Journal of the Korean Data & Information Science Society
2013, **24**(4), 931–940

http://dx.doi.org/10.7465/jkdi.2013.24.4.931
한국데이터정보과학회지

A visualizing method for investigating individual frailties using frailtyHL R-package[†]

Il Do Ha[1] · Maengseok Noh[2]

[1]Department of Asset Management, Daegu Haany University
[2]Department of Statistics, Pukyong National University

Received 3 June 2013, revised 2 July 2013, accepted 7 July 2013

Abstract

For analysis of clustered survival data, the inferences of parameters in semi-parametric frailty models have been widely studied. It is also important to investigate the potential heterogeneity in event times among clusters (e.g. centers, patients). For purpose of this analysis, the interval estimation of frailty is useful. In this paper we propose a visualizing method to present confidence intervals of individual frailties across clusters using the frailtyHL R-package, which is implemented from h-likelihood methods for frailty models. The proposed method is demonstrated using two practical examples.

Keywords: frailtyHL R-package, h-likelihood, interval estimation, multilevel frailty, shared frailty.

1. Introduction

The frailty models, Cox's proportional hazards models (Park *et al.*, 2012) allowing frailty (random effect) terms, have been widely used for the analysis of clustered survival-time data. For the inference, the marginal likelihood often involves analytically intractable integrals, particularly when modelling multilevel or correlated frailties. However, the hierarchical-likelihood (h-likelihood; Lee and Nelder, 1996, 2001) obviates the need for intractable integration over the frailty terms (Ha *et al.*, 2001, 2011; Ha and Cho, 2012). It is also important to investigate the potential heterogeneity in event times among clusters (e.g. centers, patients) in order to understand and interpret the variability in the data (Vaida and Xu, 2000). For example, despite the use of standardized protocols in multicenter randomized clinical trials, outcome may vary between centers (Rondeau *et al.*, 2008; Ha *et al.*, 2011). Such heterogeneity may alter the interpretation and reporting of the treatment. For purpose of this analysis, the interval estimation of frailty is more useful than the inference of variance components of frailty (Vaida and Xu, 2000; Lee and Nelder, 2009; Ha *et al.*,

[†] This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. 2009-0088978, No. 2010-0021165).
[1] Corresponding author: Professor, Department of Asset Management, Daegu Haany University, Gyeongsan 712-715, Korea. E-mail: idha@dhu.ac.kr
[2] Associate professor, Department of Statistics, Pukyong National University, Busan 608-737, Korea.

2011). Plots based on these confidence intervals are also useful for investigating such heterogeneity. The h-likelihood consists of data, parameters and unobserved random effects, and avoids integration over the random-effect distributions. Thus, the h-likelihood can be used directly for inference on random effects, while the marginal likelihood cannot because it eliminates them by integration. Very recently, Ha *et al.* (2012a, 2012b) have developed the **frailtyHL** R-package which implements h-likelihood (HL) procedures for semi-parametric frailty models with non-parametric baseline hazards.

However, this package does not directly provide the interval estimation of frailty. In this paper, we show how to construct the interval estimation and its plot using the **frailtyHL** package. The proposed method is illustrated using two practical examples with well-known real data sets. The paper is organized as follows. In Section 2 we review a formulation of frailty models and the h-likelihood estimation procedure including the **frailtyHL** package. In Section 3 we outline the interval estimation of frailty and present the corresponding R-codes. Our method is demonstrated using two practical examples in Section 4. Finally, we briefly discuss the proposed method in Section 5. The details of R-codes are given in Appendix.

2. The model estimation

2.1. A formulation of frailty models

Let T_{ij} $(i = 1, \ldots, q,\ j = 1, \ldots, n_i,\ n = \sum_i n_i)$ be the event time (survival time) for the jth observation in the ith cluster and let C_{ij} be the corresponding censoring time. Then the observable random variable is $y_{ij} = \min(T_{ij}, C_{ij})$ and the censoring indicator is $\delta_{ij} = I(T_{ij} \leq C_{ij})$ where $I(\cdot)$ is the indicator function. Denote by v_i a vector of unobserved log-frailty (or random effects) associated with the ith cluster. Given v_i, the conditional hazard function of T_{ij} is of the form

$$\lambda_{ij}(t|v_i) = \lambda_0(t)\exp(\eta_{ij}), \tag{2.1}$$

where $\lambda_0(\cdot)$ is the unknown baseline hazard function,

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

is the linear predictor for the log-hazard, and $x_{ij} = (x_{ij1}, \ldots, x_{ijp})^T$ is $p \times 1$ covariate vectors corresponding to fixed effects $\beta = (\beta_1, \ldots, \beta_p)^T$. We assume that the log-frailties v_i are independent and follow a distribution with frailty parameter θ. Although the results of this paper can be extended to non-normal frailties (e.g. gamma frailty), for simplicity, we assume a normal distribution $v_i \sim N(0, \theta)$ with $\theta = \sigma^2$, which is useful for modelling multi-component frailties (Ha *et al.*, 2007) including multilevel (nested) structures and/or correlated frailties including negative correlation (Rondeau *et al.*, 2008; Ha *et al.*, 2011).

Furthermore, the shared frailty model (2.1) can be easily extended to a multilevel structure in which patients, nested within hospitals, have recurrent event times as in the CGD data (Fleming and Harrington, 1991). Let T_{ijk} be the kth recurrent event time of the jth patient in the i the cluster (center). Let v_i^c be a frailty on the ith center and v_{ij}^p be that on the jth patient in the ith center. The multilevel frailty model (Yau, 2001; Ha *et al.*, 2007) is described by: $\lambda_{ijk}(t|v_i^h, v_{ij}^p) = \lambda_0(t)\exp(\eta_{ijk})$ with

$$\eta_{ijk} = x_{ijk}^T\beta + v_i^c + v_{ij}^p, \tag{2.2}$$

where $v_i^c \sim N(0, \sigma_c^2)$ and $v_{ij}^p \sim N(0, \sigma_p^2)$ are independent. Note that the extension of results from the shared model (2.1) to the multilevel model (2.2) is straightforward (Ha *et al.*, 2007, 2011).

2.2. H-likelihood procedure

In the semi-parametric frailty model (2.1), the functional form of $\lambda_0(t)$ is unknown. The non-parametric estimator 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 $\Lambda_0(t) = \sum_{k:y_{(k)} \le t} \lambda_{0k}$, where $y_{(1)} < \ldots < y_{(r)}$ are the ordered distinct event times and $\lambda_{0k} = \lambda_0(y_{(k)})$. Following Ha *et al.* (2001), the h-likelihood for the model (2.1) is given by

$$h = h\{(\beta, \lambda_0, \theta), v\} = \sum_{ij} \ell_{1ij} + \sum_i \ell_{2i}, \tag{2.3}$$

where

$$\begin{aligned} \sum_{ij} \ell_{1ij} &= \sum_{ij} \delta_{ij}\{\log \lambda_0(y_{ij}) + \eta_{ij}\} - \sum_{ij}\{\Lambda_0(y_{ij}) \exp(\eta_{ij})\} \\ &= \sum_{k=1}^{r} d_{(k)} \log \lambda_{0k} + \sum_{ij} \delta_{ij}\eta_{ij} - \sum_{k=1}^{r} \lambda_{0k} \Big\{ \sum_{(i,j)\in R_{(k)}} \exp(\eta_{ij}) \Big\}, \end{aligned}$$

$\ell_{1ij} = \ell_{1ij}(\beta, \lambda_0; (y_{ij}, \delta_{ij})|v_i)$ is the logarithm of the conditional density function for (y_{ij}, δ_{ij}) given v_i, $\ell_{2i} = \ell_{2i}(\theta; v_i)$ is the logarithm of the density function for v_i with parameter θ; if the log-frailty $v_i \sim N(0, \theta)$ with $\theta = \sigma^2$, then $\ell_{2i} = \ell_{2i}(\theta; v_i) = -(1/2)\log(2\pi\theta) - (1/2)(v^T v/\theta)$. Here, $v = (v_1, \ldots, v_q)^T$ is a vector of v_i's, $\lambda_0 = (\lambda_{01}, \ldots, \lambda_{0r})^T$, $d_{(k)}$ is the number of events at $y_{(k)}$, and $R_{(k)} = R(y_{(k)}) = \{(i,j) : y_{ij} \ge y_{(k)}\}$ is the risk set at $y_{(k)}$.

As the number of λ_{0k}s can increase with the number of events, the function $\lambda_0(t)$ is potentially of high dimension. Accordingly, for estimation of (β, v) Ha *et al.* (2001) proposed the use of the profiled h-likelihood h^* from which λ_0 in h of (2.3) is eliminated:

$$h^* = h|_{\lambda_0 = \widehat{\lambda}_0} = \sum_{ij} \ell_{1ij}^* \; + \; \sum_i \ell_{2i}, \tag{2.4}$$

where

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

are solutions of the estimating equations, $\partial h/\partial \lambda_{0k} = 0$, for $k = 1, \ldots, D$. Note here that $\sum_{ij} \ell_{1ij}^* = \sum_{ij} \ell_{1ij}|_{\lambda_0 = \widehat{\lambda}_0} = \sum_k d_{(k)} \log \widehat{\lambda}_{0k} + \sum_{ij} \delta_{ij}\eta_{ij} - \sum_k d_{(k)}$ does not depend on nuisance parameters λ_0 and that h^* in (2.4) is proportional to the penalized partial likelihood (denoted by h_p in the **frailtyHL** package) of Therneau and Grambsch (2000). Thus Lee and Nelder's (1996, 2001) h-likelihood procedure for hierarchical generalized linear models (HGLMs) can be extended to the frailty models based on h^*. Accordingly, given frailty parameter θ, the maximum h-likelihood estimators of $\tau = (\beta^T, v^T)^T$ are obtained by solving

the joint estimating equations, $\partial h^*/\partial\tau = \partial h/\partial\tau|_{\lambda_0=\hat{\lambda}_0} = 0$; in particular, $\hat{v}(\theta)$ is the solution to $\partial h^*/\partial v = 0$ for a given θ and $\hat{v}(\theta) \approx E_\theta(v|y,\delta)$ as $n^* = \min_{1\le i\le q} n_i \to \infty$ (Lee and Nelder, 2009; Ha *et al.*, 2011). Furthermore, for the estimation of θ we use an adjusted profile h-likelihood, $p_{\beta,v}(h^*)$, defined by

$$p_\tau(h^*) = \left[h^* - \frac{1}{2}\log\det\left\{\frac{J(h^*;\tau)}{(2\pi)}\right\}\right]\Bigg|_{\tau=\hat{\tau}}, \tag{2.5}$$

where $\widehat{\tau} = \widehat{\tau}(\theta) = (\widehat{\beta}^T(\theta), \widehat{v}^T(\theta))^T$ and $J(h^*;\tau) = -\partial^2 h^*/\partial\tau^2$ is an information matrix for τ with a detailed form in (3.1): see also Ha and Lee (2003).

The outline of **frailtyHL**() package (Ha *et al.*, 2012b) for fitting frailty model (2.1) is below. The main function is **frailtyHL**(). For instance,

> frailtyHL(formula=Surv(time,status)~x +(1|id), RandDist="Normal",
+ mord=0, dord=1)

fits a lognormal frailty model as a default and uses **RandDist="Gamma"** for the gamma frailty model. Here **formula** is a formula object, with the response on the left of a ~ operator, and the terms for the fixed and random effects on the right. The response is a survival object as returned by the **Surv** function (Therneau, 2010). Here, **time** and **status** denote survival time and censoring indicator having 1 (0) for uncensored (censored) observation, where **x** denotes a fixed covariate and **id** denotes the subjects for a normally distributed log frailty. The expression (**1**|**id**) ((**x**|**id**)) specify the random intercept (random slope) model. **mord** and **dord** are the orders of Laplace approximations to fit the mean parameters (**mord**= 0 or 1) and the dispersion parameters (**dord**= 1 or 2), respectively. Let HL(a,b) be the h-likelihood method using order "*a*" in **mord** and order "*b*" for **dord**. We recommend the use of HL(0,1) for the lognormal frailty and that of HL(0,2) for the gamma frailty if variance of frailty is not large: for the details see Ha *et al.* (2012b).

3. Interval estimation of frailty and R-codes

In frailty models (2.1), as in HGLMs (Lee and Nelder, 2009; Kim *et al.*, 2011), location parameters (β, λ_0, v) and frailty parameters θ are asymptotically orthogonal (Ha and Lee, 2003; Ha *et al.*, 2011). For a moment, assume that θ is known. Accordingly, we need only focus on (β, v) after eliminating λ_0 i.e. by using h^*. Following Ha and Lee (2003) and Ha (2008), the asymptotic covariances for $\hat{\beta}$ and $\hat{v} - v$ are obtained from the inverse of information matrix, $J(h^*;\beta,v) = -\partial^2 h^*/\partial(\beta,v)^2$, of β and v based on h^*:

$$J(h^*;\beta,v) = -\begin{pmatrix} \partial^2 h^*/\partial\beta\partial\beta^T & \partial^2 h^*/\partial\beta\partial v^T \\ \partial^2 h^*/\partial v\partial\beta^T & \partial^2 h^*/\partial v\partial v^T \end{pmatrix} = \begin{pmatrix} X^TW^*X & X^TW^*Z \\ Z^TW^*X & Z^TW^*Z + U \end{pmatrix}, \tag{3.1}$$

where X and Z are the $n\times p$ and $n\times q$ model matrices for β and v whose ijth row vectors are x_{ij}^T and z_{ij}^T, respectively, $U = -\partial^2\ell_2/\partial v^2$, and the weight matrix $W^* = W^*(\beta, v)$ is given in Appendix B of Ha and Lee (2003). This means that an approximated variance of $\hat{v} - v$ can be computed from the lower right-hand corner of J^{-1} in (3.1), leading to

$$\mathrm{var}(\hat{v} - v) \approx \{(Z^TW^*Z + U) \; - \; (Z^TW^*X)(X^TW^*X)^{-1}(X^TW^*Z)\}^{-1}|_{\beta=\hat{\beta},v=\hat{v}}.$$

Along the lines of Lee and Nelder (2009) and Ha *et al.* (2011), the individual $(1-\alpha)$-level h-likelihood confidence intervals (CIs) for the uni-dimensional components v_k of v are of the form

$$\hat{v}_k \pm z_{\alpha/2} \cdot \mathrm{SE}(\hat{v}_k - v_k), \tag{3.2}$$

where $\hat{v}$ maximizes the profile HL h^* in (2.4), $\mathrm{SE}(\hat{v}_k - v_k) = \sqrt{\mathrm{var}(\hat{v}_k - v_k)}$, and $z_{\alpha/2}$ is the normal quantile with probability $\alpha/2$ in the right tail (e.g. $z_{0.025} = 1.96$). Very recently, Ha *et al.* (2013) have shown via simulation studies that the HL interval in (3.2) preserves the nominal confidence level.

The R-codes for computing the 95% CIs in (3.2) of log-frailty v using the **frailtyHL** package as presented in Appendix A and B, with two practical examples.

4. Practical examples

We now propose a visualizing method for the CIs in (3.2) of individual frailties across clusters. For the illustration we re-analyze two practical examples (e.g., Ha *et al.*, 2012b) using two well-know real data sets; one is kidney infection data (McGilchrist and Aisbett, 1991) and the other is the chronic granulomatous disease (CGD) recurrent data (Fleming and Harrington, 1991).

4.1. Shared frailty model: kidney infection data

The data (McGilchrist and Aisbett, 1991) consist of times until the first and second recurrences ($n_i \equiv 2$) of kidney infection in 38 ($q = 38$) patients using a portable dialysis machine. Each survival time (**time**) is the time until infection since the insertion of the catheter. The survival times for the same patient are likely to be related because of a shared frailty describing the common patient's effect. The catheter is later removed if infection occurs and can be removed for other reasons, which we regard as censoring; about 24 per cent of the data were censored. We fit lognormal shared frailty model (2.1) with a single covariate, the sex (1 = male; 2 = female), using HL(0,1) in the **frailtyHL**(); the R-code and results are given in Appendix A.

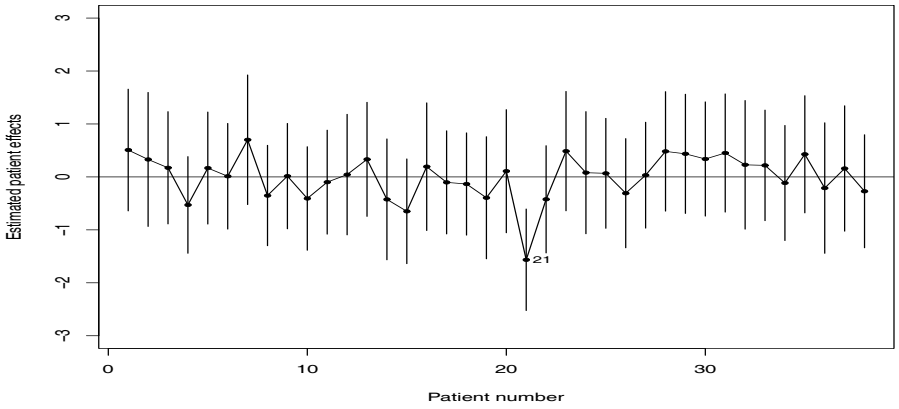

Figure 4.1 Estimated frailties of 38 patients in the kidney data and their 95% confidence intervals, under shared frailty model

The output shows that the effect of Sex is very significant (t-value = -3.214 with p-value = 0.001). That is, the female group has a significantly lower risk than the male group. Here, the variance estimate of frailty is $\hat{\sigma}^2 = 0.478$ (with SE = 0.313). Note that although we report the SE of σ^2, one should not use it for testing the absence of frailty $\sigma^2 = 0$ (Vaida and Xu, 2000). Such a null hypothesis is on the boundary of the parameter space, and hence, the critical value of an asymptotic $(\chi_0^2+\chi_1^2)/2$ distribution is $\chi^2_{1,0.10} = 2.71$ at the 5% level (Ha *et al.*, 2011, 2012b). The difference in deviance $-2p_{\beta,v}(h_p)$ in (2.5) between Cox's model without frailty and the lognormal frailty model is $369.96 - 364.68 = 5.28$ (> 2.71), indicating that the frailty effect is significant, i.e., $\sigma^2 > 0$. Here, the results from fitting Cox's model without frailty are available by adding the two arguments **varfixed=TRUE** and **varinit=c(0)** in the **frailtyHL()** function (Ha *et al.*, 2012b).

In Appendix A we present the R-codes for implementing the CIs and their plot for frailties of patients. Figure 4.1 shows that the random patient effects are heterogenous across patients; in particular, the 21th patient has a very lower frailty (i.e. less hazard) and that the corresponding CI does not include zero. This is also confirmed from the fact that the 21th patient among all patients experienced the longest 2nd infection time, 562. Thus we find that the Figure 4.1 can identify the heterogeneity of some particular patients. However, the test of frailty's variance does not provide such information.

4.2. Multilevel frailty model: CGD recurrent data

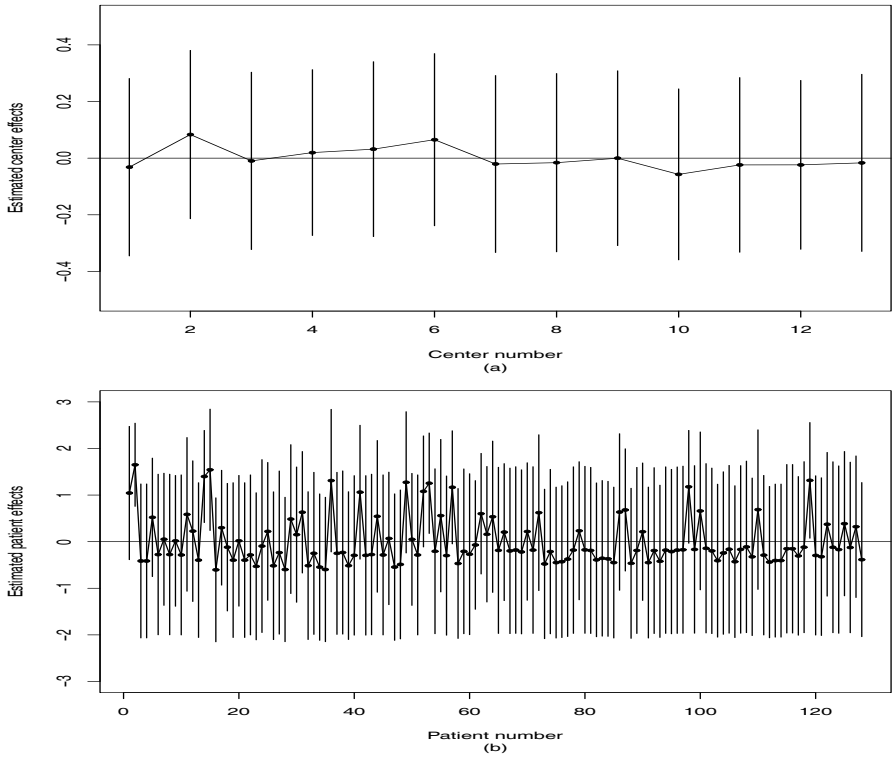

Figure 4.2 Estimated frailties of 13 centers and 128 patients in the CGD data and their 95% confidence intervals, under multilevel frailty model; (a) center effects and (b) patient effects

The CGD data set (Fleming and Harrington, 1991) consists of a placebo-controlled randomized trial of gamma interferon (rIFN-g) in the treatment of CGD. 128 patients (**id**) from 13 centers ($q_1 = 13, q_2 = 128$) were tracked for around one year. The number (i.e., cluster size) of patients per center ranged from 4 to 26. The survival times (**tstop-tstart**) are the recurrent infection times of each patient from the different centers. Censoring occurred at the last observation for all patients, except one, who experienced a serious infection on the data when he left the study; in the CGD study, about 63 per cent of the data were censored. The recurrent infection times for a given patient are likely to be correlated. However, each patient belongs to one of the 13 centers; hence, the correlation can also be attributed to a center effect.

Ignoring important random components may render many of the traditional statistical analysis techniques invalid (Ha *et al.*, 2007). We fit a multilevel lognormal frailty model (2.2) with two frailties and a single covariate, treat (rIFN-g versus placebo), using HL(0,1). Here, the two frailties are the random center and patient effects. The R codes and results are provided in Appendix B. The output shows that the effect of treatment is significant (t-value = -3.203 with p-value= 0.001), indicating that rIFN-g significantly reduces the rate of serious infection in CGD patients. The estimate of the variance of patient frailty ($\hat{\sigma}_p^2 = 0.982$) is considerably larger than the variance of center frailty ($\hat{\sigma}_c^2 = 0.026$), indicating that the random-patient effect is more heterogeneous. We find that the results above are very similar to those using HL(1,1) by Ha *et al.* (2012b). Though not reported here, we have investigated the significance of both random center and random patient effects using the difference of deviance ($-2p_{\beta,v}(h^*)$) as in Ha *et al.* (2012b); the results show the random-center effects are not necessary (i.e. $\sigma_c^2 = 0$), whereas the random-patient effects are indeed necessary (i.e. $\sigma_p^2 > 0$).

Figure 4.2 confirms these findings above; the corresponding R-codes are given in Appendix B. That is, Figure 4.2(a) shows random center effects (v^c) are very homogeneous across centers. However, Figure 4.2(b) indicates substantial variation in the random patient effects (v^p) across patients. In particular, five patients (2, 14, 15, 53, 119) noticeably stand out because their CIs do not include zero; this may indicate a possibility that the five patients are outliers: see also Noh *et al.* (2006).

5. Discussion

We have shown how to display the interval estimation of frailty based on the **frailtyHL** package. We have also found via two examples in Section 4 that the proposed method is very useful for investigating some heterogeneity across clusters. The updated R-package including interval estimation is current available from the second author. This package allows shared and multilevel frailties only. Allowance for dispersion frailties (Noh *et al.*, 2006) or correlated frailties (Ha *et al.*, 2011) is currently developing.

Appendix A: R codes and results for shared frailty model

```
############## (A1) Fitted model ###############
> library(frailtyHL)
> data(kidney)
> res<-frailtyHL(Surv(time,status)∼sex+(1|id), data=kidney)   # fitting of shared model
iteration :
48
convergence :
7.72225e-07
[1] "converged"
[1] "Results from the log-normal frailty model"
[1] "Number of data : "
[1] 76
[1] "Number of event : "
[1] 58
[1] "Model for conditional hazard : "
Surv(time, status)   sex + (1 | id)
[1] "Method : HL(0,1)"
[1] "Estimates from the mean model"
     Estimate Std. Error t-value p-value   # fixed effect's estimate and its standard error
sex -1.353 0.4209 -3.214 0.00131
[1] "Estimates from the dispersion model"
     Estimate Std. Error
id 0.4776 0.3127   # log-frailty's variance and its standard error
-2h0 -2∗hp -2∗p_{b,v}(hp)
[1,] 332.67 388.24 364.68
cAIC mAIC rAIC   # cAIC, conditional AIC; mAIC, marginal AIC; rAIC, restricted AIC
[1,] 361.77 368.79 366.68

  ############## (A2) CI for frailties ###############
> p=1; q=38
> vh<- res$vh   # estimates of log-frailties
> var<- diag(res$Hinv)[(p+1):(p+q)]   # computation of var(v̂ − v)
> se<- sqrt(var)   # their standard errors
> lb<- vh -1.96∗se   # lower bound
> ub<- vh +1.96∗se   # upper bound
> CI<- cbind(lb,ub)   # computation of their CIs
> patient<- 1:q
> plot(vh∼patient, ylim=c(-3,3), ylab="Estimated patient effects",
+      xlab="Patient number", pch=20, type="o")   # plot for interval estimation
> abline(h=0)
> for (i in 1:q){
+      x1<- c(i,i)
+      y1<- c(lb[i],ub[i])
+      lines (y1∼x1)
+      }
> text(21.8, vh[21],21, cex=.8)
```

Appendix B: R codes and results for multilevel frailty model

```
############## (B1) Fitted model ##############
> data(cgd)
 # fitting of multilevel model > res<-frailtyHL(Surv(tstop-tstart,status)~treat+(1|center)+(1|id),
data=cgd)
iteration :
170
convergence :
9.647067e-07
[1] "converged"
[1] "Results from the log-normal frailty model"
[1] "Number of data : "
[1] 203
[1]"Number of event : "
[1] 76
[1] "Model for conditional hazard : "
Surv(tstop - tstart, status)  treat + (1 | center) + (1 | id)
[1] "Method : HL(0,1)"
[1] "Estimates from the mean model"
     Estimate Std. Error t-value p-value
treatrIFN-g -1.074 0.3353 -3.203 0.001362
[1] "Estimates from the dispersion model"
     Estimate Std. Error
center 0.0262 0.1533
id 0.9817 0.5007
-2h0 -2*hp -2*p_{b,v}(hp)
[1,] 604.31 850.02 693.07
cAIC mAIC rAIC
[1,] 685.44 698.72 697.07
############## (B2) CI for frailties ##############
p=1; q1=13; q2=128; q=q1+q2
vh<- res$vh; var<- diag(res$Hinv)[(p+1):(p+q)]; se<-sqrt(var)
v1<-vh[1:q1]; se1<-se[1:q1]   #### For center
lb1<-v1-1.96*se1; ub1<-v1+1.96*se1; CI1<-cbind(lb1,ub1)
center<-1:q1
plot(v1~center, ylim=c(-0.5,0.5), ylab="Estimated center effects",
xlab="Center number", sub="(a)", pch=20, type="o")
abline(h=0)
for (i in 1:q1){ x1<-c(i,i); y1<-c(lb1[i],ub1[i]); lines (y1~x1) }
v2<-vh[(q1+1):q]; se2<-se[(q1+1):q]   #### For patient
lb2<-v2-1.96*se2; ub2<-v2+1.96*se2; CI2<-cbind(lb2,ub2)
patient<-1:q2
plot(v2~patient, ylim=c(-3,3), ylab="Estimated patient effects",
xlab="Patient number", sub="(b)",pch=20, type="o")
abline(h=0)
for (i in 1:q2){ x2<-c(i,i); y2<-c(lb2[i],ub2[i]); lines (y2~x2) }
```

References

Fleming, T. R. and Harrington, D. P. (1991), *Counting processes and survival analysis*, Wiley, New York.

Ha, I. D. (2008). A HGLM framework for meta-analysis of clinical trials with binary outcome. *Journal of the Korean Data & Information Science Society*, **19**, 1429-1440.

Ha, I. D. and Cho, G. H. (2012). H-likelihood approach for variable selection in gamma frailty models. *Journal of the Korean Data & Information Science Society*, **23**, 190-207.

Ha, I. D. and Lee, Y. (2003). Estimating frailty models via Poisson hierarchical generalized linear models. *Journal of Computational and Graphical Statistics*, **12**, 663-681.

Ha, I. D., Lee, Y. and MacKenzie, G. (2007). Model selection for multi-component frailty models. *Statistics in Medicine*, **26**, 4790-4807.

Ha, I. D., Lee, Y. and Song, J. K. (2001). Hierarchical likelihood approach for frailty models. *Biometrika*, **88**, 233-243.

Ha, I. D., Noh, M. and Lee, Y. (2012a). frailtyHL: frailty models using h-likelihood. `http://CRAN.R-project.org/package=frailtyHL`, R package version 1.1, 28-30.

Ha, I. D., Noh, M. and Lee, Y. (2012b). frailtyHL: A package for fitting frailty models with h-likelihood. *The R Journal*, **4**, 28-37.

Ha, I. D., Sylvester, R., Legrand, C. and MacKenzie, G. (2011). Frailty modelling for survival data from multi-centre clinical trials. *Statistics in Medicine*, **30**, 2144-2159.

Ha, I. D., Vaida, F. and Lee, Y. (2013). Interval estimation of random effects in proportional hazards models with frailties. *Statistical Methods in Medical Research*, Published online: 29/January/2013.

Kim, H. K., Noh, M. and Ha, I. D. (2011). A study using HGLM on regional difference of the dead due to injuries. *Journal of the Korean Data & Information Science Society*, **22**, 137-148.

Lee, Y. and Nelder, J. A. (1996). Hierarchical generalized linear models (with discussion). *Journal of the Royal Statistical Society B*, **58**, 619-678.

Lee, Y. and Nelder, J. A. (2001) Hierarchical generalised linear models: A synthesis of generalised linear models, random-effect models and structured dispersions. *Biometrika*, **88**, 987-1006.

Lee, Y. and Nelder, J. A. (2009). Likelihood inference for models with un observables: Another view (with discussion). *Statistical Science*, **24**, 255-293.

McGilchrist C.A. and Aisbett, C.W. (1991). Regression with frailty in survival analysis. *Biometrics*, **47**, 461-466.

Noh, M, Ha, I. D. and Lee, Y. (2006). Dispersion frailty models and HGLMs. *Statistics in Medicine*, **25**, 1341-1354.

Park, J. K., Oh, K. H. and Kim, M. S. (2012). Survival analysis on the business types of small business using Cox's proportional hazard regression model. *Journal of the Korean Data & Information Science Society*, **23**, 257-269.

Rondeau, V., Michiels, S., Liquet, B. and Pignon, J. P. (2008). 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*, **27**, 1894-1910.

Therneau, T. M. and Grambsch, P. M. (2000). *Modelling survival data: Extending the Cox model,* Springer, New York.

Therneau, T. M. (2010). survival: survival analysis, including penalized likelihood. `http://CRAN.R-project.org/package=survival`, R package version 2.36-2, 28-31.

Vaida, F. and Xu, R. (2000). Proportional hazards model with random effects. *Statistics in Medicine*, **19**, 3309-3324.

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