Bayesian MLIRT-based joint models for multivariate longitudinal and survival data with multiple features

Yangxin Huang^{1*}, Jiaqing Chen^{2*}, Lan Xu¹, Hanze Zhang³ and Yuanyuan Lu¹

*Correspondence: yhuang@usf.edu

CrossMark

← Click for updates

¹College of Public Health, University of South Florida, Tampa, FL 33612, USA.

²Department of Statistics, College of Science, Wuhan University of Technology, Wuhan 430070, P.R.China.

³Amgen, 1 Amgen Center Dr, Thousand Oaks, CA 91320, USA.

Abstract

Methodological development and applications of joint models for longitudinal and survival data have mostly coupled a single longitudinal outcome-based mixed-effects model with normal distribution and Cox proportional hazards model. In practice, however, (i) normality of model error in longitudinal sub-models is a routine assumption, but it may be unrealistically violating data features of subject variations. (ii) The data collected are often featured by multivariate longitudinal outcomes which are significantly correlated, ignoring their correlation may lead to biased estimation. Additionally, a parametric specification may be inflexible to capture the complicated longitudinal pattern of biomarkers. (iii) It is of importance to investigate how multivariate longitudinal outcomes are associated with an event time of interest. Multilevel item response theory (MLIRT) models have been increasingly used to analyze the multivariate longitudinal data of mixed types (e.g., continuous and categorical) in clinical studies. In this article, we develop a multivariate joint model that consists of an extended MLIRT model for the mixed types of multivariate longitudinal data and a Cox proportional hazards model, linked through random-effects. The proposed models and method are applied to analyze longitudinal-survival data arising from a primary biliary cirrhosis study. Simulation studies are conducted to evaluate the performance of the proposed models and method.

Keywords: Bayesian inference; latent variable; longitudinal-survival data; multivariate joint model; multilevel item response theory; skew-normal distribution

Introduction

Data collected in many clinical studies can be grouped into three types: (i) longitudinal (repeated) measurements of time-dependent biomarkers such as serum bilirubin (SB), serum albumin (SA), hepatomegaly (HM) and histologic stage (HS) data in primary biliary cirrhosis (PBC) clinical study [41]; (ii) event or censored times of interest; and (iii) information on some covariates that may affect the processes of both time-dependent longitudinal

exposures and time-to-event. For such studies, interest often lies in how a disease biomarker progresses with time and how it is influenced by covariates. However, there have been few studies on simultaneously accounting for longitudinal data with inherent features—skewness, correlation and mixed types along with linkage in specifying a time-to-event outcome together for joint modeling and inference. Joint modeling analysis of longitudinal and survival data is an active area of statistical research that has

received much attention recently. One of the main reasons for the increasing interest is that joint models can be used in a variety of problems with various data features and joint modeling methods may yield more precise inference [26,30,45,46,52,54]. As a result, a considerable number of statistical modeling and analysis methods have been suggested for analyzing such longitudinal-survival data. However, (i) the commonly-assumed distribution for model error is normal due to mathematical tractability and computational convenience, but this assumption may lack robustness against departures from normality and/or outliers; it may also obscure important features of between- and within-subject variations of (continuous) longitudinal data that may exhibit departures from symmetry as exemplified in (Figure 1 (top panel)). Thus, it is of practical interest to investigate models with skewed distributions including the skew-normal(SN) distribution [5,6,29,49]. In the mean time, the (continuous) longitudinal measurements-based trajectories often display irregular nonlinear time trends (Figure 1 (bottom panel)). Thus a parametric linear mixed-effects model is not appropriated to fit such data and a partially linear mixed-effects model, which inherits the advantages from both parametric and nonparametric models, may be applied. (ii) In practice, many studies are designed to collect data with multiple longitudinal exposure data of mixed types (e.g., continuous and categorical) which may be significantly

correlated or associated such as SB and SA (continuous), HM (binary) and HS (ordinal) to be used in our real data analysis, and ignoring their interrelationships may lead to biased results and reduce efficiency in estimation. (iii) Longitudinal studies, where repeated (longitudinal) measurements, an observation on a possibly censored time-to-event and additional covariates are collected for each subject, are commonplace in medical and clinical research. Interest often focuses on interrelationships between these variables [26,27,29,46,52,54]. A joint model that links the hazard to these longitudinal measures and covariates with incorporating some typical data features of interest, is becoming powerful increasingly in many applications.

To analyze such multivariate longitudinal data, multilevel item response theory (MLIRT) models have been increasingly used [16,23,25,53]. According to MLIRT, the observed multivariate outcomes are viewed as imperfect clinical manifestations of a univariate subject-specific latent variable measuring disease severity. The MLIRT model consists of two levels. The first level measurement model quantifies the relationship between a subject-specific latent disease severity and the response to the multivariate outcomes. In the second level structural model, the latent disease severity is regressed on predictors (e.g., treatment, disease duration and time) and subject-specific random-effects (describing between-subject differences) to study the overall

treatment effects [1,3,40]. The three sources of correlation are accounted for via the random-effects. Advantages of the MLIRT models include better reflection of multilevel data structure, simultaneous estimation of measurement-specific parameters and covariate effects, and accurate inference about high-level measures [20,35,39]. To obtain valid inference from the MLIRT models, marginal maximum likelihood methods [40], and Bayesian methods [19,20,21,25,33,37,42] have been widely used. To the best of our knowledge, no studies have done to explore multivariate joint models (MJM) with the SN distribution, coped with an MLIRT sub-model for multivariate longitudinal exposures with various data features and a Cox proportional hazards sub-model for survival endpoint which are linked via random-effects. Inferential procedures are dramatically complicated when multivariate longitudinal data with non-normality and their interrelationships are considered in conjunction with time-to-event into the MJM.

The rest of this article proceeds as follows. In Section 2, we discuss the setup of MJM in which an MLIRT model is used for the multivariate longitudinal outcomes, where the continuous-based longitudinal sub-model follows the SN distribution, and a Cox proportional hazards sub-model for survival endpoint. Section 3 investigate associated Bayesian inferential approach in general form so that they can be applicable to various scientific fields. The proposed Bayesian method is offered to fit MJM simultaneously for estimating all parameters in MJM under Bayesian paradigm of Markov chain Monte Carlo (MCMC). In Section 4, we describe the PBC clinical study and data structure that motivated this research, present the specific MJM formulation and apply the proposed methodology to the PBC dataset. Section 5 conducts limited simulation study to evaluate the performance of the proposed models and method. Finally, the paper concludes with some discussions in Section 6.

Setup of multivariate joint models

Let y_{ijk} be the observed outcome k from patient i at time point j , where $i = 1, \dots, N$, $j = 1, \dots, J_i$ and $k = 1, \dots, K$. Let $\mathbf{y}_{ik} = (y_{i1k}, \dots, y_{ij1k}, \dots, y_{ij1K})^T$ be the vector of observation for patient i at visit j and let $\mathbf{y}_i = (\mathbf{y}_{i1}, \dots, \mathbf{y}_{iK})^T$ be the outcome vector across visits. Throughout, all outcomes are coded so that larger observation values correspond to worse clinical conditions. Let θ_{ij} be a univariate latent variable measuring disease severity of individual i at visit j , with a higher value denoting more severe status. Let t_i be the observed event time for patient i , and c_i (1 if the event is observed and 0 otherwise) be the event indicator. We introduce an extended MLIRT sub-model for the multiple longitudinal outcomes and a Cox proportional hazard sub-model for the event time as follows.

Extended MLIRT sub-model for longitudinal data

Under the framework of traditional MLIRT model, the longitudinal continuous response is specified by the usual linear model with the normal distribution [19,20,21,25,37], where in the first level measurement model, we model the continuous outcomes, the

binary outcomes and the cumulative probabilities of ordinal outcomes by a two-parameter model [9], graded response model [19] and common factor model [36], respectively. However, the trajectories of longitudinal continuous responses often show complex nonlinear patterns and longitudinal data may exhibit non-normal feature. Thus, we consider an extended version of MLIRT model which is specified by the partially linear model-based MLIRT framework to take into account the complex non-linear effect of time on the longitudinal continuous responses, where the partially linear models are specifically composed of a linear part and a nonparametric part [9,55]. For the extended MLIRT model, in the first level measurement model, the continuous outcomes, the binary outcomes and the cumulative probabilities of ordinal outcomes are modeled by specifying the following formulations, respectively.

$$\begin{aligned} y_{ijk} &= a_k + b_k \theta_{ij} + w_k(t_{ij}) + h_{ik}(t_{ij}) + \epsilon_{ijk} \\ \text{logit}\{p(y_{ijk} = 1 | \theta_{ij})\} &= a_k + b_k \theta_{ij} \\ \text{logit}\{p(y_{ijk} \leq l | \theta_{ij})\} &= a_{kl} - b_k \theta_{ij} \text{ with } l = 1, 2, \dots, n_k - 1, \end{aligned} \quad (1)$$

where the random error for continuous outcomes $\epsilon_i = (\epsilon_{i1}, \dots, \epsilon_{iK})^T$ follows a multivariate SN distribution

$$SN_{K_1, J_i} \left(-\sqrt{2/\pi} [\Delta_{K_1} \otimes \mathbf{1}_{J_i}], \Sigma_{K_1} \otimes I_{J_i}, \Delta_{K_1} \otimes I_{J_i} \right)$$

[29,49] with unknown variance-covariance matrix $\Sigma_{K_1} = (\sigma_{kk'}^2)_{K_1 \times K_1}$ ($k, k' = 1, 2, \dots, K_1$), unknown skewness parameter matrix $\Delta_{K_1} = \text{diag}(\delta_1, \dots, \delta_{K_1})$, the vector of skewness parameters $\delta_{K_1} = (\delta_1, \dots, \delta_{K_1})^T$, and $\mathbf{1}_J = (1, \dots, 1)^T$.

Note that $-\sqrt{2/\pi} [\Delta_{K_1} \otimes \mathbf{1}_{J_i}]$ is specified here to make the SN distribution with mean zero and K_1 is the number of continuous variables. a_k is the outcome-specific 'difficulty' parameter and b_k (always positive) is the outcome-specific 'discriminating' parameter representing the discrimination of outcome k , that is, the degree to which outcome k discriminates between individuals with different disease severity θ_{ij} for patient i at time j , with higher value denoting more severe status, where θ_{ij} is a continuous latent variable. For the k th ordinal outcome with n_k categories, the order constraint $a_{k1} < \dots < a_{kl} < \dots < a_{knk} - 1$ must be satisfied and the probability that patient i is in category l on outcome k at visit j is $p(Y_{ijk} = l | \theta_{ij}) = p(Y_{ijk} \leq l | \theta_{ij}) - p(Y_{ijk} \leq l-1 | \theta_{ij})$. $w_k(t)$ and $h_{ik}(t)$ are unknown nonparametric smooth fixed-effects and random-effects functions, respectively, $h_{ik}(t)$ are iid realizations of a zero-mean stochastic process.

To fit model (1), we apply a regression spline method to $w_k(t)$ and $h_{ik}(t)$. The working principle is briefly described as follows and more details can be found in literature [13]. The main idea of regression spline is to approximate $w_k(t)$ and $h_{ik}(t)$ by using a linear combination of spline basis functions. For instance, $w_k(t)$ and $h_{ik}(t)$ can be approximated by a linear combination of basis functions $\Psi_p(t) = \{\psi_0(t), \psi_1(t), \dots, \psi_{p-1}(t)\}^T$ and $\Phi_q(t) = \{\phi_0(t), \phi_1(t), \dots, \phi_{q-1}(t)\}^T$, respectively. That is,

$$\begin{aligned} w_k(t) &\approx w_{kp}(t) = \sum_{l=0}^{p-1} \gamma_{kl} \psi_l(t) = \boldsymbol{\Psi}_p(t)^T \boldsymbol{\gamma}_{kp}, \\ h_{ik}(t) &\approx h_{ikq}(t) = \sum_{l=0}^{q-1} \xi_{ikl} \phi_l(t) = \boldsymbol{\Phi}_q(t)^T \boldsymbol{\xi}_{ikq}, \end{aligned} \quad (2)$$

where $\gamma_{kp} = (\gamma_{k0}, \dots, \gamma_{k,p-1})^T$ is a $p \times 1$ vector of fixed-effects, and $\xi_{ikq} = (\xi_{ik0}, \dots, \xi_{ik,q-1})^T$ ($q \leq p$) is a $q \times 1$ vector of random-effects ($k = 1, \dots, K_1$). We assume that $\xi_{ikq} \stackrel{iid}{\sim} N_q(0, \Sigma_\xi)$ with Σ_ξ being unrestricted covariance matrix. Based on the assumption of $h_{ik}(t)$, we can regard ξ_{ikq} as iid realizations of a zero-mean random vector. Substituting $w_k(t)$ and $h_{ik}(t)$ by their approximations $w_{kp}(t)$ and $h_{ikq}(t)$, for given $\Psi_p(t)$ and $\Phi_q(t)$ we can approximate model (1) in a compact way as follows.

$$\begin{aligned} y_{ijk} &= ak + bk\theta_{ij} + \Psi p(t_{ij})^T \gamma_{kp} + \Phi q(t_{ij})^T \xi_{ikq} + \xi_{ijk} \\ \text{logit}\{p(y_{ijk} = 1 | \theta_{ij})\} &= a_k + b_k \theta_{ij} \\ \text{logit}\{p(y_{ijk} \leq l | \theta_{ij})\} &= a_{kl} - b_k \theta_{ij} \text{ with } l = 1, 2, \dots, n_k - 1, (3) \end{aligned}$$

where we consider natural cubic spline bases with the percentile-based knots in the nonparametric part. The optimal degree of regression spline p and numbers of knots q can be also determined according to the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) [13]. In this study, the AIC/BIC values are evaluated for various models with $(p, q) = \{(1, 1), (2, 1), (2, 2), (3, 1), (3, 2), (3, 3)\}$.

In the second level structural multilevel model, the latent disease severity θ_{ij} is regressed on covariates of interest, visit time, and random-effects.

$$\theta_{ij} = (X_{i0}^T \beta_0 + u_{i0}) + (X_{i1}^T \beta_1 + u_{i1})t_{ij}, \quad (4)$$

where X_{i0} and X_{i1} are vectors of covariates associated with baseline disease severity and disease progression rate, respectively, X_{i0} may or may not be the same as X_{i1} , t_{ij} is visit time variable with $t_{i1} = 0$ for baseline, random intercept u_{i0} and random slope u_{i1} determine the subject-specific baseline disease severity and disease progression rate, respectively. The random-effects vector $u_i = (u_{i0}, u_{i1})^T$ follows $N_2(0, \Sigma_u)$, with covariance matrix Σ_u being denoted by $((1, \rho\sigma_u), (\rho\sigma_u, \sigma_u^2))$, where the variance of u_{i0} is set to 1 for identifiability, σ_u^2 is the variance of u_{i1} and ρ is the correlation coefficient. The random-effects vector u_i accounts for three sources of correlations within the same patient: (i) inter-source (different outcomes at the same visit), (ii) longitudinal (same outcome at different visits), and (iii) cross correlation (different outcomes at different visits) [43]. The regression parameter vectors β_0 and β_1 represent the covariate effects on the baseline disease severity and disease progression rate, respectively. For example, if $\theta_{ij} = (\beta_{01}x_i + u_{i0}) + (\beta_{10} + \beta_{11}x_i + u_{i1})t_{ij}$ where x_i is an indicator variable of treatment (1 if treatment, 0 otherwise), then β_{01} is the baseline group difference, and β_{10} and $\beta_{10} + \beta_{11}$ are the disease progression rates for the placebo and treatment patients, respectively. The negative significant variable β_{11} indicates that the treatment is efficacious in slowing down the disease progression. The combined level 1 and level 2 models are MLIRT model with subject-specific covariance (referred to as subject-specific MLIRT models) [12,19,20,21,51]. To use the extended MLIRT sub-model, in our PBC application later, we consider SB and SA as the continuous outcomes (i.e., $K_1=2$), the hepatomegaly as binary outcome, and the histologic

status as ordinal outcome; we use the treatment variable x_i as the only covariate in X_{i0} and X_{i1} .

Survival sub-model for time-to-event

For the time-to-event process, we assume that the distribution of T_i , the event time for the i th subject, depends on the random-effects u_i representing individual-specific longitudinal processes and covariates and/or risk factors X_i , respectively. We therefore consider a frailty model for T_i which is linked to the MLIRT sub-model (3) and (4) through the random-effects u_i . Additional covariates X_i are assumed to be associated with the event time. Specifically, under the Cox proportional hazard sub-model, the hazard of having an event at time t_i is

$$\lambda(t_i | X_i, u_i) = \lambda_0(t_i) \exp(X_i^T \alpha + u_i^T \nu), \quad (5)$$

where the function $\lambda(t)$ is the instantaneous hazard rate at time t , $\lambda_0(t)$ is an unspecified baseline hazard function. The association parameter vector $\nu = (\nu_1, \nu_2)^T$ linking the random-effects u_i measures the association between the two sub-models. Our MJM is defined by the longitudinal MLIRT sub-model (3) and (4) and the survival sub-model (5), where two sub-models are linked together via the shared random-effects u_i , which is a popular approach in joint modeling [26,28,30,32,38,46,55]. The covariate vector X_i can be the same or different from X_{i0} and X_{i1} . An alternative method can be used to approximate the Cox proportional hazards model (5) through counting process [10] which is used for our joint modeling and can obviously reduce computational burdens; the detailed discussion for the alternative method can be found in [27,28,55].

Bayesian inferential approach

Generally, estimation of a joint model for longitudinal and time-to-event data could be conducted in two ways. The first approach is based on the likelihood inferential methods, such as Expectation-Maximization (EM) algorithm and Monte Carlo Expectation-Maximization (MCEM) algorithm [47]. A simultaneous inference method based on a joint likelihood may be favorable, but the computational burden for proposed MJM with the SN distribution can be extremely intensive, even sometimes infeasible, and may have convergence problems [8,54]. The second approach, Bayesian inferential method shows the advantage. Thus, we simultaneously estimate parameters for the MJM under a fully Bayesian framework via Markov chain Monte Carlo (MCMC) procedure based on the joint likelihood of the longitudinal-survival data and specified prior distributions. Thus, the Bayesian joint modeling approach may pave a way to alleviate the computational burdens and to overcome convergence problems for such complex model setting.

We assume that ϵ_i , u_i and ξ_{ikq} ($k=1, 2$) are mutually independent of each other. According to the property and stochastic representation of SN distribution described in the publications [29,49], it can be shown that, by introducing the random vector w_i , the MJM specified by (3), (4) and (5) can be

hierarchically formulated as follows.

$$\begin{aligned} \mathbf{y}_{iK_1} &\sim N_{K_1, J_i}(\mathbf{g}_{iK_1}(t_i) + \Delta_{K_1} \otimes [\mathbf{w}_i - \sqrt{2/\pi} \mathbf{1}_{J_i}], \Sigma_{K_1} \otimes \mathbf{I}_{J_i}), \\ \mathbf{w}_i &\sim N_{J_i}(\mathbf{0}, \mathbf{I}_{J_i}) \mathbf{I}(\mathbf{w}_i > \mathbf{0}), \quad \mathbf{u}_i \sim N_2(\mathbf{0}, \Sigma_u), \\ \xi_{ikq} &\sim N_q(\mathbf{0}, \Sigma_\xi), (k = 1, \dots, K_1) \quad T_i \sim F(t_i) = \int f(\mathbf{s}_i | \mathbf{u}_i), \\ r_{ijk} &\sim \text{Bern}(P_{ijk}), \quad R_{ijk} \sim \text{Cat}(P_{ijk}), \end{aligned} \quad (6)$$

where $\mathbf{g}_{iK}(\cdot) = (\mathbf{g}_{i1}(\cdot)^T, \dots, \mathbf{g}_{iK}(\cdot)^T)^T$ and $\mathbf{g}_{ik}(\cdot) = (g_{ijk}(t_{ij}), \dots, g_{ijk}(t_{ij}))^T$ with $g_{ijk}(t_{ij}) = a_k + b_k \theta_{ij} + \Psi_p(t_{ij})^T \gamma_{kp} + \Phi_q(t_{ij})^T \xi_{ikq}$, $(k = 1, \dots, K_1)$. $P_{ijk} = P(r_{ijk} = 1)$ for binary variables and $P_{ijk} = P(R_{ijk} = l)$ ($l = 1, 2, 3$) for ordinal variables. Let $\Theta = \{\alpha, \beta, \gamma_{kp}, v, \Sigma_{K_1}, \Sigma_\xi, \Sigma_u, \delta_{K_1}, a_{kl}, b_{kl}, a_{kl} \mid l = 2, \dots, n_k - 1\}$ be the collection of unknown population parameters in the joint model. The prior distribution represents the information about the uncertain parameter vector Θ that is combined with the probability distribution of data to yield the posterior distribution for inferences involving Θ . Under the Bayesian framework, we specify prior distributions for all unknown population parameters in MJM as follows.

$$\begin{aligned} \alpha &\sim N_3(\mathbf{0}, \Lambda_1), \quad \beta \sim N_3(\mathbf{0}, \Lambda_2), \quad \gamma_{kp} \sim N_p(\mathbf{0}, \Lambda_3) (k = 1, \dots, K_1), \quad \delta_{K_1} \\ &\sim N_{K_1}(\mathbf{0}, \Lambda_4), \quad v \sim N_2(\mathbf{0}, \Lambda_5), \quad \rho \sim \text{Unif}(-\omega, \omega), \quad \sigma_u^2 \sim \text{IG}(\omega_v, \omega_2), \quad \Sigma_{K_1} \\ &\sim \text{IW}(\Lambda_{\sigma'} \omega_1), \quad \Sigma_\xi \sim \text{IW}(\Lambda_{\sigma'} \omega_2), \quad a_k \sim N(0, \sigma_k^2), \quad b_k \sim \Gamma(\omega_3, \omega_4), \quad a_{k1} \sim N(0, \sigma_{k1}^2), \\ &\eta_l \sim \Gamma(\omega_5, \omega_6), \end{aligned} \quad (7)$$

where $a_{kl} = a_{k,l-1} + \eta_l$ ($k = 1, \dots, K_1; l = 2, \dots, n_k - 1$); we assume that the Normal (N), Uniform (Unif), Inverse Gamma (IG), Inverse Wishart (IW) and Gamma (Γ) prior distributions are mutually independent. Let $f(\cdot)$, $f(\cdot|\cdot)$, $F(\cdot)$ and $\pi(\cdot)$ denote a density function, a conditional density function, a cumulative distribution function and a prior density function, respectively. Since the elements of parameter vector $\mathbf{T}$ are assumed to be independent of each other, we have $\pi(\Theta)$ is the product of prior distributions for each element of $\mathbf{T}$. Subsequently, after specifying the joint model for the observed data and the prior distributions for the unknown parameters, we can make Bayesian inference for the parameters based on their posterior distributions. Thus, the joint posterior density of Θ based on the observed data D can be given by

$$\begin{aligned} f(\Theta | D) &\propto \left(\prod_{i=1}^N \prod_{k=1}^{K_1} \int \int f(\mathbf{y}_{iK_1}) f(\mathbf{w}_i | \mathbf{w}_i > \mathbf{0}) \right. \\ &\left. f(\mathbf{s}_i | \mathbf{u}_i) f(\mathbf{r}_{ik}) f(\mathbf{R}_{ik}) f(\mathbf{u}_i) f(\xi_{ikq}) d\mathbf{u}_i d\xi_{ikq} \right) \pi(\Theta). \end{aligned} \quad (8)$$

Generally, the integrals in (8) are of high dimension and do not have a closed form. Analytic approximations to the integrals may not be sufficiently accurate. Therefore, it is prohibitive to directly calculate the posterior distribution of Θ based on the observed data. As an alternative, the MCMC procedure can be used to sample population parameters Θ and random-effects $\mathbf{u}_i$ and ξ_{ikq} ($i = 1, \dots, N; k = 1, \dots, K_1$) from posterior distributions based on (8). We adopt a hybrid MCMC algorithm to speed up calculation. This process is repeated in iterations of MCMC

procedure until convergence is reached.

Real data application PBC clinical study and data

We examine the effect of multiple longitudinal exposures on the prognosis for patients with primary biliary cirrhosis (PBC) using data collected by the Mayo Clinic from 1974 to 1984 [18,41]. This data were obtained from JM package in the R program. PBC is a chronic, fatal, but rare liver disease characterized by inflammatory destruction of the small bile ducts within the liver, which eventually leads to cirrhosis of the liver, followed by death. If PBC is not treated or reaches an advanced stage, it can lead to several major complications, including mortality. PBC is a chronic liver disease with no effective treatment other than liver transplantation [24].

Various longitudinal biomarkers and exposures such as serum bilirubin (SB), serum albumin (SA), hepatomegaly (HM) and histologic stage (HS) were collected, and interest is on examining whether these longitudinal exposures relate to the natural history of disease. This dataset has been widely analyzed using various statistical modeling methods including joint modelling approach [2,4,11,14]. The data were collected to examine the progress of PBC patients. A total of 424 PBC patients met the eligibility criteria for the randomized placebo controlled trial of the drug D-penicillamine, referred to Mayo Clinic during the ten-year interval. 312 patients, who had a baseline measurement and were followed longitudinally at 6 months and at yearly intervals thereafter, participated in the randomized trial and were randomised to receive D-penicillamine ($n=158$) or placebo ($n=154$). The additional 112 cases did not participate in the clinical trial.

Patients with PBC typically have abnormalities in several blood tests; hence, during follow-up several biomarkers associated with liver function were serially recorded for these patients. The original clinical protocol for these patients specified visits at 6 months, 1 year, and annually thereafter. This is an ideal dataset to illustrate the various features of the MLIRT-based MVM. The data set includes clinical, biochemical, and demographic risk factors for each patient. Particularly, in our application, we consider the following data variables. (i) Demographic factors (covariates) are age and sex of patients; biochemical factor is drug (D-penicillin and placebo group). (ii) Longitudinal data of mixed types: HM (liver growth status, presence of hepatomegaly 0 = No, 1 = Yes) which is binary outcome, HS (1 ≤ stage 2, 2 = stage 3 and 3 = stage 4) which is ordinal outcome, and the continuous outcomes SB (mg/dl) and SA (mg/dl) values were taken as biochemical properties. As illustrated in Section 1, the repeated SB (mg/dl) and SA (mg/dl) measurements appeared departures from symmetry (i.e., skewed distribution) and the trajectory profiles for SB and SA longitudinal exposures depicted irregular nonlinear time trends (see Figure 1). (iii) Survival data: since the patients were started to be followed, the survival time (year) were taken as the period until death. Here the failure is considered as death and other patients are expressed as cen-

sored. At the end of study, 140 of the 312 patients (44.9%) died and 172 (55.1%) was censored.

Specific MJM and model implementation

To analyze the PBC dataset, we focus on a specific MLIRT-based MJM which considers continuous out- comes SB (y_{ij1}) and SA (y_{ij2}) binary HM outcome (y_{ij3}) and ordinal HS outcome (y_{ij4}) ($i = 1, \dots, N; j = 1, \dots, J; K = 1, \dots, 4$) as multivariate longitudinal exposures, and time to death as the event endpoint. Based on the general MJM discussed in Section 2, for the nonparametric part of the continuous longi- tudinal sub-model in (3), we consider natural cubic spline bases with the percentile-based knots in the nonparametric part. The optimal degree of regression spline p and numbers of knots q can be also deter- mined according to AIC/BIC values [13]. Here, the AIC/BIC values were evaluated for various models with $(p, q) = \{(1, 1), (2, 1), (2, 2), (3, 1), (3, 2), (3, 3)\}$. Among them, we found that the model with $(p, q) = (3, 3)$ has the smallest AIC/BIC values. Thus, the specific MJM can be constructed as follows.

$$\begin{aligned} y_{ijk} &= a_k + b_k \theta_{ij} + \Psi_3(t_{ij})^T \gamma_k + \Phi_3(t_{ij})^T \xi_{ik} + \epsilon_{ijk} \quad (k = 1, 2), \\ \text{logit}\{p(y_{ij3} = 1 | \theta_{ij})\} &= a_3 + b_3 \theta_{ij} \\ \text{logit}\{p(y_{ij4} \leq l | \theta_{ij})\} &= a_{4l} - b_4 \theta_{ij} \quad \text{with } (l = 1, 2), \end{aligned} \quad (9)$$

where the random error for two continuous outcomes $\epsilon_i = (\epsilon_{i1}, \epsilon_{i2})^T$ follows a multivariate SN distribution

$$SN_{2, J_1} \left(-\sqrt{2/\pi} [\Delta_2 \otimes \mathbf{1}_{J_1}], \Sigma_2 \otimes \mathbf{I}_{J_1}, \Delta_2 \otimes \mathbf{I}_{J_1} \right) |$$

[29,49] with unknown variance-covariance matrix $\Sigma_2 = (\sigma^2 I)_{2 \times 2}$, unknown skewness parameter matrix $\Delta_2 = \text{diag}(\delta_1, \delta_2)$, and the vector of skewness parameters $\delta_2 = (\delta_1, \delta_2)^T$. $\gamma_k = (\gamma_{1k}, \gamma_{2k}, \gamma_{3k})^T$ ($k=1,2$) and $\xi_{ik} \sim N_3(\mathbf{0}, \Sigma_\xi)$ with Σ_ξ being unrestricted covariance matrix. The basis functions $\Psi_3(t) = \{\psi_0(t), \psi_1(t), \psi_2(t)\}^T$ and $\Phi_3(t) = \{\phi_0(t), \phi_1(t), \phi_2(t)\}^T$, respectively. In the second level structural multilevel model, the latent disease severity θ_{ij} is regressed on biochemical factor x_i which is an indicator variable of treatment (1 if treatment, 0 otherwise) and random-effects $\mathbf{u}_i = (u_{i0}, u_{i1})^T$.

$$\theta_{ij} = (\beta_{01} x_i + u_{i0}) + (\beta_{10} + \beta_{11} x_i + u_{i1}) t_{ij} \quad (10)$$

where β_{01} is the baseline group difference, and β_{10} and β_{11} are the disease progression rates for the placebo and treatment patients, respectively. The negative significant variable β_{11} indicates that the treatment is efficacious in slowing down the disease progression.

Under the Cox proportional hazard sub-model, the hazard of having an event at time t_i is

$$\lambda(t_i | \mathbf{X}_i, \mathbf{u}_i) = \lambda_0(t_i) \exp(\mathbf{X}_i^T \alpha + \mathbf{u}_i^T \nu), \quad (11)$$

where $\nu = (\nu_1, \nu_2)^T$ with ν_1 and ν_2 measuring the association between the two sub-models, the vector $\alpha = (\alpha_1, \alpha_2, \alpha_3)^T$ is the coefficient parameters corresponding to the risk factor vector $\mathbf{X}_i$ which includes treatment indicator variable, age at baseline, gender (male as reference).

As discussed previously, the repeated SB and SA data exhibit the asymmetric feature. Thus, it appears plausible to fit our MJM with a skew distribution to the data. Toward this end, because a normal distribution is a special case of an SN distribution when the skewness parameter is zero, we investigate how asymmetric (SN) distribution for continuously longitudinal sub-model error (Model SN) contributes to modeling results in comparison with a symmetric (normal) distribution for continuously longitudinal sub-model error (Model N). Thus, the following Model SN and Model N with specifying different distributions for continuously longitudinal sub-model are employed to compare their performance.

Model SN: An MJM where the continuously longitudinal sub-model follows the SN distribution;

Model N: An MJM where the continuously longitudinal sub-model follows the normal distribution.

To carry out a Bayesian inference, we need to specify the values of the hyper-parameters in the prior distributions. We assume weakly informative prior distributions for all the parameters. In particular, (i) the fixed-effects are taken to be independent normal distribution $N(0, 100)$ for each component of the population parameter vectors α, β, γ_k ($k = 1, 2$), δ_2 and ν , respectively. (ii) For the variance-covariance matrices Σ_2 and Σ_ξ , we assume the inverse Wishart distributions $IW(\Lambda_\xi, 3)$ and $IW(\Lambda_\gamma, 4)$, respectively, where Λ_ξ and Λ_γ are the matrices with diagonal elements being 0.01 and off-diagonal elements being 0. (iii) We use the prior distribution $\rho \sim \text{Unif}(-1, 1)$ and a noninformative inverse Gamma prior distribution $IG(0.01, 0.01)$, which has mean 1 and variance 100, for variance parameter σ^2 . (iv) a_k and a_{k1} follow normal distribution $N(0, 100)$, respectively; b_k and η_2 follow $\Gamma(0.01, 0.01)$, ($k = 1, \dots, 4$).

The MJM modeling is implemented using MCMC procedure and the program codes are available from authors upon request. When the MCMC procedure was applied to the actual data, the convergence of the generated samples was assessed using standard tools such as trace plots and Gelman-Rubin (GR) diagnostics [22]. Figure 2 shows the dynamic version of GR diagnostics as obtained from the WinBUGS output for the representative parameters based on Model SN [22]. We observe from the plots of GR diagnostics where the three curves are given: the middle and bottom curves below the dashed horizontal line (indicated by the value 1) represent the pooled posterior variance ($\hat{V}$) and average within-sample variance ($\hat{W}$), respectively, and the top curve represents their ratio ($\hat{R}$). It is seen that $\hat{R}$ is generally expected to be higher than one at the initial stage of the algorithms, but $\hat{R}$ tends to 1, and $\hat{V}$ and $\hat{W}$ stabilize as the number of MCMC iterations increase, indicating that the algorithm has reached convergence. We further monitor convergence using both autocorrelation and trace plots (data not shown to save space), confirming that convergence is ensured. Along with the convenience diagnostics observed, we proposed that the three chains were run with the following considerations. After an initial number of 30,000 burn-in iterations, every 30th MCMC sample was retained from the next 30,000 iterations

Figure 2. Convergence diagnostics with three Markov chains as obtained from the WinBUGS software for representative parameters based on Model SN: trace plots (left panel) and convergence plots of Gelman-Rubin (GR) diagnostics (right panel), where the middle and bottom curves below the dashed horizontal line (indicated by the value one) represent the pooled posterior variance ($\hat{V}$) and average within-sample variance ($\hat{W}$), respectively, and the top curve above the dashed horizontal line represents their ratio ($\hat{R}$).

for each chain. Thus, we obtained a total of 3,000 samples for each of the targeted posterior distributions of the unknown parameters for statistical inference. We also check the k -lag serial correlation of the samples for each parameter to diagnose independence of MCMC samples. We graphically checked the last 500 samples drawn from MCMC sampling scheme for each parameter (plot not shown here) and found that the consecutive samples move randomly towards different directions, indicating that MCMC is not “sticky” and MCMC samples are independent for each parameter, suggesting the convergence to the stationary distribution. The Bayesian joint modeling approach with the different scenarios is applied to fit the PBC dataset described in Section 3 for statistical inference. As a note, the computational burden for fitting the specific joint models via MCMC procedure was reasonable. For example, to fit Model SN in our application, it took about 5 hours on a Window PC with Intel Core i7-2600 CPU @ 6.80GHz and RAM of 32 GB.

Analysis results

Table 1 presents the posterior mean (PM), the corresponding standard deviation (SD) and 95% credible interval (CI) for the interesting fixed-effects (population) parameters based on Models SN and N, respectively. The following findings are obtained for estimated population parameters.

For the longitudinal MLIRT sub-model, it is observed that Models SN and N give different parameter estimates, although the same set of parameters are identified for significance by both Models. For instance, Model SN suggests that D-penicillin is associated with 0.040 unit decrease (β_{01} , 95% CI: (-0.187, 0.072)) in disease severity, comparing with the placebo, while it is associated with 0.066 unit decrease (95% CI: (-0.213, 0.065)) in disease severity in Model N. The results from Model SN indicate that the

placebo patients show significant disease progression cross time at the rate of 1.482 units per year (β_{10} , 95% CI: (0.747, 3.133)), whereas the rate of progression is 2.229 (95% CI: (0.948, 3.962)) in Model N. In comparison, Model SN suggests that the D-penicillin patients have disease progression rate of 1.261 units per year ($\beta_{10} + \beta_{11}$, 95% CI: (-0.235, 3.318)) with insignificant D-penicillin treatment effect of slowing down the disease progression rate by -0.221 per year (β_{11} , 95% CI: (-0.982, 0.185)); the similar conclusions can be made from Model N.

For the survival sub-model, it can be seen from **Table 1** that D-penicillin decreases the hazard of death by 5% [$1 - \exp(\alpha_1) = 0.05$, where $\alpha_1 = -0.054$ with 95% CI being (-0.567, 0.435)] in comparison with placebo for Model SN and 8% [$1 - \exp(\alpha_1) = 0.05$, where $\alpha_1 = -0.083$ with 95% CI being (-0.589, 0.405)] for Model N. The insignificant treatment effect is consistent with those in [2,14]. There is a 2.3% increase (HR=1.023, 95% CI:(1.003, 1.043)) in hazard for death relative to a 1-year older in age for Model SN which is consistent with the finding by Dil and Karasoy [14]; the comparable result can be made from Model N. Female patients decrease the hazard of death by 46.0% [$1 - \exp(\alpha_3) = 0.46$, where $\alpha_3 = -0.616$ with 95% CI being (-1.172, -0.048)] in comparison with male patients in Model SN, while female patients decrease the hazard of death by 46.3% [$1 - \exp(\alpha_3) = 0.463$, where $\alpha_3 = -0.621$ with 95% CI being (-1.129, -0.055)] in comparison with male patients in Model N. This finding is consistent with that in publications [14,15,44].

We also find from **Table 1** that the estimated skewness parameters of the continuous SB outcome (δ_1) and the continuous SA outcome (δ_2) from Model SN are 3.847 with 95% CI (3.692, 4.005) and -0.089 with 95% CI (-0.116, -0.061), respectively. These findings suggest that there is a significantly positive skewness in SB outcome and negative skewness in SA outcome

Table 1. Summary of estimated posterior mean (PM), standard deviation (SD) of population (fixed-effects) parameters and the corresponding 95% equal-tail credible interval (CI) as well as DIC value.

Parameter	Model SN			Model N		
	PM	SD	95% CI	PM	SD	95% CI
Parameter estimates in longitudinal sub-model						
β_{01}	-0.040	0.061	(-0.187, 0.072)	-0.066	0.070	(-0.213, 0.065)
β_{10}	1.482	0.720	(0.747, 3.133)	2.229	0.813	(0.948, 3.962)
β_{11}	-0.221	0.281	(-0.982, 0.185)	-0.334	0.335	(-1.146, 0.164)
ρ	0.061	0.056	(0.029, 0.200)	0.133	0.073	(0.016, 0.297)
σ_u^2	1.694	1.625	(0.350, 6.094)	3.632	2.549	(0.670, 10.03)
δ_1	3.847	0.079	(3.692, 4.005)	—	—	—
δ_2	-0.089	0.014	(-0.116, -0.061)	—	—	—
Parameter estimates in survival sub-model						
α_1	-0.054	0.251	(-0.567, 0.435)	-0.083	0.253	(-0.589, 0.405)
α_2	0.023	0.010	(0.003, 0.042)	0.024	0.010	(0.003, 0.045)
α_3	-0.616	0.283	(-1.172, -0.048)	-0.621	0.280	(-1.129, -0.055)
ν_1	1.776	0.762	(0.588, 3.391)	1.171	0.551	(0.405, 2.463)
ν_2	0.828	0.332	(0.306, 1.511)	0.582	0.240	(0.260, 1.162)
DIC		10036.5			14659.4	
$\bar{D}$		7922.3			13649.0	

and confirm the fact that the distribution of the repeated SB and SA measurements is skewed. Thus, it is recommended to incorporate skewness parameters in the modeling of both SB and SA data.

We observe that v_1 and v_2 are positive and significantly different from zero, [$v_1 = 1.776$ with 95% CI being (0.588, 3.391), and $v_2 = 0.828$ with 95% CI being (0.306, 1.511)] in Model SN, suggesting that the patients with worse baseline disease severity (larger u_{i0}) and faster disease progression rate (larger u_{i1}) tend to have higher hazard of death and vice versa; the similar conclusions can be made from Model N. Models SN and N give quite different estimates to the outcome-specific parameters for the continuous SB and SA, but they provide similar estimates to the outcome-specific parameters for the binary HM and ordinal HS (a and b; see Table 2 for details).

The Bayesian modeling approach through MCMC procedure has made it possible to fit increasingly complex statistical models [31,50]. To select the “best” model that fits the data adequately among candidate models, the deviance information criterion (DIC) [50] is used. With caution here, instead of being intended to identify the “correct” model, DIC is only used to find the one that fits the data better. It is seen from Table 1 that the both DIC=10036.5 and the posterior mean of the deviance ($\bar{D} = 9722.3$) in Model SN is smaller than its counterparts (DIC=14659.4 and $\bar{D} = 133649.0$) in Model N. Therefore, Model SN performs better than Model N, suggesting that it is important to take

skewness and heaviness in the longitudinal SB and SA data into account in order to achieve more reliable results, in particular if the data exhibit non-normality. With these findings, we will further report our results in detail below only for the better Model SN.

Table 1 also shows positive correlation coefficient ρ between u_{i0} and u_{i1} (0.061, 95% CI: (0.029, 0.200)) in Model SN, suggesting that the patient with worse baseline disease severity tend to have faster disease deterioration and vice versa. To obtain more insight into u_{i0} , u_{i1} , and ρ , we plot in Figure 3 the rankings of subject-specific baseline disease severity u_{i0} (upper panel) and disease progression rate u_{i1} (lower panel) with their 95% credible intervals. Each patient is ordered by her or him rank: individuals at the left corner show milder disease severity at baseline (lower rank) and slower disease progression rate (lower rank), while individuals at the right corner have poorer disease severity at baseline (higher rank) and faster disease progression rate (higher rank). To visualize the effect of the correlation coefficient ρ , we have selected two patients as examples. Patient 199 has the worst baseline disease severity and this patient ranks number 247 in the disease progression rate. Patient 227 has the fastest disease progression rate and this patient ranks number 71 in the baseline disease severity.

To visualize the difference in the disease progression rates in two groups, Figure 4 displays the estimates of the latent disease severity θ_{ij} of the patients at each visit, together with the loess smooth curves denoted by the dashed (placebo group)

Table 2. Summary of estimated posterior mean (PM), standard deviation (SD) of population (fixed-effects) parameters and the corresponding 95% equal-tail credible interval (CI) for additional parameters.

Parameter	Model SN			Model N		
	PM	SD	95% CI	PM	SD	95% CI
For SB and SA						
a_1	1.780	0.662	(0.912, 3.504)	5.662	5.293	(-4.542, 10.29)
b_1	1.577	0.712	(0.488, 3.020)	2.786	1.115	(1.221, 5.473)
a_2	3.372	0.893	(1.820, 5.096)	1.771	1.451	(-1.092, 4.696)
b_2	0.539	0.189	(0.212, 0.889)	0.418	0.157	(0.215, 0.893)
σ_{11}^2	0.077	0.020	(0.045, 0.123)	5.683	0.245	(5.209, 6.183)
σ_{12}^2	0.025	0.013	(0.001, 0.051)	-0.094	0.021	(-0.136, -0.053)
σ_{22}^2	0.101	0.004	(0.094, 0.109)	0.103	0.004	(0.096, 0.111)
γ_{11}	1.564	0.684	(-0.296, 2.482)	-3.268	5.286	(-5.750, 3.942)
γ_{21}	8.732	1.138	(6.382, 11.08)	9.112	1.916	(5.318, 12.83)
γ_{31}	1.761	1.584	(-1.940, 3.080)	-3.053	9.750	(-7.238, 5.740)
γ_{12}	0.177	0.895	(-1.535, 1.727)	1.804	1.453	(-1.117, 4.671)
γ_{22}	-0.618	0.108	(-0.835, -0.410)	-0.638	0.109	(-0.855, -0.431)
γ_{32}	-2.349	1.145	(-3.141, 2.936)	-1.023	1.180	(-3.160, 5.335)
For HM						
a_3	-0.393	0.109	(-0.603, -0.177)	-0.403	0.096	(-0.957, -0.219)
b_3	2.898	1.009	(1.169, 4.740)	2.077	0.773	(1.070, 3.943)
For HS						
a_{41}	-3.708	0.636	(-5.010, -2.568)	-3.643	0.556	(-4.742, -2.558)
a_{42}	1.836	0.606	(0.647, 2.948)	1.853	0.540	(0.835, 2.952)
b_4	8.630	1.453	(6.566, 9.180)	3.410	1.907	(1.854, 6.990)

and solid (treatment group) lines, respectively. **Figure 4** shows that the placebo individuals' PBC severity deteriorates at a faster rate than the treatment individuals before about year 11 and then the treatment group shows larger disease progression rate than the placebo group starting from about year 11, as manifested by the departure of two loess curves. This finding confirm the fact that the D-penicillin treatment group appears effective from the beginning to around year 11, but becomes negatively effective in comparison with the placebo group due to the various adverse effects on D-penicillin treatment [7,17].

Simulation studies

To assess performance of the proposed MPJM models and method, as an illustration we conduct the following simulation studies where Models SN and N are included. The design of the simulated data is similar to the real data used in Section 4. We generate the sample size $N = 300$ (150 in both treatment and placebo group), and assume that each subject has 15 scheduled longitudinal measurements. We simulate 500 datasets from specified models according to the additional specifications described below. The measurement time points used in the simulation are similar to those in the real data analysis and the true parameter values are mimic to those obtained in the real data analysis so that all simulated data make biological sense. Specifically, we simulate two continuous (y_{j1} and y_{j2}), one binary (y_{j3}) and one ordinal (y_{j4} with 3 categories) outcomes. The covariates considered which are mimic from real data include age with mean 50 years of old generated from the Uniform distribution ranged between 40 and 60, sex (male as reference) generated from Bernoulli(0.8), and treatment variable (1 if treatment and 0 if placebo) generated from Bernoulli(0.5). The true parameter values in MVJM (9)-(11) are selected as follows. $\beta = (\beta_{01}, \beta_{10}, \beta_{11})^T = (-0.10, 1.48, -0.22)^T$, $\rho = 0.10$, $\sigma^2 = 1.69$, $\alpha = (\alpha_1, \alpha_2, \alpha_3)^T = (-0.15,$

$0.20, -0.62)^T$, $v = (v_1, v_2)^T = (1.77, 0.83)^T$; $(a_1, b_1)^T = (1.8, 1.6)^T$, $(a_2, b_2)^T = (3.4, 0.5)^T$, $(a_3, b_3)^T = (-0.4, 2.9)^T$, $(a_{41}, a_{42}, b_4)^T = (-3.7, 1.8, 8.6)^T$; $\Sigma_2 = (\sigma_{kk}^2)_{2 \times 2}$ ($k, k' = 1, 2$), where $(\sigma_{11}^2, \sigma_{12}^2, \sigma_{22}^2)^T = (0.07, 0.03, 0.1)^T$; $\gamma_1 = (\gamma_{11}, \gamma_{21}, \gamma_{31})^T = (1.6, 8.7, 1.8)^T$, $\gamma_2 = (\gamma_{12}, \gamma_{22}, \gamma_{32})^T = (0.2, -0.6, -2.3)^T$; Σ_ξ values are mimic to those estimated in the real data.

In the simulation, we generated ϵ_{ijk} ($k = 1, 2$) according to $\Gamma(2, 1)$ distribution, yielding a skewed distribution with the mean $E(\epsilon_{ijk}) = 2$ and variance $Var(\epsilon_{ijk}) = 2$. Under this specification, data generated from the model with $\Gamma(2, 1)$ distribution exhibit skewed feature for two continuous (y_{j1} and y_{j2}) outcomes. Note that we adopt the same MCMC strategies to those in real data analysis in the simulation studies and the prior distributions considered are all close to non-informative; that is, prior distributions with large variances are considered. Thus, we expect the results to be somewhat robust with respect to prior distributions.

For evaluating the objective use of the criteria, the models preferred by DIC are recorded. For example, in the MCMC sampling results, none of DIC selected the normal distribution specification for any of the 100 data sets, demonstrating the ability of selection method to detect an obvious departure from non-normality which provides strong evidence of skewness. **Table 3** summarizes simulation results that include the true parameter (TP) values, average estimates (EST) of fixed-effects β , α , v , $\delta = (\delta_1, \delta_2)^T$ as well as associated the bias (quantified by relative percent bias = $100 \times (EST - TP)/|TP|$) and mean squared error (MSE) (quantified by relative percent square root of MSE = $100 \times \sqrt{MSE}/|TP|$).

To compare the numerical results of Models SN and N considered in the simulation studies, it is of interest to see that the estimated biases for β_{01} and β_{10} are negative, indicating that these parameters (the estimated baseline group difference and disease progression rate for placebo group) are underestimated,

Table 3. Simulation results of the true parameter (TP) values, average estimates (EST), Bias and MSE based on 100 simulated datasets for the interesting parameters.

TP	Model SN			Model N		
	EST	Bias	MSE	EST	Bias	MSE
Longitudinal parameter estimates						
$\beta_{01} = -0.10$	-0.11	-4.53	9.78	-0.13	-7.32	11.5
$\beta_{10} = 1.48$	1.42	-5.66	8.91	1.37	-8.04	9.87
$\beta_{11} = -0.22$	-0.20	5.31	9.05	-0.18	7.02	10.8
$\rho = 0.10$	0.11	3.87	7.53	0.14	9.55	8.91
$\sigma_u^2 = 1.69$	1.71	3.89	4.69	1.62	-10.3	11.2
δ_1	2.97	—	—	—	—	—
δ_2	1.89	—	—	—	—	—
Survival parameter estimates						
$\alpha_1 = -0.15$	-0.16	-7.33	11.5	-0.18	-9.04	12.3
$\alpha_2 = 0.20$	0.22	5.89	8.88	0.24	7.65	9.29
$\alpha_3 = -0.62$	-0.60	6.34	7.91	-0.56	9.61	9.39
$\nu_1 = 1.77$	1.75	4.32	6.88	1.73	9.01	11.5
$\nu_2 = 0.83$	0.85	6.66	4.78	0.86	7.02	6.50

while estimated biases for β_{11} is positive, suggesting that the disease progression rate in treatment group is overestimated. The average estimates of skewness parameters $\delta_1 = 2.97$ for SB and $\delta_2 = 1.89$ for SA in Model SN indicate a departure from (symmetric) normal distribution for both SB and SA longitudinal measurements. For the parameters in the survival sub-model, all parameters in Models SN and N are consistently overestimated or underestimated in general. Because the parameters v_1 and v_2 are set to be positive, the patients with worse baseline disease severity (larger u_{i0}) and faster disease progression rate (larger u_{i1}) tend to have a terminal event earlier. These results suggest that Model SN can recover the true values better than Model N in the presence of dependent event.

In summary, it is seen that, overall, Model SN outperforms Model N in terms of estimates, bias, and MSE. This suggests that adopting the assumption of (symmetric) normal distribution may lead to inaccurate inference on fixed-effects of primary interest, in particular, when data exhibit non-normal feature. Note that we investigated other combined scenarios of the sample sizes ($N = 500$) and the number of replications (1000) in our simulation studies. We found that the results are quite similar to those reported in this article. Both Models SN and N provide reasonable estimates to the difficulty and discriminating parameter vectors $\mathbf{a}$ and $\mathbf{b}$, but Model SN generally provides more accurate estimates than Model N (the results not shown here to save space).

Discussion

With more studies being conducted that repeatedly take measures over time in an effort to evaluate a patient's health status for some events, an MLIRT-based MJM approach is powerful tool to fit these complicated longitudinal-survival models with a variety of data issues ranging from evaluating non-normality, correlated multivariate longitudinal measures and others, along with uncertainty about the distributional assumptions of model errors in clinical and observational studies. In clinical trial studies, it is quite common to have various longitudinal outcomes with mixed types (continuous, binary and ordinal) subject to dependent time-to-event. Previous work of joint modeling for this type of data has been mainly focused on a single type of longitudinal outcomes accounting for the dependent censoring. In this article, we propose a multivariate joint modeling framework that consists of an extended MLIRT sub-model for the multivariate longitudinal outcomes with mixed types (continuous, binary and ordinal) and the Cox proportional hazard sub-model for time-to-event endpoint. Two sub-models are linked together via shared random-effects representing the subject-specific baseline disease severity and disease progression rate, respectively. We assume that the survival time may be associated with the multivariate longitudinal outcomes. We adopt Bayesian inference framework based on MCMC procedure for parameter estimation. This Bayesian framework provides not only accurate parameter estimates, but also various subject-specific PBC disease severity estimation. A benefit of directly

estimating parameters of key interest in our MJM, including the (latent) disease severity for all individuals at each visit through the baseline disease severity and disease progression rates is that these subject-specific quantities can be served as surrogate indicators of characteristic of longitudinal data and, in turn, incorporate into the survival sub-model as covariates to assess the effects of longitudinal profiles on the risk of time-to-event. This kind of joint modeling approach is important in many statistical application areas, allowing accurate inference of parameters while adjusting for non-normality and correlation in the multivariate longitudinal data setting.

The proposed Model SN has a better fit than Model SN in the analysis of the PBC dataset. We have found that the treatment D-penicillin is insignificant in slowing the PBC disease progression. Moreover, we have identified a significant positive correlation between the multiple longitudinal outcomes and the time to death, in addition to the positive significant correlation between the baseline disease severity and disease progression rate. The simulation studies have shown that in the presence of dependent time-to-event, Model SN successfully recovers the true parameters better than Model N which has large bias and MSE in the regression and random-effects parameters. We provide subject-specific disease severity and disease progression rate estimates for all patients at each visit, in addition to the figures for clear visualization. The proposed models have the capability to accommodate hierarchically structured data, to estimate latent disease severity at different levels, and to investigate the relationships between disease severity and predictors at different levels. The proposed model can be easily implemented using the publicly available BUGS language and can be readily accessible to, modified, and extended by applied researchers. Further, our results suggested that it is important to consider MJM with the skewed distribution in order to achieve less biased and more accurate estimates in the presence of non-normal feature in SB and SA measurements. Although this article is motivated by a PBC study, the basic concepts of the developed Bayesian joint modeling approach have generally broader applications whenever the two different sources of dependence among longitudinal measures over time, and between longitudinal exposures and survival endpoint are presented and the relevant technical specifications are met.

A couple of notes and extensions related to this study are made as follows. (i) Our modeling framework provides great flexibility for extensions. This article only considers a single-type terminal event (death). In the presence of multiple 'failure types', for example, death and dropout due to disease-related causes, the proposed MJM can be extended to accommodate competing risks survival data. (ii) In the structural multilevel model (4), the latent disease severity θ_{ij} is determined by the random-effects u_{i0} and u_{i1} , and covariate vectors X_{i0} and X_{i1} , while model (4) in [21] is more flexible allowing additional randomness by adding a random error term ϵ_i . But this flexibility does not come without paying a price; see Fox and Glas [21] for more detailed discussion of this direction. (iii) To model the nonlinear disease

progression rate, the linear time trend assumption in model (4) can be relaxed by adding higher order terms or smoothing spline terms of time [48]. Moreover, it is assumed that there exists a single (unidimensional) latent variable θ_{ij} to measure the underlying disease severity. However, there may be multiple latent variables representing multidimensional impairment caused by PBC. Expanding the unidimensional MLIRT model to the multidimensional one is an interesting direction of research. (iv) A final note to be made is that we may consider the missing data mechanism for longitudinal data by introducing a non-ignorable missing data model [27,28,34]. These interesting topics are beyond the focus of this article, but are warranted in our research pipeline under investigation.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Authors' contributions	YH	JC	LX	HZ	YL
Research concept and design	✓	✓	--	--	--
Collection and/or assembly of data	✓	--	✓	--	--
Data analysis and interpretation	✓	--	--	--	--
Writing the article	✓	✓	--	--	--
Critical revision of the article	✓	✓	✓	✓	✓
Final approval of article	✓	✓	✓	✓	✓
Statistical analysis	✓	✓	✓	✓	✓

Acknowledgment

The authors gratefully acknowledge the Editor, Section Editor and anonymous referees for their insightful comments and constructive suggestions that led to a marked improvement of the article. This work was partially supported by the National Natural Science Foundation of China grant(81671633) to Chen.

Publication history

EIC: Jimmy Efrid, East Carolina University, USA.

Received: 26-Oct-2021 Final Revised: 25-Nov-2021

Accepted: 19-Dec-2021 Published: 31-Dec-2021

References

1. R.J. Adams, M. Wilson, and M. Wu, *Multilevel item response models: An approach to errors in variables regression*, Journal of Educational and Behavioral Statistics 22 (1997), pp. 47–76.
2. P.S. Albert and J.H. Shih, *An approach for jointly modeling multivariate longitudinal measurements and discrete time-to-event data*, The Annals of Applied Statistics 4 (2010), p. 1517–1532.
3. E.B. Andersen, *Latent regression analysis based on the rating scale model*, Psychology Science 46 (2004), pp. 209–226.
4. E.R. Andrinopoulou and D. Rizopoulos, *Bayesian shrinkage approach for a joint model of longitudinal and survival outcomes assuming different association structures*, Statistics in Medicine 35 (2016), pp.4813–4823.
5. R.B. Arellano-Valle and M.G. Genton, *On fundamental skew distributions*, Journal of Multivariate Analysis 96 (2005), pp. 93–116.
6. A. Azzalini and A. Capitanio, *Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 65 (2003), pp. 367–389.
7. H.C. Bodenhimer Jr, F. Schaffner, I. Sternlieb, F.M. Klion, S. Vernace, and J. Pezzullo, *A prospective clinical trial of d-penicillamine in the treatment of primary biliary cirrhosis*, Hepatology 5 (1985), pp.1139–1142.
8. E.R. Brown and J.G. Ibrahim, *A Bayesian semiparametric joint hierarchical model for longitudinal and survival data*, Biometrics 59 (2003), pp. 221–228.
9. J. Chen and Y. Huang, *A Bayesian mixture of semiparametric mixed-effects joint models for skewed longitudinal and time-to-event data*, Statistics in Medicine 34 (2015), pp. 2820–2843.
10. D.G. Clayton, *A Monte Carlo method for Bayesian inference in frailty models*, Biometrics (1991), pp.467–485.
11. M.J. Crowther, K.R. Abrams, and P.C. Lambert, *Joint modeling of longitudinal and survival data*, The Stata Journal 13 (2013), pp. 165–184.
12. S. Curtis, *Bugs code for item response theory*, Journal of Statistical Software 36 (2010), pp. 1–34.
13. M. Davidian and D.M. Giltinan, *Nonlinear Models for Repeated Measurement Data*, Vol. 62, CRCpress, 1995.
14. E. Dil and D. Karasoy, *Joint modeling of a longitudinal measurement and parametric survival data with application to primary biliary cirrhosis study*, Pakistan Journal of Statistics and Operation Research (2020), pp. 295–304.
15. J. Ding and J.L. Wang, *Modeling longitudinal data with nonparametric multiplicative random effects jointly with survival data*, Biometrics 64 (2008), pp. 546–556.
16. J.A. Douglas, *Item response models for longitudinal quality of life data in clinical trials*, Statistics in Medicine 18 (1999), pp. 2917–2931.
17. O. Epstein and S. Sherlock, *Triethylene tetramine dihydrochloride toxicity in primary biliary cirrhosis*, Gastroenterology 78 (1980), pp. 1442–1445.
18. T. Fleming and D. Harrington, *Counting processes and survival analysis* John Wiley & Sons, Inc. New York (1991).
19. J. Fox, *Bayesian Item Response Modeling: Theory and Applications*, New York, USA: Springer-Verlag, 2010.
20. J.P. Fox, *Multilevel IRT using dichotomous and polytomous response data*, British Journal of Mathematical and Statistical Psychology 58 (2005), pp. 145–172.
21. J.P. Fox and C.A. Glas, *Bayesian estimation of a multilevel irt model using gibbs sampling*, Psychometrika 66 (2001), pp. 271–288.
22. A. Gelman and D.B. Rubin, *Inference from iterative simulation using multiple sequences*, Statistical Science (1992), pp. 457–472.
23. C.A. Glas, H. Geerlings, M.A. van de Laar, and E. Taal, *Analysis of longitudinal randomized clinical trials using item response models*, Contemporary Clinical Trials 30 (2009), pp. 158–170.
24. R.D. Gordon, S. Iwatsuki, A.G. Tzakis, C.O. Esquivel, S. Todo, L. Makowka, and T.E. Starzl, *The Denver-Pittsburgh liver transplant series*, Clinical Transplants (1987), p. 43–49.
25. B. He and S. Luo, *Joint modeling of multivariate longitudinal measurements and survival data with applications to parkinson's disease*, Statistical Methods in Medical Research 25 (2016), pp. 1346–1358.
26. R. Henderson, P. Diggle, and A. Dobson, *Joint modelling of longitudinal measurements and event time data*, Biostatistics 1 (2000), pp. 465–480.
27. Y. Huang, *Quantile regression-based bayesian semiparametric mixed-effects models for longitudinal data with non-normal, missing and mismeasured covariate*, Journal of Statistical Computation and Simulation 86 (2016), pp. 1183–1202.
28. Y. Huang and J. Chen, *Bayesian quantile regression-based nonlinear mixed-effects joint models for time-to-event and longitudinal data with multiple features*, Statistics in Medicine 35 (2016), pp. 5666–5685.
29. Y. Huang and G. Dagne, *A Bayesian approach to joint mixed-effects models with a skew-normal distribution and measurement errors in*

- covariates, *Biometrics* 67 (2011), pp. 260–269.
30. Y. Huang, G. Dagne, and L. Wu, *Bayesian inference on joint models of HIV dynamics for time-to-event and longitudinal data with skewness and covariate measurement errors*, *Statistics in Medicine* 30 (2011), pp. 2930–2946.
 31. Y. Huang, D. Liu, and H. Wu, *Hierarchical bayesian methods for estimation of parameters in a longitudinal hiv dynamic system*, *Biometrics* 62 (2006), pp. 413–423.
 32. Y. Huang, N.S. Tang, and J. Chen, *Multivariate piecewise joint models with random change-points for skewed-longitudinal and survival data*, *Journal of Applied Statistics*, doi.org/10.1080/02664763.2021.1935797 (2021).
 33. L.F. Hung and W.C. Wang, *The generalized multilevel facets model for longitudinal data*, *Journal of Educational and Behavioral Statistics* 37 (2012), pp. 231–255.
 34. J.G. Ibrahim, M.H. Chen, and S.R. Lipsitz, *Missing responses in generalised linear mixed models when the missing data mechanism is nonignorable*, *Biometrika* 88 (2001), pp. 551–564.
 35. A. Kamata, *Item analysis by the hierarchical generalized linear model*, *Journal of Educational Measurement* 38 (2001), pp. 79–93.
 36. F. Lord, M. Novick, and A. Birnbaum, *Statistical Theories of Mental Test Scores*, Boston, MA: Addison-Wesley, 1968.
 37. S. Luo, *A Bayesian approach to joint analysis of multivariate longitudinal data and parametric accelerated failure time*, *Statistics in Medicine* 33 (2014), pp. 580–594.
 38. W. M and T. A, *A joint model for survival and longitudinal data measured with error*, *Biometrics* 53 (1997), pp. 330–339.
 39. K.S. Maier, *A rasch hierarchical measurement model*, *Journal of Educational and Behavioral Statistics* 26 (2001), pp. 307–330.
 40. R.J. Mislevy, *Estimation of latent group effects*, *Journal of the American Statistical Association* 80 (1985), pp. 993–997.
 41. P.A. Murtaugh, E.R. Dickson, G.M. Van Dam, M. Malinchoc, P.M. Grambsch, A.L. Langworthy, and C.H. Gips, *Primary biliary cirrhosis: prediction of short-term survival based on repeated patient visits*, *Hepatology* 20 (1994), pp. 126–134.
 42. P. Natesan, C. Limbers, and J.W. Varni, *Bayesian estimation of graded response multilevel models using gibbs sampling: formulation and illustration*, *Educational and Psychological Measurement* 70 (2010), pp. 420–439.
 43. L.M. O'Brien and G.M. Fitzmaurice, *Analysis of longitudinal multiple-source binary data using generalized estimating equations*, *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 53 (2004), pp. 177–193.
 44. G. Papageorgiou, K. Mauff, A. Tomer, and D. Rizopoulos, *An overview of joint modeling of time-to-event and longitudinal outcomes*, *Annual Review of Statistics and Its Application* 6 (2019), pp. 223–240.
 45. D. Rizopoulos, *Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data*, *Biometrics* 67 (2011), pp. 819–829.
 46. D. Rizopoulos, *Joint models for longitudinal and time-to-event data: With applications in R*, CRC Press, 2012.
 47. D. Rizopoulos, et al., *JM: An R package for the joint modelling of longitudinal and time-to-event data*, *Journal of Statistical Software* 35 (2010), pp. 1–33.
 48. D. Ruppert, M.P. Wand, and R.J. Carroll, *Semiparametric Regression*, 12, Cambridge University Press, 2003.
 49. S.K. Sahu, D.K. Dey, and M.D. Branco, *A new class of multivariate skew distributions with applications to Bayesian regression models*, *Canadian Journal of Statistics* 31 (2003), pp. 129–150.
 50. D.J. Spiegelhalter, N.G. Best, B.P. Carlin, and A. Van Der Linde, *Bayesian measures of model complexity and fit*, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64 (2002), pp. 583–639.
 51. P. Tong and K.R. Coombes, *integrity: a method to identify genes altered in cancer by accounting for multiple mechanisms of regulation using item response theory*, *Bioinformatics* 28 (2012), pp. 2861–2869.
 52. A.A. Tsiatis and M. Davidian, *Joint modeling of longitudinal and time-to-event data: an overview*, *Statistica Sinica* 14 (2004), pp. 809–834.
 53. C. Wang, J. Douglas, and S. Anderson, *Item response models for joint analysis of quality of life and survival*, *Statistics in Medicine* 21 (2002), pp. 129–142.
 54. L. Wu, W. Liu, and X. Hu, *Joint inference on HIV viral dynamics and immune suppression in presence of measurement errors*, *Biometrics* 66 (2010), pp. 327–335.
 55. H. Zhang and 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*, *Journal of Biopharmaceutical Statistics* (2020), pp. 1–22.

Citation:

Huang Y, Chen J, Xu L, Zhang H and Lu Y. **Bayesian MLIRT-based joint models for multivariate longitudinal and survival data with multiple features.**
J Med Stat Inform. 2021; **9**:4.
<http://dx.doi.org/10.7243/2053-7662-9-4>