Journal of the Korean Statistical Society (2007), **36**: 1, pp 129–148

BAYESIAN MODEL AVERAGING FOR HETEROGENEOUS FRAILTY

ILSUNG CHANG[1] AND JOHAN LIM[2]

ABSTRACT

Frailty estimates from the proportional hazards frailty model often lead us to conjecture the heterogeneity in frailty such that the variance of the frailty varies over different covariate groups (*e.g.* male group versus female group). For such systematic heterogeneity in frailty, we consider a regression model for the variance components in the proportional hazards frailty model, denoted by the MLFM. However, in many cases, the observed data do not show any statistically significant preference between the homogeneous frailty model and the heterogeneous frailty model. In this paper, we propose a Bayesian model averaging procedure with the reversible jump Markov chain Monte Carlo which selects the appropriate model automatically. The resulting regression coefficient estimate ignores the model uncertainty from the frailty distribution in view of Bayesian model averaging (Hoeting *et al.*, 1999). Finally, the proposed model and the estimation procedure are illustrated through the analysis of the kidney infection data in McGilchrist and Aisbett (1991) and a simulation study is implemented.

AMS 2000 subject classifications. Primary 62F15; Sencondary 62N01.
Keywords. Heterogeneous frailty, kidney infection data, multi–level regression model, proportional hazards model, variance components.

1. INTRODUCTION

Analysis of multivariate failure times entails incorporating the dependence among the observed times into the proportional hazard model. The proportional hazard frailty model, denoted by PHFM hereafter, treats the dependence between multivariate failure times using the unobserved common frailty, which is assumed

Received March 2006; accepted August 2006.

[1]Johnson and Johnson Pharmaceutical Research and Development, Raritan, NJ 08869, U.S.A. (e-mail: ichang@prdus.jnj.com)

[2]Corresponding author. Department of Applied Statistics, Yonsei University, Seoul 120-749, Korea (e-mail: johanlim@yonsei.ac.kr)

to follow a specific type of distribution such as gamma and log–normal distribution. To be specific, in the PHFM, the j^{th} recurrent time of the i^{th} subject has the hazards rate at time t as

$$\lambda(t \mid v_{ij}, x_{ij}) = v_{ij}\lambda_0(t)\exp\left(x_{ij}^T\beta\right),$$

where $\lambda_0(t)$ is a baseline hazards function, x_{ij} is covariate and v_{ij} is a vector of random effects (frailty). The frailty v_{ij} is often assumed to be homogeneous which implies the frailty does not vary over time or different covariates: *e.g.*

$$v_i = v_{i1} = \cdots = v_{im}, \tag{1.1}$$

$$g(v_i|X_{i1}, X_{i2}, \ldots, X_{im}) = g(v_i), \tag{1.2}$$

where (1.1) implies the homogeneity over time and (1.2) implies the homogeneity over covariates.

The above two homogeneous assumptions are, however, easily broken in practice. First, when repeatedly measured survival times are observed, it is natural to assume that two adjacent survival times are more correlated to each other than those two which are apart to each other in time. However, the frailty model assuming (1.1) only allows a positive constant (not varying over time) correlation among repeatedly measured survival times. Second, the violation to the assumption (1.2) has long been discussed in the literature of the generalized linear model (GLM) (see the mean variance joint model in McCullough and Nelder, 1989). Although assuming the homogeneity and a specific type of distribution on frailty has been the premise in most works so far, several recent studies showed that different frailty distributions induce quite different dependence structures (Shih and Louis, 1995; Glidden, 1999). Furthermore, the mis-specification in frailty distribution may result in the bias of regression coefficient estimates (Lancaster, 1996, Section 10.4.2). Hence, one should pay careful attention to such aspects especially when the observations exhibit a strong pattern of heterogeneity; for instance, in the non–trivial presence of extraordinary large or small survival times.

To date, much effort has been made to lessen the assumptions on the frailty. First, for distributional assumption, many previous works have broadened the class of frailty distributions. Especially several heavy tail distributions have been proposed to cope with the unusual frailty estimates. For example, in the kidney data analysis by McGilchrist and Aisbett (1991), the frailty estimates from the PHFM with log–normal frailty raise the doubt on heavy tail and several wider classes of frailty distributions have been proposed; Sahu and Dey (2004) introduced a log–skewed t–distribution; Ravishanker and Dey (2000) introduced a

mixture of positive stable distributions. Second, to overcome the shortcomings from the homogeneity over time, the models with stochastically varying frailty have been proposed by several authors. Some interesting works among many are; the dynamic gamma frailty model by Yue and Chan (1996) where the multiplicative random walk is assumed for varying frailty; the autoregressive (AR) frailty model by Yau and McGilchrist (1998) where the frailty moves according to the AR process; and the similar AR model is considered by Diggle (1988) in the regression model. It should be remarked that the heterogeneous frailty can easily be confused with the heavy tail frailty, in the sense that the heterogeneous frailty, the mixture of homogeneous frailty distribution, has a heavier tail than that of each homogeneous component.

In this paper, we study the regression model for the variance components in the PHFM through the kidney data example. In the kidney data analysis, it is conjectured by several authors that the individual effects of male group have a larger variance than those of female group (Qiou *et al.*, 1999, p. 640). Motivated by such observations, we extend the mean variance joint model in the GLM (or the multi–level regression model) to the frailty model and denote it as the multi–level frailty model (MLFM). However, as we will see in the analysis of the kidney data, the observed data are often hard to provide any statistical significance between the homogeneous frailty model and the heterogeneous frailty model. For such model uncertainty from the frailty distribution, we propose a fully Bayesian approach with the reversible jump Markov chain Monte Carlo (MCMC) by Green (1995) to select the model automatically between the PHFM with homogeneous frailty and that with heterogeneous frailty. Thus, the estimate of the regression coefficient ignores the model uncertainty from the frailty distribution in the sense that it averages between the homogeneous frailty model and the heterogeneous frailty model.

A brief outline of the paper is as follows. In Section 2, we review the analysis of the kidney data in the literature and introduce the model we consider (MLFM). Section 3 introduces the Markov chain Monte Carlo procedure to estimate the proposed model. Section 4 analyzes the kidney data and implements a simulation study to see the performance of the proposed procedure for various magnitude of heterogeneity and sample size. Finally, Section 5 concludes the paper with further discussion on computing time, the extension to more general PHFMs and the accelerated failure time models.

2. KIDNEY INFECTION DATA AND MULTI–LEVEL FRAILTY MODEL

McGilchrist and Aisbett (1991) reported the recurrence times (in days) of infections of 38 kidney patients from insertion of catheter until it had to be removed owing to infection. Many PHFMs have been applied to this kidney data set and several of them pointed out the potential heterogeneity in frailty distribution between different sex groups. In particular, Qiou *et al.* (1999) addressed that the frailties for the male are rather irregularly distributed with a larger variance than those of the female. Rabishanker and Dey (2003) proposed the PHFM with a mixture of positive stable distribution as a remedy to the above problem, but they did not use the sex information in explaining the heterogeneity. In this section, we introduce a Bayesian regression model to explain the heterogeneity in frailty with the variable "sex".

2.1. Full description of the model

Let t_{ij} be the j^{th} recurrent survival time of the i^{th} subject. Then, given v_i, the hazards function of the model is

$$\lambda(t_{ij}|v_i; x_{ij}) = v_i \lambda_0(t_{ij}) \exp(x_{ij}\beta),$$

where x_{ij} is the covariate "sex", β is the regression coefficient, and $\lambda_0(\cdot)$ is the baseline hazards function.

2.1.1. Prior description for frailty. Let $\mathbf{M}_1$ denote the model where the individual effect v_i is assumed to be from the Gamma(α, α), where Gamma(α, β) denotes a gamma distribution with mean α/β and variance α/β^2. $\mathbf{M}_2$ denotes the model where the individual effect v_i is from Gamma(α_1, α_1) if $x_{ij} = x_i = 1$, which means the i^{th} subject is female and it is from Gamma(α_0, α_0) if $x_{ij} = x_i = 0$, which means the i^{th} subject is male. The gamma distribution can be replaced into any other frailty distribution including log–normal distribution and positive stable distribution. It is further assumed that as a prior distribution, α, α_1, and α_0 are independently and identically distributed (*i.i.d.*) Gamma(κ, κ), where κ is a fixed constant to be estimated or to be guessed. Gelman (2004) discussed the choice of distributions of the hyper-parameters in hierarchical models and provided an example whose final estimates strongly depend on the specification of hyper–parameter distribution. We observe that it does not apply to our case (see Section 3.4)

2.1.2. Priors for the regression coefficient. A normal prior is put on the regression coefficient, β, with zero mean and and variance b^2, for which we choose 10^3 as in Sahu *et al.* (1997).

2.1.3. Prior description for the baseline hazards function. The time period is divided into J pre-specified intervals, $I_i = (s_{i-1}, s_i)$ for $i = 1, 2, \ldots, J$, where $0 = s_0 \leq s_1 \leq s_2 \leq \cdots \leq s_J < \infty$. The baseline hazard function $\lambda_0(t)$ is assumed to be piecewise constant as

$$\lambda_0(t) = \lambda_k \quad \text{if} \quad s_k \leq t \leq s_{k+1}.$$

In this paper, we assume the piecewise independent gamma distribution as a prior for the baseline hazards function, which assumes $\underline{\lambda} = (\lambda_1, \ldots, \lambda_J)$ is from

$$f(\underline{\lambda}) = \prod_{k=1}^{J} f(\lambda_k),$$

where $f(\lambda_k)$ is Gamma(τ_k, τ_k). It is interesting to see that the proposed prior is equivalent to the independent increment gamma process prior in Clayton (1991), when τ_k is proportional to the length of the interval I_k. Although, we do not adapt it in this paper, there has been considerable amount of efforts on the correlated prior process including Arjas and Gasbarra (1994), Aslanidou *et al.* (1998) and many others.

Finally, in this paper, we empirically determine J as in Qiou *et al.* (1999), but a random choice of J can be considered using the reversible jump MCMC by Green (1995).

2.1.4. Prior distribution between models $\mathbf{M}_1$ *and* $\mathbf{M}_2$. As noted in Section 2.1.1, the model $\mathbf{M}_1$ assumes the homogeneous variance of the frailty distribution and $\mathbf{M}_2$ assumes the heterogeneous variance structure depending on the covariate ("sex" in this paper). Prior probability of each model is set to $P(\mathbf{M}_1) = P(\mathbf{M}_2) = 0.5$.

2.2. Connection of MLFM to existing models

The considered MLFM can be interpreted as extensions of some existing models (or problems) to survival data.

First, similar models are addressed in the context of the multi–level regression model in the previous literature (Heagerty and Zeger, 2000, Section 5.2), but most

of them are limited to non–survival data problems. A few of them on survival data are Yau (2001), Maples *et al.* (2002) and Zhang and Steele (2004). In particular, the random coefficient PHM by Maples *et al.* (2002) is a special case of our MLFM. Second, in the PHFM, α is the scale parameter of marginal distribution of the survival time as well as that of the frailty distribution. More specifically, in the PHFM with gamma frailty distribution, the marginal distribution of observed survival time is

$$f(t;\beta,\alpha) = \int f(t|v;\beta)g(v;\alpha)dv,$$

where $f(t|v;\beta) = \lambda_0(t)\exp(x^T\beta+\log v)\exp\{-\Lambda_0(t)\exp(x^T\beta+\log v)\}$ and $\Lambda_0(t) = \int_0^t \lambda_0(s)ds$. For example, when $g(v;\alpha)$ is Gamma(α,α) with $\lambda_0(t) = 1$ and $\beta = 0$, the marginal distribution $f(t)$ becomes $\exp\left((\alpha+1)\log\left(\alpha/(\alpha+t)\right)\right)$. Thus, as a model for survival times $\{t_{ij}, i = 1,2,\ldots,38,\ j = 1,2\}$, it can be considered as an extension of the (mean variance) joint-model of the GLM (see McCullough and Nelder, 1989) to the proportional hazard model. Finally, it also can be considered as an extension of the Behrens–Fisher problem to survival models. The Behrens–Fisher problem concerns the inference on the difference between the means of two normal populations whose ratio of variances is unknown. In our model for the kidney data, we only consider the sex variable in both the model for hazards rate, $\lambda(t_{ij})$, and the variance components, α_k, $k = 0,1$. Thus, testing the regression coefficient β (of sex) is equivalent to testing the mean of survival times between the male group and female group. Furthermore, the ratio of the variances between survival times of male group and female group is unknown.

3. FULL CONDITIONAL DISTRIBUTIONS AND MCMC

Recall that $\mathbf{M}_1$ and $\mathbf{M}_2$ denote the PHFM with homogeneous frailty and that with heterogeneous frailty, respectively. Then, the overall sampler has three components: (1) the sampler in $\mathbf{M}_1$, (2) the sampler in $\mathbf{M}_2$, and (3) reversible jump MCMC between the space of $\mathbf{M}_1$ and $\mathbf{M}_2$.

3.1. Sampling algorithms in $\mathbf{M}_1$

Given v_i, the j^{th} recurrent time of the i^{th} subject has a constant hazard of

$$h_{ij} = \lambda_k \eta_{ij} v_i$$

in the k^{th} interval $(k = 1,2,\ldots,J)$ with $\eta_{ij} = \exp(x_{ij}\beta)$. If the subject has survived beyond the k^{th} interval, *i.e.*, $t_{ij} > s_k$ for s_k defined in Section 2.1.2, the

likelihood contribution is

$$\exp\big(-\lambda_k\Delta_k\eta_{ij}v_i\big),$$

where $\Delta_k = s_k - s_{k-1}$. If the subject has failed or is censored in the k^{th} interval, *i.e.*, $s_{k-1} < t_{ij} \le s_k$, then the likelihood contribution is

$$(\lambda_k\eta_{ij}v_i)^{\delta_{ij}}\exp\big(-\lambda_k(t_{ij}-s_{k-1})\eta_{ij}v_i\big),$$

where $\delta_{ij} = 1$ if t_{ij} is not censored; otherwise, it is 0.

Let $D = (X, \delta, t, v)$ denote the complete data and $D_{\text{Obs}} = (X, \delta, t)$ denote the observed data. Then, the complete data likelihood becomes

$$\begin{aligned} &l\big(\beta, \underline{\lambda}, \underline{v}, \alpha \big| D_{\text{Obs}}\big) \\ &\propto \prod_{i=1}^{n}\prod_{j=1}^{m_i}\Bigg\{\Bigg(\prod_{k=1}^{g_{ij}}\exp\big(-\lambda_k\Delta_k\eta_{ij}v_i\big)\Bigg)\Big(\lambda_{g_{ij}+1}\eta_{ij}v_i\Big)^{\delta_{ij}} \\ &\quad\times\exp\Big(-\lambda_{g_{ij}+1}\big(t_{ij}-s_{g_{ij}}\big)\eta_{ij}v_i\Big)\Bigg\}, \end{aligned}$$

where g_{ij} is such that $t_{ij} \in (s_{g_{ij}}, s_{g_{ij}+1}] = I_{g_{ij}+1}$ and $\underline{v} = \big(v_1, \ldots, v_n\big)$. Now we specify the full conditional distributions of unknowns for the MCMC implementation.

First, with the prior Gamma(α, α), the full conditional distribution of v_i for each $i = 1, 2, \ldots, n$, becomes

$$\mathrm{P}\big(v_i \big| \beta, \underline{\lambda}, \alpha, D_{\text{Obs}}\big) \sim \text{Gamma}\big(\alpha + \sum_{j=1}^{m_i}\delta_{ij}, \alpha + S_i\big),$$

where

$$\begin{aligned} S_i &= \sum_{j=1}^{m_i}\eta_{ij}\Big(\sum_{k=1}^{g_{ij}}\lambda_k\Delta_k + \lambda_{g_{ij}+1}\big(t_{ij}-s_{g_{ij}}\big)\Big) \\ &= \sum_{j=1}^{m_i}e^{x_{ij}\beta}\Big(\sum_{k=1}^{g_{ij}}\lambda_k\Delta_k + \lambda_{g_{ij}+1}\big(t_{ij}-s_{g_{ij}}\big)\Big). \end{aligned}$$

Second, with the prior $\pi\big(\alpha\big) = \text{Gamma}(\kappa, \kappa)$, the full conditional of α is

$$\mathrm{P}\big(\alpha \big| \beta, \underline{\lambda}, \underline{v}, D_{\text{Obs}}\big) \propto \frac{\eta^{n\alpha}}{\Gamma(\alpha)^n}\exp\Big\{-\eta\sum_i\big(v_i - \log v_i\big)\Big\} \times \pi\big(\alpha\big).$$

The Metropolis–Hastings algorithm is implemented to get a sample from the posterior distribution.

Third, with the prior $\pi(\beta)$, the full conditional of β is

$$\mathrm{P}(\beta|\underline{\lambda}, \underline{v}, \alpha, D_{\mathrm{Obs}}) \propto \exp\Big(\sum_{i=1}^{n}\sum_{j=1}^{m_i}\delta_{ij}x'_{ij}\beta\Big)\exp\Big(-\sum_{i=1}^{n}v_iS_i\Big) \times \pi(\beta).$$

The sample from the posterior distribution can be obtained using the Metropolis–Hastings algorithm again.

Fourth, with the prior of piecewise independent gamma distribution, say $\pi(\lambda_k) = \mathrm{Gamma}(\tau_k, \tau_k)$, the full conditional of λ_k is

$$\mathrm{P}(\lambda_k|\beta, \underline{v}, \alpha, D_{\mathrm{Obs}}) \sim \mathrm{Gamma}(d_k + \tau_k, V_k + \tau_k),$$

where d_k is the number of failure times occurred in the interval $I_k = (s_{k-1}, s_k]$ and

$$V_k = \sum_{(i,j)\in\mathcal{R}_k}\Delta_k\eta_{ij}v_i + \sum_{(i,j)\in\mathcal{D}_k}(t_{ij} - s_{k-1})\eta_{ij}v_i. \tag{3.1}$$

R_k and D_k in (3.1) is the set of indexes of the subjects who survive longer than s_{k-1} and those of subjects who died in the interval I_k, respectively.

3.2. Sampling algorithms in $\mathbf{M_2}$

The only difference between the sampler of $\mathbf{M_2}$ from that of $\mathbf{M_1}$ is the prior distribution of v, which is, in $\mathbf{M_2}$,

$$\pi(v|Z = k) = \mathrm{Gamma}(\alpha_k, \alpha_k) \quad \text{for} \quad k = 1 \text{ or } 0,$$

where Z indicates the sex of a subject. Thus, the full conditional distributions of β and λ_k are same as those in $\mathbf{M_1}$.

Let N_k, $k = 1, 0$ be the index set of subjects whose Z value is k and $n_k = |N_k|$, the number of subjects in N_k. Then, the full conditional distribution of α_k is

$$\begin{aligned}\mathrm{P}(\alpha_k|\beta, \underline{\lambda}, \underline{v}, D_{\mathrm{Obs}}) &\propto \prod_{i\in N_k}\pi(v_i|\alpha_k)\pi(\alpha_k)\\ &= \frac{\alpha_k^{n_k\alpha_k}}{\Gamma(\alpha_k)^{n_k}}\exp\Big\{-\alpha_k\sum_{i\in N_k}(v_i - \log v_i)\Big\} \times \pi(\alpha_k).\end{aligned}$$

The samples from the full conditionals are obtained using the Metropolis-Hastings algorithm as in $\mathbf{M_1}$.

3.3. Reversible MCMC between $\mathbf{M}_1$ *and* $\mathbf{M}_2$

Among existing model choice techniques, we consider to use the reversible jump MCMC algorithm. It enables us to simultaneously estimate the posterior probabilities of several models under consideration and the parameters conditional on a specific model.

Green (1995) developed the reversible jump algorithm, which generalizes the Metropolis-Hastings algorithm into the dimension varying situation. Consider the finite mixture model, for example. The dimension of the parameter space depends on the number of components. When the number of components in the model is also unknown, the ordinary MCMC algorithm cannot be directly implemented since the dimension of all the unknowns is not fixed. The reversible jump MCMC algorithm is designed in such a way that the sampler moves across different dimensions. The following shows how to implement the reversible jump MCMC into the models we have considered.

There are 4 types of possible moves between $\mathbf{M}_1$ and $\mathbf{M}_2$; the move within $\mathbf{M}_1$, the move within $\mathbf{M}_2$, the move from $\mathbf{M}_1$ to $\mathbf{M}_2$; and the move from $\mathbf{M}_2$ to $\mathbf{M}_1$. The moves within each $\mathbf{M}_1$ and $\mathbf{M}_2$ are discussed in the previous section and the moves from one to the other model will be discussed here. Before describing the details of each move, it should be remarked that, unlike mixture cases (Richardson and Green, 1997), the allocation of the subjects to each homogeneous group in the move from $\mathbf{M}_1$ to $\mathbf{M}_2$ is pre-determined by the the covariate "sex" and *the computing time is almost same with that for the MCMC for the homogeneous frailty model.*

First, consider the move from $\mathbf{M}_1$ to $\mathbf{M}_2$. The dimension of the parameter space in $\mathbf{M}_2$ is greater than $\mathbf{M}_1$ by 1 because we have (α_1, α_0) in $\mathbf{M}_2$ for the variance of the frailty depending on the sex but only α in $\mathbf{M}_1$. For dimension matching, we need an additional continuous random variable. Let w be a random number from the exponential distribution with a rate of $\log 2$ (hence, the median is 1). The candidate values α_1 and α_0, then, can be defined as

$$\alpha_1 = w \cdot \alpha \quad \text{and} \quad \alpha_0 = \alpha / w.$$

When $y = \alpha$ and $y' = (\alpha_1, \alpha_0)$, the sampler moves from $\mathbf{M}_1$ to $\mathbf{M}_2$ with probability $\rho = \min(1, A)$, where

$$A = \frac{\pi(y' | \beta, \underline{\lambda}, \underline{v}, D_{\text{Obs}}) r_m(y')}{\pi(y | \beta, \underline{\lambda}, \underline{v}, D_{\text{Obs}}) r_m(y) q(w)} \cdot \left| \frac{\partial y'}{\partial (y, w)} \right|, \tag{3.2}$$

following the equation (7) in Section 3.1 in Richardson and Green (1997). Here, the last term is the Jacobian arising at the transformation $(y, w) \to y'$, that is,

$$\left|\frac{\partial y'}{\partial(y,w)}\right| = \begin{vmatrix} w & \alpha \\ 1/w & -\alpha/w^2 \end{vmatrix} = \frac{2\cdot\alpha}{w}$$

and the ratio of the other part for A is

$$\frac{\pi(y'|\beta,\underline{\lambda},\underline{v},D_{\text{Obs}})r_m(y')}{\pi(y|\beta,\underline{\lambda},\underline{v},D_{\text{Obs}})r_m(y)q(w)} = \frac{\prod_{k=1}^{2}\prod_{i\in N_k}\pi(v_i|\alpha_k)\pi(\alpha_k)}{\prod_{i=1}^{n}\pi(v_i|\alpha)\pi(\alpha)q(w)},$$

where $q(w)$ is the density function of the exponential distribution with the rate of 2 evaluated at w, and $r_m(y)$ is the probability of choosing move type m when in state y, which is 1 in our case.

Next, let us consider the move from $\mathbf{M}_2$ to $\mathbf{M}_1$. When the chain moves from $\mathbf{M}_2$ to $\mathbf{M}_1$, the parameters α and w are directly computed from α_1 and α_0 as:

$$\alpha = \sqrt{\alpha_1\cdot\alpha_0} \quad \text{and} \quad w = \sqrt{\alpha/\alpha_0}.$$

Subsequently, the sampler moves from $\mathbf{M}_2$ to $\mathbf{M}_1$ with probability

$$\min\left(1,\ \mathbf{A}^{-1}\right)$$

with appropriate substitutions in (3.2).

The posterior probability $p(\mathbf{M}_k|\text{ Data})$ is estimated by the relative frequency of the number of iterations the sampler visits $\mathbf{M}_k$.

3.4. Estimation of the hyper-parameters

In our full specification, there are three different hyper-parameters b, τ, and κ; b is from the prior distribution of β, τ is from the prior distribution of the baseline hazards and κ is from the prior distribution of the variance of frailty distribution. It is often observed that the posterior distribution is quite sensitive to the specification of the hyper-parameters. In such case, it is more sensible to estimate those parameters in the empirical Bayesian point of view.

Let $O = (T, \delta)$ be the observed survival times and the censoring information, where $\delta = 1$ ($\delta = 0$) if T is uncensored time (censored time). Let W be other random components including β, $\underline{\lambda}$, $\underline{v}$ which are not observed. Then, by observing

$$\frac{\partial f(O|b,\tau,\kappa)}{\partial(b,\tau,\kappa)} = f(O|b,\tau,\kappa)\int \frac{\partial \log f(O,W|b,\tau,\kappa)}{\partial(b,\tau,\kappa)} f(W|O,b,\tau,\kappa)dW,$$

maximizing marginal log–likelihood function $\log f(O|b,\kappa,\tau)$ is equivalent to maximizing the expected full log-likelihood function (expectation is with respect to the posterior distribution)

$$\int \log f(O,W|b,\tau,\kappa) f(W|O,b,\tau,\kappa)\, dW. \tag{3.3}$$

Finally, for each b, κ and τ, (3.3) can be approximated using the posterior samples $\{W^{(i)}\}_{i=1}^{B}$ by

$$\int \log f(O,W|b,\tau,\kappa) f(W|O,b,\tau,\kappa) dW \approx \frac{1}{B}\sum_{i=1}^{B} \log f(O,W^{(i)}|b,\tau,\kappa).$$

In the kidney data example from the next section, the posterior distribution of $\mathbf{M}_2$ and others are not much sensitive to the specification of hyper-parameters when τ is sufficiently small ($\tau \leq 0.1$) (see Figure 3.1). Finally, we set $b^2 = 1000$, $\kappa = 0.05$, and $\tau = 0.1$.

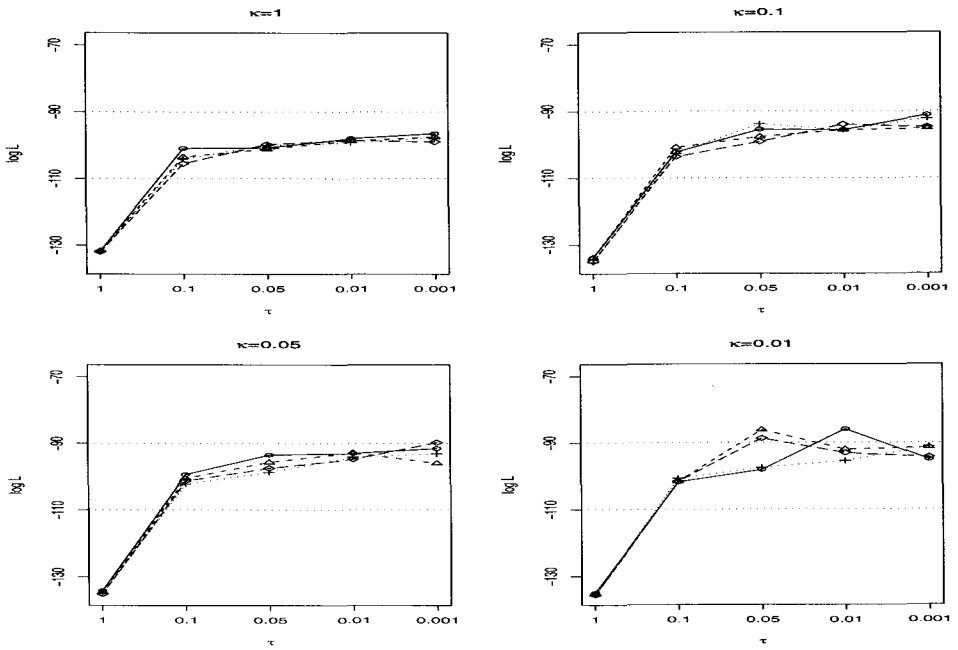

FIGURE 3.1 *Plots of mean of penalized log likelihood from MCMC samples for different values of hyper-parameters. (a)* $\kappa = 1$ *(b)* $\kappa = 0.1$ *(c)* $\kappa = 0.05$ *(d)* $\kappa = 0.01$. *In each plot, the circle indicates the case for* $b^2 = 10$, *triangle for* $b^2 = 25$, *cross for* $b^2 = 100$ *and diamond for* $b^2 = 1000$. *The x-axis indicates the inverse of variance of* $\pi(\lambda)$, *which appears in* $G(\tau,\tau)$.

4. EXAMPLES

4.1. Kidney data analysis

In this section, we applied the proposed MLFM to the kidney infection data in McGilchrist and Aisbett (1991). We generated 20,000 samples from the posterior distribution. The fast convergence of the sampler could be checked from the log-likelihood values of each Gibbs sample. Finally, 10,000 samples were selected after 10,000 burn-in period for the inference of β, $\underline{\lambda}$, $\underline{v}$, and α. Hereafter, the estimate refers the posterior sample mean.

The frailty v_is (for $i = 1, 2, \ldots, 38$) were assumed to be from Gamma(α, α) in $\mathbf{M}_1$ and they were assumed to be from Gamma(α_1, α_1) or Gamma(α_0, α_0) relying on the sex variable in $\mathbf{M}_2$. Each of the above α was assumed to be from Gamma(κ, κ). Before we report our results, it should be pointed out that in both the positive stable frailty model and the gamma frailty model, the frailty estimates of male were more variable than those of female (see Figure 4.1 and Qiou *et al.*, 1999, p. 640).

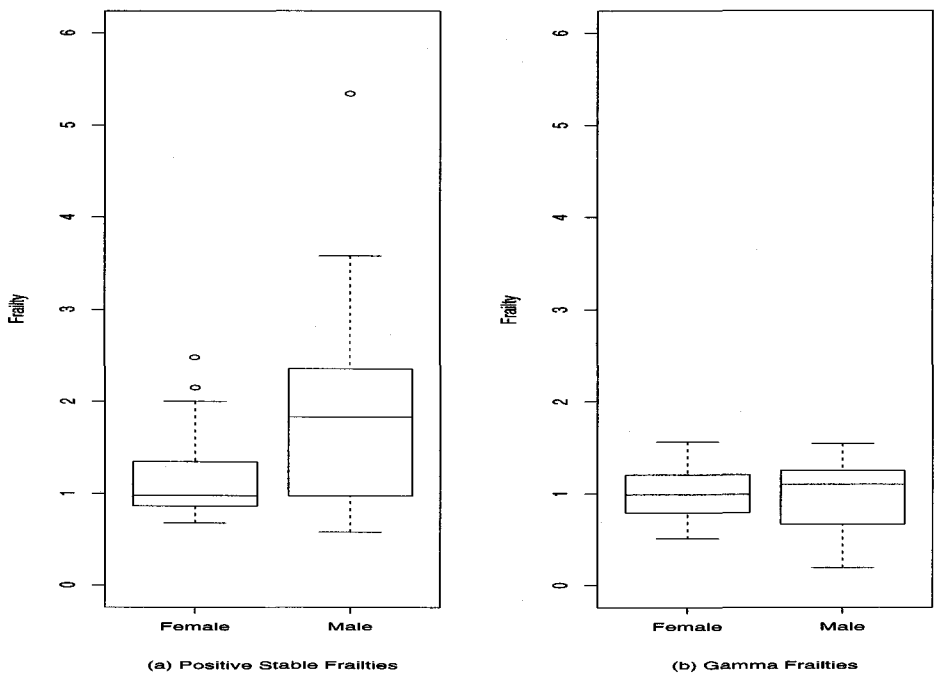

FIGURE 4.1 *Box-plots of posterior means of frailties by sex from two different frailty distributions by Qiou et al. (1999). (a) posterior means assuming positive stable distribution for frailties, (b) posterior means assuming gamma distribution for frailties.*

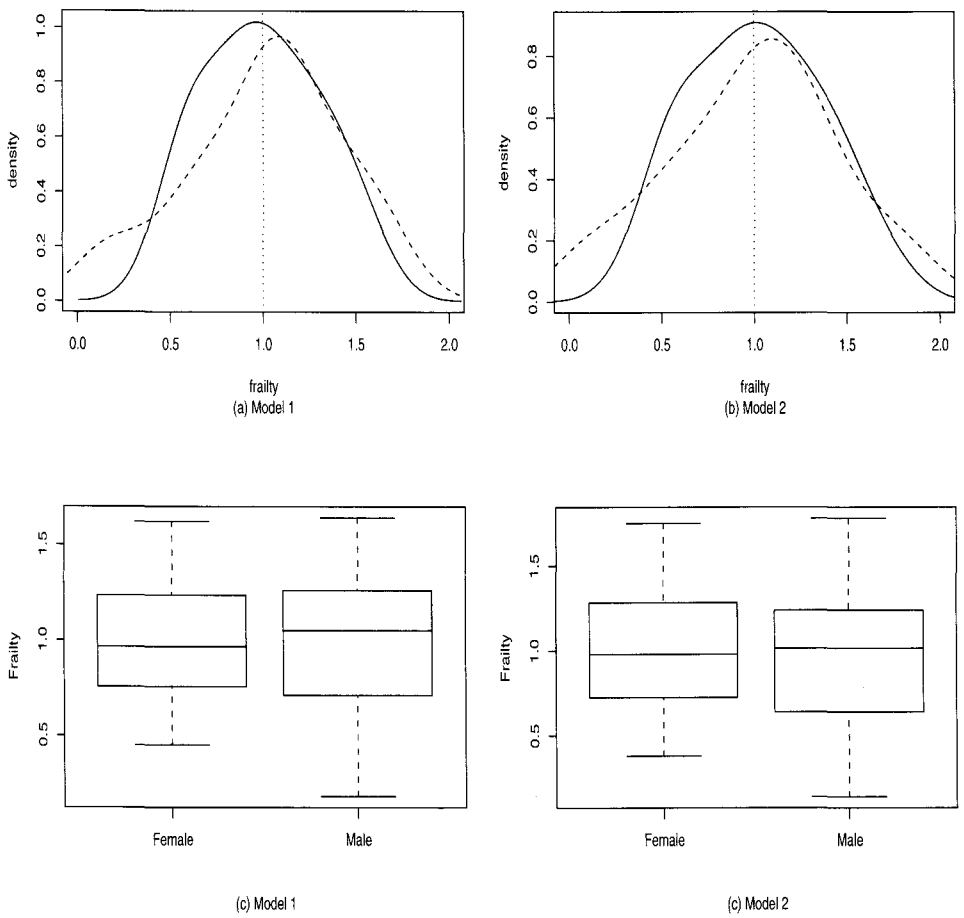

FIGURE 4.2 *Density plots (first row) and box-plots (second row) of estimated frailties over sex and models. (a) and (b) : Density plot of posterior means of frailties with solid line for female and dashed line for male. (c) and (d) : Box-plots of posterior means of frailties of female and male for each model.*

In our analysis, the frailty estimates in each group approximately had a mean of one and, in the heterogeneous frailty model, the variability of the frailty estimates differed slightly between male and female groups, but it was not much apparent as in Qiou *et al.* (1999) (see Figure 4.2). The frailty estimate of each patient is presented in Table 4.1 for $\mathbf{M}_1$ and $\mathbf{M}_2$ separately.

TABLE 4.1 *Posterior means and standard deviations of frailties by sex for two different models.* $\mathbf{M}_1$ *denotes the case of homogeneous variance on the frailty, whereas* $\mathbf{M}_2$ *denotes the case of heterogeneous variance*

$\mathbf{M}_1$				$\mathbf{M}_2$			
Female		*Male*		*Female*		*Male*	
Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
1.32	0.78	1.47	0.77	1.31	0.76	1.42	0.79
0.54	0.31	1.02	0.51	0.50	0.31	0.96	0.51
0.95	0.46	1.08	0.54	0.96	0.51	1.07	0.58
0.58	0.34	1.64	0.91	0.54	0.34	1.78	1.08
0.98	0.48	0.58	0.35	1.00	0.55	0.52	0.35
1.62	0.89	1.15	0.65	1.75	0.96	1.22	0.70
0.95	0.52	0.18	0.18	0.96	0.57	0.14	0.15
1.27	0.65	0.96	0.48	1.37	0.78	0.93	0.53
0.59	0.40	1.26	0.64	0.57	0.43	1.24	0.67
0.45	0.31	0.71	0.42	0.38	0.28	0.64	0.40
0.83	0.41			0.82	0.47		
0.80	0.40			0.78	0.44		
0.63	0.41			0.60	0.44		
1.04	0.58			1.07	0.66		
0.64	0.38			0.61	0.41		
1.49	0.79			1.54	0.98		
1.05	0.58			1.10	0.68		
0.71	0.39			0.67	0.38		
0.95	0.46			0.96	0.51		
1.47	0.78			1.56	0.87		
1.20	0.59			1.26	0.67		
1.42	0.73			1.45	0.79		
1.20	0.69			1.19	0.71		
1.12	0.55			1.17	0.59		
0.85	0.46			0.88	0.53		
1.33	0.67			1.41	0.78		
0.80	0.52			0.82	0.65		
1.13	0.64			1.18	0.77		

The posterior probability of the heterogeneous model $\mathbf{M}_2$ was estimated by the proportion of the iterations, where the sampler stayed in $\mathbf{M}_2$. It was computed as 10.06%. Thus, the observed data did not provide any statistically significant preference between the homogeneous frailty model and the heterogeneous frailty model. In the heterogeneous frailty model ($\mathbf{M}_2$), the variance estimates of the

TABLE 4.2 *Posterior means of the regression coefficient of sex and the variances (the inverse of αs) of gamma frailty over two different models: Gamma denotes the estimates from gamma frailty model listed in Table 4.1 in Qiou et al. (1999).* $\mathbf{M}_1$ *denotes the case of homogeneous variance on the frailty, whereas* $\mathbf{M}_2$ *denotes the case of heterogeneous variance*

	Gamma		$\mathbf{M}_1$		$\mathbf{M}_2$	
Parameter	*Posterior mean*	*Posterior SD*	*Posterior mean*	*Posterior SD*	*Posterior mean*	*Posterior SD*
β	-1.6200	0.4186	-1.4871	0.5027	-1.6122	0.5452
$1/\alpha$	0.3268	0.1737	0.4737	0.2961	–	–
$1/\alpha_1$	–	–	–	–	0.5482	0.2869
$1/\alpha_0$	–	–	–	–	0.5702	0.3210

TABLE 4.3 *Posterior means of baseline hazard rates denoted by* λ_j, $j = 1, \ldots, 10$ *for the kidney infection data: Gamma denotes the estimates from gamma frailty model listed in Table 4.1 in Qiou et al.(1999).* $\mathbf{M}_1$ *denotes the case of homogeneous variance on the frailty, whereas* $\mathbf{M}_2$ *denotes the case of heterogeneous variance*

	Gamma		$\mathbf{M}_1$		$\mathbf{M}_2$	
Parameter	*Posterior mean*	*Posterior SD*	*Posterior mean*	*Posterior SD*	*Posterior mean*	*Posterior SD*
λ_1	0.0015	0.0030	0.0212	0.0112	0.0232	0.0124
λ_2	0.0012	0.0025	0.0377	0.0215	0.0422	0.0255
λ_3	0.0011	0.0020	0.0833	0.0464	0.0901	0.0511
λ_4	0.0010	0.0020	0.0583	0.0439	0.0650	0.0480
λ_5	0.0012	0.0023	0.0150	0.0193	0.0163	0.0209
λ_6	0.0011	0.0024	0.0241	0.0193	0.0277	0.0212
λ_7	0.0011	0.0024	0.0514	0.0409	0.0606	0.0455
λ_8	0.0014	0.0028	0.0426	0.0526	0.0497	0.0545
λ_9	0.0056	0.0051	0.0251	0.0499	0.0307	0.0529
λ_{10}	0.3667	0.1495	0.1594	0.2290	0.1982	0.2584

frailty distribution in the male group and the female group were 0.5702 and 0.5482, respectively. However, the variance estimate of the frailty distribution in the homogeneous frailty model ($\mathbf{M}_1$) was estimated as 0.4737. It should be noted that, unexpectedly, this variance estimate was smaller than both 0.5482 and 0.5702 in $\mathbf{M}_2$. This strange phenomena was an outcome of the prior effect to the posterior distribution when the number of the observations were small; it was expected that each of the above variance estimates was the average of the prior information (forced that the variance was one) and the data information (indicated that the variance was smaller than 1).

The baseline hazard function was modelled by a piecewise gamma distribu-

tion. The break points of the time axis were chosen as $\{0, 10, 20, 30, 40, 50, 100, 200, 300, 400\}$ as in Qiou *et al.* (1999) and the prior of the baseline hazards in each interval was chosen independently and identically as Gamma(τ, τ) with $\tau = 0.1$. The estimates are reported in Table 4.3 and it is also interesting to see the the estimate for the interval $[400, \infty)$ was much larger than those for other intervals as in Qiou *et al.* (1999).

Finally, β was negatively estimated as -1.5555 whose absolute value was larger than that from the homogeneous frailty (-1.4871) and smaller than that from the heterogeneous frailty (-1.6122). the female patients had a lower risk for infection. The posterior distribution of β in $\mathbf{M}_2$ was not much different from that in M_1. The posterior samples of β are plotted in Figure 4.3. Also, parameter estimates are presented in Table 4.2.

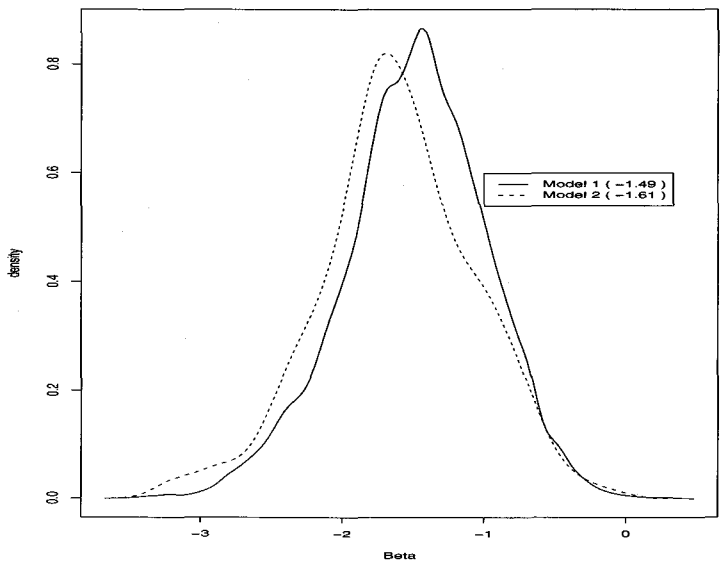

FIGURE 4.3 *Estimated densities for the regression coefficient β from 20,000 iterations after 10,000 burn-in from two different models. The straight line denotes the estimated density from $\mathbf{M}_1$ and the dashed line from $\mathbf{M}_2$. The values in the parentheses are the estimated posterior means.*

4.2. Simulated data examples

The kidney data analysis in the previous section showed the model uncertainty between the heterogeneous frailty model and its homogeneous counterpart. In

this section, we implemented a simulation study to investigate the performance of the proposed Bayesian procedure for various magnitudes of the heterogeneity in variance components and sample sizes.

The data sets were simulated from the PHFM with gamma frailty distribution. The baseline hazard function was set to be constant over time as $\lambda_0(t) = 0.01$ and the regression coefficient for the single covariate x was $\beta = 2.0$. The number of subjects (sample size) was $n = 50$ or $n = 100$, where the frailty for each subject was from the Gamma(α_1, α_1) or Gamma(α_0, α_0) according to x. We considered 4 choices of (α_1, α_0) having different magnitudes of heterogeneity; $(\alpha_1, \alpha_0) = (0.1, 10)$, $(0.2, 5)$, $(0.5, 2)$ and $(1, 1)$.

In the analysis, we used the same prior distributions with those in the kidney example. The baseline hazards function followed a piecewise exponential distribution with rate λ_k, where each λ_k was from Gamma(τ, τ). The regression coefficient β was from the Gaussian distribution with mean 0 and variance b^2. The frailty v_is were assumed to be Gamma(α, α) in $\mathbf{M}_1$ and Gamma(α_1, α_1) or Gamma(α_0, α_0) according to the sex variable in $\mathbf{M}_2$, where α, α_1 and α_0 followed from Gamma(κ, κ). The hyper-parameters were $b^2 = 1000$, $\kappa = 0.05$ and $\tau = 0.1$ as in the kidney data analysis.

Figure 4.4 displays the posterior probabilities of $\mathbf{M}_1$ for several different values of variance components and the sample sizes. It showed that the posterior probability of $\mathbf{M}_2$ increased as the frailty became more heterogeneous. However, its probability was still lower than than that of $\mathbf{M}_1$ even the frailty was very heterogeneous (*e.g.* $(\alpha_1, \alpha_0) = (0.1, 10)$) with moderate size $n = 100$.

TABLE 4.4 *Mean and standard deviation of β estimates from 50 simulations when $n = 50$ and $n = 100$*

		(α_1, α_0)			
n	*Model*	(0.1 , 10)	(0.2 , 5)	(0.5 , 2)	(1 , 1)
50	$\mathbf{M}_1$	2.239 (1.037)	2.463 (0.793)	2.441 (0.713)	2.305 (0.576)
	$\mathbf{M}_2$	2.001 (1.020)	2.279 (0.787)	2.385 (0.679)	2.316 (0.556)
	Model Avg.	2.181 (1.038)	2.427 (0.796)	2.434 (0.710)	2.305 (0.575)
100	$\mathbf{M}_1$	2.477 (0.923)	2.754 (0.512)	2.550 (0.573)	2.327 (0.314)
	$\mathbf{M}_2$	2.255 (0.929)	2.592 (0.510)	2.486 (0.542)	2.330 (0.304)
	Model Avg.	2.389 (0.946)	2.717 (0.518)	2.542 (0.571)	2.328 (0.313)

Table 4.4 reported the average posterior mean of β over 50 data sets. It shows that the model averaging estimate is adaptive in the sense that it was close to the estimate from $\mathbf{M}_1$ (or $\mathbf{M}_2$) when the data is generated from $\mathbf{M}_1$ (or $\mathbf{M}_2$).

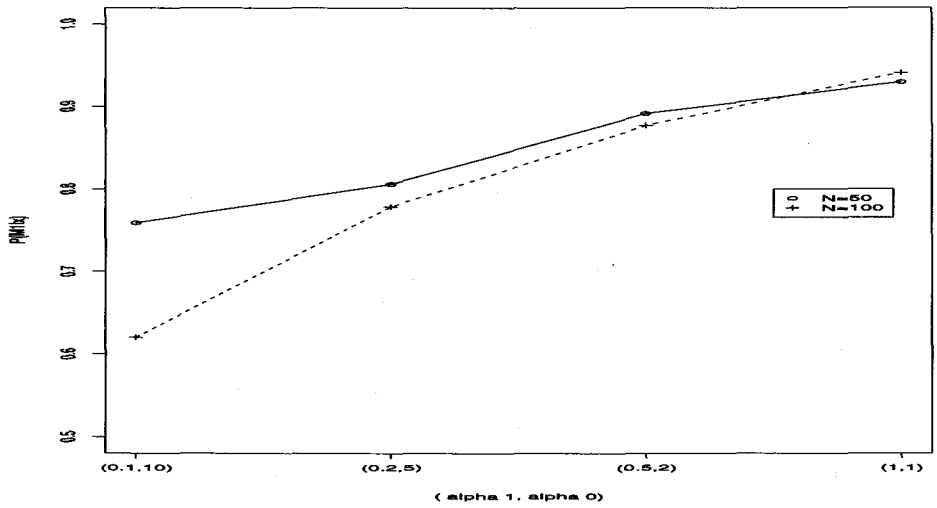

FIGURE 4.4 *Mean of estimated* $P(\mathbf{M}_1|x)$ *from 50 simulated data sets.*

5. DISCUSSION

In this paper, we consider a regression model for variance components in the PHFM and propose a fully Bayesian procedure with the reversible jump MCMC for the model uncertainty from the frailty distribution.

We conclude the paper with a few discussions not covered in the main body of the paper. First, it should be pointed out that the required computing time for the reversible jump MCMC is almost same with that of the MCMC for the homogeneous frailty model since the random allocation of subjects into two homogeneous groups is guided by the covariates (sex in the Kidney data analysis). Second, the procedure of this paper can straightforwardly be extended to multi–sample problems (the regression covariate is discrete but has more than two different levels) with a more complicate RJMCMC. The extension to more general settings such as the model with continuous covariate or the multiple regression model may be still possible but requires a specific model for the variance components as in Noh *et al.* (2006). Finally, the same issue in other survival models such as the accelerated failure time model with the frailty does not much differ from that in the PHFM.

ACKNOWLEDGMENTS

The authors are grateful to Dr. Xinlei Wang for several comments in the original version of the paper. We are also grateful to two referees and the editor for many suggestions.

REFERENCES

ARJAS, E. AND GASBARRA, D. (1994). "Nonparametric Bayesian inference from right censored survival data using the Gibbs sampler", *Statistica Sinica*, **4**, 505–524.

ASLANIDOU, H., DEY, D. K. AND SINHA, D. (1998). "Bayesian analysis of multivariate survival data using Monte Carlo methods", *The Canadian Journal of Statistics*, **26**, 38–48.

CLAYTON, D. G. (1991). "A Monte Carlo method for Bayesian inference in frailty models", *Biometrics*, **47**, 467–485.

DIGGLE, P. J. (1988). "An approach to the analysis of repeated measurements", *Biometrics*, **44**, 959–971.

GELMAN, A. (2004). "Prior distributions for variance parameters in hierarchical models", unpublished.

GLIDDEN, D. V. (1999). "Checking the adequacy of the gamma frailty model for multivariate failure times", *Biometrika*, **86**, 381–393.

GREEN, P. J. (1995). "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination", *Biometrika*, **82**, 711–732.

HEAGERTY, P. J. AND ZEGER, S. L. (2000). "Marginalized multilevel models and likelihood inference", *Statistical Science*, **15**, 1–26.

HOETING, J. A., MADIGAN, D., RAFTERY, A. E. AND VOLINSKY, C. T. (1999). "Bayesian model averaging: a tutorial", *Statistical Science*, **14**, 382–417.

LANCASTER, T. (1992). *The Econometric Analysis of Transition Data*, Cambridge University Press.

MAPLES, J. J., MURPHY, S. A. AND AXINN, W. G. (2002). "Two-level proportional hazards models", *Biometrics*, **58**, 754–763.

MCCULLAGH, P. AND NELDER, J. A. (1989). *Generalized Linear Models*, 2nd ed., Chapman & Hall /CRC.

MCGILCHRIST, C. A. AND AISBETT, C. W. (1991). "Regression with frailty in survival analysis", *Biometrics*, **47**, 461–466.

MCGILCHRIST, C. A. (1993). "REML estimation for survival models in frailty", *Biometrics*, **49**, 221–225.

NOH, M., HA, I. D. AND LEE, Y. (2006). "Dispersion frailty models and HGLMs", *Statistics in Medicine*, **25**, 1341–1354.

RAVISHANKER, N. AND DEY, D. K.(2000). "Modeling multivariate survival models with a mixture of positive stable frailties", *Methodology and Computing in Applied Probability*, **2**, 293–308.

RICHARDSON, S. AND GREEN, P. J. (1997). "On Bayesian analysis of mixtures with an unknown number of components", *Journal of the Royal Statistical Society*, Ser. B, **59**, 731–792.

SAHU, S. K. AND DEY, D. K. (2004). "On multivariate survival models with a skewed frailty and a correlated baseline hazard process", *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality* (Genton, M. G., eds), Chapman & Hall /CRC, 321–338.

SAHU, S. K., DEY, D. K., ASLANIDOU, H. AND SINHA, D. (1997). "A Weibull regression model with gamma frailties for multivariate survival data", *Lifetime Data Analysis*, **3**, 123–137.

SHIH, J. H. AND LOUIS, T. A. (1995). "Inferences on the association parameter in copula models for bivariate survival data", *Biometrics*, **51**, 1384–1399.

QIOU, Z., RAVISHANKER, N. AND DEY, D. K. (1999). "Multivariate survival analysis with positive stable frailties", *Biometrics*, **55**, 637–644.

WALKER, S. G. AND MALLICK, B. K. (1997). "Hierarchical generalized linear models and frailty models with Bayesian nonparametric mixing", *Journal of the Royal Statistical Society*, Ser. B, **59**, 845–860.

YAU, K. K. W. AND MCGILCHRIST, C. A. (1998). "ML and REML estimation in survival analysis with time dependent correlated frailty", *Statistics in Medicine*, **17**, 1201–1213.

YAU, K. K. W. (2001). "Multilevel models for survival analysis with random effects", *Biometrics*, **57**, 96–102.

YUE, H. AND CHAN, K. S. (1997). "A dynamic frailty model for multivariate survival data", *Biometrics*, **53**, 785–793.

ZHANG, W. AND STEEL, F. (2004). "A semiparametric multilevel survival model", *Journal of the Royal Statistical Society,* Ser. C, **53**, 387–404.