STATISTICS IN MEDICINE
Statist. Med. 2005; **24**:2681–2693
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sim.2113

A correlated frailty model for analysing risk factors in bilateral corneal graft rejection for Keratoconus: a Bayesian approach

Soleiman Kheiri[1,†], Mohammad Reza Meshkani[2] and Soghrat Faghihzadeh[3,*]

[1]*Faculty of Medicine, Shahrekord University of Medical Sciences, Rahmatieh, Shahrekord 571, Iran*
[2]*Department of Statistics, School of Mathematical Sciences, Shahid Beheshti University, Tehran 19839, Iran*
[3]*Department of Biostatistics, School of Medical Sciences, Tarbiat Modarres University, Tehran, Iran*

SUMMARY

There are many unknown causes that increase the rate of corneal graft rejection. In bilateral cases, some of these unknown causes are common, and some are individual factors. In this paper, we use a *correlated frailty model* to analyse risk factors for bilateral corneal graft in Keratoconus. Applying the piecewise constant baseline hazard model, we have performed a Bayesian analysis of the correlated frailty model using the Markov chain Monte Carlo method. The correlated frailty model and the shared frailty model are compared by deviance information criterion. The results show more accurate and better fit for the correlated frailty model. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: correlated frailty; random effect; Bayesian; piecewise hazard; MCMC algorithm

1. INTRODUCTION

A corneal transplant, also known as a corneal graft or as penetrating keratoplasty (PK), involves removal of the central portion of the diseased cornea and replacing it with a matched donor's button of cornea. Corneal grafts are performed on patients with damaged or scarred cornea, which prevent acceptable vision. One common indication for corneal graft is Keratoconus. Keratoconus is a non-inflammatory and usually bilateral disease of the cornea. Although corneal graft for Keratoconus is highly successful, graft failure may occur. One of the most frequent reasons for graft failure is graft rejection. Graft rejection may occur at any time but it frequently occurs within few weeks to 20 years after corneal transplant surgery [1–4].

In this paper, we explore the risk factors for graft rejection in bilateral grafts in Keratoconus. Here, the survival time is defined as the time elapsed between graft surgery and graft rejection. For those grafts for which rejection has not occurred yet, the censored survival time is the time elapsed between corneal graft surgery and the last date of examination. In the bilateral

*Correspondence to: Soghrat Faghihzadeh, Department of Biostatistics, School of Medical Sciences, Tarbiat Modarres University, Tehran, Iran.
†E-mail: kheiri@hbi.ir

Received July 2003
Accepted September 2004

cases, the more damaged eye is treated first. Usually, the second eye will require a treatment within a year or more. For bilateral grafts, many recipient-related and environmental factors are common for each graft. So we expect survival times for each subject's graft to be correlated.

There are numerous causes that increase the rate of graft rejection, many of them being unknown. The unknown causes may involve recipient-related, donor-related, surgery-related and environmental factors. For the widely used Cox proportional hazard model [5], unknown causes lead to bias in the effect of other covariates. This problem can be overcome by multiplying the baseline hazard function by a random effect or frailty component. The modified model is known as the frailty model [6, 7]. One of the most popular frailty models is the shared frailty model, which is used for modelling multivariate survival data [8]. In this model, a shared unobservable quantity in the hazard induces a positive correlation among the survival times. However, individual unknown risk factors such as donor- and surgery-related unknown factors cannot be described by a shared frailty model. Yashin *et al.* [9] extended the shared frailty model to allow different but correlated frailty among observations within a group. The extended model is known as correlated frailty model. The correlated frailty model can explain two sources of variations (variation due to the unknown shared factors and variation due to the unknown individual factors and hazard function) [10]. By using a correlated frailty model, we can obtain more accurate parameter estimates and improve the model fit to the data [9, 11].

The goal of this paper is to study the effect of important, shared and individual unknown factors on survival time of bilateral corneal graft by using a correlated frailty model. In addition, the results will be compared to those of the shared frailty model. The paper is organized as follows. In Section 2, we describe the motivating data from a retrospective study. In Section 3, we briefly describe the shared and correlated frailty models. In Section 4, we develop a Bayesian analysis of the correlated frailty model by using Markov chain Monte Carlo to estimate the parameters. In Section 5, we show the results of the analysis of data described in Section 2. Finally, Section 6 provides some discussion.

2. THE STUDY DATA

The records of 119 patients who had bilateral corneal graft at Labafinejad Hospital, Shahid Beheshti University of Medical Sciences, Tehran, between 1983 and 2002 are considered. The information included is time of operation, time of first rejection or time of last examination, status of rejection, recipient-related factors including, sex, age at time of surgery, severe ocular allergy, the extent of corneal vascularization and donor-related factors including fresh or preserved in cornea solution donor cornea. The follow-up period ranged from 1 to 221 months with a mean of 43.9 months. The survival time (time lapsed between surgery and rejection) ranged from 1.1 to 95 months with a mean of 12.5 months and median of 6.9 months. The period between the first and second graft was from 3.8 to 204 months with a mean of 44.6 months.

During the study period, there were 54 rejections, 31 rejections on the first eye and 23 rejections on the second eye. There were 18 bilateral rejections (in nine patients) and 36 unilateral rejections. Figure 1 shows the Kaplan–Meier survival curve for all 238 grafts and Figure 2 shows the Kaplan–Meier survival curve for the first and the second grafts.

Survival curves show high rate of rejection in the earlier months after surgery. There are many risk factors for graft rejection that are not included in this study. While some

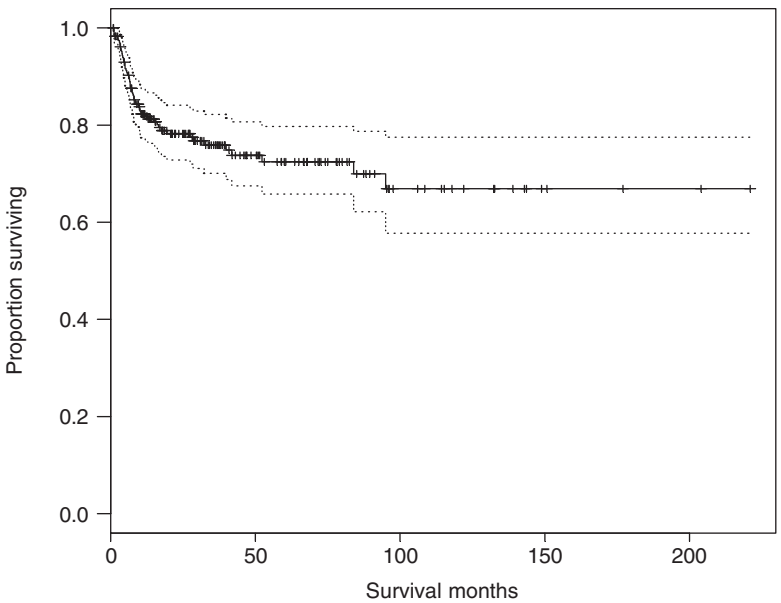

Figure 1. Kaplan–Meier survival curve for all grafts.

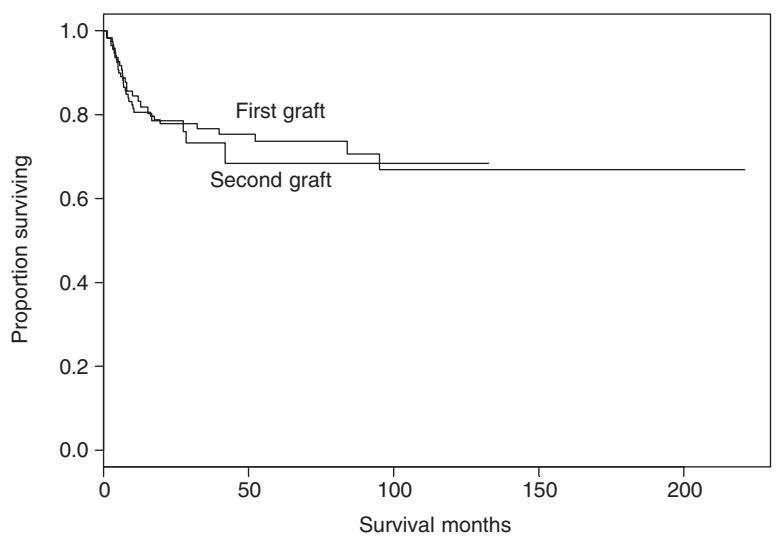

Figure 2. Kaplan–Meier survival curves for the first and second grafts.

of these could have been recorded as explanatory variables, they are not available in this retrospective study. We include an unmeasurable random variable to account for them. This random variable is considered as the sum of individual and shared random effects. In this study, the estimation of the magnitude of the individual random effects is more important than shared effects, because none of the factors in the model are surgery-related and only one of them is donor-related.

3. STATISTICAL MODEL

3.1. Shared frailty model

In this model, a shared unobservable quantity in the hazard induces a positive correlation among the survival time (time to graft rejection) of the jth subject (graft) ($j=1,2$) in the ith group (patient) ($i=1,\ldots,n$).

In a shared frailty model, the conditional hazard function of T_{ij} (rejection time for jth graft in ith patient), given the unobservable frailty random variable Y_i of the ith group (patient) and fixed observed covariate vector x_{ij}, is assumed as

$$h_{ij}(t|y_i,x_{ij})=y_i h_0(t)\exp(x'_{ij}\beta), \quad \begin{matrix} i=1,\ldots,n \\ j=1,\ldots,n_i \end{matrix}$$

where $h_0(t)$ is an unknown baseline hazard function common to every subject (graft) and β is the vector of fixed effect parameters. The shared frailty random variable Y_i is assumed to be independent and identically distributed for groups (patients), having some parametric distribution with unit mean (to obtain identifiability). The gamma distribution is most commonly used to model the frailty [12, 13].

In this paper we assume the Y_i's to be independent with

$$f_{Y_i}(y_i)=\frac{1}{\Gamma(\theta)}\theta^{\theta}y_i^{\theta-1}\exp(-\theta y_i), \quad i=1,\ldots,n$$

Thus, higher values of θ^{-1} signify larger variances for y_i, consequently greater heterogeneity among different groups (patients) and larger positive correlation between two graft rejection times for a patient.

3.2. Correlated frailty model

A bivariate extension of the shared frailty model is known as the correlated frailty model [9], wherein the random frailty varies among individuals in each group, so the conditional hazard function of T_{ij} given the unobserved frailty random variable y_{ij} and fixed observed covariate vector x_{ij} is

$$h_{ij}(t|y_{ij},x_{ij})=y_{ij}h_0(t)\exp(x'_{ij}\beta), \quad \begin{matrix} i=1,\ldots,n \\ j=1,\ldots,n_i \end{matrix}$$

where it is assumed that

$$Y_{ij} = W_i + Z_{ij}$$

and W_i, $Z_{i1}, \ldots, Z_{in_i}$ are mutually independently distributed variables.

W_i is a shared component for all subjects (grafts) in the ith group (patient i) and it generates dependence among subjects, while $Z_{i1}, \ldots, Z_{in_i}$ are individual components, generating only extra variance. In this model, the individuals in each group have different but correlated frailty. The gamma distributions have been used for all components of frailty in the applications of correlated frailty models. These distributions have been restricted to have the same scale parameter ensuring a marginal gamma distribution with mean one for Y_{ij}. In this regard, we assume the W_i's to be independent with

$$f_{W_i}(w_i) = \frac{1}{\Gamma(\varphi)} \theta^{\varphi} w_i^{\varphi-1} \exp(-\theta w_i), \quad i = 1, \ldots, n$$

and Z_{ij}'s to be independent with

$$f_{Z_{ij}}(z_{ij}) = \frac{1}{\Gamma(\theta - \varphi)} \theta^{\theta-\varphi} z_{ij}^{\theta-\varphi-1} \exp(-\theta z_{ij}), \quad \begin{matrix} i = 1, \ldots, n \\ j = 1, \ldots, n_i \end{matrix}$$

In our study, the shared component describes unknown recipient-related and environmental factors and the individual component describes unknown donor- and surgery-related factors.

4. BAYESIAN ANALYSIS OF SURVIVAL MODEL

In this section, we develop a Bayesian analysis of frailty models for the correlated frailty. Bayesian analysis of survival data using semi-parametric models requires modelling of a baseline hazard function. In this study, we use a piecewise constant hazard model for modelling the baseline hazard [14, 15]. We apply a partition of time of study into some intervals and assume the hazard function to be constant in each interval.

The full Bayesian model consists of conditional survival model considered in the previous section (correlated frailty model), and the prior distributions of the parameters and hyperparameters. We assume that the prior distributions for the fixed effects, the random effects, and the parameters of baseline hazard function are independent of each other. The joint distribution of the data and the parameters is given in the appendix, which is very complicated.

The analytical Bayes solution to the problem necessitates the determination of the posterior distribution of the parameters, including the hyperparameters, conditional on the observed data. Unfortunately, this is not possible to do analytically, nor is it practical to do numerically because of higher dimension of the parameter space. However, we can find a Markov chain that has the posterior as its long run distribution [16, 17]. Sampling from this Markov chain after an adequate burn-in period will enable us to approximate a sample from the posterior. The graphical modelling approach is used to specify the conditional distributions [18]. The hierarchical structure of the model simplifies the conditional distributions to be sampled from. By using graphical modelling in a hierarchical model, the conditional distribution of one node, given all the other nodes, is proportional to the product of the prior distribution of that

node and the conditional distribution of all its direct child nodes and coparent nodes. The Gibbs sampling, Metropolis algorithm, rejection sampling and free adaptive rejection sampling have been applied for sampling from full conditional distributions [17, 19]. The conditional distribution of the parameters and hyperparameters and a full illustration of sampling methods are given in the appendix.

In all Metropolis algorithms, the normal proposal distribution is used and its variance is manipulated for each node, in such a way that about 40 per cent of the candidates will be accepted [20]. A program written in $\mathfrak{R}$ software performs MCMC simulation, whose outputs are fed into a Bayesian output analysis (BOA) program [21] for carrying out convergence diagnostics and statistical analysis of Monte Carlo sampling outputs.

One problem arising in the implementation of the Markov chain Monte Carlo method is the assessment of convergence. We have run three parallel chains with different starting points for 10 000 iterations. Then we have calculated modified Gelman and Rubins [22] scale reduction factor R for each of the fixed effects parameters, the hyperparameters and the baseline hazard parameters from the last 5000 iterations. Comparing between chain variations and within chain variation, the scale reduction factor R measures how much improvement in the estimates would be possible by increasing the number of iterations. The scale reduction factor values were extremely close to 1. The largest R occurred for the parameter of random effect φ, and was 1.07. This indicates 5000 iterations would be a satisfactory burn-in period. We found no evidence from the multiple chains that they are not converging to the same nodes. At last we ran the chains for 20 000 iterations and used the last 15 000 samples after 5000 iterations for the burn-in period of each chains. This amounts to 45 000 iterations, for providing summaries posterior distributions of parameters. We also perform a similar Markov chain Monte Carlo simulation to estimate the parameters of the shared frailty model. We present comparison between the shared frailty model and the correlated frailty model for our data. The comparison has been carried out using deviance information criteria (DIC) for the three models. The DIC statistic introduced by Spiegelhalter *et al.* [23] is a Bayesian criterion for model comparison. Let $\underline{\theta}$ include all parameters of a model and let $D(\underline{\theta})$ be its deviance, then the DIC is defined as

$$\mathrm{DIC} = \overline{D(\underline{\theta})} + p_{\mathrm{D}}$$

where $\overline{D(\underline{\theta})}$ is the posterior mean of the deviance of the model and it is obtained by using the mean of monitored values of log likelihood after burn-in period and p_{D} is defined as

$$p_{\mathrm{D}} = \overline{D(\underline{\theta})} - D(\overline{\underline{\theta}})$$

which is the difference between the posterior mean of the deviance and the deviance of the posterior mean of parameter of interest.

The $\overline{D(\underline{\theta})}$ statistic is proposed as a Bayesian measure of fit or adequacy and p_{D} statistic is the effective number of parameters in a model and is suggested as a measure of complexity. The model with the smallest value of DIC is suggested to be the preferred one.

5. RESULTS

As it is seen in Figures 1 and 2, the survival curve of graft rejection shows a higher rejection rate in the earlier months. This calls for small intervals in the earlier months, which

Table I. Posterior summaries for the parameters of correlated frailty model.

Parameter	Mean	Median	Standard Deviation	2.5 percentile	97.5 percentile
θ	0.1488	0.1402	0.0432	0.0873	0.2691
φ	0.0576	0.0541	0.0248	0.0206	0.1254
Beta0	−3.5814	−3.6035	0.7673	−5.0459	−2.0557
Beta(Sex)	−0.4918	−0.4738	0.7258	−1.9551	0.8842
Beta(Age-mean(age))	0.0504	0.0491	0.0339	−0.0125	0.1201
Beta(Fresh cornea)	−0.5441	−0.5541	0.6521	−1.8462	0.7527
Beta(Ocular allergy)	2.5820	2.4715	1.6112	−0.1955	6.2809
Beta(Vascularization)	2.4633	2.3552	1.2514	0.3245	5.4223
Lambda1	0.6926	0.5012	0.6337	0.0921	2.4112
Lambda2	1.1275	0.9358	0.7242	0.3076	3.0246
Lambda3	1.6517	1.4229	0.9774	0.5007	4.1761
Lambda4	1.3821	1.1746	0.8746	0.3628	3.6474
Lambda5	1.1346	0.9245	0.8132	0.2441	3.2928
Lambda6	0.9738	0.7435	0.7894	0.1445	3.0664
Lambda7	0.7809	0.5643	0.7362	0.0634	2.7814

would yield more precision in modelling the baseline hazard function. Thus, we partition the study time into the following seven risk intervals measured in months: $(0 \leqslant t < 2)$, $(2 \leqslant t < 5)$, $(5 \leqslant t < 8)$, $(8 \leqslant t < 15)$, $(15 \leqslant t < 30)$, $(30 \leqslant t < 60)$, and $(60 \leqslant t < 221)$. The summaries of posterior distributions for the correlated frailty model parameters are shown in Table I. It is concluded that the extent of corneal vascularization is an important risk factor and the age of the recipient and severe ocular allergy are, however, slightly significant. The sex of reception and fresh or preserved in cornea solution of donor cornea are not important in graft rejection in Keratoconus. The estimates of baseline hazard parameters show that the rejection rate increases for the first three intervals, the highest rejection rate is in the third interval, and then the rejection continues to drop for the remaining time in the study. The posterior means of φ and θ show that the mean of shared random effect is 0.387 and the mean of individual random effect is 0.613. The variance of the shared component is 2.60 and the variance of individual component is 4.12. These values show that there is considerably more variability due to individual unknown effects than to shared unknown effects.

Summaries of posterior distributions of the parameters of the shared frailty model are also shown in Table II, which can be compared with Table I. Upon comparing the results shown in Tables I, and II, one can conclude that interval estimates of fixed effects are narrower in the shared frailty model compared to those from the correlated frailty model. The striking difference is the estimate of the coefficient of the age of recipient in the two models. The age of the recipient at surgery is a more important risk factor in the shared frailty model than in the correlated frailty model. This shows that the effect of age has been confounded with the individual random effects.

Comparing the summary of posterior distribution of θ in the two models indicates that variance of random effect in the correlated frailty model is about three times of variance

Table II. Posterior summaries for the parameters of shared frailty model.

Parameter	Mean	Median	Standard Deviation	2.5 percentile	97.5 percentile
θ	0.5230	0.4554	0.2646	0.2352	1.1736
Beta0	−4.2832	−4.2761	0.4630	−5.1320	−3.4076
Beta(Sex)	−0.3826	−0.3653	0.4578	−1.3142	0.4635
Beta(Age-mean(age))	0.0598	0.0590	0.0269	0.0085	0.1171
Beta(Fresh cornea)	−0.4879	−0.4778	0.4072	−1.3048	0.3479
Beta(Ocular allergy)	1.8656	1.7905	1.1268	−0.1933	4.4960
Beta(Vascularization)	2.1023	2.0814	0.7262	0.7386	3.6141
Lambda1	0.8945	0.6847	0.7390	0.1378	2.8624
Lambda2	1.3673	1.1770	0.7725	0.4150	3.3120
Lambda3	1.7845	1.5605	0.9579	0.5848	4.2047
Lambda4	1.2831	1.0865	0.7892	0.3543	3.3565
Lambda5	0.9224	0.7333	0.6666	0.2016	2.7002
Lambda6	0.7036	0.5293	0.5897	0.1136	2.3426
Lambda7	0.5153	0.3422	0.5329	0.0397	1.9958

Table III. Results of DIC.

Model	DIC	p_D	$D(\bar{\theta})$	$\overline{D(\underline{\theta})}$
Shared frailty	598.4	51.6	495.2	546.8
Correlated frailty	526.1	66.3	393.5	459.8

of random effect in the shared frailty model, which means that the correlated frailty model explains variability more effectively than the shared frailty model.

The result of comparison is presented in Table III and shows that the correlated frailty model has better goodness of fit and has more complexity than the shared frailty model. Overall, the DIC value for the correlated frailty model is smaller than the DIC value for the shared frailty model, implying considerably better fit for correlated frailty model. This confirms that individual unknown factors such as unknown surgery- and donor-related factors are very important in graft rejection for Keratoconus.

6. DISCUSSION

In this paper we have extended a Bayesian analysis of frailty model to the correlated frailty model, using the Markov chain Monte Carlo method with application to rejection risks of bilateral corneal grafts. We see that the correlated frailty model has better fit to our data than the shared frailty model. In the shared frailty model, it is almost impossible to have one short and one long survival time within a pair, but in the correlated frailty model, it is possible to have such situation [10]. The shared frailty model describes only positive correlation between two rejection times in the same patient and ignores the heterogeneity due to unknown individual factors, underestimating the standard errors, and lead to narrower

interval estimates than they should be. Correlated frailty model introduces the heterogeneity into the model; thus the fixed effect estimates have slightly increased standard errors and somewhat wider credible intervals than the corresponding credible intervals obtained by the shared frailty model.

In our application, a specific form for the baseline hazard is not known, and therefore a semi-parametric approach is applied. We use piecewise constant baseline hazard function, but other baseline hazard functions such as gamma process [13, 24] and correlated prior process [25] can be used. There is no clear-cut procedure for the choice of a partition. Some authors have suggested equal length intervals [15]. In this study, partitioning of the time axis has been carried out with the following consideration. The hazard function has a skewed shape touching the time axis very fast. To preserve this character, we have chosen a partition based on trial and error, which provides a piecewise constant estimate of hazard function, nearly similar to the original one. Consideration of coarser or equal length partitions would lead to an unstable estimate of hazard function. In such a case, it will not have a clear interpretation.

We have used DIC for comparing the model fitting. The most popular approach in Bayesian literature for comparing models is the Bayes Factor [26]. However, the Bayes factor is generally quite sensitive to vague priors and is not applicable for our models, since we consider some vague priors.

Although computationally intensive, Markov chain Monte Carlo is a useful technique for estimating complex Bayesian models. In this paper, we used Gibbs sampling, Metropolis algorithms, and rejection sampling for sampling from conditional distributions of the model parameters. Convergence is determined by Gelman–Rubin convergence diagnostics. The use of Monte Carlo methods in our model can provide a prediction of a future graft rejection for a particular patient. One problem in our study is that we have considered survival times of bilateral corneal grafts as parallel survival times. However, we must attend to this fact that the second graft for each patient has been carried out after a period of time lapse from the first graft. Hence further studies are needed to find a suitable model for our data.

APPENDIX A

To construct piecewise constant hazard model, we first construct a finite partition of the time axis, $0<t_1<\cdots<t_k<\infty$, with all subjects (graft rejections) being either failed or censored before t_k. Thus, we have the k intervals $I_1=[0,t_1), I_2=[t_1,t_2),\ldots,I_k=[t_{k-1},t_k)$, and we assume the baseline hazard function is constant over each interval. Suppose λ_s is the value of hazard function in the sth interval, then

$$h_0(t)=\lambda_s, \quad t\in I_s=[t_{s-1},t_s), \;\; s=1,\ldots,k$$

If g_{ij} is such that $t_{ij}\in[t_{g_{ij}},t_{g_{ij}+1})$ then, we set

$$h_0(t_{ij})=\lambda_{g_{ij}+1}$$

Let Δ_s be the length of sth interval, $\Delta_s=t_s-t_{s-1}$, then the cumulative baseline hazard is

$$H_0(t_{ij})=\sum_{s=1}^{g_{ij}}\Delta_s\lambda_s+(t_{ij}-t_{g_{ij}})\lambda_{g_{ij}+1}$$

Let δ_{ij} be the indicator variable taking value 1 if the graft is rejected during the study period and 0 if it is not rejected until the last examination. Let D_{obs} be the set of observations consisting of $(t_{ij}, \delta_{ij}, x_{ij})$ and let E be the set of random effects consisting of (w_i, z_{ij}). Assuming a non-informative censoring, then the complete likelihood is given as

$$L(\beta, \lambda | E, D_{\text{obs}}) = \prod_{i=1}^{n} \prod_{j=1}^{n_i} [y_{ij} h_0(t_{ij}) \mu_{ij}]^{\delta_{ij}} \exp(-y_{ij} H_0(t_{ij}) \mu_{ij})$$

where $\lambda = (\lambda_1, \ldots, \lambda_k)$, $H_0(t)$ is the cumulative baseline hazard function, and $\mu_{ij} = \exp(x'_{ij}\beta)$.

The joint distribution of the data and the parameters is

$$\begin{aligned}
\pi(E, D_{\text{obs}}, \beta, \theta, \varphi, \lambda) &= L(\beta, \lambda | E, D_{\text{obs}}) \cdot \left[\prod_{i=1}^{n} f(w_i | \theta, \varphi) \right] \\
&\quad \times \left[\prod_{i=1}^{n} \prod_{j=1}^{n_i} f(z_{ij} | \theta, \varphi) \right] \cdot \pi(\theta) \cdot \pi(\varphi) \cdot \pi(\beta) \cdot \pi(\lambda) \\
&= \left\{ \prod_{i=1}^{n} \prod_{j=1}^{n_i} [y_{ij} h_0(t_{ij}) \mu_{ij}]^{\delta_{ij}} \exp(-y_{ij} H_0(t_{ij}) \mu_{ij}) \right\} \\
&\quad \times \left[\prod_{i=1}^{n} \frac{1}{\Gamma(\varphi)} \theta^{\varphi} w_i^{\varphi - 1} \exp(-\theta w_i) \right] \\
&\quad \cdot \left[\prod_{i=1}^{n} \prod_{j=1}^{n_i} \frac{1}{\Gamma(\theta - \varphi)} \theta^{\theta - \varphi} z_{ij}^{\theta - \varphi - 1} \exp(-\theta z_{ij}) \right] \\
&\quad \times \pi(\theta) \cdot \pi(\varphi) \cdot \pi(\beta) \cdot \pi(\lambda)
\end{aligned}$$

We use the gamma priors with shape parameter $\theta_1 = 0.001$ and scale parameter $\theta_2 = 0.001$ for the hyperparameter of random effect θ, which is a non-informative prior and precludes large frailty effect variance. By using the graphical modelling approach, the conditional distribution of the hyperparameter θ is

$$\begin{aligned}
\pi(\theta | E) &\propto \prod_{i=1}^{n} \prod_{j=1}^{n_i} \left(\frac{1}{\Gamma(\theta)} \theta^{\theta} y_{ij}^{\theta - 1} \exp(-\theta y_{ij}) \right) \cdot \frac{1}{\Gamma(\theta_1)} \theta_2^{\theta_1} \theta^{\theta_1 - 1} \exp(-\theta \theta_2) \\
&\propto \left(\frac{1}{\Gamma(\theta)} \right)^{\sum_i n_i} \theta^{\theta \sum_i n_i + \theta_1 - 1} \left(\prod_{i=1}^{n} \prod_{j=1}^{n_i} y_{ij} \right)^{\theta} \exp\left(-\theta \left(\theta_2 + \sum_{i=1}^{n} \sum_{j=1}^{n_i} y_{ij} \right) \right)
\end{aligned}$$

This distribution is log concave in θ, so we can use the free adaptive rejection sampling of Gilks to sample from it [19].

Since the hyperparameter φ is restricted to $0 < \varphi < \theta$, we use the uniform prior distribution on interval $[0, \theta]$ for φ which is a non-informative prior. So the conditional distribution of

the hyperparameter φ is proportional to

$$\theta^{n\varphi-\sum_i n_i\varphi-1}\left(\frac{1}{\Gamma(\varphi)}\right)^n\left(\frac{1}{\Gamma(\theta-\varphi)}\right)^{\sum_{i=1}^n n_i}\left(\prod_{i=1}^n w_i^{\varphi}\right)\left(\prod_{i=1}^n\prod_{j=1}^{n_i} z_{ij}^{-\varphi}\right)$$

We find the mode of conditional distribution of φ in each iteration and use the rejection sampling for sampling from its conditional distribution. We let the envelope function be proportional to the uniform distribution on $(0,\theta)$.

The conditional distribution of the shared random effect w_i is proportional to

$$\left[\prod_{j=1}^{n_i}(w_i+z_{ij})^{\delta_{ij}}\right] w_i^{\varphi-1}\exp\left(-w_i\left(\theta+\sum_{j=1}^{n_i}H_0(t_{ij})\mu_{ij}\right)\right)$$

and the conditional distribution of individual random effect z_{ij} is proportional to

$$(w_i+z_{ij})^{\delta_{ij}} z_{ij}^{\theta-\varphi-1}\exp(-z_{ij}(\theta+H_0(t_{ij})\mu_{ij}))$$

We use the Metropolis algorithm [17] for sampling from the conditional distribution of logarithm of shared and individual random effects.

The prior distribution of the fixed effects is assumed to be multivariate normal with mean vector $\mu_0=0$ and covariance matrix $\Sigma_0=1000I$, which is a non-informative prior distribution. The conditional distribution of the fixed effects β is proportional to

$$\exp\left\{-1/2(\beta-\mu_0)'\Sigma_0^{-1}(\beta-\mu_0)+\beta'\left(\sum_{i=1}^n\sum_{j=1}^{n_i}\delta_{ij}x_{ij}\right)\right.$$
$$\left.-\sum_{i=1}^n\sum_{j=1}^{n_i} y_{ij}H_0(t_{ij})\exp(x_{ij}'\beta)\right\}$$

which can be sampled using the Metropolis algorithm [17].

For baseline hazard parameter in the sth interval (λ_s), we use gamma priors with shape parameter α_s and scale parameter $\beta_s=1$. The shape parameter α_s is assumed as the maximum likelihood estimate of hazard function in the sth interval. Let d_s be the number of rejects that occurred in the interval I_s, R_s the risk set at t_s, and D_s the number of subjects (grafts) in the interval I_s:

$$R_s = \{(i,j);\ t_{ij} > t_s\}$$
$$D_s = R_{s-1} - R_s = \{(i,j);\ t_{ij}\in I_s\}$$
$$d_s = \sum_{D_s}\delta_{ij}$$

Then the conditional distribution of λ_s, $s=1,\ldots,k$ is given as

$$\pi(\lambda_s|D_{\text{obs}},E,\lambda_{-s},\beta)\propto(\lambda_s)^{d_s}\exp(-H_{0s}\lambda_s)\cdot\lambda_s^{\alpha_s-1}\exp(-\beta_s\lambda_s)$$
$$\propto\lambda_s^{d_s+\alpha_s-1}\exp(-\lambda_s(H_{0s}+\beta_s))$$

where

$$H_{0s}=\left\{\sum_{D_s}y_{ij}\mu_{ij}(t_{ij}-t_{s-1})+\sum_{R_s}y_{ij}\mu_{ij}\Delta_s\right\}$$

So all conditional distributions of baseline hazard parameters are gamma distributions with shape parameter $d_s+\alpha_s$ and scale parameter $H_{0s}+\beta_s$, $(s=1,\ldots,k)$. These nodes are sampled directly using Gibbs sampling step [17].

ACKNOWLEDGEMENTS

This work is a part of the Ph.D. thesis of the first author at the Department of Biostatistics, Tarbiat Modarres University, Tehran, Iran. He would like to thank the faculty of the Department for their support. The authors would like to thank the referees for their valuable comments and suggestions.

REFERENCES

1. Tuft SJ, Gregory MW, Davison CR. Bilateral penetrating keratoplasty for keratoconus. *Ophthalmology* 1995; **102**:462–468.
2. Musch DC, Meyer RF. Risk of endothelial rejection after bilateral penetrating keratoplasty. *Ophthalmology* 1989; **96**:1139–1143.
3. Lim L, Pesudovs K, Coster DJ. Penetrating keratoplasty for Keratoconus; Visual outcome and success. *Ophthalmology* 2000; **107**:1125–1131.
4. Donshik PC, Cavanagh HD, Boruchoff SA, Dohlman CH. Effect of bilateral and unilateral grafts on incidence of rejections in keratoconus. *American Journal of Ophthalmology* 1979; **87**:823–826.
5. Cox DR. Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society*, *Series B* 1972; **34**:187–220.
6. Vaupel JW, Manton KG, Stallavd E. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 1979; **16**:439–454.
7. Clayton DG. A model for association in bivariate life-table and its application in epidemiological studies of chronic disease incidence. *Biometrika* 1978; **65**:141–151.
8. Clayton D, Cuzick J. Multivariate generalization of the proportional hazards model. *Journal of the Royal Statistical Society*, *Series A* 1985; **148**:82–117.
9. Yashin AI, Vaupel JW, Iachine I. Correlated individual frailty: an advantageous approach to survival analysis of bivariate data. *Mathematical Population Studies* 1995; **5**(2):1–15.
10. Hougaard P. *Analysis of Multivariate Survival Data*. Springer: New York, 2000.
11. Petersen JH. An additive frailty model for correlated life times. *Biometrics* 1998; **54**:646–661.
12. Oakes D. Bivariate survival models induced by frailty. *Journal of the American Statistical Association* 1989; **84**:487–493.
13. Clayton D. A Monte Carlo method for Bayesian inference in frailty models. *Biometrics* 1991; **47**:467–485.
14. Bolstad W, Manda OS. Investigating child mortality in Malawi using family and community random effects: a Bayesian analysis. *Journal of the American Statistical Association* 2001; **96**:12–19.
15. Ibrahim JG, Chen M, Sinha D. *Bayesian Survival Analysis*. Springer: New York, 2001.
16. Gamerman C. *Markov Chain Monte Carlo*: *Stochastic Simulation for Bayesian Inference*. Chapman & Hall: London, 1997.
17. Gilks WR, Richardson S, Spiegelhalter DJ. *Markov Chain Monte Carlo in Practice*. Chapman & Hall: London, 1996.

18. Spiegelhalter DJ. Bayesian graphical modeling: a case study in monitoring health outcomes. *Applied Statistics* 1998; **47**(1):115–133.
19. Gilks WR. Derivative-free adaptive rejection sampling for Gibbs sampling. *Bayesian Statistics*, vol. 4. Oxford University Press: Oxford, 1992; 641–649.
20. Gelman A. Inference and monitoring convergence. In *Markov Chain Monte Carlo in Practice*, Gilks WR, Richardson S, Spiegelhalter DJ (eds). Chapman & Hall: London, 1996; 75–88.
21. Smith BJ. *Bayesian Output Analysis Program* (*BOA*), *Version 1.0.0*, User's manual. http://www.public-health.uiowa.edu, June 2001.
22. Brooks S, Gelman A. General method for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 1998; **8**(4):319–335.
23. Spigelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *Journal of Royal Statistical Society B* 2002; **64**(Part 4):583–639.
24. Kalbfleisch JD. Nonparametric Bayesian analysis of survival time data. *Journal of the Royal Statistical Society*, *Series B* 1978; **40**:214–221.
25. Aslanidou H, Dey DK, Sinha D. Bayesian analysis of multivariate survival data using Monte Carlo methods. *Canadian Journal of Statistics* 1998; **26**:38–48.
26. Kass RE, Raftery AE. Bayes factor. *Journal of the American Statistical Association* 1995; **90**:773–795.

