T$. The REML estimators for ϕ are obtained by maximizing $p_\tau(h^*)$. The restricted likelihood $p_\tau(h^*)$ is the

first-order Laplace approximation to a modified marginal likelihood, which becomes exact as $N = \min_i n_i \rightarrow \infty$. We found that the increase of N rather than q reduces bias more effectively²⁸; see Lee et al.¹¹ and Ha et al.³² for further justification of asymptotic properties.

This REML method based on (17) can give zero estimates for variance components $\sigma_s^2 > 0$ ($s = 0, 1, 2, \dots, c$).¹⁵ This leads to the null confidence interval for v in (15). This issue was recognized by Morris¹⁴ in the context of linear mixed models. To extend the Morris method¹⁴, we propose the use of the adjusted likelihood p_{adj} , defined as

$$p_{\text{adj}} = p_{\tau}(h^*) + \log \det(\Sigma). \quad (18)$$

Note that the last term in (18) is asymptotically negligible (i.e. asymptotic property of the maximum p_{adj} estimator is asymptotically the same as that of the maximum $p_{\tau}(h^*)$ estimator). Furthermore,

$$\exp(p_{\text{adj}}) = \exp\{p_{\tau}(h^*)\} \det(\Sigma) \geq 0$$

and $\exp(p_{\text{adj}}) = 0$ only if $\det(\Sigma) = 0$. Thus, by adding the last term we can effectively avoid zero estimates in dispersion parameters. Morris¹⁴ proposed this adjustment to avoid the zero REML estimators in linear mixed models. Morris and Christiansen³³ proposed a similar formula to (18) for one random-component Weibull frailty model. The adjusted likelihood p_{adj} is always defined, even when the original restricted likelihood based upon the marginal likelihood is hardly available. As shown in the subsequent sections, this correction works well in practice.

In practice when we obtain a zero estimate for a certain variance component we may use the model without the corresponding random effects. In this case it is impossible to maintain the stated level of the confidence intervals (CIs). To avoid such a drawback Benjamini and Yekutieli (2005) proposed to consider false coverage rate, which is proportion of noncovering CIs among all non-null CIs. Further research are of interest how to form CIs controlling false coverage rate. If we want CIs maintaining the stated level, we should use non-negative variance component estimators.

5 Simulation study

We conducted a numerical study, based upon 500 replications of simulated data, in order to compare the operating characteristics of the EB and HL intervals. Following the setup of the Vaida and Xu² data analysis of a multicenter lung cancer clinical trial, we consider the following two frailty models, described in (7) and (9):

$$\text{M1: } \eta_{ij} = \beta_1 x_{ij1} + \beta_2 x_{ij2} + v_{i0},$$

$$\text{M2: } \eta_{ij} = (\beta_1 + v_{i1})x_{ij1} + \beta_2 x_{ij2} + v_{i0}.$$

In simulation studies allowing a covariance Σ in (10), we assume $\lambda_0(t) = 1$, $\beta_1 = -0.5$, $\beta_2 = 0.5$, $\sigma_0^2 = \sigma_1^2 = 0.2$ for smaller variances and $\sigma_0^2 = \sigma_1^2 = 1.0$ for larger variances and $\rho = -0.5$ in

M2. Even though unreported, we found similar results for $\rho = 0.5$. The binary covariates x_{ij1} and x_{ij2} are each generated from a Bernoulli distribution with success probability 0.5. In the case study described in Section 6, there are 31 institutions ($q = 31$) and the average number of patients per institution (i.e., average n_j) is 18.7. Thus, we consider the following sample sizes: $n = \sum_{i=1}^q n_i$ with $n = 150, 300, 600$ and 1200 , and $(q, n_j) = (30, 5), (60, 5), (30, 20)$ and $(60, 20)$. The censoring times were generated from an exponential distribution with empirically determined parameter values to approximately achieve two right-censoring rates, low (around 15%) and high (around 50%). All of the computations were done using SAS/IML.

The standard HL estimates for fixed parameters β and $\phi = (\sigma_0^2, \sigma_1^2, \sigma_{01})^T$ in M1 and M2 using (16) and (17) work well (not shown) as in the simulation results by Ha et al.³¹. However, the variance components (σ_0^2, σ_1^2) are sometimes estimated to be zero, leading to null intervals. Table 1 shows the percentages of zero estimates of (σ_0^2, σ_1^2) . They increase with censoring rate, which confirms the simulation results of Vu and Knuiman¹². When $\sigma_0^2 = 0.2$ in M1, the percentage of zero estimates of σ_0^2 is 4.4% under 15% censoring with small samples ($n = 150$), but it is as high as 13.4% under 50% censoring. The trends of results in M2 are similar to those evident in M1.

HL(S) and EB(S) denote the HL and EB methods using standard REML estimators for variance components using (17), whereas HL(A) and EB(A) denote those using adjusted REML estimators based on (18). In each simulation we generate $v = v^*$ firstly using $f_0(v)$. Then, we generate y_o^* using $f_0(y^*|v^*)$, where $y_o^* = (y_o, \delta_o)$ and $y^* = (y, \delta)$. Based on y_o^* , we form the EB and HL intervals for v^* . Then, we compute coverage probabilities of CIs for individual v_i^* for $i = 1, \dots, q$ based upon the 500 replications. Using q coverage probabilities we form box-plots. Figure 1 shows the coverage probabilities of the nominal 95% intervals for individual random effects (v_{j0} 's) in M1 under 15% censoring. For a small variance ($\sigma_0^2 = 0.2$) and small sample ($n_j = 5$), both HL(S) and EB(S) intervals are liberal because they often give null intervals. Figure 1 shows that the adjustments HL(A) and EB(A) adequately correct this issue. For a large variance $\sigma_0^2 = 1.0$, however, the EB(A) does not maintain the nominal level, even when n_j is large. In contrast, the HL(A) intervals maintain the nominal 95% level in all cases studied, indicating that it is necessary to correct for the uncertainty in the estimation of β . The coverage probabilities (see Figure 1.1) of intervals for all v_{j0} in M2 are consistent with simulation for M1. Although not shown here, for M2, the results for v_{j1} 's are similar to those for v_{j0} 's.

In Table 2 we summarize the means for coverage probabilities of the nominal 95% intervals of individual random effects v_{j0} in M1 and M2, with 500 replications. Overall, the HL(A) intervals work well as judging from that the mean coverage probabilities are close to the nominal 95% level. In particular, the results of the intervals for 50% censoring are similar to those of 15% censoring and they become better as $N = \min_{1 \leq j \leq q} n_j$ increases. Furthermore, we also conducted the simulations under an extreme case that the cluster size $n_j \equiv 2$. Here we considered the three sample sizes, $(q, n_j) = (50, 2), (100, 2)$ and $(200, 2)$. Table 2 shows that the proposed HL(A) intervals are improved as q increases.

In summary, we recommend the use of HL(A) for confidence intervals of realized values v^* of random effects.

6 Case study: a multicenter lung cancer clinical trial

We examine the data from the EST 1582 multicenter lung cancer trial³⁴. This trial enrolled 579 patients from 31 distinct institutions (centers). The number of patients per institution varied from 1 to 56, with a mean of 18.7 and median of 17. The subjects were randomized to one of two treatment arms, standard chemotherapy (CAV) or an alternating regimen (CAV-HEM). The primary endpoint was the time (in years) from randomization to death. The study had a high mortality rate: of the 579 patients, 569 died (censoring rate = 1.7%), with a median survival time of 0.86 years and maximum follow-up of 8.45 years. The five dichotomous covariates considered are treatment (x_{ij1} is 0 for CAV and 1 for CAV-HEM), presence(1) or absence(0) of bone metastases (x_{ij2}), presence(1) or absence(0) of liver metastases (x_{ij3}), whether the subject was ambulatory ($x_{ij4} = 1$) or confined to bed or chair ($x_{ij4} = 0$), and whether there was weight loss prior to entry (x_{ij5}).

These data were previously analyzed using a fully Bayesian approach via Gibbs sampling by Gray³ and by marginal likelihood via Monte Carlo EM by Vaida and Xu². Vaida and Xu² used EB interval estimation for random effects. Gray³ used a restricted data set (570 patients from 26 institutions) with cluster size $n_i \leq 5$ and a single covariate x_{ij1} . Such a restriction is often applied with fixed effects because random-effect estimates for centers with fewer patients are imprecise. But with random-effect model, the uncertainty is automatically accounted for. Thus, our approach can include all centers, indeed even those with only a single patient. However, all authors above assumed $\rho = 0$ for two random effects.

Let v_{j0} and v_{j1} be random center effects and random treatment effects, and let v_{j2} be random effects for presence of bone metastases. Following Vaida and Xu², we consider the following four models, which include Cox model without random effects and three frailty models (FMs):

$$\text{FM0 (Cox model): } \eta_{ij} = \beta_1 x_{ij1} + \beta_2 x_{ij2} + \beta_3 x_{ij3} + \beta_4 x_{ij4} + \beta_5 x_{ij5};$$

$$\text{FM1 (model with } v_{j0}): \eta_{ij} = v_{j0} + \beta_1 x_{ij1} + \beta_2 x_{ij2} + \beta_3 x_{ij3} + \beta_4 x_{ij4} + \beta_5 x_{ij5};$$

$$\text{FM2 (model with } v_{j0} \text{ \& } v_{j1}): \eta_{ij} = v_{j0} + (\beta_1 + v_{j1})x_{ij1} + \beta_2 x_{ij2} + \beta_3 x_{ij3} + \beta_4 x_{ij4} + \beta_5 x_{ij5};$$

$$\text{FM3 (model with } v_{j1} \text{ \& } v_{j2}): \eta_{ij} = (\beta_1 + v_{j1})x_{ij1} + (\beta_2 + v_{j2})x_{ij2} + \beta_3 x_{ij3} + \beta_4 x_{ij4} + \beta_5 x_{ij5}.$$

In two random-component models (FM2 and FM3) we always assume correlation. The HL results are listed in Table 3. The HL estimates are similar to the marginal likelihood estimates by Vaida and Xu². Ha et al.²⁴ and Plummer³⁵ pointed out that the use of the Akaike information criterion (AIC) for model selection can be problematic when there are random effects. To avoid such problem in model comparison with the common fixed effects as above, Ha et al.²⁴ proposed the use of the AIC (hAIC), based on the restricted likelihood (17):

$$\text{hAIC} = -2p_\tau(h^*) + 2d, \quad (19)$$

where the restricted likelihood $p_{\tau}(h^*)$ is a function only of φ and d is the number of frailty parameters in φ . With the hAIC, FM3 is chosen. Under FM3, $\sigma_2^2 = \text{var}(v_{i2})$ is much larger than $\sigma_1^2 = \text{var}(v_{i1})$, which is also confirmed in Figure 2.

The HL(A) random-effect estimates and their 95% intervals for each institution are plotted in Figure 2(a),(b) under a final model FM3. In particular, Figure 2(b) shows substantial institutional variation among the bone effects; the HL(A) intervals for the two institutions (22, 29) do not include zero. Overall, the HL(A) intervals are similar to the EB(S) intervals obtained by Vaida and Xu².

Table 3 shows that in FM1 and FM2, the HL estimates for the variance, $\sigma_0^2 = \text{var}(v_{i0})$, of the random center effect are very close to zero. Now we investigate the behavior of the proposed adjustment method HL(A) as compared to the standard HL(S) in the case of zero frailty variance. For this we consider the correlated model (FM2), a general model of FM1. The HL random effects and their 95% intervals for each institution under FM2 are plotted in Figure 3. For the random center effects, the HL(S) and HL(A) intervals are shown in Figures 3(a) and 3(b), respectively. Due to almost zero estimate for σ_0^2 , the HL(S) intervals shown in Figure 3(a) are substantially null, whereas the HL(A) intervals shown in Figure 3(b) have an average length of 0.57 with a standard deviation of 0.05. However, all HL(A) intervals shown in Figure 3(b) include zero, strongly indicating that random effects could be null. By having non-null intervals, HL(A) maintains the stated level of confidence. In particular, we observe that the resulting non-null HL(A) intervals and homogeneity over institutions from the HL(A) plot are similar to those from Gray's Figure 1(a) using fully Bayesian credible intervals. Next, for the random treatment effects we present only the HL(A) plot shown in Figure 3(c), because they were similar to the HL(S) intervals. Figure 3(c) shows that there are substantial variations in the treatment effect over institutions.^{2,3} For example, six institutions (13, 16, 18, 19, 28, 29) noticeably stand out. Gray³ noted that $\exp(v_{j1}) = \exp(\beta_1 + v_{j1})/\exp(\beta_1)$ is the ratio of treatment hazard rate in the j th institution to overall treatment hazard rate, so that a decreasing value of v_{j1} corresponds to an increase of treatment effect.²⁵

7 Discussion

Several methods for generating confidence intervals have been proposed for mixed effects models.¹⁵ In frailty models, HL leads to confidence intervals for random effects, improving upon the EB intervals by appropriately accounting for the variability in the estimation of the fixed effects. Numerical studies show that this improvement is substantial in terms of obtaining close to nominal coverage for these intervals, whereas the EB intervals suffer from under-coverage, especially for large variance components. However, these intervals can be null due to zero estimation of the variance components, especially for small sample sizes and/or small variance components. A correction proposed by Morris¹⁴ can be extended to avoid null intervals. These methods are straightforward to implement, and integrate seamlessly with the estimation of the general frailty model using HL.

In the literature, the word *prediction* is often used in reference to random effects to distinguish this from the estimation of the fixed effects. However, in this paper, following

Lee et al.¹¹, we use the term *estimation*, and we view the random effects as already realized unknowns. Accordingly, for simplicity, we use the term *confidence interval* for the interval that summarizes the uncertainty of estimation of random effects.

Acknowledgements

The authors thank Professor Robert J. Gray for the permission to use the lung cancer trial data from the Eastern Cooperative Oncology Group (ECOG) study (E1582). This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. 2009-0088978 and 2010-0011372).

References

1. Hougaard, P. Analysis of Multivariate Survival Data. New York: Springer; 2000.
2. Vaida F, Xu R. Proportional hazards model with random effects. *Statistics in Medicine*. 2000; 19:3309–3324. [PubMed: 11122497]
3. Gray RJ. A Bayesian analysis of institutional effects in multicenter cancer clinical trial. *Biometrics*. 1994; 50:244–253. [PubMed: 8086608]
4. Carlin, BP.; Louis, TA. Bayes and Empirical Bayes Methods for Data Analysis. 2nd edn.. London: Chapman and Hall; 2000.
5. Legrand C, Ducrocq V, Janssen P, Sylvester R, Duchateau L. A Bayesian approach to jointly estimate centre and treatment by centre heterogeneity in a proportional hazards model. *Statistics in Medicine*. 2005; 24:3789–3804. [PubMed: 16320274]
6. Komarek A, Lesaffre E, Legrand C. Baseline and treatment effect heterogeneity for survival times between centers using a random effects accelerated failure time model with flexible error distribution. *Statistics in Medicine*. 2007; 26:5457–5472. [PubMed: 17910009]
7. Lee Y, Ha ID. Orthodox BLUP versus HL methods for inferences about random effects in Tweedie mixed models. *Statistics and Computing*. 2010; 20:295–303.
8. Lee Y, Nelder JA. Hierarchical generalized linear models (with discussion). *Journal of the Royal Statistical Society, Series B*. 1996; 58:619–678.
9. Lee Y, Nelder JA. Likelihood inference for models with unobservables: another view (with discussion). *Statistical Science*. 2009; 24:255–293.
10. Meng X-L. What's the H in H-likelihood: A holy grail or an achilles' heel? (with discussion). *Bayesian Statistics*. 2011; 9:473–500.
11. Lee, Y.; Nelder, JA.; Pawitan, Y. Generalised Linear Models with Random Effects: Unified Analysis via h-Likelihood. Chapman and Hall; 2006.
12. Vu HTV, Knuiman MW. A hybrid ML-EM algorithm for calculation of maximum likelihood estimates in semiparametric shared frailty models. *Computational Statistics and Data Analysis*. 2002; 40:173–187.
13. Fay RE, Herriot RA. Estimates of income for small places: an application of James-Stein procedure to Census data. *Journal of the American Statistical Association*. 1979; 74:269–277.
14. Morris CN. Mixed model prediction and small area estimation. *Test*. 2006; 15:72–76.
15. 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]
16. Efron, B.; Tibshirani, TJ. An Introduction to the Bootstrap. Chapman and Hall; 1993.
17. Henderson CR. Best linear unbiased estimation and prediction under a selection model. *Biometrics*. 1975; 31:423–447. [PubMed: 1174616]
18. Booth JG, Hobert JP. Standard errors of prediction in generalized linear mixed models. *Journal of the American Statistical Association*. 1998; 93:262–272.
19. Morris CN. Parametric empirical Bayes inference: theory and application. *Journal of the American Statistical Association*. 1983; 78:47–59.
20. Bjørnstad JF. Predictive likelihood principle: a review (with discussion). *Statistical Science*. 1990; 5:242–265.

21. Carlin BP, Gelfand AE. Approaches for empirical Bayes confidence intervals. *Journal of the American Statistical Association*. 1990; 85:105–114.
22. Laird NM, Louis TA. Empirical Bayes confidence intervals based on bootstrap samples. *Journal of the American Statistical Association*. 1987; 82:739–750.
23. Kass RE, Steffey D. Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). *Journal of the American Statistical Association*. 1989; 84:717–726.
24. Ha ID, Lee Y, MacKenzie G. Model selection for multi-component frailty models. *Statistics in Medicine*. 2007; 26:4790–4807. [PubMed: 17476647]
25. 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]
26. Ha ID, Lee Y, Song JK. Hierarchical likelihood approach for frailty models. *Biometrika*. 2001; 88:233–243.
27. Breslow NE. Discussion of Professor Cox's paper. *Journal of the Royal Statistical Society, Series B*. 1972; 34:216–217.
28. Ha ID, Lee Y. Comparison of hierarchical likelihood versus orthodox best linear unbiased predictor approaches for frailty models. *Biometrika*. 2005; 92:717–723.
29. Ripatti S, Palmgren J. Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*. 2000; 56:1016–1022. [PubMed: 11129456]
30. Lee Y, Jang M, Lee W. Prediction interval for disease mapping using hierarchical likelihood. *Computational Statistics*. 2011; 26:159–179.
31. 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]
32. Ha ID, Noh M, Lee Y. Bias reduction of likelihood estimators in semi-parametric frailty models. *Scandinavian Journal of Statistics*. 2010; 37:307–320.
33. Morris CN, Christiansen C. Fitting Weibull duration models with random effects. *Lifetime Data Analysis*. 1995; 1:347–359. [PubMed: 9385109]
34. Ettinger DS, Finkelstein DM, Abelson MD, Ruckdeschel JC, Aisner SC, Eggleston JC. A randomized comparison of standard chemotherapy versus alternating chemotherapy and maintenance versus no maintenance therapy for extensive-stage small-cell lung cancer: a phase III study of the Eastern Cooperative Oncology Group. *Journal of Clinical Oncology*. 1990; 8:230–240. [PubMed: 2153765]
35. Plummer M. Penalized loss functions for Bayesian model comparison. *Biostatistics*. 2008; 9:523–539. [PubMed: 18209015]
36. Ha ID, Lee Y. Estimating frailty models via Poisson hierarchical generalized linear models. *Journal of Computational and Graphical Statistics*. 2003; 12:663–681.
37. Benjamini Y, Yekutieli D. False discovery rate-adjusted multiple confidence intervals for selected parameters (with discussion). *Journal of the American Statistical Association*. 2005; 100:71–93.

Figure 1.

Simulation results for coverage probabilities (y-label) of the nominal 95% (*dotted line*) EB and HL intervals of individual random effects (v_{i0} 's) in frailty model (M1) under 15% censoring, with 500 replications; boxplots for v_{i0} 's under $\sigma_0^2=0.2$ ((a), (b), (c),(d)) and $\sigma_0^2=1$ ((e), (f), (g), (h)). (30, 5), (60, 5), (30, 20), and (60, 20) in x-label indicate the sample size (q, n_j), respectively, corresponding to the total sample size $n = 150, 300, 600$, and 1200. q is the number of clusters and n_j is the cluster size.

Figure 1.1.

Simulation results for coverage probabilities (y-label) of the nominal 95% (*dotted line*) EB and HL intervals of individual random effects ($v_{\mathcal{A}}$'s) in correlated frailty model (M2) under 15% censoring, with 500 replications; boxplots for $v_{\mathcal{A}}$'s under $\sigma_0^2 = \sigma_1^2 = 0.2$ with $\rho = -0.5$ ((a), (b), (c), (d)) and $\sigma_0^2 = \sigma_1^2 = 1$ with $\rho = -0.5$ ((e), (f), (g), (h)). (30, 5), (60, 5), (30, 20), and (60, 20) in x-label indicate the sample size (q, n_i), respectively, corresponding to the total sample size $n = 150, 300, 600$, and 1200 . q is the number of clusters and n_i is the cluster size.

Figure 2.

Random effects and their 95% confidence intervals using HL(A) under correlated frailty model (FM3) allowing dependency between random treatment and random bone for the lung cancer trial; (a) random treatment effects; (b) random bone effects; 31 institutions are sorted in increasing order of number of patients.

Figure 3.

Random effects and their 95% confidence intervals using HL under correlated frailty model (FM2) allowing dependency between random center and random treatment for the lung cancer trial; (a) random center effects using HL(S); (b) random center effects using HL(A); (c) random treatment effects using HL(A); 31 institutions are sorted in increasing order of number of patients.

Table 1

Simulation results for percentages of zero HL estimates of variance components (σ_0^2, σ_1^2) over 500 replications under random center (M1) and correlated (M2) frailty models

M1				
σ_0^2	(q, n_i)	$\hat{\sigma}^2$	15% censoring	50% censoring
0.2	(30,5)	$\hat{\sigma}_0^2$	4.4	13.4
	(60,5)	$\hat{\sigma}_0^2$	0.4	2.8
	(30,20)	$\hat{\sigma}_0^2$	0	0
	(60,20)	$\hat{\sigma}_0^2$	0	0
1.0	(30,5)	$\hat{\sigma}_0^2$	0	0
	(60,5)	$\hat{\sigma}_0^2$	0	0
	(30,20)	$\hat{\sigma}_0^2$	0	0
	(60,20)	$\hat{\sigma}_0^2$	0	0
M2				
$(\sigma_0^2, \sigma_1^2, \sigma_{01})$	(q, n_i)	$\hat{\sigma}^2$	15% censoring	50% censoring
(0.2, 0.2, -0.1)	(30,5)	$\hat{\sigma}_0^2$	1.2	2.0
		$\hat{\sigma}_1^2$	3.4	3.6
	(60,5)	$\hat{\sigma}_0^2$	0.4	0.6
		$\hat{\sigma}_1^2$	1	3.6
	(30,20)	$\hat{\sigma}_0^2$	0	0
		$\hat{\sigma}_1^2$	0.6	1.2
	(60,20)	$\hat{\sigma}_0^2$	0	0
		$\hat{\sigma}_1^2$	0	0.4
(1.0, 1.0, -0.5)	(30,5)	$\hat{\sigma}_0^2$	0	0.2
		$\hat{\sigma}_1^2$	0.6	0.6

Author Manuscript

(60,5)	$\hat{\sigma}_0^2$	0	0
	$\hat{\sigma}_1^2$	0	0.6
(30,20)	$\hat{\sigma}_0^2$	0	0
	$\hat{\sigma}_1^2$	0	0
(60,20)	$\hat{\sigma}_0^2$	0	0
	$\hat{\sigma}_1^2$	0	0

HL, h-likelihood; q , the number of clusters; n_j , cluster size.

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Simulation results for means of coverage probabilities of the nominal 95% EB(empirical Bayes) and HL(h-likelihood) intervals of individual random effects (v_{ij} 's) in random center (M1) and correlated (M2) frailty models, with 500 replications

M1	σ_0^2	(q, n_j)	15% censoring			50% censoring		
			EB(S)	HL(S)	EB(A)	HL(A)	EB(S)	HL(A)
0.2		(30,5)	0.860	0.868	0.939	0.947	0.774	0.778
		(60,5)	0.908	0.911	0.942	0.946	0.869	0.871
		(30,20)	0.930	0.944	0.935	0.950	0.927	0.936
		(60,20)	0.938	0.946	0.940	0.948	0.938	0.943
		(30,5)	0.909	0.936	0.913	0.943	0.907	0.924
1.0		(60,5)	0.927	0.941	0.929	0.943	0.926	0.934
		(30,20)	0.877	0.948	0.878	0.953	0.901	0.945
		(60,20)	0.913	0.948	0.914	0.949	0.922	0.946
	0.2	(50,2)	0.802	0.808	0.960	0.968	0.740	0.744
		(100,2)	0.848	0.850	0.961	0.964	0.783	0.785
1.0		(200,2)	0.872	0.873	0.949	0.950	0.811	0.811
		(50,2)	0.901	0.914	0.920	0.938	0.875	0.884
		(100,2)	0.922	0.928	0.931	0.938	0.900	0.904
		(200,2)	0.933	0.936	0.938	0.941	0.913	0.915
							0.915	0.930
M2	$(\sigma_0^2, \sigma_1^2, \sigma_{01})$	(q, n_j)	15% censoring			50% censoring		
			EB(S)	HL(S)	EB(A)	HL(A)	EB(S)	HL(A)
(0.2, 0.2, -0.1)		(30,5)	0.834	0.842	0.957	0.964	0.838	0.843
		(60,5)	0.858	0.860	0.953	0.956	0.840	0.842
		(30,20)	0.909	0.920	0.936	0.947	0.894	0.901
		(60,20)	0.931	0.934	0.942	0.945	0.927	0.930
		(30,5)	0.882	0.901	0.919	0.937	0.867	0.882
(1.0, 1.0, -0.5)		(60,5)	0.909	0.917	0.927	0.935	0.897	0.902
		(30,20)	0.899	0.940	0.902	0.945	0.908	0.934
							0.916	0.942

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

(60,20)	0.926	0.944	0.927	0.946	0.929	0.940	0.932	0.943
(0.2, 0.2, -0.1)	(50,2)	0.833	0.957	0.968	0.840	0.843	0.974	0.982
	(100,2)	0.855	0.972	0.976	0.842	0.844	0.986	0.988
	(200,2)	0.875	0.965	0.966	0.839	0.840	0.970	0.971
(1.0, 1.0, -0.5)	(50,2)	0.860	0.931	0.947	0.835	0.843	0.953	0.963
	(100,2)	0.887	0.942	0.948	0.853	0.857	0.945	0.952
	(200,2)	0.911	0.945	0.947	0.866	0.867	0.938	0.940

EB(S) & HL(S), standard EB & HL; EB(A) & HL(A), adjusted EB & HL; q_i No. of clusters; n_i cluster size.

Table 3

HL analyses for lung cancer data under four frailty models (FMs)

Model	FM0	FM1	FM2	FM3
	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)
Trt β_1	-0.25 (0.09)	-0.25 (0.09)	-0.22 (0.10)	-0.23 (0.10)
Bone β_2	0.22 (0.09)	0.22 (0.09)	0.22 (0.10)	0.26 (0.13)
Liver β_3	0.43 (0.09)	0.43 (0.09)	0.42 (0.09)	0.39 (0.09)
PS β_4	-0.60 (0.10)	-0.60 (0.10)	-0.64 (0.11)	-0.65 (0.11)
WL β_5	0.20 (0.09)	0.20 (0.09)	0.22 (0.09)	0.21 (0.09)
Center σ_0^2	–	0.000	0.001	–
Center-Trt σ_1^2	–	–	0.059	0.053
Center-Bone σ_2^2	–	–	–	0.135
$\rho[\sigma_{01} \text{ or } \sigma_{12}]$	–	–	0.974[0.006]	0.187[0.016]
$-2p_{\tau} (h^*)$	6112.2	6112.2	6108.2	6100.5
hAIC	5.7	7.7	7.7	0

FM0, Cox model without random effects;

FM1, frailty model with one random-center effect;

FM2, correlated frailty model with random center and random treatment-by-center interaction (Center-Trt);

FM3, correlated frailty model with random treatment-by-center (Center-Trt) and random bone-by-center interaction (Center-Bone);

Trt, treatment; PS, ambulatory performance status; WL, weight loss;

 $p_{\tau} (h^*)$, the restricted likelihood given in (17);

hAIC, AIC differences in (19) where the smallest AIC is adjusted to be zero.