Joint Frailty Mixing Model for Recurrent Event Data with an Associated Terminal Event: Application to Hospital Readmission Data

Goutam Barman[1], Babulal Seal[2], Shreya Bhunia[2] and Proloy Banerjee[2,*]

[1]*Department of Statistics, Krishnagar Government College, West Bengal, India*

[2]*Department of Mathematics and Statistics, Aliah University, West Bengal, India*

Abstract: Recurrent events like repeated hospitalization, cancer tumour recurrences, and many others occur frequently. The follow-up on recurrent events may be stopped by a terminal event like death. It is obvious that if the frequencies of recurrent events are more, then it may lead to a terminal event and in this case terminal event becomes 'dependent'. In this article, we study a joint modelling and analysis of recurrent events with a dependent terminal event. Here, the proportional intensity model for the recurrent events process and the proportional hazard model for the terminal event time are taken. To account for the association between recurrent events and terminal events, mixing frailty or random effect is studied rather than available pure frailty. In our case, the distribution of frailty is introduced as a mixture of folded normal distribution and gamma distribution rather than using pure gamma distribution. An estimation procedure in the joint frailty model is applied to estimate the parameters of the model. This method is close to the method of minimum chi-square rather than a complicated one. An extensive simulation study has been performed to estimate the model parameters and the performances are evaluated based on bias and MSE criteria. Further from an application point of view, the method is illustrated to a hospital readmission data for colorectal cancer patients.

Keywords: Frailty, Proportional hazard model, Proportional intensity model, Mixture distribution, Recurrent events.

1. INTRODUCTION

Sometimes the event of interest per subject can occur more than once and such outcomes have been termed as recurrent events. Examples include cancer tumour recurrences, repeated drug use, repeated hospitalization, and many others. Various methods based on modelling the intensity or rate functions have been considered for the analysis of recurrent event data. Prentice *et al*. [1] studied the regression analysis of multivariate failure time data when there are a fairly large number of study subjects. In the context of a single failure time variable, Pepe and Cai [2] suggested some rate functions when analysing recurrent failure time data or when the effect of a categorical time-dependent covariate is of interest. Based on the Nelson's method for estimating the cumulative mean function for identically distributed processes of recurrent events, Lawless and Nadeau [3] suggested a similar method with more general models, including regression. A class of mixed models for recurrent event data was proposed by Sun *et al*. [4].

However, the recurrent events and the follow-up for a particular subject may be stopped by terminal events like death. For example, patients may experience cancer tumour recurrences which are terminated by death. Usually, this terminal event is expected to be related and also may be strongly related to the recurrent events of interest, and that is to be accounted for in the analysis. In the last few years, joint analysis of recurrent events with informative terminal events has become more popular. For more details, one may see [5-8] etc.

The existing methods for the analysis of recurrent event data in the presence of a terminal event are generally classified into two approaches: frailty methods and marginal methods. Frailties or random effects are used in the frailty method to account for the relation between recurrent and terminal events [9]. Huang and Wang [10] provided a shared frailty model with proportional intensity for recurrent events and proportional hazards for terminal events. Ye *et al*. [11] discussed a semiparametric method to jointly model the recurrent and terminal event processes, incorporating shared gamma frailty in both the recurrent event rate and terminal event hazard function to account for their interdependence. In the marginal method, focus is given on the marginal rates of the recurrent and terminal events and leaving their correlation unspecified [12-14]. Frailty models are seen to be the extensions of the Cox proportional hazards model [15] and can be used to analyse such data and provide explicit measures about the dependency between the events [16-17]. While determining the relationship between the time of occurrence and one of the independent variables, the Cox proportional hazards model is used [29]. Most of the time, the proportional hazard model is used for recurrent events. It is known that when proportional hazard assumptions are not met or violated then the proportional hazards model may not fit survival data well and, in this case, the additive hazard model is one such alternative [18].

In many applications, Monte Carlo Expectation-Maximization (EM) algorithm is issued to

*Address correspondence to this author at the Department of Mathematics and Statistics, Aliah University, West Bengal, India; E-mail: proloy.stat@gmail.com

E-ISSN: 1929-6029/23

estimate the hazard functions and model parameters [19-20]. Simulation by this algorithm takes much time and also, we cannot directly estimate the smooth hazard function. Joly *et al*. [21] introduced a penalized likelihood approach to estimate the model parameters of arbitrarily censored and truncated data. Later on, Rondeau *et al*. [22] used a semiparametric estimation procedure by using penalized likelihood to estimate the parameter of the joint model. In this approach robust Marquardt optimization algorithm which is a combination of the Newton-Raphson algorithm and the steepest decent algorithm, was used to estimate the parameters [23].

In this article, we propose a joint model for recurrent events and terminal events by a subject-specific common frailty. The frailty is in our cases taken as a mixture of a folded normal and gamma distribution. Some earlier authors including Mazroui *et al*. [24], Toenges, and Jahn-Eimermacher [25] worked with gamma frailty. It is reasonable as it generalises the basic exponential life distribution, which is frequently helpful for simulating positive data [26]. We try to generalise the case of concentrated distribution and also skewed distribution. That is why, here a mixture of folded normal and gamma has been taken in a more general form. Also, the normal distribution has been folded to make things positive. Similar to Liu *et al*. [19], the proportional intensity model is used for recurrent event processes, and for modelling terminal event time, we use the proportional hazard model. A general estimation procedure in the joint frailty model is applied to estimate the parameters of the model. Based on the numerical results, we observe that the MSEs of all the estimators are very small and these are reduced with sample sizes. So, our proposed method performs reasonably well. Here our interest is to see whether a simpler method may be tried in such cases and also to generalize the frailty structure.

The article is organized in the following manner. In Section 2, we describe the joint frailty models and explain the estimating procedure of the model parameters in Section 3. Some results from the simulation study are reported in Section 4 and in Section 5, the method is applied to hospital readmission data for colorectal cancer patients, and summaries in this respect are given. Finally, a concluding discussion is presented in Section 6. The construction of full log-likelihoods, mathematical derivations, and parameter estimations are included in the Appendix.

2. JOINT FRAILTY MODELFOR RECURRENT EVENTS AND A TERMINAL EVENT

We denote X_{ij} as the $j^{th}(j = 1, 2, \dots\dots, n_i)$ recurrent event time for the individual i $(i = 1,2, \dots\dots, N)$, C_i as the right-censoring time and D_i as the death time. Each follow-up time or event time for the individual i is denoted by $T_{ij} = \min(X_{ij}, C_i, D_i)$ and also we denote the last follow-up time for the individual i by $T_i^* = \min(C_i, D_i)$. We define the recurrent events indicator as $\delta_{ij} = 0$ when either $T_{ij} = C_i$ or D_i and $\delta_{ij} = 1$ when $T_{ij} = X_{ij}$ and death indicator as $\delta_i^* = 0$ when $T_i^* = C_i$ and $\delta_i^* = 1$ when $T_i^* = D_i$.

Let $N_i^{R^*}(t) =$ Number of recurrent events for i^{th} individual over the interval $(0, t]$. We observe the process $N_i^R(t) = N_i^{R^*}(\min(T_i^*, t))$ which counts the observed number of recurrent events. Similarly, we denote the actual death indicator by $N_i^{D^*}(t) = I_{(D_i \le t)}$ and observed death indicator by $N_i^D(t) = I_{(T_i^* \le t, \delta_i^* = 1)}$. Furthermore, let $Y_i(t) = I_{(T_i^* \ge t)}$ denote whether or not the individual i is at risk at time t. The number of recurrent events that occur over the interval $[t, t + dt)$ for the individual i is $dN_i^{R^*}(t) = N_i^{R^*}((t + dt)^-) - N_i^{R^*}(t^-)$ and we have $dN_i^R(t) = Y_i(t) dN_i^{R^*}(t)$. The process history of i^{th} individual up to time t is

$$H_{it} = \sigma\{Y_i(k), N_i^R(k), N_i^D(k), Z_i(k); 0 \le k \le t\}, i = 1, 2, \dots\dots, N,$$

where $Z_i(k)$ is a vector of covariates. We denote the following σ fields

$$\mathcal{F}_{it} = \sigma\{H_{it}, \omega_i\}, i = 1, 2, \dots\dots, N.$$

The random effect ω_i links the recurrent event intensity process and the terminal event intensity process for i^{th} individual. We assume that recurrent, terminating, and censoring processes are continuous. We consider that death happens first in the small interval $[t, t + dt)$. The observation of new recurrent events precludes death but censoring for end of study or loss of follow-up, does not interrupt the occurrence of new recurrent events.

The recurrent event intensity process at time t is expressed from the above σ fields as

$Y_i(t) r_i(t) dt = P(dN_i^R(t) = 1 | \mathcal{F}_{it^-})$, where $r_i(t)dt = P(dN_i^{R^*}(t) = 1 | Z_i(t), \omega_i, t \le D_i)$ in general form

and the terminal intensity process at time t is

$Y_i(t) \lambda_i(t) dt = P(dN_i^D(t) = 1 | \mathcal{F}_{it^-})$, where $\lambda_i(t)dt = P(dN_i^{D^*}(t) = 1 | Z_i(t), \omega_i, t \le D_i)$.

Now the above general form is studied through hazard function as in the following. Similar to Liu *et al*. [19] and Peng *et al*. [27], the joint model for the hazard functions for recurrent events and terminal events are respectively as follows in time scale:

$$\begin{cases} r_i(t|\omega_i) = \omega_i r_0(t) \exp(\beta_1 Z_i(t)) \text{ (recurrent)} \\ \lambda_i(t|\omega_i) = \omega_i^{\alpha} \lambda_0(t) \exp(\beta_2 Z_i(t)) \text{ (terminal)} \end{cases} \quad (1)$$

and

$$\begin{cases} r_i(t|\omega_i) = \omega_i r_0(t)\exp\big(\beta_1 Z_i(t)\big) \text{ (recurrent)} \\ \lambda_i(t|\omega_i) = \omega_i^{\alpha}\{\lambda_0(t) + \beta_2 Z_i(t)\} \text{ (terminal)} \end{cases} \quad (2)$$

where $r_0(t)\big(\lambda_0(t)\big)$ is the baseline hazard function for recurrent (terminal) events and $\beta_1(\beta_2)$ is the regression coefficient associated with the covariate $Z_i(t)$. The random effect ω_i (frailties) associate the recurrent event process with the terminal event for i^{th} individual and can be generated from a mixture of folded normal distribution and gamma distribution. For simulation purposes, we are taking folded normal with mean 0 and standard deviation $1/3$ and gamma with unit mean and variance 0.5. Variations can be studied.

The probability density function for frailty ω is given by

$$f(\omega) = p\xi(\omega) + (1-p)g(\omega)$$

$$= p.\frac{1}{\sigma\sqrt{2\pi}}\left[e^{-\frac{(\omega-\mu)^2}{2\sigma^2}} + e^{-\frac{(\omega+\mu)^2}{2\sigma^2}}\right] + (1-p)\frac{\beta^{\theta}}{\Gamma(\theta)}\omega^{\theta-1}e^{-\beta\omega},$$

where $\omega > 0, \mu \in \mathbb{R}, \sigma, \theta, \beta > 0, 0 < p < 1$.

When we assume $\alpha = 0$ then the terminal event will be non-informative, that is recurrent event and terminal event are not associated. When $\alpha = 1$, the effect of frailty is identical to both events.

3. ESTIMATION PROCEDURES

Let us denote $\phi = (r_0(.), \lambda_0(.), \beta, \alpha, \theta)$. The full marginal log-likelihood function in timescale is given by (details construction of the likelihood function is given in Appendix)

$$l(\phi) = \sum_{i=1}^{N}\Bigg[\sum_{j=1}^{n_i}\delta_{ij}\log r_i\big(T_{ij}\big) + \delta_i^*\log\lambda_i(T_i^*) + \log\int_0^{\infty}\omega^{N_i^R(T_i^*)+\alpha\delta_i^*}\left[\frac{p}{\sigma\sqrt{2\pi}}\exp\Big(-\omega\sum_{j=1}^{n_i}\int_{Ti(j-1)}^{Tij}dR_i(t) - \omega^{\alpha}\int_0^{\infty}Y_i(t)\,d\Lambda_i(t)\Big)\left\{e^{-\frac{(\omega-\mu)^2}{2\sigma^2}} + e^{-\frac{(\omega+\mu)^2}{2\sigma^2}}\right\} + \frac{(1-p)\beta^{\theta}}{\Gamma(\theta)}\omega^{\theta-1}\exp\Big(-\omega\sum_{j=1}^{n_i}\int_{Ti(j-1)}^{Tij}dR_i(t) - \omega^{\alpha}\int_0^{\infty}Y_i(t)\,d\Lambda_i(t) - \beta\omega\Big)\right]d\omega\Bigg] \quad (3)$$

where $\Lambda_i(t) = \int_0^t\lambda_i(u)\,du$ and $R_i(t) = \int_0^t r_i(u)\,du$ are the cumulative hazard functions for death and recurrent events respectively and $T_{i0} = 0$ and $T_{in_i} = T_i^*$.

There are scopes to estimate the parameters from the likelihood function by using penalized likelihood estimation method (Rondeau *et al*. [22]). In this article, we estimate the parameters $\widehat{\beta_1}$, $\widehat{\beta_2}$ and $\hat{\alpha}$ from the dataset by using the method of having a minimum L_2 norm and hence we obtain the estimated parameters by solving the following equations successively (details are given in the Appendix).

$$\widehat{e^{\beta_1}} = \frac{\sum_{i=1}^{n}\omega_i\big[r_i(\widehat{T_{i1}}) + r_i(\widehat{T_{i2}}) + \cdots + r_i(\widehat{T_{in_i}})\big]}{\sum_{i=1}^{n}n_i\omega_i^2} \quad (4)$$

$$\widehat{e^{\beta_2}} = \frac{\sum_{i=1}^{n}\widehat{\lambda_i(T_i^*)}\omega_i^{\hat{\alpha}}}{\sum_{i=1}^{n}\omega_i^{2\hat{\alpha}}} = \frac{\sum_{i=1}^{n}\omega_i^{\hat{\alpha}}}{\sum_{i=1}^{n}\omega_i^{2\hat{\alpha}}} \quad (5)$$

and

$$\sum_{i=1}^{n}\left(T_{in_i} + n_i\left[\frac{\sum_{i=1}^{n}\omega_i^{\hat{\alpha}}}{\sum_{i=1}^{n}\omega_i^{2\hat{\alpha}}}\right]^2\right)\log_e\omega_i\,.\,\omega_i^{2\hat{\alpha}} + \sum_{i=1}^{n}\big[\omega_i^{\hat{\alpha}} - 1\big]\log_e\omega_i\,.\,\omega_i^{\hat{\alpha}} = 0. \quad (6)$$

So, for $\omega_1, \omega_2, \ldots., \omega_n$ drawn from the mixture distribution of folded normal distribution and gamma distribution, we get a numeric solution for $\hat{\alpha}$ from (6) and then we get $\widehat{\beta_2}$ from (5). This method is simpler.

4. SIMULATION STUDIES

In this section, we conduct a simulation study to evaluate the performance of the estimators of the joint frailty mixing model. The performances of the estimators are evaluated based on the bias and mean squared error (MSE) criteria. The simulations are performed in R software (version 4.2.3). The numerical outcomes are presented in Tables **1-6**, where the estimates and the corresponding bias and MSE values are displayed. We generate 1000 simulated datasets, each with 200, 350, and 500 subjects or samples. The algorithm of the simulation study is given below.

For each subject i,

- we generate the covariate Z_i from a Bernoulli distribution with $P(Z = 1) = 0.5$.
- The frailty ω_i is generated from a mixture of folded normal distribution with mean 0 and standard deviation $1/3$ and gamma distribution with mean 1 and variance 0.5.
- We generate terminal event time D_i by using a proportional hazard model $\lambda_i(t|\omega_i) = \omega_i^{\alpha}\lambda_0(t)\exp(\beta_2 Z_i(t))$ with $\lambda_0(t) = 1.0$.
- We set censoring time $C_i = \min\{e_i + 0.5, 2.5\}$ where e_i follows an exponential distribution with a mean of 1.
- We generate the recurrent event times from a Poisson process with an intensity function $r_i(t|\omega_i) = \omega_i r_0(t)\exp(\beta_1 Z_i)$ with $r_0(t) = 1.0$.
- The data generation continued until the observed time reached to $\min(D_i, C_i)$.

Table 1: Simulation Results for the Estimation of the Parameters for $\beta_1 = -1.5, \beta_2 = 0.2$ when $\alpha = 0$

n	p	Parameter	Est	BIAS	MSE
200	0.3	α	0.3099	−0.0388	0.0036
		β_1	−1.5105	0.0105	0.0142
		β_2	0.0782	0.1218	0.0151
	0.5	α	0.3949	0.0412	0.0045
		β_1	−1.5065	0.0065	0.0174
		β_2	0.1407	0.0593	0.0041
	0.7	α	0.5110	0.0177	0.0058
		β_1	−1.4747	−0.0253	0.0301
		β_2	0.2383	−0.0383	0.0028
350	0.3	α	0.3098	0.0131	0.0014
		β_1	−1.5027	0.0027	0.0079
		β_2	0.0792	0.1208	0.0148
	0.5	α	0.3940	−0.0248	0.0026
		β_1	−1.4963	−0.0037	0.0114
		β_2	0.1402	0.0598	0.0039
	0.7	α	0.5027	−0.0834	0.0102
		β_1	−1.4724	−0.0276	0.0170
		β_2	0.2351	−0.0351	0.0020
500	0.3	α	0.3099	0.0061	0.0010
		β_1	−1.5064	0.0064	0.0056
		β_2	0.0791	0.1209	0.0147
	0.5	α	0.3910	0.0172	0.0017
		β_1	−1.4979	−0.0021	0.0075
		β_2	0.1394	0.0606	0.0039
	0.7	α	0.5019	0.1134	0.0151
		β_1	−1.4636	−0.0364	0.0113
		β_2	0.2355	−0.0355	0.0018

Table 2: Simulation Results for the Estimation of the Parameters for $\beta_1 = -1.5, \beta_2 = 0.2$ when $\alpha = 0.5$

n	p	Parameter	Est	BIAS	MSE
200	0.3	α	0.3420	0.0306	0.0029
		β_1	−1.5618	0.0618	0.0211
		β_2	0.0770	0.1230	0.0154
	0.5	α	0.4202	−0.0651	0.0070
		β_1	−1.5415	0.0415	0.0259
		β_2	0.1358	0.0642	0.0047
	0.7	α	0.5389	−0.1551	0.0294
		β_1	−1.4987	−0.0013	0.0334
		β_2	0.2342	−0.0342	0.0025
350	0.3	α	0.3371	−0.0059	0.0012
		β_1	−1.5603	0.0603	0.0139
		β_2	0.0754	0.1246	0.0157
	0.5	α	0.4182	0.0252	0.0023
		β_1	−1.5398	0.0398	0.0138
		β_2	0.1348	0.0652	0.0046
	0.7	α	0.5311	−0.0518	0.0056
		β_1	−1.4975	−0.0025	0.0174
		β_2	0.2310	−0.0310	0.0017

(Table 2). Continued.

n	p	Parameter	Est	BIAS	MSE
500	0.3	α	0.3354	−0.0024	0.0008
		β_1	−1.5569	0.0569	0.0100
		β_2	0.0748	0.1252	0.0158
	0.5	α	0.4185	0.0059	0.0014
		β_1	−1.5370	0.0370	0.0112
		β_2	0.1346	0.0654	0.0045
	0.7	α	0.5312	0.0602	0.0060
		β_1	−1.4871	−0.0129	0.0127
		β_2	0.2309	−0.0309	0.0014

Table 3: Simulation Results for Estimation of the Parameters for $\beta_1 = -1.5, \beta_2 = 0.2$ when $\alpha = 1.0$

n	p	Parameter	Est	BIAS	MSE
200	0.3	α	0.3572	0.0223	0.0023
		β_1	−1.6279	0.1279	0.0368
		β_2	0.0720	0.1280	0.0166
	0.5	α	0.4416	0.0206	0.0032
		β_1	−1.6009	0.1009	0.0400
		β_2	0.1330	0.0670	0.0051
	0.7	α	0.5539	−0.0593	0.0086
		β_1	−1.5454	0.0454	0.0408
		β_2	0.2286	−0.0286	0.0020
350	0.3	α	0.3555	−0.0075	0.0011
		β_1	−1.6281	0.1281	0.0290
		β_2	0.0718	0.1282	0.0166
	0.5	α	0.4393	−0.0152	0.0018
		β_1	−1.5950	0.0950	0.0232
		β_2	0.1314	0.0686	0.0050
	0.7	α	0.5528	0.0235	0.0038
		β_1	−1.5186	0.0186	0.0204
		β_2	0.2270	−0.0270	0.0015
500	0.3	α	0.3548	−0.0008	0.0007
		β_1	−1.6266	0.1266	0.0241
		β_2	0.0716	0.1284	0.0166
	0.5	α	0.4359	−0.0165	0.0015
		β_1	−1.5957	0.0957	0.0195
		β_2	0.1305	0.0695	0.0051
	0.7	α	0.5480	0.1145	0.0153
		β_1	−1.5288	0.0288	0.0153
		β_2	0.2276	−0.0276	0.0013

We consider two sets of (β_1, β_2) such as $\beta_1 = -1.5, \beta_2 = 0.2$ and $\beta_1 = -1.3, \beta_2 = 0.3$. Logically, β_1 should not be highly negative as the effect will be negligible in that case. Also, β_1 should not be positive from the structure of the model. Similarly, β_2 should be positive and too high. For each set, we consider the following three settings for α:

1. Setting I corresponding to $\alpha = 0$
2. Setting II corresponding to $\alpha = 0.5$
3. Setting III corresponding to $\alpha = 1.0$

For each setting of α, we consider the mixing parameter $p = 0.3, 0.5, 0.7$. The estimate of the parameters β_1, β_2 and α are obtained successively by using (4), (5), and (6) respectively.

From the simulation results, it has been observed that, as the sample size increases, the MSE values decrease and thus the consistency property of all the estimators holds. Based on mixing parameter p, we

Table 4: Simulation Results for Estimation of the Parameters for $\beta_1 = -1.3, \beta_2 = 0.3$ when $\alpha = 0$

n	p	Parameter	Est	BIAS	MSE
200	0.3	α	0.3108	0.0077	0.0021
		β_1	−1.4997	0.1997	0.0534
		β_2	0.0800	0.2200	0.0487
	0.5	α	0.3957	0.0083	0.0029
		β_1	−1.4794	0.1794	0.0507
		β_2	0.1412	0.1588	0.0258
	0.7	α	0.5135	0.0837	0.0138
		β_1	−1.4603	0.1603	0.0543
		β_2	0.2373	0.0627	0.0054
350	0.3	α	0.3105	−0.0003	0.0015
		β_1	−1.4876	0.1876	0.0428
		β_2	0.0786	0.2214	0.0492
	0.5	α	0.3911	0.0277	0.0025
		β_1	−1.4785	0.1785	0.0420
		β_2	0.1398	0.1602	0.0260
	0.7	α	0.5040	0.0545	0.0060
		β_1	−1.4551	0.1551	0.0392
		β_2	0.2364	0.0636	0.0048
500	0.3	α	0.3079	−0.0053	0.0011
		β_1	−1.4919	0.1919	0.0423
		β_2	0.0778	0.2222	0.0495
	0.5	α	0.3901	−0.0100	0.0016
		β_1	−1.4824	0.1824	0.0403
		β_2	0.1388	0.1612	0.0262
	0.7	α	0.5030	−0.0152	0.0027
		β_1	−1.4546	0.1546	0.0346
		β_2	0.2358	0.0642	0.0047

Table 5: Simulation Results for Estimation of the Parameters for $\beta_1 = -1.3, \beta_2 = 0.3$ when $\alpha = 0.5$

n	p	Parameter	Est	BIAS	MSE
200	0.3	α	0.3397	0.0117	0.0020
		β_1	−1.5379	0.2379	0.0726
		β_2	0.0758	0.2242	0.0505
	0.5	α	0.4244	−0.0079	0.0028
		β_1	−1.5141	02141	0.0692
		β_2	0.1368	01632	0.0272
	0.7	α	0.5369	−0.0778	0.0116
		β_1	−1.4797	0.1797	0.0596
		β_2	0.2323	0.0677	0.0058
350	0.3	α	0.3380	0.0201	0.0016
		β_1	−1.5389	0.2389	0.0661
		β_2	0.0747	0.2253	0.0509
	0.5	α	0.4204	−0.0194	0.0021
		β_1	−1.5213	0.2213	0.0608
		β_2	0.1353	0.1647	0.0274
	0.7	α	0.5317	0.0747	0.0087
		β_1	−1.4748	0.1748	0.0474
		β_2	0.2310	0.0690	0.0055

(Table 5). Continued.

n	p	Parameter	Est	BIAS	MSE
500	0.3	α	0.3377	0.0337	0.0020
		β_1	−1.5393	0.2393	0.0633
		β_2	0.0751	0.2249	0.0507
	0.5	α	0.4172	0.0124	0.0014
		β_1	−1.5195	0.2195	0.0562
		β_2	0.1346	0.1654	0.0276
	0.7	α	0.5291	−0.0208	0.0026
		β_1	−1.4712	0.1712	0.0403
		β_2	0.2294	0.0706	0.0055

Table 6: Simulation Results for Estimation of the Parameters for $\beta_1 = -1.3, \beta_2 = 0.3$ when $\alpha = 1.0$

n	p	Parameter	Est	BIAS	MSE
200	0.3	α	0.3603	0.0330	0.0032
		β_1	−1.6143	0.3143	0.1197
		β_2	0.0734	0.2266	0.0516
	0.5	α	0.4421	0.0131	0.0029
		β_1	−1.5878	0.2878	0.1128
		β_2	0.1328	0.1672	0.0285
	0.7	α	0.5570	−0.0494	0.0079
		β_1	−1.5148	0.2148	0.0804
		β_2	0.2284	0.0716	0.0064
350	0.3	α	0.3553	−0.0180	0.0014
		β_1	−1.6125	0.3125	0.1087
		β_2	0.0718	0.2282	0.0522
	0.5	α	0.4376	−0.0328	0.0027
		β_1	−1.5802	0.2802	0.0930
		β_2	0.1313	0.1687	0.0288
	0.7	α	0.5491	0.0169	0.0032
		β_1	−1.5133	0.2133	0.0646
		β_2	0.2263	0.0737	0.0061
500	0.3	α	0.3564	−0.0461	0.0029
		β_1	−1.6069	0.3069	0.1018
		β_2	0.0716	0.2284	0.0523
	0.5	α	0.4371	−0.0356	0.0024
		β_1	−1.5763	0.02763	0.0863
		β_2	0.1306	0.1694	0.0289
	0.7	α	0.5456	−0.0241	0.0028
		β_1	−1.5084	0.2084	0.0573
		β_2	0.2259	0.0741	0.0060

see that when it increases, MSE of $\hat{\alpha}$ increases whereas the MSE of $\widehat{\beta_2}$ decreases. Also, when p increases, for the first set of (β_1, β_2), i.e. when $\beta_1 = -1.5$ and $\beta_2 = 0.2$, the MSE of $\widehat{\beta_1}$ increases, however for the second set, i.e. when $\beta_1 = -1.3, \beta_2 = 0.3$ the MSE of $\widehat{\beta_1}$ decreases. Further, based on the different choices of α, it is observed that α increases, MSE of both the estimators $\widehat{\beta_1}$ and $\widehat{\beta_2}$ increase in most of the cases for fixed sample size n and the mixture parameter p.

5. APPLICATION

For illustration purposes, we consider a dataset published by González *et al*. [28] regarding the sex differences in hospital readmission among colorectal cancer patients. The study took place in the Hospital de Bellvitge, Barcelona, Spain. A total 523 patients from January 1996 to December 1998 were identified with incident colorectal cancer and among them 403 patients had operations. In our study, we consider the

Table 7: Number of Hospital Readmissions and Death according to Chemotherapy Received or Not

Chemotherapy	No. of Patients	No. of death	No. of readmission since 1st discharge						
			1	2	3	4	5	6	≥ 7
Treated	216	57	124	51	22	7	4	4	4
Non-treated	187	51	75	54	23	14	11	4	6

data of these 403 patients. The remaining 120 (23%) patients were excluded because either they died before the study started they were refused to participate in the study or due to their lack of information. The study began from the date of surgery and follow-up continued till June 2002. Consequently, the follow-up period for each patient is different and depends on their surgery date. After surgery for colorectal cancer, patients may have several hospital readmissions. In general, more readmission of cancer patients leads to a higher mortality rate. That is, the terminal event (death) will be strongly correlated with recurrent events (hospital readmission) of interest. Here, we apply the proposed method to analyse jointly the recurrent events and death and thus focus on the effects of receiving chemotherapy on hospital readmission, and death is evaluated.

Information on patients receiving chemotherapy in the follow-up period and the number of readmissions (recurrent events) or deaths is given in Table **7**. A total of 108 (26.8%) patients died during the study and all subjects had at least one recurrent event. Table **7** shows that the patients receiving chemotherapy have a lower death rate and also rate of recurrences is decreased in nature.

The survival functions following hospital recurrent admission are presented in Figure **1**. This figure does not show a clear trend about the risk of recurrence.

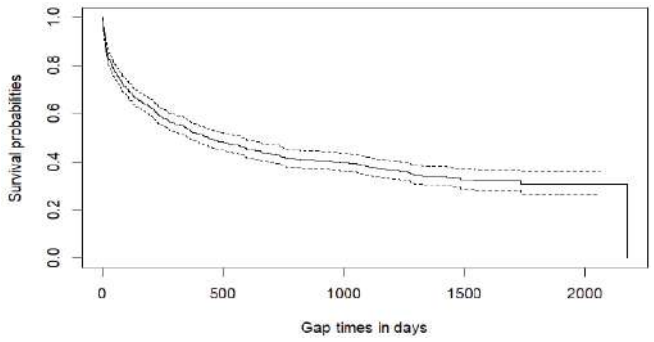

Figure 1: Survival functions for successive recurrences.

For individual i, let Z_i be a binary indicator of treatment (chemotherapy received $= 1$, chemotherapy not received $= 0$). We model the joint distribution of the survival times and hospital readmission (joint model (1)). Subject-specific frailty term represents to account of the effects of unobserved factors on the chances of both hospital readmission and death. Based on the methodology discussed above, we obtain the estimated values of the unknown parameters by varying the mixture parameter p. Table **8** shows the result for the recurrence of the hospital admission data.

Table 8: Application Results for Recurrent Events and Death

p	Parameters	Est
0.3	α	0.3982
	β_1	−0.9435
	β_2	0.0641
0.5	α	0.5776
	β_1	−0.8172
	β_2	0.0678
0.7	α	0.7494
	β_1	−0.6724
	β_2	0.1587

The negative value of $\widehat{\beta_1}\,(-0.8172)$ makes the hazard rate of recurrent events smaller and the positive value of $\widehat{\beta_2}\,(0.0678)$ makes the hazard rate of terminal events higher, which means that the rate of hospital readmission decreases for the cancer patients who are receiving chemotherapy and the survival risk higher for them. It is clear that the joint model helps understand the effect of chemotherapy for hospitalization and also gives about their survival. Also, the positive value of $\hat{\alpha} = 0.5776$ in the joint model indicates that the incidence of hospital readmission is positively associated with terminal events.

6. CONCLUSION

In literature, earlier works were done on the joint frailty model where the authors considered the frailty as gamma or uniform distribution only. But usually, uniform distribution is not preferable over the whole positive range. Here, we explore the more general form of frailty distribution i.e., a mixture of a folded normal and gamma distribution with associated weight p and $1 - p$ respectively. The advantage of taking a mixture distribution is that we can study the behaviour of the estimators over a range of mixing parameters and sometimes this may be close to gamma which is usually the general form of life distribution or a peaked

distribution like normal. An extensive simulation study has been carried out to estimate the parameters of the considered joint mixing frailty model. Also, the performances of the estimator are studied based on bias and MSE criteria. It has been observed that the method is consistent as the MSEs of all the estimators are reduced with the sample size. So, we observe that for different combinations of sample size n while varying the mixing parameter p, the method can estimate the regression coefficients of the joint mixing frailty model to some extent. We also illustrate the method through a study of patients in hospital readmission with colorectal cancer. As the theoretical findings from the simulation seem to be reasonable, the estimation using the hospital readmission data is expected to be quite accurate.

In addition, from this work, we see that α should be in between (0,1) and lower α values give more hazard to terminal events. So using this model in a prior study of such disease, we can estimate α value. Then by new recurrent observation or v_i, we can get the hazard of the terminal event v_i^{α}. Thus we can predict the terminal case of a patient using v_i^{α}.

APPENDIX

Construction of Full Log-Likelihood Function for the Joint Frailty Model with Calendar Timescale

The conditional distribution of the survival times given ω_i is the product of the individual contributions is given by

$$L_i(\phi|\omega_i) = \prod_{j=1}^{n_i}\left[dR_i(T_{ij}|\omega_i)^{\delta_{ij}} \times \exp\left(-\omega\sum_{j=1}^{n_i}\int_{T_{ij-1}}^{T_{ij}} dR_i(t)\right)\right] \times d\Lambda_i(T_i^*|\omega_i)^{\delta_i^*} \times \exp\left(-\omega^{\alpha}\int_0^{\infty} Y_i(t)\, d\Lambda_i(t)\right)$$

$$= \prod_{j=1}^{n_i} d\left[\omega R_i(T_{ij})\right]^{\delta_{ij}} \times d[\omega^{\alpha}\Lambda(T_i^*)]^{\delta_i^*} \times \exp\left(-\omega\sum_{j=1}^{n_i}\int_{T_{i(j-1)}}^{T_{ij}} dR_i(t) - \omega^{\alpha}\int_0^{\infty} Y_i(t)\, d\Lambda_i(t)\right)$$

$$= \prod_{j=1}^{n_i} \omega^{\delta_{ij}}\, dR_i(T_{ij})^{\delta_{ij}} \times \omega^{\alpha\delta_i^*}\, d\Lambda_{\mathrm{i}}(T_i^*)^{\delta_i^*} \times \exp\left(-\omega\sum_{j=1}^{n_i}\int_{T_{i(j-1)}}^{T_{ij}} dR_i(t) - \omega^{\alpha}\int_0^{\infty} Y_i(t)\, d\Lambda_i(t)\right)$$

$$= \omega^{N_i^R(T_i^*)} \times \prod_{j=1}^{n_i} dR_i(T_{ij})^{\delta_{ij}} \times \omega^{\alpha\delta_i^*}\, d\Lambda_{\mathrm{i}}(T_i^*)^{\delta_i^*} \times \exp\left(-\omega\sum_{j=1}^{n_i}\int_{T_{i(j-1)}}^{T_{ij}} dR_i(t) - \omega^{\alpha}\int_0^{\infty} Y_i(t)\, d\Lambda_i(t)\right)$$

The probability density function for frailty ω is given by

$$f(\omega) = p\xi(\omega) + (1-p)g(\omega)$$

$$= p.\frac{1}{\sigma\sqrt{2\pi}}\left[e^{-\frac{(\omega-\mu)^2}{2\sigma^2}} + e^{-\frac{(\omega+\mu)^2}{2\sigma^2}}\right] + (1-p)\frac{\beta^{\theta}}{\Gamma(\theta)}\omega^{\theta-1}e^{-\beta\omega},$$

where $\omega > 0, \mu \in \mathbb{R}, \sigma, \theta, \beta > 0, 0 < p < 1$

The marginal contribution to the likelihood for subject i is

$$L_i(\phi) = \int L_i(\phi|\omega_i) f(\omega)\, d\omega$$

$$= \int_0^{\infty} \omega^{N_i^R(T_i^*)} \times \prod_{j=1}^{n_i} dR_i(T_{ij})^{\delta_{ij}} \times \omega^{\alpha\delta_i^*}\, d\Lambda_{\mathrm{i}}(T_i^*)^{\delta_i^*} \times \exp\left(-\omega\sum_{j=1}^{n_i}\int_{T_{i(j-1)}}^{T_{ij}} dR_i(t) - \omega^{\alpha}\int_0^{\infty} Y_i(t)\, d\Lambda_i(t)\right) \times$$

$$\left[p.\frac{1}{\sigma\sqrt{2\pi}}\left[e^{-\frac{(\omega-\mu)^2}{2\sigma^2}} + e^{-\frac{(\omega+\mu)^2}{2\sigma^2}}\right] + (1-p)\frac{\beta^{\theta}}{\Gamma(\theta)}\omega^{\theta-1}e^{-\beta\omega}\right] d\omega$$

$$= \prod_{j=1}^{n_i} dR_i(T_{ij})^{\delta_{ij}} \times d\Lambda_i(T_i^*)^{\delta_i^*}$$

$$\times \int_0^\infty \omega^{N_i^R(T_i^*)+\alpha\delta_i^*} \left[\frac{p}{\sigma\sqrt{2\pi}} \exp\left(-\omega \sum_{j=1}^{n_i} \int_{Ti(j-1)}^{Tij} dR_i(t) - \omega^\alpha \int_0^\infty Y_i(t)\, d\Lambda_i(t) \right) \left\{ e^{-\frac{(\omega-\mu)^2}{2\sigma^2}} + e^{-\frac{(\omega+\mu)^2}{2\sigma^2}} \right\} \right.$$

$$\left. + \frac{(1-p)\beta^\theta}{\Gamma(\theta)} \omega^{\theta-1} \exp\left(-\omega \sum_{j=1}^{n_i} \int_{Ti(j-1)}^{Tij} dR_i(t) - \omega^\alpha \int_0^\infty Y_i(t)\, d\Lambda_i(t) - \beta\omega \right) \right] d\omega$$

Then the full log-likelihood function is given

$$l(\phi) = \log \prod_{i=1}^{N} L_i(\phi)$$

$$= \sum_{i=1}^{N} \log L_i(\phi)$$

$$= \sum_{i=1}^{N} \left[\sum_{j=1}^{n_i} \delta_{ij} \log r_i(T_{ij}) + \delta_i^* \log \lambda_i(T_i^*) \right.$$

$$+ \log \int_0^\infty \omega^{N_i^R(T_i^*)+\alpha\delta_i^*} \left[\frac{p}{\sigma\sqrt{2\pi}} \exp\left(-\omega \sum_{j=1}^{n_i} \int_{Ti(j-1)}^{Tij} dR_i(t) - \omega^\alpha \int_0^\infty Y_i(t)\, d\Lambda_i(t) \right) \left\{ e^{-\frac{(\omega-\mu)^2}{2\sigma^2}} + e^{-\frac{(\omega+\mu)^2}{2\sigma^2}} \right\} \right.$$

$$\left. \left. + \frac{(1-p)\beta^\theta}{\Gamma(\theta)} \omega^{\theta-1} \exp\left(-\omega \sum_{j=1}^{n_i} \int_{Ti(j-1)}^{Tij} dR_i(t) - \omega^\alpha \int_0^\infty Y_i(t)\, d\Lambda_i(t) - \beta\omega \right) \right] d\omega \right]$$

Estimation of Parameters of Joint Frailty Model

Suppose, data of recurring of n patients are

$$T_{11}, T_{12}, T_{13}, \dots\dots\dots\dots, T_{1n_1}$$

$$T_{21}, T_{22}, T_{23}, \dots\dots\dots\dots, T_{2n_2}$$

$$\vdots$$

$$T_{n1}, T_{n2}, T_{n3}, \dots\dots\dots\dots, T_{nn_n}$$

We can obtain data-based hazard function, $\widehat{r_i(t)} (i = 1,2, \dots\dots, n)$ as

$$\widehat{r_i(t)} = \sum_{y_{(j)} \le t} \frac{\delta_{(j)}}{n - i + 1}$$

where $y_{(j)} \le t \Rightarrow$ j many observations are less than or equal to t.

Similarly, for terminal data, we have T_1^*

$$T_2^*$$

$$\vdots$$

$$T_n^*$$

From this, we can get data-based estimate as $\widehat{\lambda_i(t)}\ (i = 1,2, \dots\dots, n)$ [mentioned above].

From recurring data, we estimate $\widehat{\beta_1}$ using the method of having minimum L_2 norm and similarly we estimate $\widehat{\beta_2}$ from terminal data.

Now, let us first obtain $\widehat{\beta_1}$ from the following:

$\widehat{r_i(t)} \cong \omega_i r_0(t) e^{\beta_1 Z_i(t)}\ \forall t$ & over all i

and then we minimize

$\sum_{i=1}^{n} \int_0^{\infty} \left(\widehat{r_i(t)} - \omega_i r_0(t) e^{\beta_1 Z_i(t)}\right)^2 dt.$

To do that,

$$\frac{\partial}{\partial \beta_1} \sum_{i=1}^{n} \int_0^{\infty} \left(\widehat{r_i(t)} - \omega_i r_0(t) e^{\beta_1 Z_i(t)}\right)^2 dt = 0$$

$$\Rightarrow \sum_{i=1}^{n} \int_0^{\infty} \left(\widehat{r_i(t)} - \omega_i r_0(t) e^{\beta_1 Z_i(t)}\right) \omega_i r_0(t) Z_i(t) e^{\beta_1 Z_i(t)}\, dt = 0 \quad -------(7)$$

Now,

$$\omega_i \sim p\xi + (1-p)g$$

where ξ~folded normal having mean 0, standard deviation $\frac{1}{3}$ and $g\sim$ Gamma with mean 1, variance 0.5. These means and variances can be taken other also.

For theoretical simulation, we take the values of the parameters of the distributions and that of p. So, these are known for our purpose, and from these we draw $\omega_1, \omega_2, \ldots\ldots, \omega_n$. Also $r_0(t)$ may be taken as 1, because the standard hazard function from natural level is uniformly 1. Also, we have $Z(T_{i1}), Z(T_{i2}), \ldots\ldots, Z(T_{in_i})$ from independent Bernouli distribution with known parameter. This indicates whether recurrent events occur or not. Thus, values are 1 at these points and at others are 0. So, we have from (7),

$$\sum_{i=1}^{n} \int_0^{\infty} \left(\widehat{r_i(t)} - \omega_i e^{\beta_1 Z_i(t)}\right) \omega_i Z_i(t) e^{\beta_1 Z_i(t)}\, dt = 0$$

$$\Rightarrow \sum_{i=1}^{n} \int_0^{T_{in_i}} \left(\widehat{r_i(t)} - \omega_i e^{\beta_1 Z_i(t)}\right) \omega_i Z_i(t) e^{\beta_1 Z_i(t)}\, dt = 0$$

$$\Rightarrow \sum_{i=1}^{n} \sum_{T_{i1}, T_{i2}, .., T_{in_i}} \left(\widehat{r_i(t)} - \omega_i e^{\beta_1}\right) \omega_i e^{\beta_1} = 0$$

$$\Rightarrow \sum_{i=1}^{n} \sum_{T_{i1}, T_{i2}, .., T_{in_i}} \left(\widehat{r_i(t)} - \omega_i e^{\beta_1}\right) \omega_i = 0$$

$$\Rightarrow \sum_{i=1}^{n} \left[\sum_{T_{i1}, T_{i2}, \ldots, T_{in_i}} \widehat{r_i(t)} \omega_i \right] = \sum_{i=1}^{n} \left[\sum_{T_{i1}, T_{i2}, \ldots, T_{in_i}} \omega_i^2 e^{\beta_1} \right]$$

$$\Rightarrow \sum_{i=1}^{n} \omega_i \left[\widehat{r_i(T_{i1})} + \widehat{r_i(T_{i2})} + \cdots + \widehat{r_i(T_{in_i})} \right] = e^{\beta_1} \sum_{i=1}^{n} n_i \omega_i^2$$

$$\Rightarrow \widehat{e^{\beta_1}} = \frac{\sum_{i=1}^{n} \omega_i \left[\widehat{r_i(T_{i1})} + \widehat{r_i(T_{i2})} + \cdots + \widehat{r_i(T_{in_i})} \right]}{\sum_{i=1}^{n} n_i \omega_i^2}$$

Similarly for $\widehat{\beta_2}$ we have,

$$\frac{\partial}{\partial \beta_2}\sum_{i=1}^{n}\int_0^{\infty}\left(\widehat{\lambda_i(t)} - \omega_i^{\alpha}e^{\beta_2 Z_i(t)}\right)^2 dt = 0$$

$$\Rightarrow \sum_{i=1}^{n}\int_0^{T_i^*}\left(\widehat{\lambda_i(t)} - \omega_i^{\alpha}e^{\beta_2 Z_i(t)}\right)\omega_i^{\alpha}Z_i(t)e^{\beta_2 Z_i(t)} = 0$$

$$\Rightarrow \sum_{i=1}^{n}\left(\widehat{\lambda_i(T_i^*)} - \omega_i^{\alpha}e^{\beta_2}\right)\omega_i^{\alpha}e^{\beta_2} = 0$$

$$\Rightarrow \sum_{i=1}^{n}\left(\widehat{\lambda_i(T_i^*)} - \omega_i^{\alpha}e^{\beta_2}\right)\omega_i^{\alpha} = 0$$

$$\Rightarrow \sum_{i=1}^{n}\widehat{\lambda_i(T_i^*)}\omega_i^{\alpha} = \left(\sum_{i=1}^{n}\omega_i^{2\alpha}\right)e^{\beta_2}$$

$$\Rightarrow \widehat{e^{\beta_2}} = \frac{\sum_{i=1}^{n}\widehat{\lambda_i(T_i^*)}\omega_i^{\hat{\alpha}}}{\sum_{i=1}^{n}\omega_i^{2\hat{\alpha}}} = \frac{\sum_{i=1}^{n}\omega_i^{\hat{\alpha}}}{\sum_{i=1}^{n}\omega_i^{2\hat{\alpha}}}$$

Similarly for $\hat{\alpha}$, we have

$$\frac{\partial}{\partial \alpha}\sum_{i=1}^{n}\int_0^{\infty}\left(\widehat{\lambda_i(t)} - \omega_i^{\alpha}e^{\beta_2 Z_i(t)}\right)^2 dt = 0$$

$$\Rightarrow -\sum_{i=1}^{n}\int_0^{T_i^*}\left(\widehat{\lambda_i(t)} - \omega_i^{\alpha}e^{\beta_2 Z_i(t)}\right)\log_e \omega_i \,.\, e^{\beta_2 Z_i(t)}.\,\omega_i^{\alpha}\, dt = 0$$

$$\Rightarrow -\sum_{i=1}^{n}\left[(0-\omega_i^{\alpha})\log_e \omega_i \,.\, \omega_i^{\alpha}.T_{in_i} + n_i\left(0-\omega_i^{\alpha}e^{\beta_2}\right)\log_e \omega_i \,.\, e^{\beta_2}\omega_i^{\alpha} + (0-\omega_i^{\alpha})\log_e \omega_i \,.\, \left(T_i^* - T_{in_i}\right)\omega_i^{\alpha} + (1-\omega_i^{\alpha})\log_e \omega_i \,.\, \omega_i^{\alpha}\right] = 0$$

[Note that $\widehat{\lambda_i(t)} = 0$ when $t < T_i^*$ and $\widehat{\lambda_i(t)} = 1$ when $t = T_i^*$ and naturally $Z_i(T_i^*) = 0$]

$$\Rightarrow \sum_{i=1}^{n}\left(T_{in_i} + n_i e^{2\beta_2}\right)\log_e \omega_i \,.\, \omega_i^{2\alpha} + \sum_{i=1}^{n}\left[\left(T_i^* - T_{in_i}\right)\omega_i^{\alpha} - (1-\omega_i^{\alpha})\right]\log_e \omega_i \,.\, \omega_i^{\alpha} = 0$$

$$\Rightarrow \sum_{i=1}^{n}\left(T_{in_i} + n_i\left[\frac{\sum_{i=1}^{n}\omega_i^{\alpha}}{\sum_{i=1}^{n}\omega_i^{2\alpha}}\right]^2\right)\log_e \omega_i \,.\, \omega_i^{2\alpha} + \sum_{i=1}^{n}\left[\left(T_i^* - T_{in_i} + 1\right)\omega_i^{\alpha} - 1\right]\log_e \omega_i \,.\, \omega_i^{\alpha} = 0$$

$$\Rightarrow \sum_{i=1}^{n}\left(T_{in_i} + n_i\left[\frac{\sum_{i=1}^{n}\omega_i^{\alpha}}{\sum_{i=1}^{n}\omega_i^{2\alpha}}\right]^2\right)\log_e \omega_i \,.\, \omega_i^{2\alpha} + \sum_{i=1}^{n}\left[\omega_i^{\alpha} - 1\right]\log_e \omega_i \,.\, \omega_i^{\alpha} = 0$$

REFERENCES

[1] Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika 1981; 68: 373-379. https://doi.org/10.1093/biomet/68.2.373

[2] Pepe MS, Cai J. Some graphical displays and marginal regression analysis for recurrent failure times and time-dependent covariates. J Am Statist Assoc 1993; 88: 811-820. https://doi.org/10.2307/2290770

[3] Lawless JF, Nadeau C. Some simple robust methods for the analysis of recurrent events. Technometrics 1995; 37: 158-168. https://doi.org/10.2307/1269617

[4] Sun L, Zhao X, Zhou J. A class of mixed models for recurrent event data. Canad J Statist 2011; 39: 578-590. https://doi.org/10.1002/cjs.10132

[5] Ghosh D, Lin DY. Nonparametric analysis of recurrent events and death. Biometrics 2000; 56: 554-562. https://doi.org/10.1111/j.0006-341x.2000.00554.x

[6] Miloslavsky M, Keles S, Van der Laan MJ, *et al*. Recurrent events analysis in the presence of time-dependent covariates and dependent censoring. J R Statist Soc 2004; 66B: 239-257. https://doi.org/10.1111/j.1467-9868.2004.00442.x

[7] Zeng D, Lin DY. Semiparametric transformation models with random effects for joint analysis of recurrent and terminal events. Biometrics 2009; 65: 746-752. https://doi.org/10.1111/j.1541-0420.2008.01126.x

[8] Zeng D, Cai J. A semiparametric additive rate model for recurrent events with informative terminal event. Biometrika 2010; 97: 699-712. https://doi.org/10.1093/biomet/asq039

[9] Wang MC, Qin J, Chiang CT. Analyzing recurrent event data with informative censoring. J Am Statist Assoc 2001; 96: 1057-1065. https://doi.org/10.1198/016214501753209031

[10] Huang CY, Wang MC. Joint Modeling and estimation of recurrent event process and failure time. J Am Statist Assoc 2004; 99: 1153-1165. https://doi.org/10.1198/016214504000001033

[11] Ye Y, Kalbfleisch JD, Schaubel DE. Semiparametric analysis of correlated recurrent and terminal events. Biometrics 2007; 63: 78-87. https://doi.org/10.1111/j.1541-0420.2006.00677.x

[12] Cook RJ, Lawless JF. Marginal analysis of recurrent events and a terminating event. Stat Med 1997; 16: 911-924. https://doi.org/10.1002/(SICI)1097-0258(19970430)16:8<911::AID-SIM544>3.0.CO;2-I

[13] Ghosh D, Lin DY. Marginal regression models for recurrent and terminal events. Statist Sinica 2002; 12: 663-668. https://www.jstor.org/stable/24306989

[14] Cook RJ, Lawless JF, Lakhal CL, *et al*. Robust estimation of mean functions and treatment effects for recurrent events under event-dependent censoring and termination: Application to skeletal complications in cancer metastatic to bone. J Am Statist Assoc 2009; 104: 60-75. https://doi.org/10.1198/jasa.2009.0004

[15] Cox DR. Regression models and life-tables, Journal of the Royal Statistical Society. Series B (Methodological) 1972; 34(2): 187-220. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

[16] Hougaard P. Analysis of multivariate survival data, Springer Verlag: New York 2000.

[17] Duchateau L, Janssen P. The frailty model, Springer: New York, 2008.

[18] Che X, Angus J. A new joint model of recurrent event data with the additive hazards model for the terminal event time. Metrica 2016. https://doi.org/10.1007/s00184-016-0577-9

[19] Liu L, Wolfe RA, Huang X. Shared frailty models for recurrent events and a terminal event. Biometrics 2004; 60: 747-756. https://doi.org/10.1111/j.0006-341X.2004.00225.x

[20] Huang X, Liu L. A joint frailty model for survival and gap times between recurrent events. Biometrics 2007; 63(2): 747-756. https://doi.org/10.1111/j.1541-0420.2006.00719.x

[21] Joly P, Commenges D, Letenneur L. A penalized likelihood approach for arbitrarily censored and truncated data: application to age-specific incidence of dementia. Biometrics 1998; 54(1): 185-194. https://doi.org/10.2307/2534006

[22] Rondeau V, Mathoulin-Pelissier S, Jacqmin-Gadda H, Brouste V, Soubeyran P. Joint frailty models for recurring events and death using maximum penalized likelihood estimation: application on cancer events. Biostatistics 2007; 8(4): 708-721. https://doi.org/10.1093/biostatistics/kxl043

[23] Marquardt D. An algorithm for least-squares estimation of nonlinear parameters. SIAM Journal on Applied Mathematics 1963; 11: 431-441. https://www.jstor.org/stable/2098941

[24] Mazroui Y, Mathoulin-Pelissier S, Soubeyran P and Rondeau V. General joint frailty model for recurrent event data with a dependent terminal event: application to follicular lymphoma data. Statistics in Medicine 2012. https://doi.org/10.1002/sim.4479

[25] Toenges G, Jahn-Eimermacher A. Computational issues in fitting joint frailty models for recurrent events with an associated terminal event. Computer Methods and Programs in Biomedicine 2020; 188: 1-13. https://doi.org/10.1016/j.cmpb.2019.105259

[26] Banerjee P, Goswami A, Bhunia S, Basu S. Determination of Causal Relationship Between Bilirubin and Other Liver Biomarker in Case of Hepatitis C. Biomed Stat Informatics 2021; 6(2): 23-31. https://doi.org/10.11648/j.bsi.20210602.11

[27] Peng Y, Jiajia D, Jun Z. Joint analysis of recurrent event data with a dependent terminal event. J Syst Sci Complex 2017; 30: 1443-1458. https://doi.org/10.1007/s11424-017-6097-5

[28] González JR, Fernandez E, Moreno V, Ribes J, Peris M, Navarro M, Cambray M, Borrás JM. Sex differences in hospital readmission among colorectal cancer patients. Journal of Epidemiology and Community Health 2005; 59: 506-511. https://doi.org/10.1136/jech.2004.028902

[29] Fathoni M, Gunardi G, Adi-Kusumo F, Hutajulu SH, I Purwanto. Cox Proportional Hazard Regression Interaction Model and Its Application to Determine The Risk of Death in Breast Cancer Patients after Chemotherapy. Int J Stats Med Res 2022; 11: 105-113. https://doi.org/10.6000/1929-6029.2022.11.13

Received on 26-09-2023 Accepted on 23-10-2023 Published on 24-11-2023

https://doi.org/10.6000/1929-6029.2023.12.25