Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 2, 2021 (New Series), pp 199-216

Cause Specific Shared Frailty Proportional Reversed Hazards Models

Gleeja V. L[1]. and P. G. Sankaran[2]
1. *Department of Statistics, College of Veterinary and Animal Sciences, Mannuthy-680651, Kerala, India*
2. Department of Statistics, Cochin University of Science and Technology, Cochin-682022, Kerala, India

Received: 22 May 2020; Revised: 17 December 2020; Accepted: 20 December 2020

Abstract

In family studies, usually, information on ages at onset of diseases is collected and the resulting data are often left censored. When there is a possibility of multiple types of events (disease) in a left censored family (clustered) data, the concept of cause specific reversed hazard models and the concept of frailty are needed for modeling and analysis of the data. Hence, in this paper, for the analysis of clustered multiple event data with left censored observations, frailty models in terms of cause specific reversed hazard rates are introduced. The shared gamma frailty reversed hazards model for bivariate multiple event data are developed. The first model is developed for the analysis of data without the presence of covariates. In the second model, covariates are included and regression coefficients are assumed to be different for different type of events. The estimation of the parameters of the models by maximum likelihood method, using EM algorithm, is presented. The properties of the estimates are also discussed. Finally, the models are applied to real data sets.

Key words: Reversed hazard rate; Competing risk; Frailty models.

1. Introduction

In family studies on diseases with ages at onset, assessing the familial association is often the problem of interest. When age at onset is considered, the data is often left censored. Left censored observations occur when the exact value of a response has not been observed and instead, an upper bound on that response is observed. Such observations also arise if a measuring instrument lacks the sensitivity needed to measure the observations below a known threshold. Then the measurement is taken and if the signal is below the instrument threshold, all which is known is that measurement is less than the threshold. Left censored observations also occur in studies determining the age at which a child learns to perform a specified task. Often, some children can already perform the task when they enter to the study. Such lifetimes are considered as left censored. The modeling and analysis of such left censored lifetime data is carried out using reversed hazard rate. The concept of reversed hazard rate (RHR) has been proposed as dual to hazard rate by Barlow *et al.* (1963) and is defined for a nonnegative random variable T as $(t) = \lim_{\Delta t \to 0} \frac{P(t-\Delta t < T \leq t+\Delta t)}{\Delta t}$. That is, in a small interval, the product of the RHR function and the length of the interval is the approximate probability of failure in the interval given failure before the end of the interval. RHR was used for the estimation of the survival

Corresponding Author: Gleeja V. L.
E-mail: gleeja@gmail.com

function in the presence of left censored observations by Ware and DeMets (1976) for a baboon descent data. Later RHR was used for characterization of life distributions by Shaked and Shantikumar (1994), for investigating the properties for k out of n systems by Block *et al.* (1998) and for developing nonparametric estimators for right truncated data by Lawless (2003). Different authors, Chandra and Roy (2001), Gupta and Nanda (2001), Gupta and Wu (2001), Kalbfleisch and Lawless (1989), Nair, Sankaran and Asha (2005), Sankaran and Gleeja (2006), Bartoszewicz and Skolimowska (2006) and Faith (2017), extensively studied and presented results related to RHR.

Sometimes in studies involving family or subgroups, lifetimes of individuals within the subgroup may be related. For modeling association between individual lifetimes within subgroups, the notion of frailty was introduced by Vaupel *et al.* (1979). The model assumes frailty as a common random effect that acts multiplicatively on the hazard rates of all subgroup members. The most widely used frailty model is shared frailty model with gamma distribution as frailty distribution. It has been discussed in Vaupel *et al.* (1979), Clayton and Cuzick (1985), Klein (1992) and Andersen *et al.* (2003). Some other distributions for frailty like positive stable, Weibull, lognormal *etc.* are investigated in Hougaard (2000). The estimation of the parameters of shared frailty model using maximum likelihood method via the EM algorithm is developed in Nielson *et al.* (1992) and the asymptotic normality and efficiency of the estimators are studied and proved in Murphy (1994, 1995).

The concept of frailty as a common random effect that acts multiplicatively on RHR has been introduced in Sankaran and Gleeja (2008). Let (T_1, T_2) be the lifetimes of two related individuals. Then $m_j(Z, t_j) = Z m_{0j}(t_j)$ be conditional individual RHRs given frailty Z where $m_{0j}(t_j)$, $j = 1, 2$ are the baseline reversed hazards. Assume that lifetimes (T_1, T_2) are conditionally independent given frailty Z and Z follows a gamma distribution with mean one and variance θ. Then shared gamma frailty models is introduced by Sankaran and Gleeja (2011) as the distribution function of (T_1, T_2), $F(t_1, t_2) = \left[\theta M_{01}(t_1) + \theta M_{02}(t_2) + 1\right]^{-(1/\theta)}$ where $M_{01}(t_1)$ and $M_{02}(t_2)$ are the cumulative baseline reversed hazard function and $\theta \geq 0$.

When time to failure of paired organs like kidney, lungs, eyes, ears, dental implants *etc.* are considered, it is more appropriate to model using shared frailty models. The estimation was done using maximum likelihood method via EM algorithm. Later, estimation of parameters involved in the shared frailty model by the Bayesian estimation procedure using the Markov chain Monte Carlo (MCMC) technique was discussed in Hanagal *et al.* (2014). The most commonly used frailty distribution is Gamma distribution, because of its mathematical convenience. Other distributions can be used as frailty distribution and Hanagal and Pandey (2015) developed three parametric shared frailty models with inverse Gaussian frailty using RHR. Gamma frailty models with different baseline distributions are discussed in Hanagal and Bhambure (2017) and Hanagal and Pandey (2017). Inverse Gaussian correlated frailty model with different baseline distributions are discussed in Hanagal (2020) and Hanagal and Pandey (2020). The shared frailty models are attracting recent interest of researchers and extensive research is being conducted on these models. While analyzing family data on age at onset of a particular disease, shared frailty models using RHR is very useful.

But in certain studies on age at onset of diseases, individuals may be susceptible to more than one type of diseases or in some survival studies death can occur due to any one of the two

or more causes. When there are multiple types of causes for the event, the concept of competing risks facilitates analysis. Modeling and analysis of lifetime data with multiple type of events under right censoring is discussed in Crowder (2001), Kalbfleisch and Prentice (2002) and Lawless (2003). The analysis of competing risks data under left censoring using RHR has been carried out in Sankaran and Anjana (2014). Specifically, they presented the analysis of left censored data with multiple types of events using cause-specific RHR functions. Let (T, J) be a pair of random variables where T is possibly a censored lifetime and J represents cause of event. J takes values on the set {1,2,...,*r*}. These *r* causes are mutually exclusive and exhaustive, so that the individual can have at most one realized lifetime with an identifiable cause. Then cause specific RHR of *T* is defined as $m_j(t) = \lim_{\Delta t \to 0} \frac{P(t - \Delta t < T \le t, J = j / T \le t)}{\Delta t}$, *j* = 1,2,...,*r*. Thus $m_j(t)$ specifies the instantaneous rate of failure of an individual at time *t* due to cause *j* given that it failed before time *t*. Then the marginal RHR of T was given as $m(t) = \sum_{j=1}^{r} m_j(t)$. Sankaran and Anjana (2016) introduced a proportional cause specific RHR model for modeling and analysis of left censored competing risks data in the presence of covariates. The model was given as $m_j(t \mid \mathbf{x}) = m_{0j}(t)\exp(\boldsymbol{\beta}'\mathbf{x})$, *j*=1,2,...,*r* where $m_j(t \mid \mathbf{x})$ is the cause-specific RHR due to cause *j* in the presence of covariate **x** and **x** is a vector of p covariates, $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)'$ is the vector of p regression parameters, and $m_{0j}(t)$ is the baseline cause-specific RHR due to cause j. The vector of regression parameters **β** measures the effect of the covariate vector on the cause-specific RHR. But these models are not appropriate for clustered data like family data, as it does not consider the association exist between members of the family.

Thus, in order to analyze a left censored family (clustered) data with multiple types of events (diseases), a frailty-based competing risks models using RHR is needed. Motivated by this, in this paper, a shared gamma frailty model in terms of cause specific RHR is developed.

The paper is organized as follows. In section 2, cause specific shared frailty proportional RHR model is developed with and without the presence of covariates. The estimation and the asymptotic properties of the parameters of models are studied in section 3. In section 4, the model is illustrated with data sets from Ying and Wei (1994) and McGilchrist and Aisbett (1991). Finally conclusions and discussions are given in section 5.

2. Cause Specific Shared Frailty Proportional RHR Model

The model is constructed to deal with a clustered or family data with multiple causes of event. The time to event is the variable of interest and let us consider bivariate situation.

2.1. Cause specific shared frailty proportional RHR model without covariates

Let $T = (T_1, T_2)$ be the pair of lifetimes of two related individuals defined on a common probability space (Ω, F, P) with absolutely continuous distribution function. Let $F(t_1, t_2)$ and $F_j(t_j)$ respectively denote the joint distribution function of T and the marginal distribution function of T_j, *j=1,2*. Let the support of T be $D = [0, b_1] \times [0, b_2]$ where (b_1, b_2) is such that

$b_j = \inf\left\{t \mid F_j(t) = 1\right\}$, $j = 1,2$. Assume that each of the pair (T_1, T_2) is subject to multiple causes of event. Let $C = (C_1, C_2)$ denote the cause of event for T. Suppose that there are *r* causes for the event for each individual in the process. Assume that C_j is a unique element of the set {1,2,…,*r*}, *j*=1,2. We assume that individual can have at most one realized lifetime with an identifiable cause. Observations from the same cluster or family may share common environment or some other factors. Hence it is assumed that the pair of lifetimes shares a common unobserved frailty Z. First define the cause specific RHR of T_j for given frailty *Z* as

$$m_{jk}(Z, t_j) = Z m_{0jk}(t_j),\ j=1,2 \text{ and } k=1,2,\ldots,r$$

where $m_{0jk}(t_j)$ is baseline cause specific RHR function of T_j, *j*=1,2 and Z is an unobservable random variable having a probability density function $g(z)$. The marginal RHR of T_j for given frailty *Z* is obtained as

$$m_j(Z, t_j) = \sum_{k=1}^{r} m_{jk}(Z, t_j) = Z \sum_{k=1}^{r} m_{0jk}(t_j),\ j=1,2.$$

We assume that lifetimes (T_1, T_2) are conditionally independent given frailty Z. Then the distribution function of (T_1, T_2) given frailty Z is

$$F(t_1, t_2 \mid Z) = \exp\left\{-\sum_{k=1}^{r} \int_{t_1}^{\infty} m_{1k}(Z, u)\, du - \sum_{k=1}^{r} \int_{t_2}^{\infty} m_{2k}(Z, v)\, dv\right\}.$$

Let $g(z)$ be the joint density function of Z. Then the bivariate distribution function of (T_1, T_2) is

$$\begin{aligned} F(t_1, t_2) &= \int_0^{\infty} F(t_1, t_2 \mid Z)\, g(z)\, dz \ = E\left(F(t_1, t_2 \mid Z)\right) \\ &= E\left(\exp\left\{-Z\left(\sum_{k=1}^{r} M_{01k}(t_1) + \sum_{k=1}^{r} M_{02k}(t_2)\right)\right\}\right) \end{aligned}$$

where $M_{01k}(t_1)$ and $M_{02k}(t_2)$ are the cumulative baseline cause specific reversed hazard function.

The marginal distribution function of T_j is

$$F_j(t_j) = \int_0^{\infty} F_j(t_j \mid z)\, g(z)\, dz = E\left(\exp\left\{-Z \sum_{k=1}^{r} M_{01k}(t_1)\right\}\right),\ j = 1,2.$$

Suppose that *Z* is *i.i.d.* random variable with the following gamma density function

$$g(z)=\frac{z^{(1/\theta)-1}\exp\{-z/\theta\}}{\theta^{(1/\theta)}\Gamma(1/\theta)},\qquad \theta\geq 0.$$

The mean value of Z is 1 and variance is θ. Then bivariate distribution function of (T_1,T_2) is obtained as

$$F(t_1,t_2)=\left[\theta\sum_{k=1}^{r}M_{01k}(t_1)+\theta\sum_{k=1}^{r}M_{02k}(t_2)+1\right]^{-(1/\theta)}\qquad \text{where } \theta\geq 0. \tag{1}$$

Thus, cause specific shared gamma frailty proportional RHR model can be represented by (1).

The marginal distribution function T_j is then obtained as

$$F_j(t_j)=\left[\theta\sum_{k=1}^{r}M_{0jk}(t_j)+1\right]^{(-1/\theta)},\ j=1,2.$$

Therefore, the bivariate distribution function of (T_1,T_2) can be represented in terms of marginal distribution functions as

$$F(t_1,t_2)=\left[F_1(t_1)^{-\theta}+F_2(t_2)^{-\theta}-1\right]^{-(1/\theta)}\qquad \text{where } \theta\geq 0. \tag{2}$$

Remark 1: T_1 and T_2 are independent, when $\theta = 0$.

Remark 2: The model given in (2) is identifiable. Let $F(t_1,t_2)$ be a known distribution function given as in (2), and $\theta>0$, and let $m_i(b_i)=\sum_{k=1}^{r}m_{ik}(b_i)\neq 0$, $i=1,2$. $k=1,2,...,r$ for r different causes.

We obtain the joint density function as

$$f(t_1,t_2)=\frac{(1+\theta)\sum_{k=1}^{r}m_{1k}(t_1)\sum_{k=1}^{r}m_{2k}(t_2)F_1(t_1)^{-\theta}F_2(t_2)^{-\theta}}{\left[F_1(t_1)^{-\theta}+F_2(t_2)^{-\theta}-1\right]^{\left(\frac{1}{\theta}\right)+2}}.$$

Since $m_i(b_i)=\sum_{k=1}^{r}m_{ik}(b_i)\neq 0$ and $F_i(b_i)=1$, $i=1,2$, we have

$$\theta=\frac{f(b_1,b_2)}{m_1(b_1)m_2(b_2)}-1.$$

From the above expression it is clear that the identified value is unique. Then the model is identifiable.

2.2. Cause specific shared frailty proportional RHR model with covariates

Often the lifetime of individual is influenced by age, gender, history or severity of diseases. If information is available about such factors, then the heterogeneity in a population arising from the influence of those factors can be incorporated in models by specifying them as covariates. Accordingly, cause specific RHR function of T_j in presence of covariates and frailty is defined as

$$m_{jk}\left(Z,t_j\right)=Zm_{0jk}\left(t_j\right)\exp(\boldsymbol{\beta}_k{}'\mathbf{x}_j)$$

for j = 1,2 and k = 1,2,…,r, where and $\mathbf{x}_j=\left(x_{j1},x_{j2},...,x_{jp}\right)'$, j = 1,2 is a $p\times1$ vector of covariates and $\boldsymbol{\beta}_k=\left(\beta_{1k},\beta_{2k},...,\beta_{pk}\right)'$ is the vector of regression coefficients and are assumed to be different for different causes of events.

Proceeding as in Section 2.1, bivariate distribution function of $\left(T_1,T_2\right)$ is obtained as

$$F(t_1,t_2)=E\left(\exp\left\{-Z\left(\sum_{k=1}^{r}M_{01k}\left(t_1\right)\exp(\boldsymbol{\beta}_k{}'\mathbf{x}_1)+\sum_{k=1}^{r}M_{02k}\left(t_2\right)\exp(\boldsymbol{\beta}_k{}'\mathbf{x}_2)\right)\right\}\right). \qquad (3)$$

Assuming that frailty variable Z follows gamma distribution with mean one and variance θ, (3) reduces to

$$F\left(t_1,t_2\right)=\left[\theta\sum_{k=1}^{r}M_{01k}\left(t_1\right)\exp(\boldsymbol{\beta}_k{}'\mathbf{x}_1)+\theta\sum_{k=1}^{r}M_{02k}\left(t_2\right)\exp(\boldsymbol{\beta}_k{}'\mathbf{x}_2)+1\right]^{-(1/\theta)} \text{ where } \theta>0. \qquad (4)$$

The marginal distribution function T_j is then obtained as

$$F_j\left(t_j\right)=\left[\theta\sum_{k=1}^{r}M_{0jk}\left(t_j\right)\exp(\boldsymbol{\beta}_k{}'\mathbf{x}_j)+1\right]^{(-1/\theta)},\ j=1,2.$$

When the bivariate distribution function of $\left(T_1,T_2\right)$ is represented in terms of marginal distribution functions, (4) reduces to (2).

The parameters of the model could be estimated from observed data only if the model is identifiable. The identifiable property of the models follows from Sankaran and Gleeja (2011).

3. Estimation

The estimation procedures are developed for cause specific shared gamma frailty proportional RHR model when the data is left censored. Let $T=\left(T_1,T_2\right)$ be the pair of lifetimes

of two related individuals and $U=(U_1,U_2)$ be a pair of corresponding censoring times defined on a common probability space (Ω,F,P) with absolutely continuous distribution function. Under bivariate censoring, one could observe $(T_1^*,T_2^*,C_1,C_2,\delta_1,\delta_2)$ where $T_j^*=\max(T_j,U_j)$ and $\delta_j=I(T_j=T_j^*)$, j=1,2 with $I(.)$ as usual indicator function. Suppose that $(T_{i1}^*,T_{i2}^*,C_{i1},C_{i2},\delta_{i1},\delta_{i2})$, i=1,2,…,n are n independent and identically distributed observations of $(T_1^*,T_2^*,C_1,C_2,\delta_1,\delta_2)$. Define $N_{ijk}(t)=I\{T_{ij}^*\le t,\delta_{ij}=1,C_{ij}=k\}$, $Y_{ijk}(t)=I\{T_{ij}^*\le t,C_{ij}=k\}$, for i=1,2,…,n, j=1,2, k=1,2,…,r. Define the predictable process $\boldsymbol{Y}(\boldsymbol{t})$ as

$$\boldsymbol{Y}(\boldsymbol{t})=\left(Y_{ijk}(t),\ i=1,2,...,n,\ j=1,2,\ k=1,2,...,r\right)$$

and

$$\boldsymbol{N}(\boldsymbol{t})=\left(N_{ijk}(t),\ i=1,2,...,n,\ j=1,2,\ k=1,2,...,r\ \right)$$

as a multivariate counting process with components N_{ijk}, where components with the same value of the first index i share the same frailty variable Z_i. Further it is assumed that conditional on Z, T and U are independent.

3.1. Estimation of cause specific shared frailty proportional RHR model without covariates

For the case without covariates, let the cause specific RHR of T_j for given frailty Z for the ith cluster is $m_{ijk}(t)=Z_iY_{ijk}(t)m_{0jk}(t)$, $i=1,2,...,n,\ j=1,2,\ k=1,2,...,r$, where Z_i is unobservable *i.i.d.* random variable from Gamma $(1/\theta,\ 1/\theta)$ distribution. The estimate of parameter θ and cumulative baseline cause specific RHR function $M_{0jk}(t)=\int_t^b m_{0jk}(s)ds$ is to be obtained. Assume that conditional on $Z=z$, censoring is non-informative. So the partial conditional likelihood based on $\boldsymbol{N}(\boldsymbol{t})$ is given by the product integral

$$dP=\prod_t\left\{\prod_i\prod_j\prod_k\left[m_{ijk}(t)dt\right]^{\Delta N_{ijk}(t)}\left[1-m_{\cdot}(t)dt\right]^{1-\Delta N_{\cdot}(t)}\right\},\qquad(5)$$

where $N_{\cdot}(t)=\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{r}N_{ijk}(t)$ and $m_{\cdot}(t)=\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{r}m_{ijk}(t)$. Considered as a function of Z, (5) is proportional to conditional density of $(\boldsymbol{N}(\boldsymbol{t}),\boldsymbol{Y}(\boldsymbol{t}))$ given $Z=z$. Substituting the specification of $m_{ijk}(t)$ and evaluating the product integral, $L(\theta)$ is obtained as

$$L(\theta)=\prod_i\left\{\frac{z_i^{(1/\theta)-1}\exp\{-z_i/\theta\}}{\theta^{(1/\theta)}\Gamma(1/\theta)}\prod_j\prod_k\exp\left[-z_i\int_0^b Y_{ijk}(s)dM_{0jk}(s)\right]\prod_t\left(z_iY_{ijk}(t)dM_{0jk}(t)\right)^{\Delta N_{ijk}(t)}\right\}\qquad(6)$$

$$=\prod_i\frac{z_i^{(1/\theta)+N_{i\cdot\cdot}(b)-1}\exp\{-z_i(\frac{1}{\theta}+\sum_j\sum_k\int_0^b Y_{ijk}(s)dM_{0jk}(s))\}}{\theta^{(1/\theta)}\Gamma(1/\theta)}\prod_j\prod_k\prod_t\left(Y_{ijk}(t)dM_{0jk}(t)\right)^{\Delta N_{ijk}(t)}$$

where $N_{i..}(b) = \sum_{j=1}^{2}\sum_{k=1}^{r} N_{ijk}(b)$. Conditional on data, Z_i are still independent and gamma distributed with parameters $(1/\theta) + N_{i..}(b)$ and $(1/\theta) + \sum_j \sum_k \int_0^b Y_{ijk}(s)\, dM_{0jk}(s)$. Integrating out Z in (6), the marginal partial likelihood is obtained as

$$L(\theta) = \prod_i \left\{ \frac{\left(\Gamma\left((1/\theta) + N_{i..}(b)\right) \prod_j \prod_k \prod_t \left(Y_{ijk}(t)\, dM_{0jk}(t) \right)^{\Delta N_{ijk}(t)} \right)}{\left(\theta^{(1/\theta)} \Gamma(1/\theta) \left[(1/\theta) + \sum_{j=1}^{2}\sum_{k=1}^{r} \int_0^b Y_{ijk}(s)\, dM_{0jk}(s) \right]^{(1/\theta)+N_{i..}(b)} \right)} \right\}. \tag{7}$$

EM algorithm is used to maximize (7). The estimates of parameters which maximizes (7) maximizes (6) also. The E step is to estimate

$$\hat{z}_i = \frac{(1/\theta) + N_{i..}(b)}{(1/\theta) + \sum_{j=1}^{2}\sum_{k=1}^{r} \int_0^b Y_{ijk}(s)\, dM_{0jk}(s)}.$$

The M step is then to calculate $\hat{\theta}$, the maximum likelihood estimator for θ from (7), and to estimate cumulative baseline cause specific RHR function

$$\hat{M}_{0jk}(t) = \int_t^b \frac{dN_{.jk}(s)}{\sum_i \hat{z}_i Y_{ijk}(s)}, \text{ where } N_{.jk}(s) = \sum_{i=1}^{n} N_{ijk}(s).$$

The initial estimates of $\hat{z}_i$ and $\hat{M}_0(t)$ are obtained by taking $\theta = 0$. By general theory of EM algorithm, if this algorithm converges, it converges to a stationary point of $\log L(\theta)$.

3.2. Estimation of cause specific shared frailty proportional RHR model with covariates

For the model with covariates, let the vector observed be $\left(T_{i1}^{*}, T_{i2}^{*}, C_{i1}, C_{i2}, \delta_{i1}, \delta_{i2}, \mathbf{x}_{i1}, \mathbf{x}_{i2}\right)$. Then cause specific proportional RHR in presence of covariates and frailty is represented as

$$m_{ijk}(t) = z_i Y_{ijk}(t)\, m_{0jk}(t) \exp(\boldsymbol{\beta}_k' \mathbf{x}_{ij}) \tag{8}$$

for j = 1,2 and k = 1,2,…,r, where $\boldsymbol{\beta}_k = (\beta_{1k}, \beta_{2k}, ..., \beta_{pk})'$ is the vector of regression coefficients and $\mathbf{x}_{ij} = (X_{ij1}, X_{ij2}, ..., X_{ijp})'$, j=1,2 is a px1 vector of covariates. Then the likelihood function conditional on the covariate $\mathbf{x}_{ij}$ and frailty Z_i for the model (8) is given as

$$L(\boldsymbol{\beta}, Z_i, \mathbf{x}_{ij}) = \prod_i \left\{ \prod_j \prod_k \prod_t \left(z_i Y_{ijk}(t) \exp(\boldsymbol{\beta}_k' \mathbf{x}_{ij}) dM_{0jk}(t) \right)^{\Delta N_{ijk}(t)} \exp\left[-z_i \exp(\boldsymbol{\beta}_k' \mathbf{x}_{ij}) \int_0^b Y_{ijk}(s) dM_{0jk}(s) \right] \right\}.$$

Let Z_i be unobservable *i.i.d.* random variable from Gamma $(1/\theta, 1/\theta)$ distribution. Then $L(\theta, \boldsymbol{\beta})$ is obtained as

$$L(\theta, \boldsymbol{\beta}) = \prod_i \frac{z_i^{(1/\theta)+N_{i..}(b)-1} \exp\{-z_i (\frac{1}{\theta} + \sum_j \sum_k \exp(\boldsymbol{\beta}_k' \mathbf{x}_{ij}) \int_0^b Y_{ijk}(s) dM_{0jk}(s))\}}{\theta^{(1/\theta)} \Gamma(1/\theta)} \prod_j \prod_k \prod_t \left(\exp(\boldsymbol{\beta}_k' \mathbf{x}_{ij}) Y_{ijk}(t) dM_{0jk}(t) \right)^{\Delta N_{ijk}(t)}.$$

Then loglikelihood function can be written as $\log L(\theta, \boldsymbol{\beta}) = l_1(\theta) + l_2(\boldsymbol{\beta})$ where

$$l_1(\theta) = \sum_i \frac{1}{\theta} \log \frac{1}{\theta} - \log \Gamma(1/\theta) + \left[(1/\theta) + N_{i..}(b) - 1 \right] \log z_i - \frac{z_i}{\theta}, \tag{9}$$

and

$$l_2(\boldsymbol{\beta}) = \sum_i (-z_i) \sum_j \sum_k \exp(\boldsymbol{\beta}_k' \mathbf{x}_{ij}) \int_0^b Y_{ijk}(s) dM_{0jk}(s) + \log \left[\prod_j \prod_k \prod_t \left(\exp(\boldsymbol{\beta}_k' \mathbf{x}_{ij}) Y_{ijk}(t) dM_{0jk}(t) \right)^{\Delta N_{ijk}(t)} \right]. \tag{10}$$

Proceeding as in the case without covariates and using EM algorithm, estimate of Z_i is obtained as

$$\hat{z}_i = \frac{(1/\theta) + N_{i..}(b)}{(1/\theta) + \sum_{j=1}^{2} \sum_{k=1}^{r} \exp(\boldsymbol{\beta}_k' \mathbf{x}_{ij}) \int_0^b Y_{ijk}(s) dM_{0jk}(s)}.$$

The M step is then to calculate $\hat{\theta}$ and $\hat{\boldsymbol{\beta}}_k$, the maximum likelihood estimator for θ and $\boldsymbol{\beta}_k$ from (9) and (10) respectively and to estimate cumulative baseline cause specific RHR function

$$\hat{M}_{0jk}(t) = \int_t^b \frac{dN_{.jk}(s)}{\sum_i \hat{z}_i \exp(\boldsymbol{\beta}_k' \mathbf{x}_{ij}) Y_{ijk}(s)}.$$

For testing of independence of variables T_1 and T_2, likelihood ratio test can be used. The case of shared frailty model with no covariates and with standard conditions on the censoring distribution is discussed in Maller and Zhou (2003). They obtained that the likelihood ratio statistic has an asymptotic null distribution which is an equal mixture of a point mass at zero and a chi-square distribution with one degree of freedom. For testing $H_0 : \theta = 0$, the likelihood ratio test statistic is -2log Q = 2 (log $L(\hat{\theta})$-log $L(0)$). When -2log $Q > 0.5(\chi^2_{0,\alpha} + \chi^2_{1,\alpha})$, the null hypothesis is rejected at 5% level of significance.

The asymptotic properties of the estimators follow from Sankaran and Gleeja (2011). The consistency of the estimators is established in Theorem 1.

Let T_1 be the first jump of N, θ_0 lies in a known interval $[0, S]$ and true cumulative baseline reversed hazard M_{0jk} be strictly decreasing and continuous on $[0, b]$ for $b<\infty$.

Theorem 1: Assume that

i. Y is a non-decreasing step function and $P\left(Y\left(t\right)\geq 1\right)$ has at most finite number of discontinuities in $t\in\left(0,b\right)$,

ii. $\underset{u\in(0,b)}{Inf}\ E\left(Y\left(u\right)\right)>0$,

iii. $P\left(Y\left(T_1\right)\geq 1\right)<1$,

then $\underset{t\in(0,b)}{Sup}\left|\hat{M}_{0jk}\left(t\right)-M_{0jk}\left(t\right)\right|\to 0$ almost surely (a. s.) and $\left|\hat{\theta}-\theta_0\right|\to 0$ a. s.

Proof: The assumption (i) is used to prove that $\hat{M}_{0jk}\left(t\right)$ does not diverge to infinity, (ii) is used to ensure that counting process N has sufficient activity on the entire interval so as to estimate the parameters, and (iii) excludes the possibility of N having at most only one jump. The model becomes unidentifiable if all N_i have only one jump. The rest of the proof is similar to the one given in Murphy (1994).

The asymptotic normality of the estimators can be established in the following way. Set $M_{0jkt}\left(0\right)=\int_0^b 1+th_1\left(u\right)d\hat{M}_{0jk}\left(u\right)$ and $\theta_t=th_2+\hat{\theta}$ for h_1 a function and h_2 a scalar, and differentiate at $t=0$ to get $F_n\left(\hat{M}_{0jk},\hat{\theta}\right)\left(h_1,h_2\right)$. Then, if $\left(\hat{M}_{0jk},\hat{\theta}\right)$ maximizes $\log L\left(\theta\right)$, then $F_n\left(\hat{M}_{0jk},\hat{\theta}\right)\left(h_1,h_2\right)=0$ for all $\left(h_1,h_2\right)$. The form of F_n is given by $F_n=F_{n1}+F_{n2}$, where

$$F_{n1}\left(M_{0jk},\theta\right)\left(h_1\right)=n^{-1}\sum_{i=1}^{n}\int_0^b h_1dN_{ijk}-\frac{\theta^{-1}+N_{ijk}\left(b\right)}{\theta^{-1}+\int_0^b Y_{ijk}dM_{0jk}}\int_0^b h_1Y_{ijk}dM_{0jk}$$

and $$F_{n2}\left(M_{0jk},\theta\right)\left(h_2\right)=h_2n^{-1}\sum_{i=1}^{n}\int_0^b\frac{N_{ijk}\left(u\right)}{1+\theta N_{ijk}\left(u\right)}dN_{ijk}\left(u\right)$$

$$+\theta^{-2}\left(\log\left(1+\theta\int_0^b Y_{ijk}dM_{0jk}\right)-\frac{\theta^{-1}+N_{ijk}\left(b\right)}{\theta^{-1}+\int_0^b Y_{ijk}dM_{0jk}}\theta\int_0^b h_1Y_{ijk}dM_{0jk}\right)$$

For $\theta=0$, the last term is taken as its limit as θ approaches zero to get $\left(\left(\int_0^b Y_{ijk}dM_{0jk}\right)^2\Big/2\right)+N_{ijk}\left(b\right)\int_0^b Y_{ijk}dM_{0jk}$. The class of h is taken to be the space of bounded variation cross the reals. Define the norm to be $\left\|h\right\|_H=\left\|h_1\right\|_v+\left|h_2\right|$, where $\left\|h_1\right\|_v$ is absolute value of $h_1\left(0\right)$ plus the total variation of h_1 on the interval $\left[0,b\right]$. Define H_p to be the product space

of bounded variation functions on $[0,b]$ and real valued scalars with norm $\|h\|_H = \|h_1\|_v + |h_2| \le p$. If $p = \infty$, then the inequality is strict. In the following p is assumed to be finite unless stated otherwise. Define $\left(M_{0jk}, \theta\right)(h) = \int_0^b h_1 dM_{0jk} + h_2\theta$. Then the parameter space Ψ can be considered to be a subset of $l^\infty\left(H_p\right)$, which is the space bounded by real valued functions on H_p under the supremum norm $\|U\| = \sup_{h \in H_p} |U(h)|$. The score function F_n is a random map from Ψ to $l^\infty\left(H_p\right)$ for all finite p.

Theorem 2: Assume that

i. $\underset{t\in(0,b)}{Sup}\left|\hat{M}_{0jk}(t) - M_{0jk}(t)\right| \to 0$ a.s. and $\left|\hat{\theta} - \theta_0\right| \to 0$ a. s.,

ii. There exist a constant K for which $\|Y\| \le K$ and $N(b) \le K$ a.s. ,

iii. $\underset{u\in(0,b)}{Inf}\ E\left(Y(u)\right) > 0$,

iv. $P\left(Y(T_1) \ge 1\right) < 1$,

then $\sqrt{n}\left|\hat{M}_{0jk}(t) - M_{0jk}(t)\right| . \sqrt{n}\left|\hat{\theta} - \theta_0\right| \Rightarrow \mathrm{G}$ on $l^\infty\left(H_p\right)$; G is a tight Gaussian process on $l^\infty\left(H_p\right)$ with mean zero and covariance process $Cov\left(G(h), G(h')\right) = \int_0^b h_1 \tilde{\sigma}_{(1)}^{-1}(h') dM_{0jk} + h_2 \tilde{\sigma}_{(2)}^{-1}(h')$, where $\tilde{\sigma} = \left(\tilde{\sigma}_1, \tilde{\sigma}_2\right)$ is a continuously invertible linear operator from H_∞ onto H_∞ with inverse $\tilde{\sigma}^{-1} = \left(\tilde{\sigma}_{(1)}{}^{-1}, \tilde{\sigma}_{(2)}{}^{-1}\right)$. The form of $\tilde{\sigma}$ is as follows:

$$\tilde{\sigma}_1(h)(u) = h_1(u) E\left(ZY(u)\right) - \sum_{j=1}^{2}\sum_{k=1}^{r} E\left(\frac{\theta_0 \int_0^b Y_{jk} h_1 dM_{0jk}}{1 + \theta_0 \int_0^b Y_{jk} dM_{0jk}} ZY_{jk}(u)\right)$$

$$- h_2 E\left(\frac{Y(u)}{1 + \theta_0 \int_0^b Y dM_{0jk}} \left(\int_0^b ZY dM_{0jk} - N(b)\right)\right)$$

and

$$\tilde{\sigma}_2(h) = h_2 \sum_{j=1}^{2}\sum_{k=1}^{r} E\left(-\frac{\partial^2 \log L(\theta, M_{0jk})}{\partial(\theta)^2}\Big|_{(\theta_0, M_{0jk})}\right) - \sum_{j=1}^{2}\sum_{k=1}^{r} E\left(\frac{\int_0^b Y h_1 dM_{0jk}}{1 + \theta_0 \int_0^b Y dM_{0jk}} \left(\int_0^b ZY dM_{0jk} - N(b)\right)\right)$$

where

$$-\frac{\partial^2 \log L(\theta, M_{0jk})}{\partial(\theta)^2}\Big|_{(\theta_0, M_{0jk})} = n^{-1}\sum_{i=1}^{n}\int_0^b \left(\frac{N_{ijk}(u)}{1+\theta_0 N_{ijk}(u)}\right)^2 dN_{ijk}(u) - N_{ijk}(b)\left(\frac{\int_0^b Y_{ijk}\,dM_{0jk}}{1+\theta_0\int_0^b Y_{ijk}\,dM_{0jk}}\right)^2$$

$$+2\theta_0^{-3}\left(\ln\left(1+\theta_0\int_0^b Y_{ijk}\,dM_{0jk}\right) - \frac{\theta_0\int_0^b dM_{0jk}}{1+\theta_0\int_0^b Y_{ijk}\,dM_{0jk}} - \frac{1}{2}\left(\frac{\theta_0\int_0^b dM_{0jk}}{1+\theta_0\int_0^b Y_{ijk}\,dM_{0jk}}\right)^2\right).$$

When $\theta_0 = 0$, the last term above is defined by its limit, which is $\frac{2}{3}\sum_{j=1}^{2}\sum_{k=1}^{r}\left(\int_0^b Y_{ijk}\,dM_{0jk}\right)^3$.

Proof: Proof of the theorem follows from Murphy (1995).

4. Data Analysis

The analysis of the proposed model is illustrated with data concerning the times to tumor appearance or death of mice from the same litter in a tumor genesis experiment by Mantel and Ciminera (1979), reported in Ying and Wei (1994). The observations from drug treated rat (T_1) and its litter matched control (T_2) which were either dead ($C_i = 1, i = 1,2$) or appeared with tumor ($C_i = 0, i = 1,2$) are considered for the analysis. The observations with the value 60 indicate left censored observations. The indicator function $\delta_i = 0,\ i = 1,2$ if observation is left censored, $\delta_i = 1,\ i = 1,2$ if it is not censored. The data consist of 22 pairs. The analysis of data is carried out with cause specific shared frailty proportional RHR model without covariates. Then the maximum likelihood estimate for θ is obtained as 0.7557 and is significantly greater than zero (p<0.001) using likelihood ratio test. Hence the pairs are not independent. The value of frailty variable estimated is given in Table 1.

Table 1: Estimates of frailty variable for mice in the same litter

Drug (T_1)	Control (T_2)	δ_1	δ_2	C_1	C_2	Z_i	Drug (T_1)	Control (T_2)	δ_1	δ_2	C_1	C_2	Z_i
60	60	0	0	0	0	0.1558	77	79	1	1	1	1	0.9270
60	60	0	0	0	0	0.1558	89	90	1	1	0	1	1.2745
81	69	1	1	1	1	0.3670	102	80	1	1	0	1	1.3389
60	77	0	1	1	1	0.4267	86	94	1	1	0	1	1.3499
67	68	1	1	0	0	0.4560	104	77	1	1	0	0	1.3830
80	73	1	1	0	1	0.4686	103	91	1	1	0	1	1.6442
76	74	1	1	1	1	0.4959	92	102	1	1	0	0	1.6924
73	66	1	1	0	0	0.5160	88	99	1	1	1	1	1.7022
70	77	1	1	0	1	0.5600	91	92	1	1	1	1	1.7225
80	76	1	1	0	1	0.6460	103	84	1	1	0	0	1.7461
76	78	1	1	1	0	0.7655	93	103	1	1	1	1	2.2075

If realization of Z is less than one, then members of the group tend to experience the event earlier. Hence most fragile ones have values less than one. The value of Z increases with increase in lifetime. The estimates of cumulative baseline cause specific reversed hazard function for time to event for drug treated rat and its litter matched control is shown in Figure1 and Figure 2. It can be seen that in drug treated rat, based on cumulative baseline reversed hazard function, tumor appearance is more than death without tumor. But in case of control, based on cumulative baseline reversed hazard function, tumor appearance is less than death without tumor.

Figure 1: Cumulative baseline cause specific RHR for time to event of drug treated rats

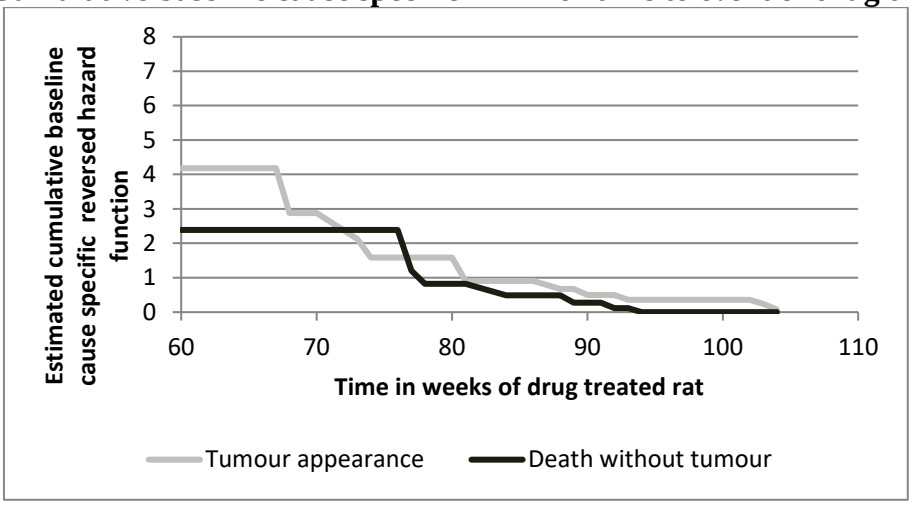
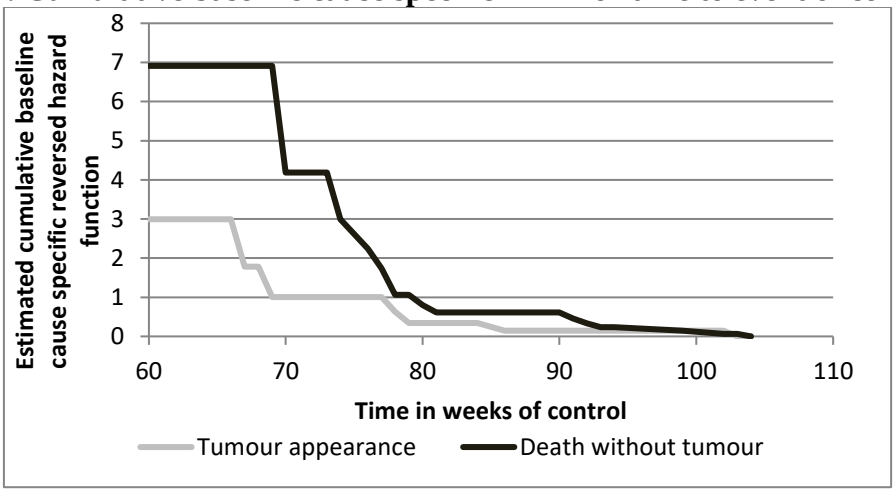


Figure 2: Cumulative baseline cause specific RHR for time to event of control rats



For illustration of the model with covariates, excerpt of the bivariate data set given in McGilchrist and Aisbett (1991), is being used. This data shows the infection times at the point of insertion of catheter for kidney patients using portable dialysis equipment. The observations with value 10 indicate left censored observations. Data for the first two occurrences of infection are given. Let T_1 and T_2 represents the first and second occurrences of infection. Disease types, glomerulo nephritis = (0), acute nephritis = (1), polycystic kidney disease = (2), others = (3), are treated as four different causes for infection. Let C_1 is the variable denoting cause for first occurrence and C_2 is the variable denoting cause for second occurrence. It takes value 0, 1, 2, or 3, depending on the disease type causing infection. Gender is considered as the

covariate for the study, 1=male and 0=female. In the model, it is assumed that regression coefficients associated with covariate gender are different for different types of diseases. Let β_1, β_2, β_3 and β_4 denote the regression coefficients associated with covariate for different disease types glomerulo nephritis, acute nephritis, polycystic kidney disease and others respectively. Estimates of parameters of the model are given in Table 2. Estimates are significant (p<0.001) using likelihood ratio test. Being a male increase the risk of getting infected at earlier time compared to that of female for all disease types. Males with disease type glomerulo nephritis and polycystic kidney disease are more prone to infection. As θ is significant, pairs are not independent. The value of frailty variable estimated is given in Table 3.

Table 2: Estimates of parameters

Parameter	θ	β_1	β_2	β_3	β_4
Estimate	0.0069	-1.1952	-0.4877	-1.1952	-0.9258
Standard error	0.0012	0.0661	0.0485	0.0661	0.0655

Table 3: Estimates of frailty variable for kidney patients

T_1	T_2	δ_1	δ_2	X_1	C_1	C_2	Z
10	16	0	1	1	3	3	0.9929
22	28	1	1	1	3	3	1.0037
447	318	1	1	0	3	3	1.0119
30	12	1	1	1	3	3	1.0024
24	245	1	1	0	3	3	1.0017
511	30	1	1	1	0	0	1.0046
53	196	1	1	0	1	1	1.0091
15	154	1	1	0	0	0	1.0071
10	333	0	1	1	1	1	0.9917
96	38	1	1	0	1	1	1.0002
185	177	1	1	0	3	3	1.0070
292	114	1	1	0	3	3	1.0066
15	108	1	1	0	3	3	0.9871
152	562	1	1	0	2	2	1.0090
13	66	1	1	1	1	1	1.0010
12	40	1	1	0	1	1	0.9996
132	156	1	1	1	0	0	1.0077
34	30	1	1	0	1	1	0.9899
10	25	0	1	0	0	0	0.9914
130	26	1	1	1	0	0	0.9928
27	58	1	1	0	1	1	1.0020
152	30	1	1	0	2	2	0.9909
119	10	1	0	0	3	3	0.9811

If realization of Z is less than one, then members of the group tend to experience the event earlier. Hence most fragile ones have values less than one. The value of Z increases with increase in lifetime. The estimates of cumulative baseline cause specific reversed hazard function for time to first and second occurrence of infection are presented in Figure 3 and Figure 4.

Figure 3: Cumulative baseline cause specific RHR for time to first occurrence of infection

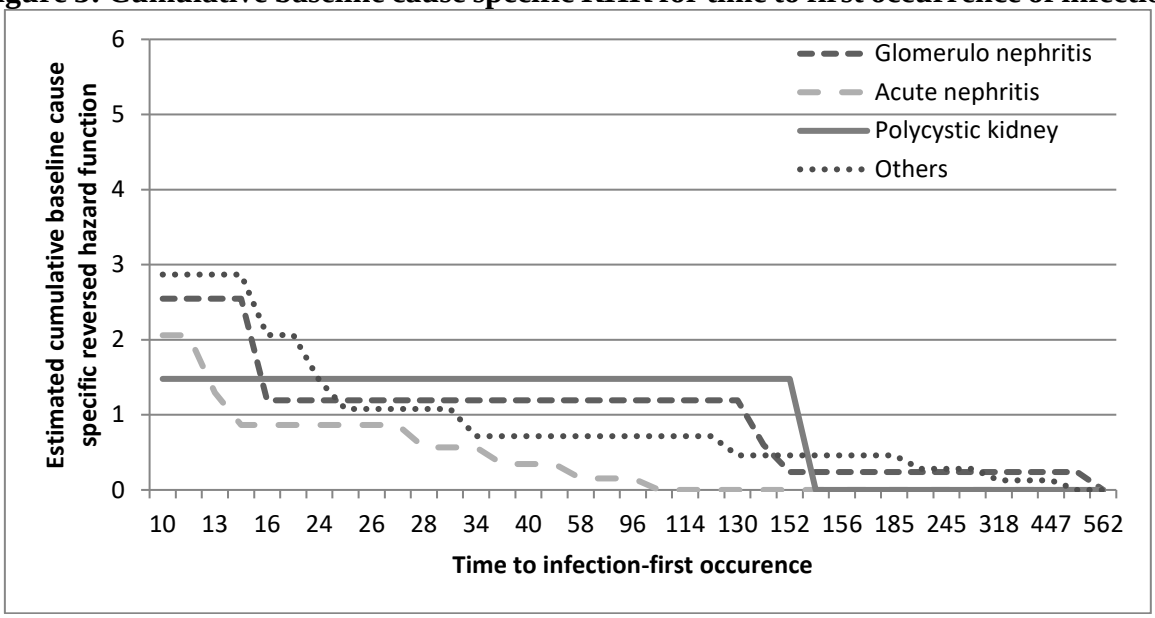
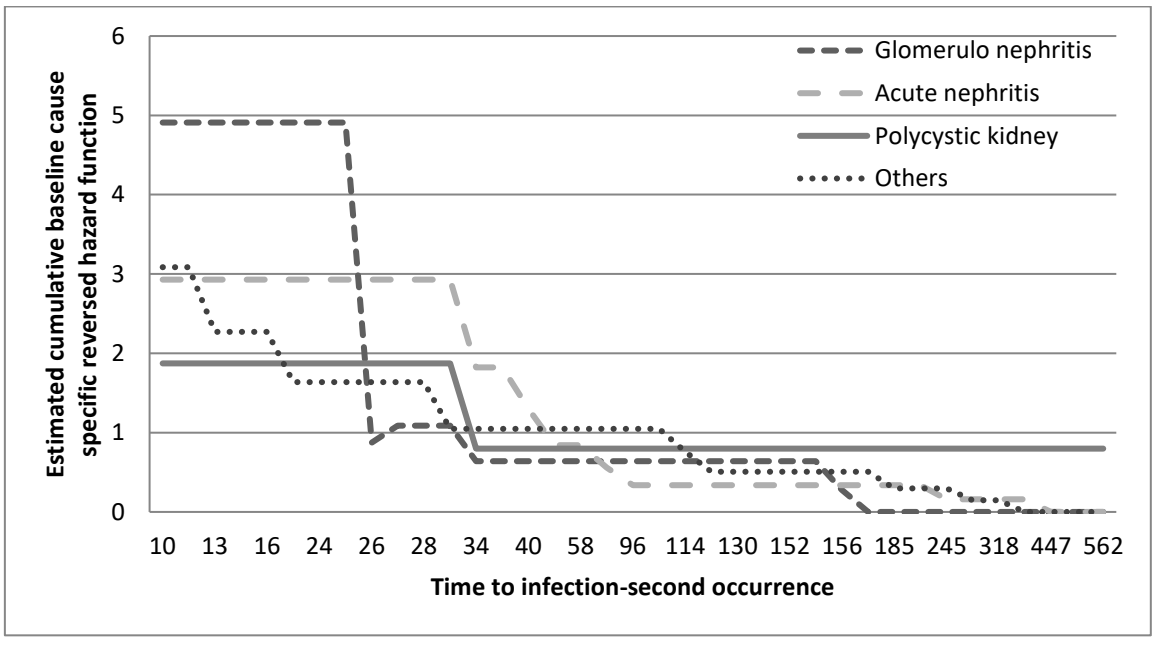


Figure 4: Cumulative baseline cause specific RHR for time to second occurrence of infection



5. Conclusion and Discussion

In this paper, a shared gamma frailty model in terms of cause specific RHR has been introduced for the analysis of competing risks data under left censoring. The gamma distribution with mean one and variance θ is chosen as distribution of the frailty random

variable. The model is discussed with and without the presence of covariates. The parameters of the models were estimated by maximum likelihood method, using EM algorithm, and discussed the properties of the estimators. The proposed models were applied to real life data sets. The data in Mantel and Ciminera (1979) was analyzed for checking the adequacy of gamma frailty distribution with marginal proportional hazard model by Cui and Sun (2004). They obtained the estimate of parameter of gamma distribution as 0.888 which is very close to the value obtained by the present model. Existence of strong association and dependency in litter matched pairs is reported in Anisha (2012). The present model helps to quantify the strength of association in litter matched pairs for left censored data. The data in McGilchrist and Aisbett (1991) were analysed by several authors. The report on analysis by Hanagal and Dabade (2013) and Hanagal (2020) depicts gender as the significant covariate and observes that females are at lower risk. The same result holds for the present model and different regression coefficients are estimated for different causes in the present model. It was observed that more fragile individuals are having realization of frailty variable as less than one and those who experience the event of interest at a later stage are having the value greater than one. The models discussed in Anisha (2012), Cui and Sun (2004) and Hanagal and Dabade (2013) were able to consider only right censored or complete observations. Those models were not dealing with left censored data. So, in order to analyse a left censored family data with multiple type of diseases, shared gamma frailty model in terms of cause specific reversed hazard rates is more appropriate and is recommended.

References

Andersen, P. K., Borgan O., Gill, R. D., and Keilding, N. (1993). *Statistical Models Based on Counting Processes.* Berlin, Springer- Verlag, New York.

Anisha, P. (2012) *Modelling and Analysis of Recurrent Event Data With Multiple Causes*. Ph.D. Thesis. Cochin University of Science and Technology, Kerala, India.

Barlow, R. E., Marshall, A. W., and Proshan, F. (1963). Properties of probability distributions with monotone hazard rate. *The Annals of Mathematical Statistics*, **34**, 375-389.

Bartoszewicz , J., and Skolimowska, M. (2006). Preservation of classes of life distributions and stochastic orders under weighting. *Statistics and Probability Letters*, **76**, 587-596.

Block, H. W, Savits, T. H., and Singh, H. (1998). The reversed hazards rate function. *Probability in Engineering and Informational Sciences*, **12**, 69-90.

Chandra, N. K., and Roy, D. (2001). Some results on reversed hazard rate. *Probability in Engineering and Informational Sciences,* **15**, 95-102.

Clayton, D., and Cuzick, J. (1985). Multivariate generalizations of the proportional hazards model. *Journal of the Royal Statistical Society, Series A*, **148** (**2**), 82-117.

Crowder, M. J. (2001). *Classical Competing Risks*. CRC Press, London.

Cui S., and Sun Y. (2004). Checking for the gamma frailty distribution under the marginal proportional hazards frailty model. *Statistica Sinica,* **14**, 249-267.

Faith K. (2017). The E-Bayesian and hierarchical Bayesian estimations for the proportional reversed hazard rate model based on record values. *Journal of Statistical Computation and Simulation,* **87**(**11**), 2253-2273.

Gupta, R. D and Nanda, A. K. (2001). Some results on reversed hazard rate ordering. *Communications in Statistics - Theory and Methods*, **30**, 2447-2457.

Gupta, R. C., and Wu, H. (2001). Analyzing survival data by proportional reversed hazard model. *International Journal of Reliability and Applications,* **2**(**1**), 1-26.

Hanagal, D. D. (2020) Analysis of kidney infection data using correlated Inverse Gaussian frailty model. *Statistics and Applications*, **18**(**1**), 1-19.

Hanagal, D. D., Arvind, P., and Sankaran, P. G. (2014). Shared frailty model based on reversed hazard rate for left censored data. *Communications in Statistics - Simulation and Computation,* **46**(**1**), 230-243.

Hanagal, D., and Bhambure, S. M. (2017). Shared gamma frailty models based on reversed hazard rate for modeling Australian twin data. *Communications in Statistics - Theory and Methods,* **46**(**12**), 5812-5826.

Hanagal, D. D., and Dabade, A. D. (2013). Compound negative binomial shared frailty models for bivariate survival data. *Statistics and Probability Letters*, **83**, 2507–2515.

Hanagal, D. D., and Pandey, A. (2015). Inverse Gaussian Shared Frailty Models with Generalized Exponential and Generalized Inverted Exponential as baseline distributions. *Journal of Data Science,* **13**, 569-602.

Hanagal, D. D., and Pandey, A. (2017). Correlated gamma frailty models for bivariate survival data based on reversed hazard rate. *International Journal of Data Science,* **2**(**4**)**,** 301-324.

Hanagal, D. D., and Pandey, A. (2020). Correlated inverse Gaussian frailty models for bivariate survival data. *Communications in Statistics - Theory and Methods*, **49**(**4**), 845–863.

Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. Springer-Verlag, New York.

Kalbfleisch, J. D., and Lawless, J. F. (1989). Inference based on retrospective ascertainment: an analysis of the data based on transfusion-related AIDS. *Journal of the American Statistical Association,* **84**, 360-372.

Kalbfleisch, J. D., and Prentice, R. L. (2002). *The Statistical Analysis of Failure time data*. John Wiley and Sons, New York,

Klein, J. P. (1992). Semi parametric estimation of random effects using Cox model based on EM algorithm. *Biometrics*, **48**, 795-806.

Lawless J. F. (2003). *Statistical Models and Methods for Lifetime data*. Wiley, New York,

Maller, R. A., and Zhou X. (2003). Testing for individual heterogeneity in parametric models for event history data. *Mathematical Methods of Statistics*, **12**, 276–304.

Mantel, N., and Ciminera, J. L. (1979). Use of logrank scores in the analysis of litter-matched data on time to tumor appearance. *Cancer Research*, **39**, 4308-4315.

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

Murphy, S. A. (1994). Consistency in a proportional hazards model incorporating a random effect. *The Annals of Statistics*, **22** (**2**), 712-731.

Murphy, S. A. (1995). Asymptotic theory for the frailty model. *The Annals of Statistics*, **23**(**1**), 182-198.

Nair, N. U., Sankaran, P. G., and Asha, G. (2005). Characterizations of distributions using reliability concepts. *Journal of Applied Statistical Science*, **14,** 237-241.

Nielsen, G. G., Gill, R. D., Andersen, P. K., and Sorensen, T. I. A. (1992). A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian Journal of Statistics*, **19**, 25-43.

Sankaran, P. G., and Anjana, S. (2014). A class of tests for the equality of cause- specific reversed hazard rates in competing risks models. *Journal of the Indian Statistical Association*, **52**, 161-176.

Sankaran, P. G., and Anjana, S. (2016). Nonparametric estimation of cumulative cause specific reversed hazard rates under masked causes of failure. *Journal of Biostatistics and Biometric Applications,* **1**(**2**), 201-211.

Sankaran, P. G., and Gleeja, V. L. (2006). On bivariate reversed hazard rates. *Journal of Japan Statistical Society,* **36**(**2**), 213-224.

Sankaran, P. G., and Gleeja, V. L. (2008). Proportional reversed hazard and frailty models. *Metrika*, **68**, 333–342.

Sankaran, P.G., and Gleeja, V. L. (2011). On proportional reversed hazards frailty models. *Metron*, **69**, 151-173.

Shaked, M., and Shantikumar, J. G. (1994). *Stochastic Orders and Their Applications.* Academic Press, New York.

Vaupel, J. W., Manton, K. G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, **16** (**3**), 439-454.

Ware, J. H., and DeMets, D. L. (1976). Reanalysis of some baboon descent data. *Biometrics*, **32**, 459-463.

Ying, Z., and Wei, L. J. (1994). The Kaplan-Meier estimator for dependent failure time observations. *Journal of Multivariate Analysis*, **50**,17-29.