ISSN 2282-0930 • Epidemiology Biostatistics and Public Health - 2023, Volume 18, Number 1

Joint frailty model for recurrent events and a terminal event in the presence of cure fraction

Zahra Arab Borzu, PhD[(1)], Ahmad Reza Baghestani[(2)], Elaheh Talebi Ghane, PhD[(3)], Ali Akbar Khadem Maboudi, PhD[(4)], Ali Akhavan[(5)], Anahita Saeedi[(6)]

(1) Assistant Professor, Department of Biostatistics & Epidemiology, School of Public Health, Zahedan University of Medical Sciences, Zahedan, Iran
(2) Associate Professor of Biostatistics Department, Physiotherapy Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
(3) Assistant Professor of Biostatistics, Modeling of Noncommunicable Diseases Research Center, Hamadan University of Medical Sciences, Hamadan, Iran
(4) Associate Professor of Biostatistics Department, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
(5) Assistant Professor of Radiation Oncology, Isfahan University of Medical Sciences, Isfahan, Iran
(6) Department of Biostatistics, School of public Health & Health Sciences, University of Massachusetts, Amherst, MA, USA

CORRESPONDING AUTHOR: *Ahmad Reza Baghestani, Associate Professor of Biostatistics Department, Physiotherapy Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Email: baghestani.ar@gmail.com*

SUMMARY

The observations of repeated or recurrent events occur in many longitudinal studies. Furthermore, sometimes there may exist a terminal event such as death, which is strongly correlated with recurrent events. In many situations, a fraction of subjects who will never experience the event of interest during a long follow-up period is considered to be cured. In this article, we proposed a joint frailty model in the presence of cure fraction. The dependency is modeled by shared frailty that is contained in both the recurrent and terminal events hazard functions. It allows to estimate two separate sets of parameters on the recurrent, death, and cure model. We applied the maximum likelihood method under a piecewise constant hazard function for model fitting. The proposed model is evaluated by simulation studies and an application to a breast cancer data is provided.

Keywords: Breast cancer; Cure model; Joint model; Recurrent event; Terminal event.

INTRODUCTION

In many clinical or epidemiological studies, there are situation in which subjects are measured repeatedly over a fixed time. For instance, repeated episodes of hospitalization or experience asthma attacks, tumor recurrences. Many methodologies have been considered for the analysis of recurrent event data [1-6]. In many settings exists a terminal event such as death. Therefore, the terminal event may be strongly correlated with recurrent events. More explicitly, if the rate of the recurrent event is unusually low (high) in a subject, that subject is also subject to decreased (increased) rate of death. For example, recurrent asthma attacks during a follow-up, which can lead to death. In this case, the ordinary assumption of independent censoring can be violated and lead to biased estimates [4]. There are two major approaches to analyze recurrent events in the presence of a terminal event: The marginal models and the frailty models. Marginal models attend on the marginal rates of the recurrent and terminal events that can not specify the dependence between recurrent and terminal events [7-10]. Frailty models mostly apply a latent variable to account for the correlation between the recurrent and terminal events so that the two event processes

DOI: 10.54103/2282-0930/20639
Accepted: 14th May 2023

are independent given the frailty. For example, Huang and Wolf proposed a general joint frailty model to account for the informative censoring [11]. Liu et al., introduced a nonparametric maximum penalized likelihood method for estimating hazard function in a joint frailty model with right censoring and delayed entry [4]. Mazroui et al., suggested a joint frailty model to analyze recurrent events and death. They used two gamma-distributed frailties to allow for both the inter-recurrences dependence and the dependence between the recurrences and the survival times [12].

In recent years, the development of new drugs and treatment regimens has resulted in the significant number of patients in the population who are not susceptible to the event and live longer with diseases such as cancer; consequently, a cured fraction of the population exists. The use of standard survival models, for example, the Cox proportional hazard model for such data may be inappropriate since these models are based on the assumption that all the subjects experience the event with probability one so that the overall survivor function descends to zero, approximately. This assumption cannot be used in recent clinical trials and medical researches, because many subjects may never experience the event of interest if the follow-up period is sufficiently long. In such cases, cure models are widely applied. In this paper, we had a motivating example of patients with breast cancer (BC). A total of 357 patients received surgery to remove tumors. Two hundred and fifty-seven (72%) patients had no recurrence and death due to BC. We showed the Kaplan–Meier curve of disease-free survival (time to the first recurrence or death, whichever happened first) for patients with BC in Figure 1. There were very few events after 5 years of follow-up period, denoting the existence of a large proportion of cured patients. Ignoring the existence of "cured" patients leads to underestimation of the hazard and consequently overestimation of the overall survival of non-cured patients [13].

Many studies have been done on cure models [14-17]. In the context of recurrent event data, Rondeau et al., proposed a frailty model for the recurrent events in the presence of cure fraction [13]. Zhao et al., introduced a new model for recurrent with terminal events which can incorporated zero recurrence subjects [17]. Kim proposed a joint model for recurrent with a terminal event in the presence of cure fraction. The suggested model applied two types of deaths for the cure and susceptible groups, which would be regarded as competing risk with a missing cause [18]. Liu et al., proposed a joint frailty model for zero-inflated recurrent events in the presence of a terminal event. In that model, the frailty effect on recurrent and death rates is the same. In this article, we presented a joint frailty model in the presence of cure fraction for recurrent events and terminal event (death) by a shared gamma frailty in which the frailty can have different effects on recurrent events and death rates [19]. Thus, our model combined the features of the Liu et al. (2016) for patients who had no chance of experiencing the recurrent or death events from breast cancer, "cured patients", and the Liu et al. (2004) for the joint frailty analysis of recurrent and terminal events; the frailty effect on recurrent and death rates is the different. One advantage of our model is that it can estimate the effect of covariates on the recurrence and death times, and the cured probability, simultaneously. It can also reveal the degree of dependency between disease recurrence and death.

The remainder of the article is organized as follows. In Section 2, we introduced the joint frailty model in the presence of cure fraction and the estimation method. In Section 3, we presented the simulation studies and their results. In Section 4, we applied the proposed model to the analysis of a real dataset and a concluding discussion is presented in Section 5.

THE MODEL

Notations

We define notations and definitions that are used in the model. Let $T_{ij} = \min(X_{ij}, C_i, D_i)$ be the observed follow-up time so that X_{ij}, C_i and D_i correspond to the jth recurrent event time for ith subject $(i = 1,\ldots,N, j = 1,\ldots,n_i)$, the right-censoring time and the death time. Similarly, the terminal time denote by$T^* = \min(C_i, D_i)$. We consider a binary indicator for recurrent event as $\delta_{ij} = I(T_{ij} = X_{ij})$ so that if $n_i > 0$ then $\delta_{ij} = 1$ and a binary indicator for terminal event as $\Delta_i = I(T^*_i = D_i)$. S_{ij} indicate gap times (the time interval from previous to next recurrent event) so that $S_{ij} = T_{ij} - T_{i(j-1)}$ are independent with conditional on frailties and covariates. The observation for subject i is $O_i(t) \equiv \{S_{ij}, T^*_i, \delta_{ij}, \Delta_i\}$. Based on the theory of multivariate counting processes [4,14], $N^{D^*}_i(t) = I(D_i \le t)$ and $N^D_i(t) = I(X_i \le t, \Delta_i = 1)$ are the actual and the observed death indicator by time t, respectively. Similarly, we denote by $N^{R^*}_i(t)$ and $N^R_i(t) = N^{R^*}_i(\min(X_i, t))$ the actual and observed number of recurrent events, respectively. Let $Y_i(t) = 1_{(t \le T^*_i)}$ the at-risk indicator of subject i at time t. The observed and the actual number of recurrent events that occurs for ith in $[t, t+dt)$ is respectively $dN^{R^*}_i(t) = N^{R^*}_i((t+dt)^-) - N^{R^*}_i(t^-)$ and $N^R_i(t) = Y_i(t)dN^{R^*}_i(t)$. The process history of subject i up to time t, is represented as $H_{it} = \sigma\{Y(h), N^R_i(h), N^D_i(h), Z_i(h), \omega_i(h), 0 \le h \le t\}$. Where $Z_i(h)$ is the vector of covariates and $\omega_i(h)$ is shared frailty for subject i. Furthermore, recurrent event processes, death and censoring times assume to be continuous, therefore, in the simultaneous occurrences of recurrent and death events, we assume that death

happens first. The death event and the recurrent events intensity processes at t are $Y_i(t)h_i(t)dt = P(dN_i^R(t) = 1 | \mathcal{F}_{it^-})$ and $Y_i(t)\lambda_i(t)dt = P(dN_i^D(t) = 1 | \mathcal{F}_{it^-})$, respectively, where $h_i(t)dt = P(dN_i^{R^*}(t) = 1 | Z_i(t), \omega_i, D_i \geq t)$ and $\lambda_i(t)dt = P(dN_i^{D^*}(t) = 1 | Z_i(t), \omega_i, D_i \geq t)$.

Model for recurrent events and a terminal event

Following the model of Liu et al (2004), the joint model for the recurrent and terminal events given by:

$$\begin{cases} \lambda_i(t \mid \omega_i) = \omega_i \lambda_0(t) \exp(\beta' Z_i(t)) = \omega_i \lambda_i(t) \\ h_i(t \mid \omega_i) = \omega_i^{\alpha} h_0(t) \exp(\beta^* Z_i(t)) = \omega_i^{\alpha} h_i(t) \end{cases} \quad (2.1)$$

Where $\lambda_0(t)$ and $h_0(t)$ are baseline hazard functions for recurrent events and death respectively. The parameters β and β^* are regression coefficients vector associated with the covariate vector Z_i for recurrent event and death rates that could be different. The random effect ω_i takes into account the dependence between recurrent times and the death time. We assume ω_i have the gamma distribution with mean 1 and variance θ. When $\theta = 0$ implies that the random effects ω_i's are exactly 1, i.e., and heterogeneity in both recurrent and terminal events is only explained by Z_i. In the proposed model (2.1), the degree of dependence between recurrent and death times showed by α. The assumption is that $\alpha = 0$ that is $h_i(t)$ does not depend on ω_i, and terminal event (death) is non-informative for the recurrent events $\lambda_i(t)$, so that two rates $h_i(t)$ and $\lambda_i(t)$ are independent. When $\alpha = 1$, the effect frailty on recurrent events and death is the same. When $\alpha > 1$ the recurrent and death rates are positively correlated; higher frailty will result in earlier death. Inversely, $\alpha < 1$ demonstrates that subjects with higher frailty will be less likely to death.

Joint cure model for recurrent events and a terminal event

Let U be a binary variable that a subject will eventually $(U_i = 1)$ or never experience the event of interest $(U_i = 0)$. The survival function of T given by $S(t \mid z) = p\, S_u(t \mid z) + (1 - p)$. Where $S_u(t \mid z)$ is survival function for uncured subject and $p = Pr(U = 1)$.

In order to assess the relationship between Z_i and the probability of cure, a logit link function is used:

$$\log it(p_i) = \gamma^T Z_i. \quad (2.2)$$

Where γ is a parameter that is associated with the cure rate through covariate Z.

Following the model of Liu et al. (2016), the frailty proportional hazard model for recurrent events for subjects that are susceptible or not cured is:

$$\lambda_i(t \mid \omega_i, U_i = 1) = \lambda_0(t) \exp(\beta Z_i + \omega_i). \quad (2.3)$$

Similarly, hazard model for terminal event is:

$$h_i(t \mid \omega_i, U_i = 1) = h_0(t) \exp(\beta^* Z_i + \omega^{\alpha}{}_i). \quad (2.4)$$

Combining equations (2.2), (2.3) and (2.4) we have a joint model of the recurrent and terminal events with a cure fraction. In this case, a subject cured cannot experience any recurrent events, nor death due to the disease. Conditional likelihood for subject i th can be written as:

$$L(O_i \mid \omega_i) = L_{i1}^{I(n_i>0,\Delta_i=1)} L_{i2}^{I(n_i=0,\Delta_i=1)} L_{i3}^{I(n_i>0,\Delta_i=0)} L_{i4}^{I(n_i=0,\Delta_i=0)}$$

Where

L_{i1} is the likelihood of observing recurrent events $(n_i > 0)$ and death $(\Delta_i = 1)$,

L_{i2} is the likelihood of observing no recurrent events $(n_i = 0)$ and death $(\Delta_i = 1)$,

L_{i3} is the likelihood of observing recurrent events $(n_i > 0)$ and no death $(\Delta_i = 0)$,

L_{i4} is the likelihood of observing no recurrent events $(n_i = 0)$ and no death $(\Delta_i = 0)$.

That L_{i4} is cure on recurrent and terminal (death) events.

We can write:

$$L_{i1} = (1 - p_i) S_i^R(t_i \mid \omega_i, U_i = 1) \prod_{j=1}^{n_i} \lambda_i(t_{ij} \mid \omega_i, U_i = 1)$$
$$\times h_i(t^* \mid \omega_i, U_i = 1)^{\Delta_i} S^D{}_i(t^*{}_i \mid \omega_i, U_i = 1),$$

$$L_{i2} = (1 - p_i) S_i^R(t_i \mid \omega_i, U_i = 1) \times h_i(t_i^* \mid \omega_i, U_i = 1)^{\Delta_i}$$
$$\times S^D{}_i(t_i^* \mid \omega_i, U_i = 1),$$

$$L_{i3} = (1 - p_i) S_i^R(t_i \mid \omega_i, U_i = 1) \times \prod_{j=1}^{n_i} \lambda_i(t_{ij} \mid \omega_i, U_i = 1)$$
$$\times S^D{}_i(t_i^* \mid \omega_i, U_i = 1),$$

$$L_{i4} = p_i + (1 - p_i) S_i^R(t_i \mid \omega_i, U_i = 1) \times S^D{}_i(t_i^* \mid \omega_i, U_i = 1),$$

Where $S_i^R(t \mid \omega_i, U_i = 1)$ and $S_i^D(t^* \mid \omega_i, U_i = 1)$ are survival functions for the recurrent and death times for those not cured:

$$S_i^R(t \mid \omega_i, U_i = 1) = \exp(-\exp(\beta Z_i + \omega_i) \Lambda_0(t)),$$
$$S_i^D(t^* \mid \omega_i, U_i = 1) = \exp(-\exp(\beta^* Z_i + \omega^{\alpha}{}_i) H_0(t))$$

The $\Lambda_0(t)$ and $H_0(t)$ are cumulative baseline hazard function for the recurrent event and death respectively. The full loglikelihood is:

$$l(O) = \ln \prod_{i=1}^{N} \int_0^{\infty} L(O_i \mid \omega_i) \pi_{\theta}(\omega_i) d\omega_i \quad (2.5)$$

Where $\pi_{\theta}(\omega_i)$ is density function for frailty shared.

Estimation

To obtain the parameters estimation in proposed model, we utilize maximize likelihood technique to estimate different parameter $\Phi = (h_0(.), \lambda_0(.), \beta, \beta^*, \alpha, \theta, \gamma)$ due to the difficulty of solving the integral in the full log-likelihood (2.5), we used approach Gauss–Laguerre quadrature which is a numerical approximation of an integral using a weighted average of the integrand computed at M predetermined quadrature points $u_m (m = 1, 2, \ldots, M)$ over random effect ω_i. This the numerical approximation can be as such, $L(O) \approx \sum_{m=1}^{M} L(O_i \mid u_m)\pi_\theta(u_m)\nu_m$, with $u_m = \sqrt{2z_m}$ and $\nu_m = \sqrt{2\eta_m}\exp(z^2{}_m)$ Where η_m and z_m can be obtained from tables or algorithms, details of the procedure presented by [20,21]. Further, we apply a piecewise constant baseline hazard function for the estimation of baseline hazard functions in our estimation method. In the piecewise constant hazard function, we first divided the follow–up duration for recurrent events in to 5 intervals by 5th quantile (denoted by knots $Q_1^{\lambda}, Q_2^{\lambda}, \ldots, Q_5^{\lambda}$ and $Q_0^{\lambda} = 0$ or the smallest recurrent event time). We have:

$\tilde{\lambda_0}(t) = \lambda_{0k}$ for $Q^{\lambda}{}_{k-1} < t \le Q^{\lambda}{}_k$ where $k = 1, 2, \ldots, 5$ or

$$\lambda_0(t) = \sum_{k=1}^{5} \lambda_{0k} I(Q^{\lambda}{}_{k-1} < t \le Q^{\lambda}{}_k)$$

The cumulative baseline hazard function is

$$\Lambda_0(t) = \sum_{k=1}^{5} \lambda_{0k} \max(0, \min(Q^{\lambda}{}_k - Q^{\lambda}{}_{k-1}, t - Q^{\lambda}{}_{k-1}))$$

Following the similar procedure, we can create the piecewise constant baseline hazard function for death, denoted by $\tilde{h_0}(t)$ and $\tilde{H_0}(t)$ for cumulative baseline death hazard.

We use $\hat{H}^{-1}$ as a variance estimator, where H is the converged Hessian matrix of the log likelihood. Moreover, due to positively constraints on the parameter $(\theta > 0)$, we utilize the exponential transformation and their standard error calculated by the delta method [22].

After replacing cumulative baseline hazards in log-likelihood (2.5), the resulting log-likelihood can be maximized by the Gauss–Laguerre quadrature with implementation in R software.

SIMULATION

In this study, six hundred replicate datasets were generated, each with sample size (n=250, 500, 1000) to investigate the effect of increased sample size in parameters estimation. The simulation results of the parameters estimation are provided in Tables 1-4, which includes the Estimation parameter (Est), the empirical standard errors (SE), the mean square error (MSE), and the 95% empirical coverage probabilities (CP). The AIC mean and the number of propriety for the proposed and reduced models, which was the result of the minimum AIC value, were also reported, we considered the right-censored and utilized calendar time scale representation.

Generating Data

For each subject i, we generated binary explanatory variables $Z_i (i = 1, 2)$, from a Bernoulli distribution with probability 0.5. The random variables ω_i was generated from gamma distribution so that $\omega_i \sim gamma\left(\frac{1}{\theta}, \frac{1}{\theta}\right)$ with $\theta = 0.5$. A fixed right-censoring time was taken as $C_i = 6 + Unif(0, 6)$. We generated the gap times X_{ik} from $\lambda_i(s \mid \omega_i, U_i = 1) = \omega_i \lambda_0(s)\exp(\beta_1 Z_1 + \beta_1 Z_{2i})$ where $\lambda_0(t) = 0.65t^{0.25}$ and death time D_i generated from $h_i(t \mid \omega_i) = \omega^{\alpha}{}_i h_0(t)\exp(\beta_2 Z_{1i})$ where $h_0(t) = 0.4t^{0.25}$.

A death time D_i was generated from the hazard function $h_i(t \mid \omega_i)$.

If observed time was a death time $D_i \le C_i$ then $T_i^* = D$ and $\Delta_i = 1$.

If $D_i > C_i$ individual was censored then $T_i^* = C_i$ and $\Delta_i = 0$.

We used a logistic regression for probability of cure so that: $p_i = \frac{1}{1 + \exp(\gamma_0 + \gamma_1 Z_{i1})}$ and set $\alpha_0 = -0.5$ and $\alpha_1 = 1$.

We generated a random variable u_i from uniform distribution $[0, 1]$. The individual was cured (any recurrent nor death) if $u_i < p_i$ and individual was non-cured if $u_i \ge p_i$. The calendar times created from

$$T_{ij} = \min(C_i, D_i, \sum_{k=1}^{j} X_{ik})$$

If $T_{ij} < T_i^*$ then the observed time can be a recurrent event time and $\delta_{ij} = 1$. The data generating continues until $T_{ij} < T_i^*$.

If $T_{ij} \ge T_i^*$ individual was censored then $T_{ij} = T_i^*$ and $\delta_{ij} = 0$.

We set $\beta_1 = 1, \beta_2 = -0.5, \beta_1^* = 0.7, \alpha = 2, \gamma_0 = -0.5, \gamma_1 = 1$.

To compare the proposed model with two reduced models, we considered four different settings of $\alpha, \gamma_0, \gamma_1$ as following.

In setting I, we generated joint frailty model without cure fraction $(\gamma_0 = \gamma_1 = 10, \alpha = 2)$ Since $p_i = \frac{1}{1 + \exp(\gamma_0 + \gamma_1 Z_{i1})}$ for $(Z_{i1} = 0, 1)$, we had a mean

of cure percentage (p_i) close to zero. The estimates of parameters in the proposed model can be compared with the model of Liu et al., (2004).

In setting II, we generated joint frailty model in the presence of cure fraction ($\gamma_0 = -0.5\gamma_1 = 1, \alpha = 1$). For the situation, mean of cure percentage (p_i) close to 0.5 and the frailty effect on recurrent and terminal event rates is the same. The estimates of parameters in the proposed model can be compared with model of Liu et al., (2016).

In setting III, we generated joint frailty model with $\alpha > 1$, so that the recurrent rate and death are positively associated ($\gamma_0 = -0.5\gamma_1 = 1, \alpha = 2$). We can compare the estimates of parameters in the proposed model with the two reduced models (Liu et al., (2004) and Liu et al., (2016)).

In setting IV, we generated joint frailty model with $\alpha < 1$, so that the recurrent rate and death are negatively associated ($\gamma_0 = -0.5\gamma_1 = 1, \alpha = -2$). We can compare the estimates of parameters in the proposed model with the two reduced models (Liu et al., (2004) and Liu et al., (2016)).

Results of the simulation studies

The average numbers of deaths were 68% to 78%, the average numbers of recurrent events (among all 600 subjects) were 0.25 to 0.69 with a maximum fixed number of eight. The mean cure percentage was 50% in setting II and III.

In setting I, the mean cure percentage (p_i) was close to zero, so there was no cure fraction in datasets. In this case, both the joint frailty model (proposed model) and the reduced model (Liu et al, 2004) were equivalent. The mean of the estimates for γ_0 and γ_1 by the joint model are 9.937 and 10 respectively, which are very close to the true values.

It can be seen that the mean square errors and biases of parameters decreased with an increase in the sample size. In addition, AIC mean in proposed model was about four units more than the AIC mean in the reduced model, which was due to two extra parameters in the proposed model. Also, AIC percentage in the reduced model was lower than the proposed model in more 98.82% of cases. This indicates that even when the cure fraction does not exist, it is still valid to use the proposed model for data analysis.

In setting II, we had $\alpha = 1$ so there was a same correlation between recurrent and terminal event. The result showed that both the cure joint frailty model (proposed model) and the reduced model (II) were equivalent. The parameter estimates from these two models were virtually similar, with almost the same accuracy and precision. The mean of the estimates for α by the proposed model is 1.061, which is very close to the true value of $\alpha = 1$ in sample size 250. Thus, by increasing the sample size, α is underestimated.

We obtained a clear improvement in the estimates of parameters and mean square errors with increasing sample size. AIC mean in the proposed model was about one unit more than the AIC mean in the reduced model that by increasing the sample size, the difference raised to two. Furthermore, based on AIC percentage of all 600 replicate datasets, model (II) was preferred at least 81.8% times. This shows that when dependence between recurrent and terminal events is same, proposed model and model (II) are equivalent.

In setting III, we generated data from proposed model and set $\gamma_0 = -0.5, \gamma_1 = 1$. so that the mean of cure percentage was close to 0.5. We assumed $\alpha = 2$ which indicates significant positive dependence between recurrent and the death rates. In this case, proposed model is compared with two reduced models (Liu et al., (2004) and Liu et al., (2016)). The results of our model are summarized in the first panel in Table 3. The mean parameter estimates by new proposed joint frailty model were very close to their true values. There was a good agreement between the empirical and estimated standard errors of these parameter estimates, and the coverage probabilities were close to the nominal level of 95%. Moreover, the results show an underestimate for death risk and α which does not get better by increasing the sample size.

This can be due to the positivity constraint on the variance parameter. In comparison, we fit the model without the cure fraction. The results are reported in the third panel of Table 3. The results show that the absent of the cure fraction led to significant in biases and mean square errors in the estimate of parameters and very poor coverage probabilities. The estimate of the variance of the random effect in model without cure was much larger than that in our model (1.781 vs. 0.45). This shows that the new proposed cure joint frailty model can effectively capture the heterogeneity. Additionally, the lowest AIC mean and the high AIC percentage (98%) in the new proposed model suggests a better fit than two reduced models.

In setting IV, in order to assess a negative association between recurrent events and death rates we considered ($\alpha = -2$). Findings illustrate that the new proposed model offers very accurate parameter estimates and powerful coverage probabilities (Table 4).

APPLICATION-BREAST CANCER STUDY

Breast cancer (BC) is the most commonly diagnosed disease among females and includes 23% of total cancer cases with 14% risk of death. The cycle of this disease is usually determined by a response to initial treatment, followed by relapses. Moreover, relapse of breast cancer may increase the risk of death, which indicates an association between relapse and death.

In recent years, the improvement in treatment has led to 70-80% of patients being cured of BC. Common statistical models are not suitable for analyzing these data [23]. We applied the joint frailty model to analyze breast cancer (BC) with the new proposed model and two reduced models. Our real example is obtained from Shahid Ramezanzadeh Radiotherapy Center between April 2004 to March 2012; the patients were followed until April 2016. There were 357 females with BC included in the analysis. Among them, 77(21.6%) died, 69(19.3%) patients experienced recurrence of BC. The maximum number of recurrences for a patient was three. The numbers of patients with one, two and three recurrent events were 50(14%), 18(5%) and 1(0.3%), respectively. Two hundred and fifty-seven (72%) cases were cured, meaning that they experienced neither a recurrent event nor death due to BC. In this study, we considered four baseline covariates for each patient: Lymphovascular invasion (positive versus negative), age (50 years or older versus younger than 50 years), Lymph node status (positive versus negative) and tumor size (II, III versus I). Then we used the proposed joint model to analyze the effect of prognostic factors on recurrent and death times in the presence of cure fraction. For comparison, we also applied the joint frailty model without cure fraction and joint frailty model with the same frailty model in the presence of cure fraction, as introduced by Liu et al., in the years 2004 and 2016, respectively. In three models, the baseline hazard function is assumed to be piecewise constant for recurrent and terminal events, each with 5 intervals. The estimation results are shown in Table 5. We can see that the tumor size was significant in the cure model (P=0.013). The patients with larger tumor size were less likely to be cured. For illustration, hazard ratio of tumor size III and II were 0.77 and 0.48, respectively. Among those who were "not cured", tumor size was not significant. Patients with larger tumor sizes were more likely to experience recurrences. The hazard ratio of the patients with tumor size II and III were 1.012 and 1.008, respectively. In contrast, patients with larger tumor sizes had a lower morality rate than patients with tumor size I. Furthermore, we considered the same association between recurrent and terminal events leads to reduced model introduced by Liu et al., (2016), as shown in the second panel of Table 5. In this reduced model, sign and effect of variables were similar to those in our model except for the Lymph node status in death model, which showed that the patients with positive lymph node status were associated with a decreased risk of death (HR=0.704, P=0.304). We also fit another reduced model, which is a joint model without cure fraction introduced by Liu et al., (2004) as indicated in the third panel of Table 5. We noticed that the parameter estimates and their significance were different from those in the presence of cure fraction. The estimate of frailty variance without a cure fraction was more than that in the cure fraction model (variance estimate of θ increased from 0.904 to 1.288). This suggests that ignoring cure fraction leads to more heterogeneity for recurrent events in the reduced model. The positive values of α =1.357 to 1.8 show that the recurrence of disease and death rates were positively associated (P<0.001). The cured probability in our model and reduced model (II) was 77% and 85%, respectively. We obtained the cured probability in the data about 72%, indicating a more accurate estimate in our model. The Akaike information criterion (AIC) was also calculated, the AIC values indicated that the proposed model had a better fit than reduced models with the lowest value AIC=2062.

DISCUSSION AND CONCLUSION

In this paper, we introduced a joint frailty model in the presence of cure fraction. Our proposed model has two main advantages: on the one hand, the new joint frailty model can take into account a cure component. In this situation, the cured subject experience neither the recurrent events, nor death due to the diseases. On the other hand, our proposed model can evaluate the degree of dependence between recurrent and death times through the estimation parameter α. We have shown by simulation that using our joint frailty model in the presence of cure fraction led to unbiased regression coefficients, smaller mean square error, better coverage probabilities and less AIC in comparison with two reduced models. The simulation results show that in the presence of cure fraction, if $\alpha > 1$ and we falsely consider $\alpha = 1$, an underestimation of the recurrent and death rates occurs. In contrast, if $\alpha < 1$ and we falsely consider $\alpha = 1$, then recurrent and death rates is overestimated. The simulation results demonstrate that our proposed model is valid, even when there is the same dependence between recurrent and death times or there is non-cure fraction in the dataset. The proposed model was applied to a breast cancer dataset, and we showed that a positive association exists between recurrent and death rates. In this case, higher frailty implies an expected real death. In this article, we used gamma distribution for frailty. Other distributions can be used as well, e.g., Gaussian distribution (Liu et al., 2016). We have assumed piecewise for $\lambda_0(t)$ and $h_0(t)$. We can consider semi-parametric modeling (using spline function) for baseline hazard functions for recurrence and death, which provide more flexibility and reliable estimates of the cure fraction [12,13,24]. For the future works, our model can be more complex by considering longitudinal biomarkers and the joint with the recurrent model.

ACKNOWLEDGMENTS

The authors thank the reviewers and associate editor for their careful reading and valuable comments.

This study was approved by the Ethics Committees of the Shahid Beheshti University of Medical Science, Tehran, Iran (no: IR.SBMU.RETECH.REC.1400.697).

REFERENCES

1. Cai J, Schaubel DE. Marginal means/rates models for multiple type recurrent event data. Lifetime data analysis. 2004;10(2):121-38.
2. Cook RJ, Lawless JF. Analysis of repeated events. Statistical Methods in Medical Research. 2002;11(2):141-66.
3. Little RJ. Modeling the drop-out mechanism in repeated-measures studies. Journal of the american statistical association. 1995;90(431):1112-21.
4. Liu L, Wolfe RA, Huang X. Shared frailty models for recurrent events and a terminal event. Biometrics. 2004;60(3):747-56.
5. Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika. 1981;68(2):373-9.
6. Sun J, Song PXK. Statistical analysis of repeated measurements with informative censoring times. Statistics in medicine. 2001;20(1):63-73.
7. Li QH, Lagakos SW. Use of the Wei–Lin–Weissfeld method for the analysis of a recurring and a terminating event. Statistics in medicine. 1997;16(8):925-40.
8. Lin D, Wei L, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2000;62(4):711-30.
9. Schaubel DE, Zhang M. Estimating treatment effects on the marginal recurrent event mean in the presence of a terminating event. Lifetime data analysis. 2010;16(4):451-77.
10. Wei L-J, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. Journal of the American statistical association. 1989;84(408):1065-73.
11. Huang X, Wolfe RA. A frailty model for informative censoring. Biometrics. 2002;58(3):510-20.
12. Mazroui Y, Mathoulin-Pelissier S, Soubeyran P, Rondeau V. General joint frailty model for recurrent event data with a dependent terminal event: application to follicular lymphoma data. Statistics in medicine. 2012;31(11-12):1162-76.
13. Rondeau V, Schaffner E, Corbiere F, Gonzalez JR, Mathoulin-Pelissier S. Cure frailty models for survival data: application to recurrences for breast cancer and to hospital readmissions for colorectal cancer. Statistical methods in medical research. 2013;22(3):243-60.
14. Andersen PK, Borgan O, Gill RD, Keiding N. Statistical models based on counting processes: Springer Science & Business Media; 2012.
15. Peng Y, Dear KB. A nonparametric mixture model for cure rate estimation. Biometrics. 2000;56(1):237-43.
16. Sy JP, Taylor JM. Estimation in a Cox proportional hazards cure model. Biometrics. 2000;56(1):227-36.
17. Zhao X, Wang J, Zhou X, Zhu Z. Recurrent events analysis in the presence of terminal event and zero-recurrence subjects. Communications in Statistics-Theory and Methods. 2015;44(4):710-25.
18. Kim YJ. Joint model for recurrent event data with a cured fraction and a terminal event. Biometrical Journal. 2020;62(1):24-33.
19. Liu L, Huang X, Yaroshinsky A, Cormier JN. Joint frailty models for zero-inflated recurrent events in the presence of a terminal event. Biometrics. 2016;72(1):204-14.
20. bramowitz M, Stegun IA. Handbook of mathematical functions: with formulas, graphs, and mathematical tables: Dover publications New York; 1972.
21. Golub GH, Welsch JH. Calculation of Gauss quadrature rules. Mathematics of computation. 1969;23(106):221-30.
22. Knight K. Mathematical Statistics. Chapman and Hall/CRC Texts in Statistical Science: New York. 2000.
23. Ewertz M, Jensen AB. Late effects of breast cancer treatment and potentials for rehabilitation. Acta Oncologica. 2011;50(2):187-93.
24. Zhangsheng Y, Liu L. A joint model of recurrent events and a terminal event with a nonparametric covariate function. Statistics in medicine. 2011;30(22):2683-95.

APPENDIX

Appendix A: construction of the log-likelihood for the proposed joint frailty model with calendar timescale.

In this appendix, we explain the structure of full likelihood $L(O \mid \omega_i) = \prod_i \int L(O_i \mid \omega_i) f(\omega_i) d\omega_i$, where $L_i(h_0(.), \lambda_0(.), \beta, \beta^*, \alpha, \theta) = L_i(O \mid \omega)$ for subject i and $(j = 1, 2, \ldots, n_i)$ such as, $\delta_{(i, n_i+1)} = 0$.

we calculated the conditional likelihood for the patients who experience the occurrence of disease and death $(n_i > 0, \Delta_i = 1)$, we have:

$$S_i^R(t_i \mid \omega_i) = \exp(-\omega_i \sum_{j=1}^{n_i+1} \int_{T_{i(j-1)}}^{T_{ij}} Y_i(t)\lambda_i(t)dt),$$

$$S_i^D(t_i^* \mid \omega_i) = \exp(-\omega_i^{\alpha} \int_0^{T_i^*} Y_i(t)h_i(t)dt)$$

$$L_{i1}(O \mid \omega_i) = (1 - p_i)S_i^R(t_i \mid \omega_i, U_i = 1) \prod_{j=1}^{n_i} \lambda_i(t_{ij} \mid \omega_i, U_i = 1) \times h_i(t_i^* \mid \omega_i, U_i = 1)^{\Delta_i} S^D{}_i(t_i^* \mid \omega_i, U_i = 1)$$

$$= (1 - p_i)\omega_i^{n_i} \prod_{j=1}^{n_i} \lambda_i(t_{ij})^{\delta_{ij}} \times \exp(-\omega_i \sum_{j=1}^{n_i+1} \int_{T_{i(j-1)}}^{T_{ij}} Y_i(t)\lambda_i(t)dt) \times (\omega_i h_i(t_i^*))^{\alpha\delta_i^*} \times \exp(-\omega_i^{\alpha} \int_0^{T_i^*} Y_i(t)h_i(t)dt)$$

2) We consider $\omega_i \sim G(\frac{1}{\theta}, \frac{1}{\theta})$ with probability density $f(\omega) = \frac{\omega^{(1/\theta - 1)} \exp(-\omega/\theta)}{\Gamma(1/\theta)\theta^{(\frac{1}{\theta})}}$

The contribution of marginal likelihood is obtained by integrating out the random effect (ω_i)

$$l_{i1}(O) = \frac{(1 - p_i) \times \prod_{j=1}^{n_i+1} \lambda(t_{ij})^{\delta_{ij}} \times (h_i(T_i^*))^{\alpha\delta_i^*}}{\theta^{(\frac{1}{\theta})}\Gamma(\frac{1}{\theta})} \times \int_0^{\infty} \omega^{(\alpha\delta_i^* + n_i + \frac{1}{\theta} - 1)} \times \exp(-\omega_i \sum_{j=1}^{n_i+1} \int_{T_{i(j-1)}}^{T_{ij}} Y_i(t)\lambda_i(t)\,dt)$$

$$-\omega_i^{\alpha} \int_0^{T_i^*} Y_i(t)h_i(t)dt(-\frac{\omega_i}{\theta})d\omega_i$$

The contribution of marginal log-likelihood for individual i is:

$$l_{i1}(O) = \log(1 - p_i) + \delta_{ij} \sum_{j=1}^{n_{i+1}} \log(\lambda_{ij}) + \alpha\delta_i^* \log(h_i(t_i^*)) - \frac{1}{\theta}\log(\theta) - \log(\Gamma(\frac{1}{\theta})) - \frac{1}{\theta}\log(\Gamma(\frac{1}{\theta}))$$

$$+ \log\left[\int_0^{\infty} \omega^{(\alpha\delta_i^* + n_i + \frac{1}{\theta} - 1)} \times \exp(-\omega_i \sum_{j=1}^{n_i+1} \int_{T_{i(j-1)}}^{T_{ij}} Y_i(t)\lambda_i(t)\,dt) - \omega_i^{\alpha} \int_0^{T_i^*} Y_i(t)h_i(t)dt(-\frac{\omega_i}{\theta})d\omega_i\right]$$

In situation 2, we have subjects that do not experience the recurrent event $n_i = 0$ and observing death $\Delta_i = 1$. the contribution of marginal log-likelihood for individual i can write:

$$l_{i2}(O \mid \omega_i) = \log(1 - p_i) + \alpha\delta_i^* \log(h_i(t_i^*)) - \theta\log(\frac{1}{\theta}) - \log(\Gamma(\frac{1}{\theta}))$$

$$+ \log\left[\int_0^{\infty} \omega_i^{(\alpha\delta_i^* + \frac{1}{\theta} - 1)} \exp(-\omega_i \sum_{j=1}^{n_i+1} \int_{T_{i(j-1)}}^{T_{ij}} Y_i(t)\lambda_i(t)\,dt) - \omega_i^{\alpha} \int_0^{T_i^*} Y_i(t)h_i(t)\,dt - \frac{\omega_i}{\theta} d\omega_i\right]$$

In situation 3, we have subjects that experience the recurrent event $\delta_{ij} = 1$ but no observing death $\Delta_i = 0$. the

contribution of marginal log-likelihood for individual i can write:

$$l_{i3}(O \mid \omega_i) = \log(1-p_i) + \delta_{ij}\sum_{j=1}^{n_i}\log(\lambda(t_{ij})) - \frac{1}{\theta}\log(\theta) - \log(\Gamma(\frac{1}{\theta}))$$

$$+\log\left[\int_0^{\infty} -\omega_i \sum_{j=1}^{n_i+1}\int_{T_{i(j-1)}}^{T_{ij}} Y_i(t)\lambda_i(t)dt - \omega_i{}^{\alpha}\int_0^{T_i^*} Y_i(t)h_i(t)dt - \frac{\omega_i}{\theta}d\omega_i\right]$$

In situation 4, we have subjects that experience neither recurrence nor death from the disease

$$l_{i4}(O \mid \omega_i) = \log\left\{\frac{[p_i + (1-p_i)]\times\int_0^{\infty}\omega_i{}^{(\frac{1}{\theta}-1)}\exp(-\omega_i\sum_{j=1}^{n_i+1}\int_{T_{i(j-1)}}^{T_{ij}} Y_i(t)\lambda_i(t)dt - \omega_i{}^{\alpha}\int_0^{T_i^*} Y_i(t)h_i(t)dt - \frac{\omega_i}{\theta}d\omega_i}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})}\right\}$$

We can obtain full log likelihood by sum of the four marginal contribution of log-likelihood for subject i as follows:

$$l(O) = \sum_{i=1}^{4} l_i(O)$$

We can obtained the log-likelihood for gap times with replace T_{ij} by S_{ij} and $\int_{T_{i(j-1)}}^{T_{ij}}$ by $\int_0^{S_{ij}}$ in expression of log-likelihood.

Table 1. *Simulation Results for a generated joint frailty model with different frailty effect in absent of cure fraction*

Sample size	Parameter	Proposed model					Model by Liu et al. (2016)				
		Est	SE emp	SE $(\sqrt{\hat{H}^{-1}})$	MSE	CP	Est	SE emp	SE $(\sqrt{\hat{H}^{-1}})$	MSE	CP
N=250	$\beta_1 = 1$	0.964	0.152	0.155	0.024	0.934	0.969	0.15	0.155	0.023	0.938
	$\beta_2 = -0.5$	-0.507	0.131	0.131	0.017	0.941	-0.506	0.131	0.131	0.017	0.939
	$\beta_1^* = 0.7$	0.617	0.239	0.246	0.064	0.941	0.622	0.239	0.246	0.063	0.941
	$\theta = 0.5$	0.46	0.105	0.101	0.013	0.944	0.463	0.103	0.101	0.012	0.943
	$\alpha = 2$	1.833	0.452	0.457	0.232	0.958	1.826	0.437	0.453	0.222	0.943
	$\gamma_0 = 10$	9.826	1.264	44.079	1.628	0.963	-	-	-	-	-
	$\gamma_1 = 10$	10	0	32424.04	0	0.98	-	-	-	-	-
	mean_AIC			1690.986					1687.116		
	Percent_AIC			1.52%					98.48%		
N=500	$\beta_1 = 1$	0.976	0.103	0.109	0.011	0.946	0.978	0.103	0.109	0.011	0.942
	$\beta_2 = -0.5$	-0.507	0.093	0.092	0.009	0.949	-0.507	0.093	0.092	0.009	0.949
	$\beta_1^* = 0.7$	0.614	0.166	0.171	0.035	0.912	0.616	0.165	0.17	0.034	0.915
	$\theta = 0.5$	0.465	0.074	0.072	0.007	0.921	0.467	0.074	0.071	0.007	0.927
	$\alpha = 2$	1.767	0.276	0.307	0.13	0.874	1.765	0.275	0.305	0.131	0.87
	$\gamma_0 = 10$	9.872	0.965	37.245	0.947	0.978	-	-	-	-	-
	$\gamma_1 = 10$	10	0	14439.45	0	0.987	-	-	-	-	-
	mean_AIC			2135.385					3357.693		
	Percent_AIC			15.5%					98.48%		
N=1000	$\beta_1 = 1$	0.978	0.078	0.077	0.006	0.941	0.979	0.077	0.077	0.006	0.943
	$\beta_2 = -0.5$	-0.503	0.064	0.065	0.004	0.944	-0.503	0.064	0.065	0.004	0.944
	$\beta_1^* = 0.7$	0.618	0.12	0.12	0.021	0.887	0.619	0.12	0.12	0.021	0.889
	$\theta = 0.5$	0.467	0.052	0.051	0.004	0.896	0.468	0.052	0.051	0.004	0.899
	$\alpha = 2$	1.732	0.199	0.212	0.111	0.712	1.731	0.198	0.211	0.111	0.705
	$\gamma_0 = 10$	9.937	0.858	28.489	0.74	0.975	-	-	-	-	-
	$\gamma_1 = 10$	10	0.001	9335.436	0	0.992	-	-	-	-	-
	mean_AIC			6692.079					6688.152		
	Percent_AIC			1.18%					98.82%		

AIC, Akaike information criterion; CP, coverage probability; MSE, mean square error; SE, standard error

ISSN 2282-0930 • Epidemiology Biostatistics and Public Health - 2023, Volume 18, Number 1

Table 2. *Simulation Results for a generated joint frailty model with same frailty in presence of cure fraction*

Sample size	Parameter	Proposed model					Model by Liu et al. (2016)				
		Est	SE emp	SE $(\sqrt{\hat{H}^{-1}})$	MSE	CP	Est	SE emp	SE $(\sqrt{\hat{H}^{-1}})$	MSE	CP
N=250	$\beta_1 = 1$	0.981	0.259	0.253	0.068	0.945	0.983	0.259	0.252	0.068	0.948
	$\beta_2 = -0.5$	-0.511	0.207	0.201	0.043	0.955	-0.512	0.207	0.200	0.043	0.953
	$\beta_1^* = 0.7$	0.687	0.318	0.297	0.101	0.937	0.675	0.308	0.281	0.095	0.943
	$\theta = 0.5$	0.462	0.174	0.169	0.032	0.955	0.452	0.164	0.157	0.029	0.957
	$\alpha = 1$	1.061	0.431	0.429	0.189	0.948	-	-	-	-	-
	$\gamma_0 = -0.5$	-0.501	0.219	0.208	0.048	0.942	-0.501	0.219	0.208	0.048	0.943
	$\gamma_1 = 1$	1.018	0.292	0.281	0.086	0.953	1.019	0.293	0.281	0.086	0.953
	mean_AIC			1078.811					1077.80		
	Percent_AIC			15.17%					84.83%		
N=500	$\beta_1 = 1$	0.98	0.182	0.178	0.034	0.945	0.981	0.182	0.177	0.034	0.942
	$\beta_2 = -0.5$	-0.517	0.145	0.142	0.021	0.953	-0.517	0.144	0.14	0.021	0.952
	$\beta_1^* = 0.7$	0.675	0.209	0.203	0.044	0.947	0.678	0.206	0.199	0.043	0.948
	$\theta = 0.5$	0.477	0.123	0.123	0.016	0.957	0.462	0.112	0.112	0.014	0.948
	$\alpha = 1$	0.969	0.244	0.259	0.061	0.945	-	-	-	-	-
	$\gamma_0 = -0.5$	-0.505	0.151	0.147	0.023	0.96	-0.506	0.15	0.146	0.023	0.96
	$\gamma_1 = 1$	1.015	0.203	0.198	0.042	0.947	1.015	0.203	0.198	0.042	0.947
	mean_AIC			2135.385					2134.354		
	Percent_AIC			15.5%					85.5%		
N=1000	$\beta_1 = 1$	0.973	0.125	0.126	0.016	0.942	0.974	0.125	0.125	0.016	0.945
	$\beta_2 = -0.5$	-0.508	0.101	0.101	0.010	0.947	-0.508	0.101	0.099	0.010	0.947
	$\beta_1^* = 0.7$	0.657	0.142	0.141	0.022	0.942	0.668	0.142	0.140	0.021	0.935
	$\theta = 0.5$	0.485	0.091	0.088	0.008	0.948	0.468	0.081	0.08	0.008	0.932
	$\alpha = 1$	0.93	0.165	0.172	0.032	0.922	-	-	-	-	-
	$\gamma_0 = -0.5$	-0.503	0.109	0.103	0.012	0.953	-0.505	0.109	0.103	0.012	0.952
	$\gamma_1 = 1$	1.008	0.149	0.139	0.022	0.947	1.009	0.148	0.139	0.022	0.947
	mean_AIC			4255.608					4254.756		
	Percent_AIC			18.2%					81.8%		

AIC, Akaike information criterion; CP, coverage probability; MSE, mean square error; SE, standard error

Table 3. *Simulation Results for a generated joint frailty model with different frailty* ($\alpha > 0$) *on recurrent and death rate in presence of cure fraction*

Sample size	Parameter	Proposed model										Liu et al(2004)				
		Est	SE emp	SE $(\sqrt{\hat{H}^{-1}})$	MSE	CP	Est	SE emp	SE $(\sqrt{\hat{H}^{-1}})$	MSE	CP	Est	SE emp	SE $(\sqrt{\hat{H}^{-1}})$	MSE	CP
N=250	$\beta_1 = 1$	0.98	0.263	0.262	0.069	0.95	0.975	0.262	0.271	0.069	0.955	1.724	0.303	0.261	0.616	0.34
	$\beta_2 = -0.5$	-0.515	0.187	0.190	0.035	0.938	-0.518	0.189	0.210	0.036	0.937	-0.534	0.229	0.200	0.054	0.95
	$\beta_1^* = 0.7$	0.656	0.419	0.407	0.177	0.943	0.501	0.314	0.299	0.138	0.905	1.399	0.389	0.385	0.640	0.582
	$\theta = 0.5$	0.442	0.181	0.178	0.036	0.943	0.495	0.217	0.191	0.047	0.958	1.765	0.086	0.187	1.608	0.000
	$\alpha = 2$	2.010	0.726	0.768	0.527	0.955	-	-	-	-	-	1.426	0.245	0.301	0.390	0.300
	$\gamma_0 = -0.5$	-0.501	0.228	0.219	0.052	0.947	-0.493	0.232	0.222	0.054	0.948	-	-	-	-	-
	$\gamma_1 = 1$	1.016	0.296	0.290	0.088	0.947	1.019	0.298	0.292	0.089	0.948	-	-	-	-	-
	mean_AIC			1121.014										1190.82		
	Percent_AIC			63.33%										0.00%		
N=500	$\beta_1 = 1$	0.976	0.184	0.185	0.035	0.958	0.970	0.183	0.191	0.034	0.96	1.719	0.218	0.184	0.565	0.080
	$\beta_2 = -0.5$	-0.517	0.135	0.134	0.019	0.953	-0.520	0.136	0.147	0.019	0.950	-0.532	0.164	0.140	0.028	0.947
	$\beta_1^* = 0.7$	0.636	0.279	0.275	0.082	0.940	0.503	0.215	0.211	0.085	0.857	1.380	0.276	0.264	0.539	0.312
	$\theta = 0.5$	0.451	0.132	0.130	0.02	0.943	0.501	0.142	0.136	0.02	0.953	1.775	0.061	0.132	1.630	0.000
	$\alpha = 2$	1.836	0.421	0.474	0.204	0.960	-	-	-	-	-	1.383	0.164	0.201	0.408	0.053
	$\gamma_0 = -0.5$	-0.510	0.150	0.154	0.023	0.952	-0.504	0.152	0.155	0.023	0.955	-	-	-	-	-
	$\gamma_1 = 1$	1.019	0.204	0.204	0.042	0.935	1.021	0.205	0.205	0.042	0.932	-	-	-	-	-
	mean_AIC			2218.537										2360.208		
	Percent_AIC			81%										0.00%		
N=1000	$\beta_1 = 1$	0.973	0.133	0.130	0.018	0.957	0.966	0.132	0.134	0.019	0.958	1.711	0.159	0.130	0.531	0.005
	$\beta_2 = -0.5$	-0.512	0.093	0.094	0.009	0.957	-0.515	0.094	0.103	0.009	0.957	-0.526	0.111	0.099	0.013	0.940
	$\beta_1^* = 0.7$	0.617	0.188	0.190	0.042	0.928	0.495	0.149	0.148	0.064	0.718	1.354	0.189	0.183	0.463	0.060
	$\theta = 0.5$	0.450	0.093	0.093	0.011	0.927	0.501	0.097	0.097	0.009	0.947	1.781	0.044	0.093	1.642	0.000
	$\alpha = 2$	1.783	0.287	0.319	0.129	0.870	-	-	-	-	-	1.366	0.113	0.138	0.415	0.003
	$\gamma_0 = -0.5$	-0.508	0.105	0.108	0.011	0.948	-0.502	0.106	0.109	0.011	0.948	-	-	-	-	-
	$\gamma_1 = 1$	1.009	0.142	0.143	0.020	0.947	1.012	0.143	0.144	0.021	0.95	-	-	-	-	-
	mean_AIC			4419.620					4427.928					4705.045		
	Percent_AIC			98.16%					1.84%					0.00%		

AIC, Akaike information criterion; CP, coverage probability; MSE, mean square error; SE, standard error

Table 4. *Simulation Results for a generated joint frailty model with different frailty on recurrent and death rate* ($\alpha < 0$) *in presence of cure fraction*

Sample size	Parameter	Proposed model					Liu et al. (2016)					Liu et al. (2004)				
		Est	SE emp	SE $(\sqrt{\hat{H}^{-1}})$	MSE	CP	Est	SE emp	SE $(\sqrt{\hat{H}^{-1}})$	MSE	CP	Est	SE emp	SE $(\sqrt{\hat{H}^{-1}})$	MSE	CP
N=250	$\beta_1 = 1$	0.998	0.199	0.198	0.04	0.943	1.094	0.188	0.152	0.044	0.921	1.705	0.315	0.251	0.597	0.385
	$\beta_2 = -0.5$	-0.5	0.157	0.166	0.025	0.942	-0.492	0.174	0.136	0.03	0.946	-0.469	0.276	0.223	0.077	0.943
	$\beta_1^* = 0.7$	0.649	0.348	0.342	0.124	0.948	0.401	0.22	0.212	0.138	0.735	0.934	0.227	0.233	0.107	0.831
	$\theta = 0.5$	0.438	0.192	0.165	0.041	0.95	0.055	0.04	0.035	0.2	0	2.061	0.096	0.189	2.445	0
	$\alpha = -2$	-1.883	0.724	0.598	0.537	0.972	-	-	-	-	-	0.569	0.071	0.092	6.603	0
	$\gamma_0 = -0.5$	-0.498	0.188	0.188	0.035	0.953	-0.495	0.187	0.187	0.035	0.953	-	-	-	-	-
	$\gamma_1 = 1$	0.998	0.264	0.264	0.07	0.952	0.995	0.264	0.263	0.07	0.951	-	-	-	-	-
	mean_AIC			718.103					739.415					900.029		
	Percent_AIC			99%					0.5%					0. 5%		
N=500	$\beta_1 = 1$	0.998	0.144	0.139	0.021	0.946	1.084	0.134	0.095	0.025	0.903	1.681	0.217	0.155	0.51	0.123
	$\beta_2 = -0.5$	-0.506	0.113	0.116	0.013	0.961	-0.498	0.128	0.085	0.016	0.951	-0.483	0.201	0.138	0.041	0.953
	$\beta_1^* = 0.7$	0.623	0.229	0.234	0.058	0.951	0.398	0.146	0.132	0.112	0.448	0.921	0.153	0.144	0.073	0.717
	$\theta = 0.5$	0.453	0.139	0.126	0.022	0.948	0.064	0.028	0.03	0.191	0	2.076	0.067	0.119	2.489	0
	$\alpha = -2$	-1.728	0.402	0.399	0.236	0.903	-	-	-	-	-	0.562	0.05	0.056	6.567	0
	$\gamma_0 = -0.5$	-0.495	0.129	0.131	0.017	0.956	-0.494	0.13	0.116	0.017	0.953	-	-	-	-	-
	$\gamma_1 = 1$	0.991	0.183	0.185	0.034	0.951	0.986	0.183	0.163	0.034	0.951	-	-	-	-	-
	mean_AIC			1411.985					1465.70					1785.917		
	Percent_AIC			100%					0.00%					0.00%		
N=1000	$\beta_1 = 1$	0.999	0.102	0.099	0.01	0.952	2.21	1.14	2.289	1.31	0.98	1.697	0.157	0.072	0.51	0.014
	$\beta_2 = -0.5$	-0.505	0.083	0.082	0.007	0.96	-0.121	2.45	1.861	6.14	0.971	-0.46	0.145	0.064	0.022	0.957
	$\beta_1^* = 0.7$	0.616	0.159	0.164	0.032	0.913	1.77	9.51	1.972	91.6	0.971	0.916	0.106	0.067	0.058	0.475
	$\theta = 0.5$	0.471	0.098	0.094	0.011	0.945	0.064	0.023	0.002	0.19	0	2.083	0.048	0.055	2.507	0
	$\alpha = -2$	-1.665	0.261	0.277	0.18	0.736	-	-	-	-	-	0.565	0.035	0.026	6.581	0
	$\gamma_0 = -0.5$	-0.499	0.087	0.093	0.008	0.943	-0.023	1.68	0.284	2.89	0.968	-	-	-	-	-
	$\gamma_1 = 1$	0.993	0.129	0.131	0.017	0.945	1.22	1.61	0.469	2.65	0.971	-	-	-	-	-
	mean_AIC			2841.235					2948.954					3587.462		
	Percent_AIC			100%					0.00%					0.00%		

AIC, Akaike information criterion; CP, coverage probability; MSE, mean square error; SE, standard error

Table 5. *Application results*

Variables	Modalities	proposed model			Reduced Model 1 (Cure with same frailty)			Reduced Model 2 (Without cure with different frailty)		
		Est (SE)	HR	P-value	Est (SE)	HR	P-value	Est (SE)	HR	P-value
Recurrent events										
Age (ref:≤50)	>50	0.013(0.007)	1.013	0.64	0.014(0.008)	1.014	0.086	0.012(0.005)	1.012	0.018
Lymphovascular (ref:negative)	positive	0.008 (0.004)	1.008	0.067	0.008 (0.005)	1.008	0.091	0.007(0.003)	1.007	0.019
Lymph node Status(ref:negative)	positive	0.007 (0.004)	1.007	0.07	0.007 (0.004)	1.007	0.088	0.006(0.003)	1.006	0.023
Tumor size (ref:I)	II	0.012(0.007)	1.012	0.068	0.012(0.007)	1.012	0.076	0.01(0.004)	1.01	0.021
	III	0.008(0.005)	1.008	0.08	0.009(0.005)	1.009	0.083	0.006(0.002)	1.006	0.029
Cancer death										
Age (ref:≤50)	>50	0.375 (0.359)	1.45	0.231	0.355(0.346)	1.426	0.236	0.616(0.346)	1.852	0.081
Lymphovascular (ref:negative)	positive	0.938(0.503)	2.55	0.07	0.625(0.382)	1.868	0.105	0.419(0.374)	1.521	0.213
Lymph node status(ref:negative)	positive	0.621 (0.606)	1.89	0.236	-0.352 (0.477)	0.704	0.304	-0.624(0.449)	0.536	0.152
Tumor size (ref:I)	II	-0.725(0.446)	0.48	0.106	-0.259(0.442)	0.772	0.336	-1.024(0.438)	0.359	0.026
	III	-1.227(0.704)	0.29	0.87	-0.344(0.888)	0.709	0.37	-1.73(0.655)	0.177	0.012
Cure Logistic Model		Est (SE)	OR	P-value	Est (SE)	OR	P-value			
Intercept		0.867(0.742)	2.38	0.201	0.991(1.141)	2.693	0.274	-----	-----	-----
Tumor size	II	-0.651(0.657)	0.52	0.244	-1.067(0.833)	0.344	0.176	-----	-----	-----
	III	-1.47 (0.561)	0.23	0.013	-1.736 (1.059)	0.176	0.102	-----	-----	-----
θ =		0.904(0.332)	------	0.383	1.099(0.941)	------	0.397	1.288(0.173)	------	<0.1
α =		1.357(0.361)	------	<0.001	------	------	------	1.8(0.351)	------	<0.001
AIC		2062.394			2063.579			2080.42		

AIC, Akaike information criterion; HR, hazard ratio; SE, standard error

ISSN 2282-0930 • Epidemiology Biostatistics and Public Health - 2023, Volume 18, Number 1

Fig 1. Kaplan-Meier curve for the cancer free survival. The censoring time is denoted by "+".