Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2023, 2023:41
https://doi.org/10.28919/cmbn/7940
ISSN: 2052-2541

A MIXED-EFFECTS JOINT MODEL WITH SKEW-T DISTRIBUTION FOR LONGITUDINAL AND TIME-TO-EVENT DATA: A BAYESIAN APPROACH

MELKAMU M. FEREDE[1,2,*], SAMUEL M. MWALILI[3], GETACHEW A. DAGNE[4]

[1]Pan African University Institute for Basic Sciences, Technology and Innovation, Nairobi 62000-00200, Kenya

[2]Department of Statistics, University of Gondar, Gondar 196, Ethiopia

[3]Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi 62000-00200, Kenya

[4]Department of Epidemiology and Biostatistics, College of Public Health, University of South Florida, 13201 Bruce B. Downs, Tampa, FL 33612, USA

Abstract. Modelling longitudinal biomarkers and time-to-event processes jointly is becoming essential in medical research and other follow-up studies in order to evaluate their association, obtain unbiased results, and make valid statistical inferences. This study was motivated by follow-up data on chronic kidney disease (CKD), which is a major global health problem. Numerous studies have been conducted in the literature to analyse and assess the kidney function of CKD patients using cross-sectional data. However, joint models on CKD follow-up data have not been extensively studied in the literature. In the construction of joint models on CKD data, most previous studies proposed mixed-effects submodels with Gaussian distributions for longitudinal outcomes. However, longitudinal outcomes may have asymmetric (skewed) distributions. Proposing a normal distribution for skewed longitudinal data may yield biased results and invalid statistical inferences. In this paper, therefore, we propose a mixed-effects joint model with a skew-t distribution for longitudinal and time-to-event data under the Bayesian approach. We assessed the performance of the proposed joint model using simulation studies and applied the model to real CKD

[*]Corresponding author

E-mail address: melkamum2m@gmail.com

Received March 04, 2023

data. The proposed joint model with a skew-t distribution was compared with joint models with skew-normal and normal distributions of model errors. The findings of the simulation and application studies showed that the proposed joint model with skew-t distribution performed well.

Keywords: time-to-event data; longitudinal data with skewness; Bayesian joint modelling; skew-t distribution; chronic kidney disease.

2020 AMS Subject Classification: 62-08, 60E05, 62F15, 62F35, 62N02, 92C50.

1. Introduction

In follow-up studies, a group of subjects can be followed over time to determine the outcome of exposures, processes, or effects of a characteristic. In such studies, longitudinal and event-time data can be recorded simultaneously. In the literature, there are numerous approaches for modelling and evaluating the time-to-event and longitudinal processes separately. When the two processes are related, however, modelling them separately may give biased results. A joint modelling technique is needed to obtain an unbiased result, make valid statistical inferences, and assess the association between them [1].

Modelling the time-to-event and longitudinal data jointly is therefore becoming essential in many follow-up studies. In medical research, joint models of longitudinal biomarkers and time-to-event data have been a crucial statistical methodology and active research area [2, 3, 4, 5, 6, 7, 8]. In the formulation of joint models, the most common approaches postulated as submodels are mixed-effects models [1, 9, 10, 11, 7] for longitudinal measurements and a Cox's proportional hazard submodel [12, 13, 14, 15, 16] for survival data.

This study was motivated by a follow-up data on chronic kidney disease (CKD), which is a major global health problem that affects around half a billion people worldwide [17]. The majority of the prevalence of CKD is in low- and middle-income countries, and according to Shiferaw et al. [18], an estimated 35.52 percent of people with diabetes in Ethiopia had CKD prevalence. Several biomarkers of the progress of a patient's renal function can be measured over time, and event times can also be recorded. For instance, serum creatinine, albuminuria, and other biomarkers can be measured at each follow-up visit, and the glomerular filtration rate (GFR) of the patient can be estimated. In addition, the time to an event, i.e., time to end-stage renal disease (ESRD), death, or transplant, can be recorded if the patient experiences an event during the follow-up period. Here, the longitudinal outcome (eGFR, for example) and the survival outcome (time-to-ESRD) are biologically related. That is, the chance of developing ESRD can rise when a CKD patient's eGFR declines. As a result, a joint modelling approach can be more appropriate to model the two outcomes and evaluate their associations.

Numerous studies have been conducted in the literature to analyse and assess the kidney function of CKD patients using cross-sectional data. However, joint models on CKD follow-up data have not been extensively studied in the literature. Armero et al. [19] developed a joint model for longitudinal measurements and competing event times to analyse CKD data in children. They proposed Cox PH and linear mixed-effects submodels in the construction of their joint model. A study conducted by Yang et al. [20] proposed a joint model with mixed-effects and linear regression submodels for longitudinal and event-time CKD data, respectively. Teixeira et al. [21] also postulated a joint model for longitudinal and competing-events peritoneal dialysis data. The majority of these earlier studies postulated mixed-effects models with normal distributions for longitudinal outcomes.

However, longitudinal outcomes in some applications may have asymmetric (skewed) distributions. For instance, in this paper's application data, the longitudinal outcome eGFR exhibits a skew distribution (Figure 1). Figure 1 also demonstrates the existence of between- and within-patient variations and the asymmetric distribution of eGFR.

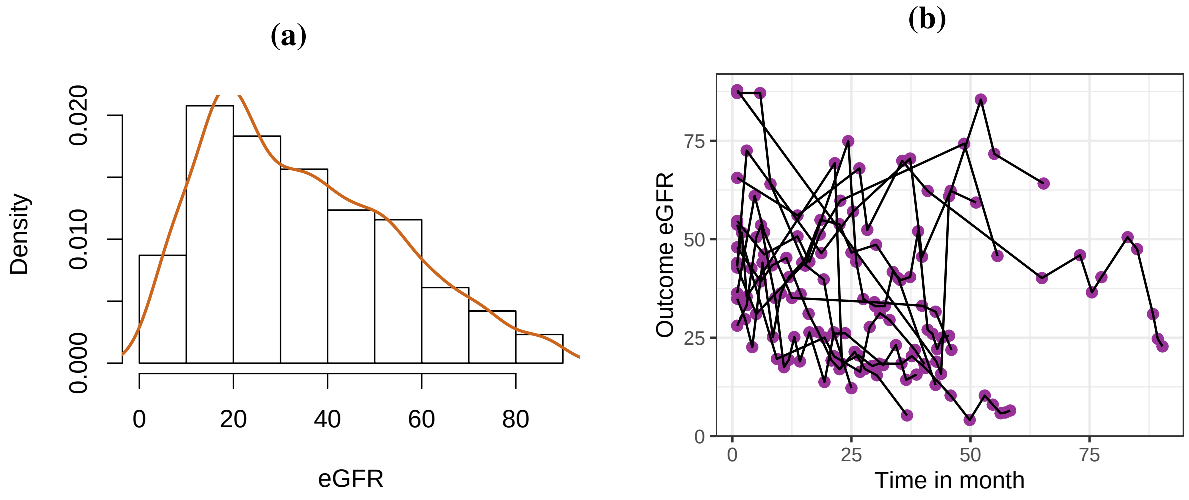

FIGURE 1. (**a**) histogram with pdf of eGFR and (**b**) trajectories of eGFR for some representative CKD patients.

Proposing a normal distribution for such skewed longitudinal data may yield biased results and invalid statistical inference [22]. As a result, recent studies suggest that more flexible distributional assumptions of model errors may be needed in order to accurately describe and model such complex longitudinal outcomes and make a valid statistical inference [23, 24]. In this paper, therefore, we propose a mixed-effects joint model with a skew-t distribution for longitudinal and time-to-event data under the Bayesian approach.

The reminder of this paper is organized as follows. Section 2 briefly describes the methods. In this section, the construction of the joint model, parameter estimation and inference, and model comparison are briefly described. In Section 3, the simulation studies are presented. Analysis of the CKD data is presented in Section 4. Section 5 includes conclusion and suggestions for future work.

2. Methods

2.1. Notations and the Joint model. Let y_{ij} denotes the longitudinal outcome for subject i measured at the j^{th} time t_{ij}: $i = 1,...,m,\ j = 1,...,m_i$. For convenience, the longitudinal outcome and measurement time can be rewritten in vector form as $\mathbf{y}_i = (y_{i1},...,y_{ij},...,y_{im_i})^T$ and $\mathbf{t}_i = (t_{i1},...,t_{ij},...,t_{im_i})^T$, respectively. Let S_i, T_i and C_i denote the observed event-time, true event-time, and censoring time, respectively, for the i^{th} subject. Where T_i can be computed as $T_i = min(S_i,C_i)$. Let ρ be an event indicator that can be 1, the event of interest, or 0, the censoring event.

The construction of the joint model consists of the longitudinal submodel (1) and the event-time submodel (2). A skewed mixed-effects submodel is proposed for the longitudinal outcome $\mathbf{y}_i$ and is defined by

$$\begin{aligned} &\mathbf{y}_i = \mathbf{X}_i^T\beta + \mathbf{R}_i^T\xi_i + \varepsilon_i, \qquad i = 1,...,m \\ &\varepsilon_i \sim ST_{m_i,\vartheta_\varepsilon}(\mu_\varepsilon,\Sigma_\varepsilon,\delta_\varepsilon), \quad \xi_i \sim N_l\left(\mathbf{0},\Sigma_\xi\right) \end{aligned} \tag{1}$$

Where $\mathbf{X}_i = (\mathbf{x}_{1i},...,\mathbf{x}_{pi})^T$ and $\mathbf{R}_i = (\mathbf{r}_{1i},...,\mathbf{r}_{li})^T$ represent fixed-effects and random-effects covariate design matrices, respectively. Where p and l are dimensions of the design matrices. β and ξ_i are associated parameter vectors of the fixed- and random-effects covariates. Due to its computational simplicity, in most previous studies, the model errors were considered to follow the commonly used Gaussian distribution. However, as shown in the introduction section (Figure 1), the real longitudinal outcome data of this study follow an asymmetric distribution. As a result, we propose a multivariate skew-t distribution [25] for model errors $\varepsilon_i = (\varepsilon_{i1}, ..., \varepsilon_{ij}, ..., \varepsilon_{im_i})^T \sim ST_{m_i,\vartheta_\varepsilon}(\mu_\varepsilon, \Sigma_\varepsilon, \delta_\varepsilon)$. Where ϑ_ε, $\mu_\varepsilon = \mathbf{0}$, $\Sigma_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_{m_i}$, and $\delta_\varepsilon = \delta_\varepsilon \mathbf{1}_{m_i}$ represent the degree freedom , mean vector, covariance matrix, and skewness vector of ε_i, respectively. Σ_ξ represents the covariance matrix of ξ_i; and $\mathbf{1}_{m_i} = (1, ..., 1)^T$ represents an identity vector.

For the time-to-event process, we propose an extended proportional hazard model:

$$\lambda_i(t; \mathbf{Z}_i, \xi_i) = \lambda_0(t) \exp\left\{\gamma^T \mathbf{Z}_i + \eta^T \xi_i\right\} \tag{2}$$

Where ξ_i represents the effects of subject-specific longitudinal outcomes (random effects), $\mathbf{Z}_i$ is the baseline covariates vector with coefficient vector γ, $\lambda_i(t;.)$ represents the risk rate of the event of interest at time t, and $\lambda_0(t)$ is the baseline hazard function. Given ξ_i, the longitudinal outcome $\mathbf{y}_i$ and the event-time T_i are assumed to be conditionally independent. η is a parameter vector that denotes the level of association between the hazard rate of the event of interest and the subject-specific longitudinal outcome.

The survival function, the likelihood that subject i won't experience the event of interest after time t, is given by:

$$S(t) = \exp\left\{-\int_0^t \lambda_0(v) \exp\left(\gamma^T \mathbf{Z}_i + \eta^T \xi_i\right) dv\right\} \tag{3}$$

In this paper, the piecewise constant function is used to specify the baseline hazard function $\lambda_0(t)$ and model the event-time more flexibly [26]. By partitioning the event-time into P intervals, $0 = l_0 < l_1 < \ldots < l_{p-1} < l_P < \infty$, the baseline hazard can be specified as $\lambda_0(t) = \lambda_p,\ \ for\ \ t \in (l_{p-1}, l_p],\ \ \ p = 1, ..., P.$

2.2. Estimation and Inference.

The likelihood (ML) and Bayesian methods are the two approaches that are most commonly employed for parameter estimation in the literature. Since we proposed a joint model with multivariate skew-t distribution, estimating the parameters from the joint likelihood using the ML-method can be quite time-consuming. Hence, in this paper, to estimate all the parameters simultaneously, we adopted the Bayesian technique, which can reduce computational load and allow for the inclusion of prior knowledge for the parameters.

For Markov chain Monte Carlo (MCMC) computation, it is necessary to specify the proposed mixed-effects longitudinal submodel (1) with a skew-t distribution. As a result, to represent the skew-t distribution based on the stochastic representation [25], we introduced a random vector $\mathbf{W}_{\varepsilon i} = (W_{\varepsilon i1}, \ldots, W_{\varepsilon i m_i})^T$ and a random variable v_ε. Thus, the hierarchical reformulation of the joint model can be given by:

$$\begin{aligned} \mathbf{y}_i | \xi_i, \mathbf{W}_{\varepsilon i}, v_{\varepsilon i}; \theta_y &\sim N_{m_i}\left(\mathbf{X}_i^T \beta + \mathbf{R}_i^T \xi_i + \delta_\varepsilon \mathbf{W}_{\varepsilon i}, v_{\varepsilon i}^{-1} \sigma_\varepsilon^2 \mathbf{1}_{m_i}\right), \\ \mathbf{W}_{\varepsilon i} | v_{\varepsilon i} &\sim N_{m_i}(\mathbf{0}, v_{\varepsilon i}^{-1} \mathbf{I}_{m_i}) I(\mathbf{W}_{\varepsilon i} > \mathbf{0}) \\ v_{\varepsilon i} | \vartheta_\varepsilon &\sim \Gamma(\vartheta_\varepsilon/2, \vartheta_\varepsilon/2) \\ \xi_i | \Sigma_\xi &\sim N_l\left(\mathbf{0}, \Sigma_\xi\right), \\ T_i | \xi_i; \lambda, \gamma, \eta &\sim \int^{T_i} f(t | \xi_i; \lambda, \gamma, \eta) dt \end{aligned} \tag{4}$$

Where $\theta_y = \{\beta, \sigma_\varepsilon^2, \Sigma_\xi, \delta_\varepsilon\}$, $N_{m_i}(.)$ is m_i-variate normal distribution, and $\Gamma(.)$ is a gamma distribution.

Let $\mathscr{D}$ be the overall observed longitudinal and event-time data, and $\Omega = \{\beta, \sigma_\varepsilon^2, \Sigma_\xi, \delta_\varepsilon, \vartheta_\varepsilon, \lambda, \gamma, \eta\}$ be the set of all parameters of the joint model. Then, the joint likelihood function of $\mathscr{D}$ is given by

$$\begin{aligned} f(\mathscr{D}|\Omega) = \prod_{i=1}^{m} \int_{\xi_i} & f(\mathbf{y}_i | \xi_i, \mathbf{X}_i, \mathbf{W}_{\varepsilon i}, v_{\varepsilon i}; \theta_y) \\ & \times f(\xi_i | \Sigma_\xi) f(\mathbf{W}_{\varepsilon i} | v_{\varepsilon i}, \mathbf{W}_{\varepsilon i} > \mathbf{0}) f(v_{\varepsilon i}) \\ & \times f(T_i, \rho_i | \xi_i, \mathbf{Z}_i; \theta_t) d\xi_i \end{aligned} \tag{5}$$

The prior distribution for each parameter, hence, must be specified in order to approximate the posterior distribution of the hierarchically constructed joint model. Independent normal $N_p(\beta_0, \Omega_\beta)$, $N_{p'}(\gamma_0, \Omega_\gamma)$, $N_{q'}(\eta_0, \Omega_\eta)$, and $N(0, \kappa_{\delta_\varepsilon})$ prior distributions are assumed for β, γ, η, and δ_ε, respectively. Inverse-Wishart $IW_l(\mathbf{D}_\xi, v_\xi)$ prior is assumed for Σ_ξ and Inverse-Gamma $IG(\rho_{\varepsilon 1}, \rho_{\varepsilon 2})$ prior is assumed for σ_ε^2. Independent gamma $G(\phi_1, \phi_2)$ priors are assumed for λ_p. For the degree freedom ϑ_ε, a truncated exponential $Exp(\vartheta_{\varepsilon 0}) I(\vartheta_\varepsilon > 3)$ prior distribution is assumed. The hyperparameter matrices $\Omega_\beta, \mathbf{D}_\xi, \Omega_\gamma$ and Ω_η are assumed diagonal for convenient implementation.

Given the joint likelihood $f(\mathscr{D}|\Omega)$ and joint prior distribution $\pi(\Omega)$, the joint posterior density can be derived as

$$
\begin{aligned}
\pi(\Omega|\mathscr{D}) &\propto f(\mathscr{D}|\Omega) \times \pi(\Omega) \\
&\propto \prod_{i=1}^{m} \int_{\xi} \Bigg\{ \left(\sigma_{\varepsilon}^{2}\right)^{-m_i/2} exp\left(-\frac{1}{2}\left(\mathbf{y}_i-\mu_y\right)^T\left(\frac{\sigma_{\varepsilon}^{2}\mathbf{I}_{m_i}}{v_{\varepsilon i}}\right)^{-1}\left(\mathbf{y}_i-\mu_y\right)\right) \\
&\times |\Sigma_{\xi}|^{-\frac{1}{2}} exp\left(-\frac{1}{2}\xi_i^T\Sigma_{\xi}^{-1}\xi_i\right) \times \exp\left(-\frac{1}{2}v_{\varepsilon i}\mathbf{W}_{\varepsilon i}^{T}\mathbf{W}_{\varepsilon i}\right) \\
&\left\{\frac{1}{\Gamma(\vartheta_{\varepsilon}/2)(\vartheta_{\varepsilon}/2)^{\vartheta_{\varepsilon}/2}} v_{\varepsilon i}^{\frac{\vartheta_{\varepsilon}}{2}-1}\right\} exp\left(-\frac{2}{\vartheta_{\varepsilon}}v_{\varepsilon i}\right) \\
&\times \lambda_0(t)^{\rho_i}\exp\left(\left(\gamma^T\mathbf{Z}_i+\eta^T\xi_i\right)\rho_i\right) \\
&\times \exp\left(-\int_0^t \lambda_0(v)\exp\left(\gamma^T\mathbf{Z}_i+\eta^T\xi_i\right)ds\right)\Bigg\} d\xi_i \\
&\times exp\left(-\frac{1}{2}(\beta-\beta_0)^T\Omega_{\beta}^{-1}(\beta-\beta_0)\right) \\
&\times (\sigma_{\varepsilon}^{2})^{-\rho_{\varepsilon 1}-1} exp(-\rho_{\varepsilon 2}/\sigma_{\varepsilon}^{2}) \\
&\times |\Sigma_{\xi}|^{-\frac{(v_{\xi}+1)}{2}} exp\left(-\frac{1}{2}\mathrm{tr}\left(\Omega_{\xi}\Sigma_{\xi}^{-1}\right)\right) \\
&\times exp\left(-\frac{1}{2\kappa_{\delta_{\varepsilon}}}\delta_{\varepsilon}^{2}\right) \times exp\left(-\vartheta_{\varepsilon 0}\vartheta_{\varepsilon}\right\} \\
&\times exp\left(-\frac{1}{2}(\gamma-\gamma_0)^T\Omega_{\gamma}^{-1}(\gamma-\gamma_0)\right) \\
&\times exp\left(-\frac{1}{2}(\eta-\eta_0)^T\Omega_{\eta}^{-1}(\eta-\eta_{q0})\right) \times \lambda_p^{\phi_{p1}-1} exp\left(-\phi_{p2}\lambda_p\right)
\end{aligned}
\tag{6}
$$

where $\pi(\Omega)$ is the product of the prior distributions of the parameters, $\mu_y = \mathbf{X}_i\beta + \mathbf{H}_i\xi_i + \delta_{\varepsilon}\mathbf{W}_{\varepsilon i}$.

To draw a sample from the conditional posterior distribution and estimate the posterior mean and standard deviation of each of the parameters, we utilised the Metropolis-Hastings algorithm within Gibbs sampler and implemented the MCMC algorithm in WinBUGS software.

2.3. Model comparison.

We compare joint models by considering different distributional specifications of the model errors from the longitudinal submodel:

- **JModel I**: A joint model with multivariate skew-t (ST) distribution of model errors ε_i.

- **Model II**: A joint model with multivariate skew-normal (SN) distribution of ε_i.
- **Model III**: A joint model with multivariate normal (N) distributions of ε_i.

3. Simulation Studies

To assess and compare the performance of the joint models, simulation studies are carried out. To conduct the simulation, 400 individuals with eleven measurement times (equally-spaced), i.e., a total of 4,400 observations, are taken into account. The mixed-effects longitudinal submodel (1) with two binary covariates was used to simulate the longitudinal outcome data. In order to obtain skewed longitudinal data, each component of the model error vector ε_i is first generated from a gamma distribution $G(2,1)$, and then deducted by two [17]. We set $\beta = (4.70, -0.25, -0.27, -0.24)^T$, $\gamma = (-0.01, 0.20, 1.35, 2.43)$ and $\eta = (-3.40, -1.45)$.

The Cox proportional hazard model (2) with constant baseline hazard ($\lambda_0(t) = 0.2$) and four predictors (one continuous and three binary predictors) was used to simulate event-time data. Censoring time C_i is generated from an $exp(0.5)$ distribution. We generated the random effects vector ξ_i from a four-variate normal distribution with a mean of $\mathbf{0}$ and an identity diagonal covariance matrix.

Weakly-informative prior distributions for the parameters were taken into consideration when performing the Bayesian inference. That is, for each component of β, δ_e, γ, λ, and η, a $N(0,100)$ prior is assumed. For σ_e^2, Σ_ξ, and ϑ_ε, an $IG(0.01, 0.01)$, an $IW(diag(0.01, 0.01), 2)$, and $Exp(0.5)$ prior distributions, respectively, are assumed.

In order to run the MCMC procedure in the Bayesian framework, three chains, each with 135,000 iterations and a burn-in of 60,000, were utilised in R2WinBUGS in R. By keeping every 50^{th} MCMC sample from the next 75,000, we obtained 4,500 simulated samples of the unknown parameters from each joint model. The Brooks-Gelman-Rubin (BGR) diagnostics plot, an autocorrelation plot, and a trace plot were all used to evaluate convergence. To assess the behaviour of the estimators in each joint model and to compare the models, we computed the relative bias (RB), 95% coverage probability (CP), root-mean-square (RMS) error, and deviance information criterion (DIC). A model with smaller values of DIC, RMS, and RB and a larger value of CP can be considered a better-performing joint model.

The simulation results, i.e., the posterior mean estimates with the corresponding RB, RMS, CP and DIC for each parameters of the joint models, are shown in Table 1. Thus, we found that the JModel I (a joint model with skew-t distribution) has performed better.

Para	**TPV**	**Method**	**Jmodel I**	**Jmodel II**	**Jmodel III**
β_1	4.70	RB	-0.066	-0.367	0.013
		RMS	1.031	1.726	0.083
		CP	100.00	50.00	80.70
β_2	-0.25	RB	-0.017	-0.095	-0.080
		RMS	0.070	0.066	0.074
		CP	94.95	93.08	93.97
β_3	-0.27	RB	0.002	-0.093	-0.082
		RMS	0.054	0.053	0.059
		CP	94.80	91.25	92.98
β_4	-0.24	RB	-0.027	-0.268	0.066
		RMS	0.042	0.066	0.024
		CP	7.00	72.30	84.97
σ_ε^2	0.70	RB	1.444	-0.284	3.107
		RMS	1.340	0.204	2.176
		CP	89.45	61.55	50.00
$\sigma_{\xi_1}^2$	0.07	RB	0.023	-0.148	0.074
		RMS	0.023	0.018	0.023
		CP	94.82	90.67	94.43
$\sigma_{\xi_2}^2$	0.02	RB	0.188	0.166	0.194
		RMS	0.006	0.005	0.006
		CP	84.67	85.50	85.02
γ_1	-0.01	RB	0.126	0.225	0.070
		RMS	0.008	0.008	0.009
		CP	94.82	93.57	94.90

γ_2	0.20	RB	-0.790	-0.845	-0.871
		RMS	0.271	0.289	0.293
		CP	88.65	88.13	88.93
γ_3	1.35	RB	0.050	0.017	0.045
		RMS	0.282	0.261	0.288
		CP	93.45	94.97	94.60
γ_4	2.43	RB	0.149	0.136	0.147
		RMS	0.564	0.549	0.523
		CP	86.42	89.48	85.60
η_1	-3.40	RB	-0.580	0.881	0.267
		RMS	6.497	3.938	4.272
		CP	95.75	81.23	96.52
η_2	-1.45	RB	-0.173	0.415	1.536
		RMS	2.513	1.998	2.938
		CP	94.07	93.53	83.38
DIC			23640	27140	25360

Table 1: Simulation results: The true parameter value (TPV), RB, CP, and RMS for each parameter of the joint models and the DIC value.

4. Application: Analysis of the CKD Data

In this paper, eight years of follow-up data, between June 2014 and June 2022, on chronic kidney disease (CKD) are used to apply the proposed joint model. The data was collected from the University of Gondar Comprehensive Specialised Hospital in Ethiopia. Medical records and patients' profiles (or charts) were used as sources of the data. The data comprises 198 CKD patients' baseline characteristics, comorbidities, repeatedly measured kidney function biomarkers, and time to event (s). The estimated glomerular filtration rate (eGFR) is used as a longitudinal response variable that measures the progressive loss of kidney function. In order to adequately capture a wide range of possible trajectories of a patient's kidney function over time and model it properly, patients with an eGFR value of less than ninety are included in the analysis. The CKD patients were around 55 years old on average, and 56.6% of them were men. Among the CKD patients, baseline prevalences of hypertension and diabetes were 34.4% and 23.81%, respectively.

The event of interest in this study was an end-stage renal disease (ESRD) event, and 31.2% of patients experienced it over the follow-up period. Figure 2 demonstrates the plots of the survival probabilities of CKD patients. The figure clearly shows a fast decline in the likelihood that CKD patients would be free of ESRD beyond a given time.

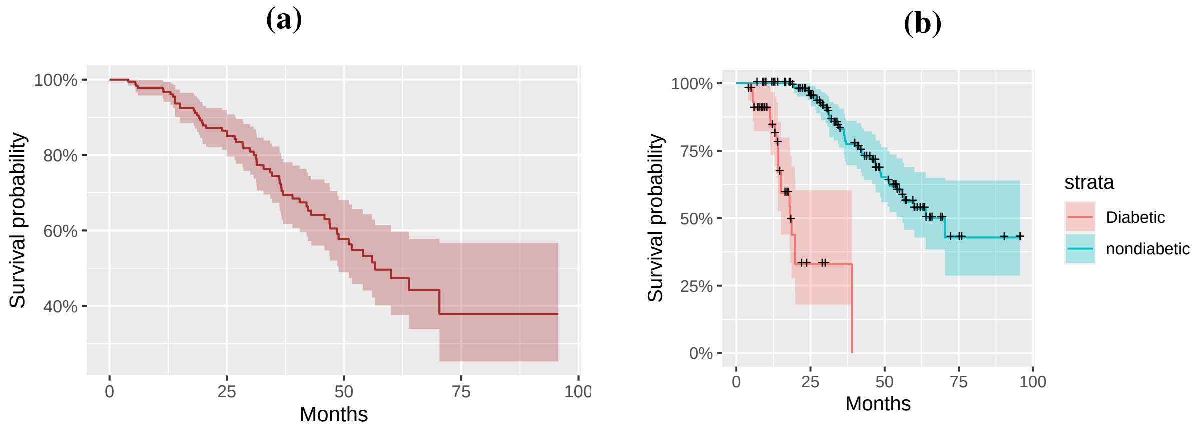

FIGURE 2. Plots of the survival probability of CKD patients. Plot (**a**) is the survival provability curve and shows a decreasing probability that a patient survives from ESRD beyond a specific follow-up time. Plot (**b**) demonstrates that CKD patients without diabetes had a higher chance of surviving than those with diabetes.

To specify and apply the proposed models to the CKD data, log-transformed eGFR is used as the longitudinal outcome, and diabetes, hypertension, and log-transformed measurement time are taken into account as associated covariates for the longitudinal submodel (1). Both the random intercept and random slope of time-effects are also considered. For the time-to-ESRD joint model (2), the baseline covariates age, sex, diabetes, hypertension, and the subject-specific eGFR process (the random effects) are taken into account to predict the hazard rate of ESRD for each patient. The time that a CKD patient does not yet experience an ESRD event until the end of the data collection period or withdraws from the follow-up is considered as a right-censoring time. We considered four intervals based on the quantiles of the observed event time to specify the baseline hazard function using piecewise constant functions.

In order to fit and compare the three joint models using the real CKD data, we employed the same specifications of the priors, MCMC computation techniques, and convergence assessment tools as in the simulation section (Section 3). The summary results of the fitted joint models' parameters with different distributions of model errors are presented in Table 2.

TABLE 2. The posterior mean estimates (PME), standard deviation (SD), 95% credible interval (CI), and DIC for the parameters from the proposed joint models with different distributions.

Prs	**JModel I**			**JModel II**			**JModel III**		
	PME	**SD**	**CI**	**PME**	**SD**	**CI**	**PME**	**SD**	**CI**
β_1	4.69	0.05	(4.58, 4.80)	4.70	0.05	(4.60, 4.80)	4.37	0.05	(4.27, 4.47)
β_2	-0.24	0.07	(-0.37, -0.10)	-0.26	0.07	(-0.39, -0.11)	-0.22	0.07	(-0.36, -0.08)
β_3	-0.29	0.02	(-0.33, -0.24)	-0.28	0.03	(-0.34, -0.23)	-0.33	0.03	(-0.38, -0.28)
β_4	-0.23	0.02	(-0.26, -0.20)	-0.23	0.02	(-0.26, -0.20)	-0.24	0.02	(-0.27, -0.21)
σ^2_ε	0.008	0.002	(0.004, 0.012)	0.024	0.004	(0.016, 0.032)	0.117	0.005	(0.108, 0.127)
$\sigma^2_{\xi_1}$	0.156	0.028	(0.108, 0.217)	0.162	0.026	(0.116, 0.219)	0.174	0.023	(0.112, 0.203)
$\sigma_{\xi_{12}}$	-0.004	0.008	(-0.02, 0.01)	0.000	0.008	(-0.02, 0.014)	0.006	0.007	(-0.01, 0.021)
$\sigma^2_{\xi_2}$	0.030	0.004	(0.023, 0.039)	0.030	0.004	(0.023, 0.039)	0.032	0.004	(0.024, 0.041)
δ_ε	-0.54	0.03	(-0.60, -0.48)	-0.51	0.02	(-0.54, -0.47)	–	–	–
ϑ_ε	3.76	0.70	(3.02, 5.56)	–	–	–	–	–	–
γ_1	-0.01	0.01	(-0.03, 0.01)	-0.01	0.01	(-0.03, 0.01)	-0.01	0.01	(-0.02, 0.01)
γ_2	0.16	0.30	(-0.41, 0.76)	0.19	0.30	(-0.40, 0.77)	0.18	0.30	(-0.41, 0.76)
γ_3	1.74	0.51	(0.77, 2.76)	1.73	0.49	(0.76, 2.70)	1.65	0.52	(0.64, 2.70)
γ_4	2.33	0.52	(1.26, 3.31)	2.32	0.53	(1.26, 3.35)	2.30	0.53	(1.24, 3.34)
η_1	-1.13	0.44	(-2.01, -0.26)	-1.21	0.41	(-2.02, -0.42)	-1.19	0.42	(-2.00, -0.35)
η_2	-4.92	1.46	(-8.04, -2.32)	-4.65	1.44	(-7.67, -2.04)	-4.84	1.45	(-7.95, -2.17)
DIC	4842			6138			5579		

Since the 95% credible intervals of most of the parameters do not include zero, as we can clearly see from Table 2, each joint model results in slightly different but statistically significant estimates. In general, compared to the skew-t joint models (JModels-I), the Gaussian joint model (JModels-III) yield larger parameter estimates. In particular, the last two joint models yields relatively large estimates of the within- and between-subject variations of the longitudinal outcome. For example, the estimated values of σ_{ε}^{2} (within-patient variation of eGFR) and $\sigma_{\xi_1}^{2}$ (inter-patient variation of eGFR) are relatively large in JModel-III.

The estimate of the skewness parameter (Table 2) is statistically significantly different from zero ($\hat{\delta}_{\varepsilon} = -0.54;\ 95\%CI : -0.60, -0.48$), confirming that the joint model with a skew-t distribution is more appropriate than the Gaussian joint model. In addition, we also used DIC to select the best-fitting joint model. As we can see from Table 2, the proposed joint model (JModel-I), in comparison to the other joint models, has a relatively lower DIC value. Thus, the smaller variance estimates, lower DIC value, and existence of significant skewness lead us to conclude that JModel-I is the best Bayesian joint model that fits the CKD data well, and we use its findings to interpret the results and draw conclusions.

Convergence diagnostic checking was done before interpreting the fitted chosen joint model's results and drawing conclusions. Figure 3 presents the trace plots, and Figure 4 shows the plots of the BGR, ACF, and density of the parameters from the proposed joint model with skew-t distribution. All the figures clearly show convergence.

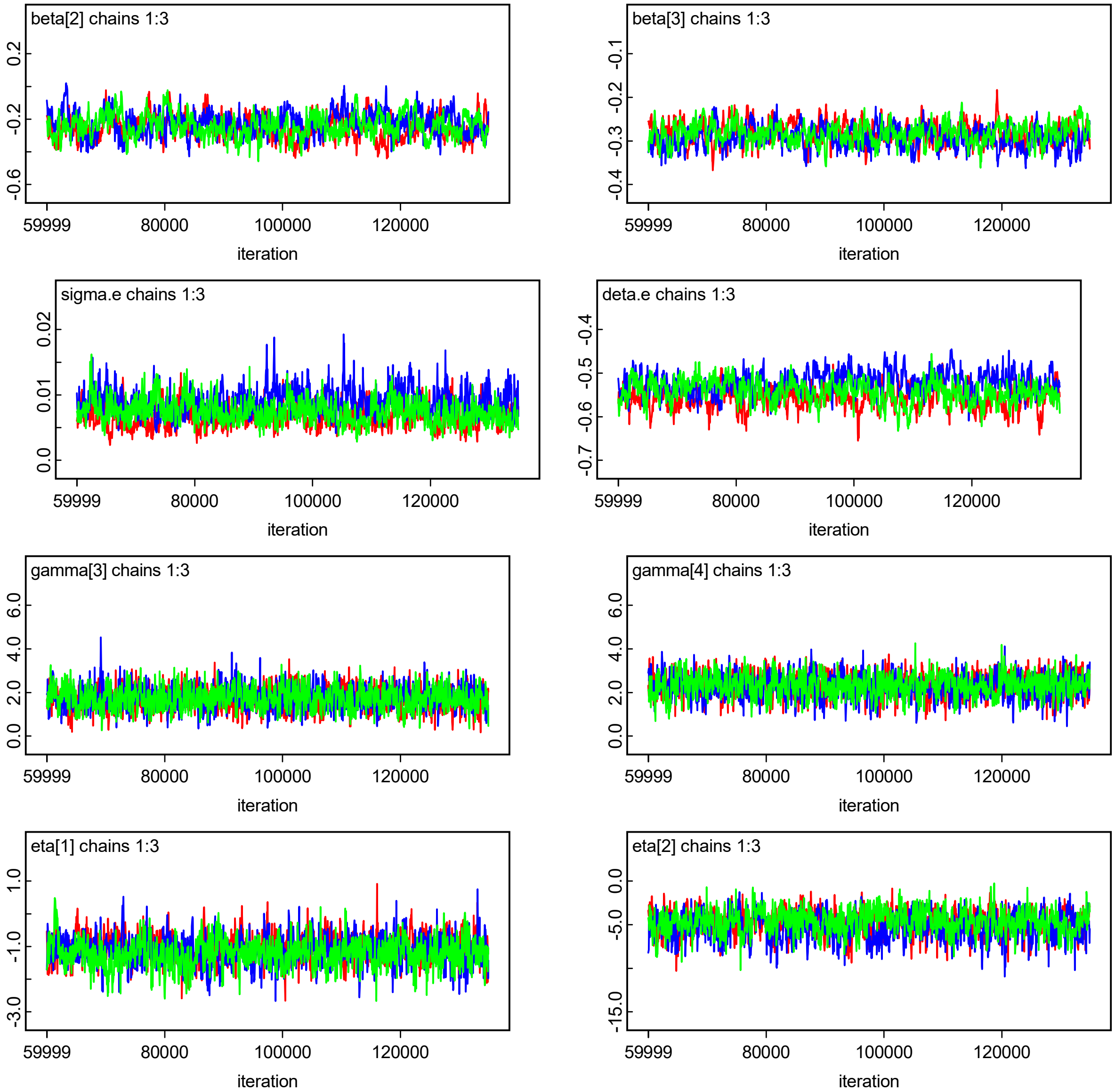

FIGURE 3. Trace plots of some JModel-I parameters. The plots demonstrate that the three chains of the MCMC, for each parameter, are well mixed, indicating convergence.

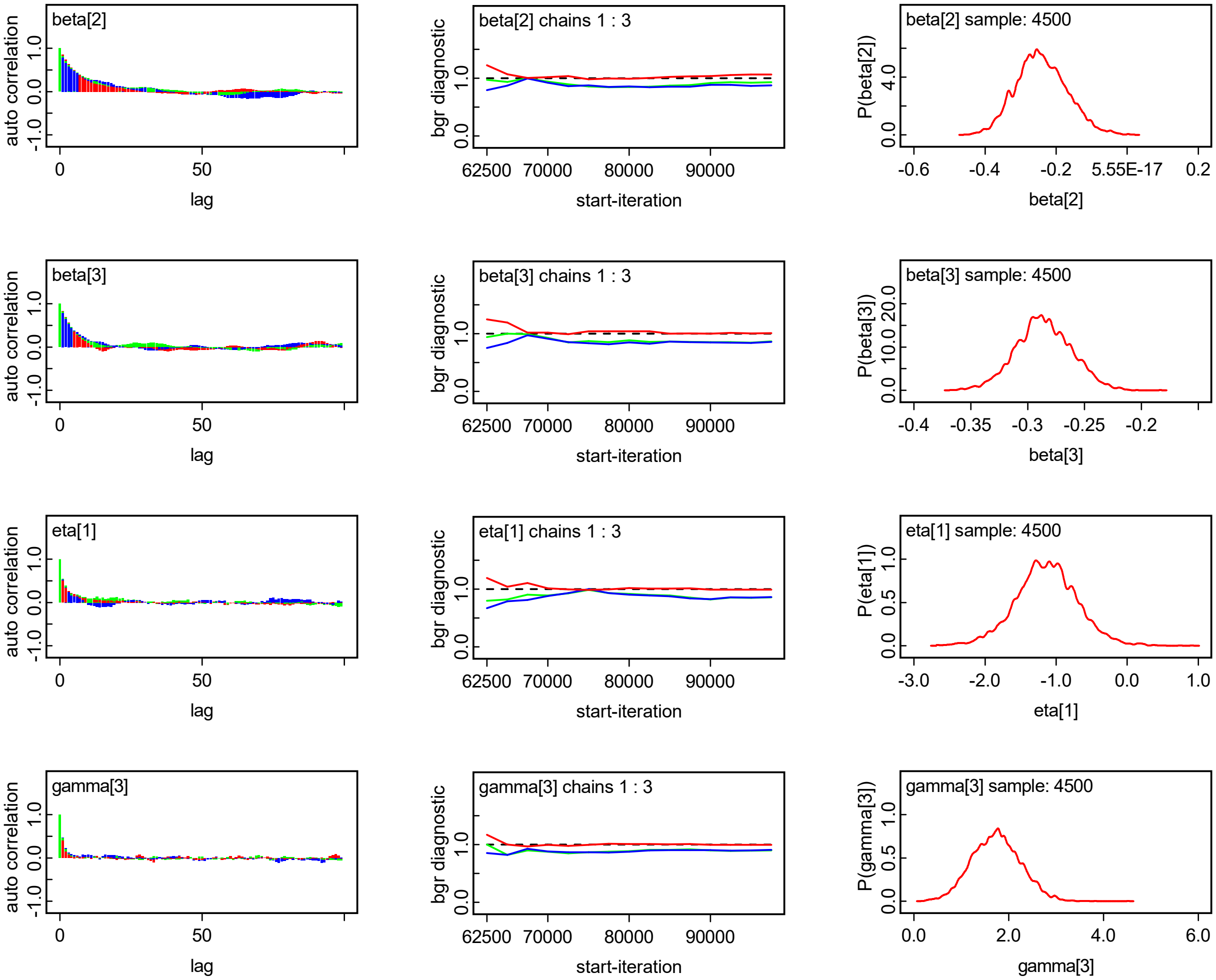

FIGURE 4. (**a**) ACF plots, (**b**) BGR plots, and (**c**) posterior density plots of some JModel-I parameters. For each parameter, the chains' correlation with their successive lags is low, and the ratio of the BGR plots approaches 1, indicating convergence.

The association parameters of the subject-specific longitudinal outcome eGFR and the time-to-ESRD processes resulting from JModel-I are statistically significant ($\hat{\eta}_1 = -1.13;\ 95\%CI : -2.01, -0.26$; and $\hat{\eta}_2 = -4.92;\ 95\%CI : -8.04, -2.32$). This demonstrates that proposing a joint modelling approach for the two processes is reasonable. $\hat{\eta}_1 = -1.13[HR = 0.32]$, for instance, can be interpreted as the risk of a CKD patient developing ESRD being decreased by 68% when the patient-specific mean log-eGFR increases by one unit.

The findings of the JMode-I in Table 2 also show that the estimates of parameters for the covariates hypertension and diabetes (β_2 and β_3) in the longitudinal outcome eGFR submodel and (γ_3 and γ_4) in the event-time submodel are significantly different from zero. The baseline patients' age and gender, however, do not significantly predict the hazard rate of ESRD. This is due to the fact that the 95% credible intervals of the estimates of their parameters (γ_1 and γ_2) contain zero.

The negative coefficients of hypertension, diabetes and log-measurement time (β_4) indicate that the mean log of the longitudinal outcome eGFR has a negative association with them. For instance, $\hat{\beta}_2 = -0.24$ (95% CI: $[-0.37, -0.10]$) indicates that, compared to a CKD patient without diabetes, the mean log-eGFR value of a CKD patient with diabetes can be declined by $0.24\ mL/min/1.73\ m2$, assuming the effect of other covariates as constant. The decline of the mean log-eGFR value of a CKD patient with diabetes can be up to $0.37\ mL/min/1.73\ m2$. The estimates of β_3 and β_4 can be interpreted similarly as β_2.

As stated above, diabetes and hypertension are also significantly and strongly associated with the instantaneous rate of ESRD. For example, according to the estimate of the hypertension coefficient, $\hat{\gamma}_4 = 2.33$ $[HR = 10.28]$, a CKD patient who is hypertensive is 10.28 times more likely than a non-hypertensive patient to develop ESRD.

5. Conclusion and Suggestions

According to current literature suggestions, joint modelling of complex longitudinal and event-time clinical data is an active medical research field. The main objective of this research was to develop a joint model for skewed longitudinal and event-time data and make a valid statistical inference using the Bayesian approach. To flexibly model the skewed longitudinal eGFR data, a mixed-effects submodel with multivariate skew-t distribution was proposed. To model the event-time (time-to-ESRD) data and accommodate more hazard shapes, a Cox proportional hazard submodel with a piecewise constant baseline hazard function was postulated.

We assessed the proposed joint model's performance using simulation studies and applied the model to real chronic kidney disease data. The proposed joint model with a skew-t distribution was compared with joint models with skew-normal and normal distributions of model errors. The relative bias, root mean square error, coverage probability, and DIC were used as performance evaluation tools. A joint model with skew-t distribution better fitted the data compared to the other joint models. We then evaluated the association between or the impact of the patient-specific eGFR process on the time to ESRD and other covariates and interpreted the results.

The association parameters of the subject-specific longitudinal outcome eGFR and the time-to-ESRD processes were statistically significant, indicating that proposing a joint modelling approach for the two processes is reasonable. The findings of this study also suggest that the specifications of the distributional assumptions of model errors require special attention. According to the application's findings, diabetes, hypertension, and measurement time were significant predictors of and had a negative association with kidney function. Hypertension and diabetes are also significantly associated with high risks of experiencing end-stage renal disease.

We note that one can use our methodology by considering other additional biomarkers of kidney function and assessing their associations with the time-to-events and making valid statistical inferences and predictions (which are out of the scope of this study). In addition to the motivating CKD follow-up data, our methodology has broader applications whenever continuous outcomes and associated biomarkers are repeatedly measured, the time-to-event is recorded, and the basic submodels and joint model specifications are met. Our simulation and application studies revealed that our work contributed to this interesting study area by making use of a more flexible methodology to model skewed longitudinal and time-to-event data.

The methodology proposed in this paper has some extensions for future research. (i) In this paper, we considered only the time-to-ESRD as an event of interest, assuming independent censoring. That is, other failure events, such as a patient's death, are taken into account as independent censoring of the time-to-ESRD process. However, death can be regarded as a competing risk for ESRD because its occurrence may prevent the occurrence of ESRD. Thus, our methodology can be expanded by taking into account multiple failure types and proposing competing risk failure time submodels for the time-to-event processes. (ii) For the longitudinal outcome, eGFR, we proposed a fully parametric (linear) mixed-effects submodel. However, in some applications, the exact form of the relationship between the longitudinal outcomes and the time effects may be non-linear (irregular). For instance some CKD patients in our application data have non-linear trajectories of the outcome eGFR over time. As a result, a fully parametric modelling approach may not be flexible enough to model such types of complex longitudinal data. Thus, by considering non-parametric smoothing functions of time, our modelling approach can be extended to a more flexible semi-parametric modelling approach and ensure future work. We note that the investigation of these two issues has also been completed and is currently available. (iii) The methodology of this work could also be extended to a multivariate setting to accommodate multiple longitudinal outcomes that are repeatedly measured for each subject.

ABBREVIATIONS

ACF	Autocorrelation Function
CKD	Chronic Kidney Disease
CI	Credible Interval
DIC	Deviance Information Criterion
eGFR	estimated Glomerular Filtration Rate
ESRD	End Stage Renal Disease
GFR	Glomerular Filtration Rate
HR	Hazard Ratio
MCMC	Markov chain Monte Carlo
MDRD	Modification of Diet in Renal Disease
RMSE	Root Mean Squared Error
SCr	Serum Creatinine
SN	Skew-Normal
ST	Skew-T

ACKNOWLEDGEMENTS

We would like to express our appreciation to the Pan African University Institute of Basic Sciences, Technology and Innovation (PAUSTI) for the support. Our appreciation also goes to the University of Gondar Comprehensive Specialized Hospital for the data access.

ETHICS APPROVAL

Ethical approval for this study was obtained from both the Jomo Kenyata University of Agriculture and Technology (JKUAT) Institutional Ethics Review Committee in Kenya and the University of Gondar (UOG) Ethical Review Board in Ethiopia.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] A.A. Tsiatis, M. Davidian, Joint modeling of longitudinal and time-to-event data: An overview, Stat. Sinica. 14 (2004), 809-834. https://www.jstor.org/stable/24307417.

[2] Y.Y. Chi, J.G. Ibrahim, Joint models for multivariate longitudinal and multivariate survival data, Biometrics. 62 (2005), 432-445. https://doi.org/10.1111/j.1541-0420.2005.00448.x.

[3] L. Wu, W. Liu, G.Y. Yi, Y. Huang, Analysis of Longitudinal and Survival Data: Joint Modeling, Inference Methods, and Issues, J. Probab. Stat. 2012 (2012), 640153. https://doi.org/10.1155/2012/640153.

[4] S. Luo, J. Wang, Bayesian hierarchical model for multiple repeated measures and survival data: an application to Parkinson's disease, Stat. Med. 33 (2014), 4279-4291. https://doi.org/10.1002/sim.6228.

[5] A.H. Dessiso, Bayesian joint modelling of longitudinal and survival data of HIV/AIDS patients: A case study at Bale Robe General Hospital, Ethiopia, Amer. J. Theor. Appl. Stat. 6 (2017), 182-190. https://doi.org/10.11648/j.ajtas.20170604.13.

[6] V. Leiva-Yamaguchi, D. Alvares, A two-stage approach for Bayesian joint models of longitudinal and survival data: correcting bias with informative prior, Entropy. 23 (2020), 50. https://doi.org/10.3390/e23010050.

[7] K. Mauff, E. Steyerberg, I. Kardys, et al. Joint models with multiple longitudinal outcomes and a time-to-event outcome: a corrected two-stage approach, Stat. Comput. 30 (2020), 999-1014. https://doi.org/10.1007/s11222-020-09927-9.

[8] A. Bhattacharjee, G.K. Vishwakarma, S. Banerjee, Joint modeling of longitudinal and time-to-event data with missing time-varying covariates in targeted therapy of oncology, Commun. Stat.: Case Stud. Data Anal. Appl. 6 (2020), 330-352. https://doi.org/10.1080/23737484.2020.1782286.

[9] R.M. Elashoff, G. Li, N. Li, A joint model for longitudinal measurements and survival data in the presence of multiple failure types, Biometrics. 64 (2007), 762-771. https://doi.org/10.1111/j.1541-0420.2007.00952.x.

[10] K. Das, A semiparametric Bayesian approach for joint modeling of longitudinal trait and event time, J. Appl. Stat. 43 (2016), 2850-2865. https://doi.org/10.1080/02664763.2016.1155108.

[11] J. Mwanyekange, S. Mwalili, O. Ngesa, Bayesian inference in a joint model for longitudinal and time to event data with Gompertz baseline hazards, Mod. Appl. Sci. 12 (2018), 159-172. https://doi.org/10.5539/mas.v12n9p159.

[12] X. Song, M. Davidian, A.A. Tsiatis, A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data, Biometrics. 58 (2002), 742-753. https://doi.org/10.1111/j.0006-341x.2002.00742.x.

[13] E.R. Brown, Assessing the association between trends in a biomarker and risk of event with an application in pediatric HIV/AIDS, Ann. Appl. Stat. 3 (2009), 1163-1182. https://doi.org/10.1214/09-aoas251.

[14] D. Rizopoulos, Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data, Biometrics. 67 (2011), 819-829. https://doi.org/10.1111/j.1541-0420.2010.01546.x.

[15] J. Wang, S. Luo, L. Li, Dynamic prediction for multiple repeated measures and event time data: An application to Parkinson's disease, Ann. Appl. Stat. 11 (2017), 1787-1809. https://doi.org/10.1214/17-aoas1059.

[16] H. Zhang, Y. Huang, Quantile regression-based Bayesian joint modeling analysis of longitudinal–survival data, with application to an AIDS cohort study, Lifetime Data Anal. 26 (2019), 339-368. https://doi.org/10.1007/s10985-019-09478-w.

[17] J.W. Stanifer, A. Muiru, T.H. Jafar, U.D. Patel, Chronic kidney disease in low- and middle-income countries, Nephrol. Dial. Transplant. 31 (2016), 868–874. https://doi.org/10.1093/ndt/gfv466.

[18] W.S. Shiferaw, T.Y. Akalu, Y.A. Aynalem, Chronic kidney disease among diabetes patients in Ethiopia: A systematic review and meta-analysis, Int. J. Nephrol. 2020 (2020), 1-15. https://doi.org/10.1155/2020/8890331.

[19] C. Armero, A. Forte, H. Perpiñán, et al. Bayesian joint modeling for assessing the progression of chronic kidney disease in children, Stat. Methods Med. Res. 27 (2018), 298-311. https://doi.org/10.1177/0962280216628560.

[20] W. Yang, D. Xie, Q. Pan, et al. Joint modeling of repeated measures and competing failure events in a study of chronic kidney disease, Stat. Biosci. 9 (2016), 504-524. https://doi.org/10.1007/s12561-016-9186-4.

[21] L. Teixeira, I. Sousa, A. Rodrigues, et al. Joint modelling of longitudinal and competing risks data in clinical research, REVSTAT-Stat. J. 17 (2019), 245-264. https://doi.org/10.57805/REVSTAT.V17I2.267.

[22] R.M. Elashoff, X. Huang, G. Li, A joint model of longitudinal and competing risks survival data with heterogeneous random effects and outlying longitudinal measurements, Stat. Interface. 3 (2010), 185-195. https://doi.org/10.4310/sii.2010.v3.n2.a6.

[23] A. Azarbar, Y. Wang, S. Nadarajah, Simultaneous Bayesian modeling of longitudinal and survival data in breast cancer patients, Commun. Stat. - Theory Methods. 50 (2019), 400-414. https://doi.org/10.1080/03610926.2019.1635701.

[24] H. Zhang, Y. Huang, Bayesian joint modeling for partially linear mixed-effects quantile regression of longitudinal and time-to-event data with limit of detection, covariate measurement errors and skewness, J. Biopharm. Stat. 31 (2020), 295-316. https://doi.org/10.1080/10543406.2020.1852248.

[25] S. Lee, G.J. McLachlan, Finite mixtures of multivariate skew t-distributions: some recent and new results, Stat. Comput. 24 (2012), 181-202. https://doi.org/10.1007/s11222-012-9362-4.

[26] E. Alvares, V. Lázaro, C. Gómez-Rubio, et al. Bayesian survival analysis with BUGS, Stat. Med. 40 (2021), 2975-3020. https://doi.org/10.1002/sim.8933.