Bayesian Joint Modeling of Multivariate Longitudinal and Survival Data With an Application to Diabetes Study

Yangxin Huang[1][*][†], Jiaqing Chen[2][†], Lan Xu[1] and Nian-Sheng Tang[3]

[1] College of Public Health, University of South Florida, Tampa, FL, United States, [2] Department of Statistics, College of Science, Wuhan University of Technology, Wuhan, China, [3] Department of Statistics, Yunnan University, Kunming, China

Joint models of longitudinal and time-to-event data have received a lot of attention in epidemiological and clinical research under a linear mixed-effects model with the normal assumption for a single longitudinal outcome and Cox proportional hazards model. However, those model-based analyses may not provide robust inference when longitudinal measurements exhibit skewness and/or heavy tails. In addition, the data collected are often featured by multivariate longitudinal outcomes which are significantly correlated, and ignoring their correlation may lead to biased estimation. Under the umbrella of Bayesian inference, this article introduces multivariate joint (MVJ) models with a skewed distribution for multiple longitudinal exposures in an attempt to cope with correlated multiple longitudinal outcomes, adjust departures from normality, and tailor linkage in specifying a time-to-event process. We develop a Bayesian joint modeling approach to MVJ models that couples a multivariate linear mixed-effects (MLME) model with the skew-normal (SN) distribution and a Cox proportional hazards model. Our proposed models and method are evaluated by simulation studies and are applied to a real example from a diabetes study.

Keywords: Bayesian inference, longitudinal and survival data, Markov Chain Monte Carlo, multivariate joint models, skew-normal distribution

1. INTRODUCTION

In epidemiologic and clinical studies, a lot of attention is focused on developing the specific patterns of the longitudinal measurements and the associations between those patterns and the time to a certain event, such as diagnosis of disease, time to transplantation, or death. Those studies have been in a highly active research area (Henderson et al., 2000; Brown and Ibrahim, 2003; Tsiatis and Davidian, 2004; Rizopoulos, 2011, 2012). For example, in diabetes studies, repeated measures of continuous exposures such as the children's growth (height and weight) and time to type 1 diabetes (T1D) are collected.

During the last two decades, the research on joint modeling of longitudinal and time-to-event data has been received rapid and considerable development. In literature, various joint models and associated statistical methods have been introduced to analyze such longitudinal and survival data. However, the following issues may stand out. (i) Most joint models focus on a single longitudinal variable associated with a time-to-event outcome (Henderson et al., 2000; Brown and Ibrahim, 2003; Tsiatis and Davidian, 2004; Rizopoulos, 2011, 2012). However, in practice, many

studies often collect multiple longitudinal outcomes (Lin et al., 2002; Brown et al., 2005; Chi and Ibrahim, 2006; Fieuws and Verbeke, 2006; Albert and Shih, 2010; Rizopoulos and Ghosh, 2011; Kim and Albert, 2016; Chen and Wang, 2017; Tang et al., 2017a,b; Proudfoot et al., 2018; Chen et al., 2021) which may be significantly correlated. For example, the weight and height repeated measures presented in **Figure 1** (left and middle panels) show significant correlation and it may lead to biased estimation if their correlation is ignored. In addition, time-to-event such as the time to T1D depicted in **Figure 1** (right panel) may be dependent on the longitudinal weight and height measures. (ii) In traditional linear mixed-effects models, within subject measurement errors are often under a normality assumption due to the mathematical tractability and computational convenience. However, the normality assumption may not be realistic. Alternatively, the skew-elliptical (SE) distributions including skew-normal (SN) distribution (Sahu et al., 2003) should be more appropriate to model the skewed data (Azzalini and Capitanio, 2003; Sahu et al., 2003; Arellano-Valle and Genton, 2005; Huang and Dagne, 2011). Although a few studies investigated multivariate joint (MVJ) models (Chi and Ibrahim, 2006; Albert and Shih, 2010; Rizopoulos and Ghosh, 2011; Kim and Albert, 2016; Chen and Wang, 2017; Tang et al., 2017a,b; Chen et al., 2021), they have not considered non-normal features of longitudinal data.

In statistical literature, longitudinal data analysis has focused on developing models to capture only specific aspects of the motivating studies. Inferential procedures may be complex dramatically if one considers multiple longitudinal outcomes data with correlation and skewed distributions in conjunction with an event time in MVJ models. Our Bayesian approach enables the fitting of such models efficiently and the convergence problem can be solved.

The rest of the article is organized as follows. In Section 2, we introduce MVJ models with the SN distribution and discuss the associated Bayesian infer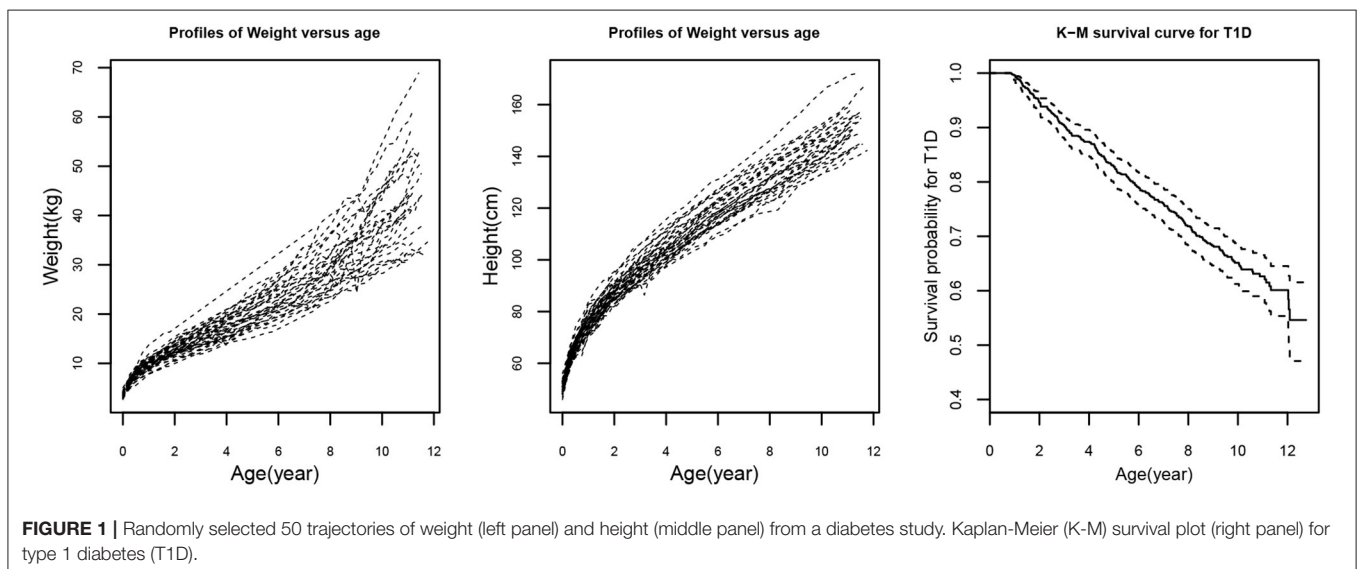ential method. In Section 3, we present the data set from a diabetes study that motivated this research, apply the specific MVJ model to the data, and report the results. Section 4 conducts limited simulation studies to evaluate the performance of the proposed models and method. Finally, a general discussion and conclusion are presented in Section 5.

2. JOINT MODELS AND ASSOCIATED BAYESIAN APPROACH

This section presents the MVJ model and related Bayesian modeling method in full generality for multiple longitudinal data with non-normality and correlation and survival endpoint with censoring to illustrate that our modeling method can be applied in various applications. To relax the normality assumption, the multivariate longitudinal model with SN distribution is assumed. Let y_{ijk} denote an observation of the kth longitudinal variable ($k = 1, 2, \ldots, K$) for the ith subject ($i = 1, 2, \ldots, n$) at the jth visit time t_{ijk} ($j = 1, 2, \ldots, n_i$). Let $\boldsymbol{Y}_i = (\boldsymbol{Y}_{i1}^T, \ldots, \boldsymbol{Y}_{iK}^T)^T$ be the K-variate vector of continuous longitudinal responses, where $\boldsymbol{Y}_{ik} = (y_{i1k}, \ldots, y_{in_ik})^T$. Similarly, we can define $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$. Let the vector of population parameters $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, \ldots, \boldsymbol{\beta}_K^T)^T$ with $\boldsymbol{\beta}_k^T = (\beta_{0k}, \beta_{1k}, \beta_{2k}, \ldots, \beta_{pk})^T$ associated with the kth longitudinal variable. The vector of subject-specific parameters by $\boldsymbol{b}_i = (\boldsymbol{b}_{i1}^T, \ldots, \boldsymbol{b}_{iK}^T)^T$ with $\boldsymbol{b}_{ik}^T = (b_{i0k}, b_{i1k}, b_{i2k}, \ldots, b_{iqk})^T$. We denote T_i^* as the 'event' time, C_i as the censoring time and $T_i = min(T_i^*, C_i)$ as the observed time for subject i. Let ρ_{ij} denote the indicator for an event, i.e., $\rho_{ij} = 0$ when censoring occurs and $\rho_{ij} = 1$ ($T_i^* \leq C_i$) when the event is observed. $\boldsymbol{x}_i$ is a covariate vector that may be associated with an event time.

2.1. Multivariate Linear Mixed Effects Models With SN Distribution
We consider a general multivariate linear mixed-effects (MLME) model with SN distribution as follows.

FIGURE 1 | Randomly selected 50 trajectories of weight (left panel) and height (middle panel) from a diabetes study. Kaplan-Meier (K-M) survival plot (right panel) for type 1 diabetes (T1D).

$$Y_{ik} = X_{ik}\boldsymbol{\beta}_k + Z_{ik}\boldsymbol{b}_{ik} + \boldsymbol{\epsilon}_{ik},$$

$$\boldsymbol{b}_i = (\boldsymbol{b}_{i1}^T, \ldots, \boldsymbol{b}_{iK}^T)^T \overset{\text{iid}}{\sim} N_{Kq}(\boldsymbol{0}, \boldsymbol{\Sigma}_b),$$

$$\boldsymbol{\epsilon}_i = (\boldsymbol{\epsilon}_{i1}^T, \ldots, \boldsymbol{\epsilon}_{iK}^T)^T \overset{\text{iid}}{\sim} \qquad\qquad\qquad (1)$$

$$SN_{Kn_i}\left(-\sqrt{2/\pi}[\boldsymbol{\Delta}_K \otimes \boldsymbol{1}_{n_i}], \boldsymbol{\Sigma}_K \otimes \boldsymbol{I}_{n_i}, \boldsymbol{\Delta}_K \otimes \boldsymbol{I}_{n_i}\right),$$

where the random error vector $\boldsymbol{\epsilon}_i$ follows a multivariate SN distribution with unknown variance-covariance matrix $\boldsymbol{\Sigma}_K = (\sigma_{kk'}^2)_{K \times K}$ ($k, k' = 1, 2, \ldots, K$, unknown skewness parameter matrix $\boldsymbol{\Delta}_K = \text{diag}(\delta_1, \ldots, \delta_K)$, the skewness parameter vector $\boldsymbol{\delta}_K = (\delta_1, \ldots, \delta_K)^T$, and $\boldsymbol{1}_{n_i} = (1, \ldots, 1)^T$. Note that $-\sqrt{2/\pi}[\boldsymbol{\Delta}_K \otimes \boldsymbol{1}_{n_i}]$ is considered here in order to make the SN distribution with mean zero; refer to (Sahu et al., 2003; Huang and Dagne, 2011; Xu, 2021) for a detailed discussion of SN distribution. The vector of random-effects $\boldsymbol{b}_i$ follows $N_{Kq}(\boldsymbol{0}, \boldsymbol{\Sigma}_b)$ with $\boldsymbol{\Sigma}_b$ being a covariance matrix. In the application below, we are interested in the height and weight longitudinal data. Let $\boldsymbol{\Delta}_2 = \text{diag}(\delta_1, \delta_2)$, and δ_1 and δ_2 quantify skewness of height and weight, respectively, which is the case represented in this article.

2.2. Cox Proportional Hazard Model

In survival analysis, the semiparametric Cox hazard model (Cox, 1972; Xu, 2021) has been commonly adopted to explore the association between survival time and one or more covariates in medical research. To account for the association between the multiple longitudinal exposures, we assume that the distribution of T_i, the time to diagnosis of T1D for subject i, depends on the random-effects of individual-specific longitudinal processes $\boldsymbol{b}_{ik}$, and other covariates $\boldsymbol{x}_i$, respectively. The survival model for T_i here is linked to the multivariate longitudinal model (6) through the random-effects $\boldsymbol{b}_i$. In addition, assuming that the covariates $\boldsymbol{x}_i$ are associated with the event time. In particular, the conditional hazard rate of T_i at time t_i for the survival component is given by Xu (2021).

$$\lambda(t_i|\boldsymbol{b}_i, \boldsymbol{x}_i) = \lambda_0(t_i)\exp(\boldsymbol{\Upsilon}^T\boldsymbol{b}_i + \boldsymbol{\alpha}^T\boldsymbol{x}_i) = \lambda_0(t_i)\exp(\boldsymbol{\gamma}^T\boldsymbol{d}_i), \quad (2)$$

where $\lambda_0(t_i)$ is the baseline hazard function, $\boldsymbol{d}_i = (\boldsymbol{b}_i^T, \boldsymbol{x}_i^T)^T$, $\boldsymbol{\gamma} = (\boldsymbol{\Upsilon}^T, \boldsymbol{\alpha}^T)^T$, $\boldsymbol{\Upsilon}$ and $\boldsymbol{\alpha}$ are unknown parameters linked with the covariates $\boldsymbol{x}_i$ and random-effects $\boldsymbol{b}_i$ to the conditional hazard rate, respectively. The association parameter vector $\boldsymbol{\Upsilon}$ linking the random-effects $\boldsymbol{b}_i$ measures the association between the two sub-models. An alternative method can be used to approximate the Cox proportional hazards model (2) through the counting process (Clayton, 1991) which is adopted for our joint modeling and can obviously reduce computational burdens; a detailed discussion of the alternative method can be found in Huang (2016), Huang and Chen (2016), and Zhang and Huang (2021).

2.3. Simultaneous Bayesian Inferential Approach

Generally, different approaches are applied to link the longitudinal and survival submodels. The first approach is

under the framework of the likelihood inferential methods, such as the Expectation-Maximization (EM) algorithm and Monte Carlo Expectation-Maximization (MCEM) algorithm (Rizopoulos et al., 2010; Farcomeni and Viviani, 2015; Xu, 2021). A simultaneous inferential method through a joint likelihood may be favorable, but the computational burden for proposed MVJ modelings can be very intensive, even sometimes infeasible, and may cause problems of algorithm convergence (Brown and Ibrahim, 2003; Wu et al., 2010). The second approach is Bayesian inference, Bayesian joint modeling method shows the advantage. Thus, the parameters can be estimated simultaneously for the MVJ models implemented by Markov Chain Monte Carlo (MCMC) techniques for the skew-normal MVJ model. The simultaneous statistical inference on all unknown model parameters will capture the underlying association between the longitudinal exposures and the event time data.

First, $\boldsymbol{b}_{ik}$ and $\boldsymbol{\epsilon}_{ik}$ are assumed mutually independent of each other. To specify the Model (1) for MCMC computational techniques, by introducing the n_i-dimensional random vector $\boldsymbol{w}_i$ based on the stochastic representation of SN distribution detailed in the publication (Huang and Dagne, 2011), we can hierarchically formulate the MVJ model, which consists of the MLME Model (1) and Cox proportional hazard Model (2) as follows.

$$\boldsymbol{Y}_i|\boldsymbol{b}_i, \boldsymbol{w}_i \sim N\left(X_i\boldsymbol{\beta} + Z_i\boldsymbol{b}_i + \boldsymbol{\Delta}_K \otimes [\boldsymbol{w}_i - \sqrt{2/\pi}\boldsymbol{1}_{n_i}], \boldsymbol{\Sigma}_K \otimes \boldsymbol{I}_{n_i}\right),$$

$$\boldsymbol{w}_i \sim N_i(\boldsymbol{0}, \boldsymbol{I}_{n_i})I(\boldsymbol{w}_i > \boldsymbol{0}), \quad \boldsymbol{b}_i \sim N_{Kq}(\boldsymbol{0}, \boldsymbol{\Sigma}_b),$$

$$T_i \sim F(t_i|\boldsymbol{d}_i, \lambda_0) = \int f(\boldsymbol{\rho}_i|\boldsymbol{b}_i, \boldsymbol{x}_i), \qquad\qquad (3)$$

Then, under the Bayesian framework, we need to specify $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_b, \boldsymbol{\Sigma}_K, \boldsymbol{\delta}_K)$ as all unknown population parameters in the joint Model (1) and (2), where $\boldsymbol{\delta}_K = (\delta_1, \ldots, \delta_K)^T$ is the vector of skewness parameters. We assign weakly informative priors to ensure the property of posteriors. Thus, the prior distributions for all of the unknown parameters are specified as follows.

$$\boldsymbol{\beta} \sim N\left(\boldsymbol{\beta}_0, \boldsymbol{\Omega}_1\right), \quad \boldsymbol{\gamma} \sim N(\boldsymbol{\gamma}_0, \boldsymbol{\Omega}_2), \quad \boldsymbol{\Sigma}_b \sim IW(\boldsymbol{\Omega}_3, \omega_1),$$
$$\boldsymbol{\Sigma}_K \sim IW(\boldsymbol{\Omega}_4, \omega_2), \quad \boldsymbol{\delta}_K \sim N(\boldsymbol{0}, \boldsymbol{\Omega}_5), \qquad (4)$$

where the mutually independent Normal (N) and Inverse Wishart (IW) prior distributions are chosen to facilitate computations. The super-parameter matrices $\boldsymbol{\Omega}_1$, $\boldsymbol{\Omega}_2$, $\boldsymbol{\Omega}_3$, $\boldsymbol{\Omega}_4$ and $\boldsymbol{\Omega}_5$ can be assumed to be diagonal for convenient implementation.

Subsequently, let $f(\cdot)$, $f(\cdot|\cdot)$, $F(\cdot|\cdot)$, and $\pi(\cdot)$ denote a density function, a conditional density function, a cumulative density function (c.d.f), and a prior density function, respectively. As the elements of $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_b, \boldsymbol{\Sigma}_K, \boldsymbol{\delta}_K\}$ are assumed to be independent of each other, we have $\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta})\pi(\boldsymbol{\gamma})\pi(\boldsymbol{\Sigma}_b)\pi(\boldsymbol{\Sigma}_K)\pi(\boldsymbol{\delta}_K)$. After we specify the MVJ model for the observed data and the prior distributions for the unknown parameters, we can make the Bayesian inference for the parameters based on their posterior distributions. Thus,

the joint posterior density of $\boldsymbol{\theta}$ based on the observed data $\mathcal{D} = \{\boldsymbol{Y}_i, \boldsymbol{b}_i, \boldsymbol{\rho}_i\}$ can be represented by

$$f(\boldsymbol{\theta}|\mathcal{D}) \propto \{\prod_{i=1}^{n} \int f(\boldsymbol{Y}_i|\boldsymbol{b}_i, \boldsymbol{w}_i)f(\boldsymbol{b}_i) \qquad (5)$$
$$f(\boldsymbol{w}_i|\boldsymbol{w}_i > 0)f(\boldsymbol{\rho}_i|\boldsymbol{b}_i, \boldsymbol{x}_i)d\boldsymbol{b}_i\}\pi(\boldsymbol{\theta}).$$

In general, the integrals in (5) do not have a closed form and are high dimensional. Since the approximated analysis of the integrals may not be sufficiently accurate, it is prohibitive to directly compute the posterior distribution of $\boldsymbol{\theta}$ from the observed data. Alternatively, the MCMC technique can be used for sampling random-effects $\boldsymbol{b}_i$ and population parameters $\boldsymbol{\theta}$ from conditional posterior distributions based on (5), by employing the Gibbs sampler with the Metropolis-Hastings (M-H) algorithm together. We repeat this process in iterations of the MCMC algorithm until convergence is achieved by adopting the publicly available WinBUGS package (Lunn et al., 2000). An advantage of the WinBUGS package is that it does not require explicitly deriving the full conditional posterior distributions for unknown parameters. Although their derivations are straightforward based on the joint posterior (5), they are not presented here to save space due to some cumbersome algebra.

3. APPLICATION

3.1. Motivating Data Set
The motivated data set is from a diabetes study, which is a prospective multinational (U.S., Finland, Sweden, and Germany) cohort to investigate the environmental determinants of T1D (TEDDY Study Group, 2007; Larsson et al., 2008; Chen et al., 2021). This study recruited both the first degree relative (FDR) children and the general population to be screened for genetic predisposition for T1D-related Human Leukocyte Antigen-antigen D and isotype R (HLA-DR) genotypes at the time of birth when they are eligible. The details on the characteristics of families of the diabetes cohort have been reported (Lernmark et al., 2011; Baxter et al., 2012). The participants enrolled in this diabetes study are followed prospectively from birth to 15 years old, with study visits beginning at 3 months of age, then every 3 months until 4 years of age, and every 6 months thereafter depending on the development of T1D. The details of screening and follow-up have been published previously (Kiviniemi et al., 2007; Hagopian et al., 2011). The collected data at each visit time include repeated height and weight measurements, time-to-event outcomes such as the first sign of islet autoimmunity (IA) and clinical diagnosis of T1D, biological data, dietary records, demographic and health histories for the children, and psychological measurements (TEDDY Study Group, 2007). T1D is a common pediatric chronic disease and is preceded by a preclinical period of IA in the presence of islet autoantibodies.

In this real data application here, we consider the dataset of 732 children from all subjects who have developed IA which is the preclinical sign for potentially clinical diagnosis of T1D and have repeated weight and height values from birth to

age at diagnosis of T1D or most recent visit. The confirmed T1D is defined as confirmed positive antibodies to insulin, insulinoma antigen 2, or glutamic acid decarboxylase, which are analyzed by a radiobinding assay at least 2 consecutive visit times (Larsson et al., 2008; Elding Larsson et al., 2014; Chen et al., 2021). For children with an event that occurred, only repeated measurements up to the date of diagnosis of T1D are included in the analysis, while for subjects with censoring, repeated measurements up to the age of 15 are used. A child's growth trajectory in early life shows a quadratic pattern approximately as displayed in **Figure 1** (left and middle panels). The pattern of a subject may be an important clinical implication because of its association with the T1D risk. **Figure 1** (right panel) shows the Kaplan-Meier (K-M) survival curve for T1D as the event. Among the 732 subjects, 246 (33.61%) children are progressed to T1D. The main risk factors used in the analysis include gender (women vs. men), country of residence (Finland, Germany, Sweden as compared to the United States FDR status (yes or no), and HLA genotype (HLA-DR3/4 genotype compared with others) which are the most important genetic and environmental factors in this diabetes study (Larsson et al., 2008; Chen et al., 2021).

3.2. Model Implementation
We illustrate our models and method for the part of longitudinal data described in Section 3.1. We used an SN MLME model with random intercept, random slope, and quadric of age and gender for the longitudinal submodel and adjusted for random intercept and random slope, gender, HLA genotype, and FDR status at baseline in the survival submodel. We consider the following specific bivariate linear mixed-effects models for height and weight:

$$y_{ijk} = (\beta_{0k} + b_{i0k}) + (\beta_{1k} + b_{i1k})\text{Age}_{ij}$$
$$+ \beta_{2k}\text{Age}_{ij}^2 + \beta_{3k}\text{Gender}_i + e_{ijk}, \qquad (6)$$
$$\text{for } k = 1, 2,$$

Specifically, where $k = 1$ and 2 correspond to the respective height and weight responses. y_{ij1} and y_{ij2} are the respective standardized height and weight observations for the *ith* subject at time t_{ij}; the random-effects b_{i0k} and b_{i1k} represent a subject-specific random intercept and a subject-specific random slope, respectively. In this model, the mean baseline value (intercept), mean change rate (slope), and quadratic of age are assumed to be different between men and women.

The survival analysis of the joint model is explained in Section 2.3. The Cox proportional hazard model applied in our study is specified as follows.

$$\lambda(t_i|\boldsymbol{b}_i, \boldsymbol{x}_i) = \lambda_0(t_i)exp(\upsilon_1 b_{0i1} + \upsilon_2 b_{1i1} + \upsilon_3 b_{0i2}$$
$$+ \upsilon_4 b_{1i2} + \alpha_1\text{Finland}_i$$
$$+ \alpha_2\text{Germany}_i + \alpha_3\text{Sweden}_i + \alpha_4\text{Gender}_i \qquad (7)$$
$$+ \alpha_5\text{HLA}_i + \alpha_6\text{FDR}_i).$$

where $\Upsilon = (\upsilon_1, \upsilon_2, \upsilon_3, \upsilon_4)$ is the parameters corresponding to the random-effects $\boldsymbol{b}_i = (b_{i01}, b_{i11}, b_{i02}, b_{i12})$ and other

risk factors are included in the survival model. Additionally, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_6)$ is corresponding to the risk factors including country of residence (Finland, Germany, and Sweden), gender (female = 1), HLA genotype, and FDR status.

In the diabetes study data, height and weight variables exhibit skewness and outliers. Based on the nature of the diabetes study data, the two statistical models with different distributions are implemented to compare their performance as follows.

- **Model N:** MVJ model with the normal distribution for random errors.
- **Model SN:** MVJ model with the SN distribution for the random errors.

Because a normal distribution is a special case of an SN distribution when the skewness parameter becomes zero, we explore how the MVJ model with an SN distribution contributes to modeling results and parameter estimates in comparison with that with a symmetric normal distribution.

In order to perform the Bayesian inferential method, we need to specify the values of the hyper-parameters in the prior distributions (4). Due to the absence of historical data, we apply weakly informative prior distributions for the parameters in MVJ models. In particular, (i) fixed-effects are taken to be independent normal distribution $N(0, 0.01)$ for each element of the population parameter vectors $\boldsymbol{\beta}$, $\boldsymbol{\nu}$ and $\boldsymbol{\alpha}$; (ii) the priors for the variance covariance matrices $\boldsymbol{\Sigma}_K$ and $\boldsymbol{\Sigma}_b$ follow IW distributions $IW(\text{diag}(0.01, 0.01), 2)$ and $IW(\text{diag}(0.01, 0.01, 0.01, 0.01), 4)$; (iii) for each of the skewness parameters δ_1 and δ_2, which represent the skewness of height and weight, respectively, independent normal distributions $N(0, 0.01)$ is chosen.

The MCMC sampler is implemented and the program codes are provided in **Appendix**. When the MCMC algorithm is applied to the diabetes study data, the convergence of the generated samples is assessed using standard tools such as Gelman-Rubin (GR) diagnostics (Gelman and Rubin, 1992) and trace plots. **Figure 2** shows the dynamic version of GR diagnostic plots and the trace plots based on Model SN for the representative parameters β_{01}, β_{02}, δ_1, ν_4, α_1 and α_6. We can see from trace plots (left panel) that the lines of three different chains mix or cross in the trace, indicating that the algorithmic convergence is achieved. For the plots of GR diagnostics (right panel) where the three curves are given. The top curve indicates the ratio ($\hat{R}$) of the middle curve and the bottom curve below the dashed horizontal line (indicated by the value 1) which represent, respectively the pooled posterior variance ($\hat{V}$) and average within-sample variance ($\hat{W}$). It can be seen that $\hat{R}$ is generally higher than one at the initial stage of the algorithms, but it tends to 1 eventually, and $\hat{V}$ and $\hat{W}$ tend to stabilize as the number of iterations increases, suggesting that the MCMC algorithm has approached convergence.

When these criteria indicate the algorithm convergence of chains, we propose that, after an initial number of 10,000 burn-in iterations of three chains of length 30,000, every 20th MCMC sample was retained from the next 10,000 for each chain. Thus, we obtain a total of 3,000 samples for targeted

posterior distributions of the unknown parameters for statistical inference. Even though this is a high-dimensional computation working load, the MCMC algorithm has no problem regarding the convergence of a solution for the inverse of matrices and parameter estimates in this application.

3.3. Data Analysis Results

Bayesian joint modeling approach based on MLME models is used to fit height and weight, as well as time-to-event data jointly. From the model fitting results, we have seen that, in general, the longitudinal sub-model provides a reasonably good fit to the observed data for most participants in the study; in particular, **Figure 3** shows the three randomly selected individual fitting curves of the height and weight trajectories estimated by the MVJ modeling approach based on Models N and SN. We can see that the estimated individual curves for Model SN, where the model error follows the SN distribution, fit the observed data more closely than those for Model N where the model error is normally distributed. The following findings are obtained from MVJ modeling.

To access the goodness-of-fit of the two models based on the MVJ modeling approach, the diagnostic plots of observed values vs. fitted values of height and weight based on Models N and SN are represented in **Figure 4**. It is shown from **Figure 4** that Model SN provides a much better fit to the observed values of height and weight, as compared to Model N.

Table 1 presents the population posterior mean (PM), the corresponding standard deviation (SD), and 95% equal-tail credible interval (CI) for the fixed-effects parameters and Cox proportional hazard model parameters based on Models SN and N. The following are our findings for the results of estimated parameters. First, in the multivariate longitudinal model (6), the results present that these estimates are different from zero since 95% of CIs do not contain zero. In particular, for the parameters that are of interest, the estimates of β_{11} and β_{12} for the growth rate of height and weight, respectively, for Model SN, are slightly smaller than their counterparts for Model N, while the baseline estimates of β_{01} and β_{02} from Model SN are slightly larger than those from Model N. Second, the estimates of the within-subject variances σ_{11}^2, σ_{22}^2, and covariance σ_{12}^2 for Model SN are smaller than their counterparts for Model N. This is expected because of high variability, heaviness of the tails, and skewness are pertinent to certain criteria. The estimates of the skewness parameters $\delta_1 = -4.68$ and $\delta_2 = -1.82$ are significantly negative for Model SN. The results provide evidence of obvious left-skewness exists in our data. Consequently, Model SN containing the skewness parameters is recommended. Third, there is an interesting finding in the Cox proportional hazard model (7) for the time-to-event process. Model SN indicates that the estimated two association parameters of the height [-0.27 with 95% CI ($-0.33 - 0.21$) and 0.95 with 95% CI (0.69, 1.22)], and one association parameter of weight [0.42 with 95% CI $0.42(0.25, 0.62)$] are significantly associated with the risk of T1D. In the comparison of Model SN and Model N, there is not too much difference in the parameter estimates of the Cox proportional hazard model. The estimated results also show that it is not directly associated with the covariates of
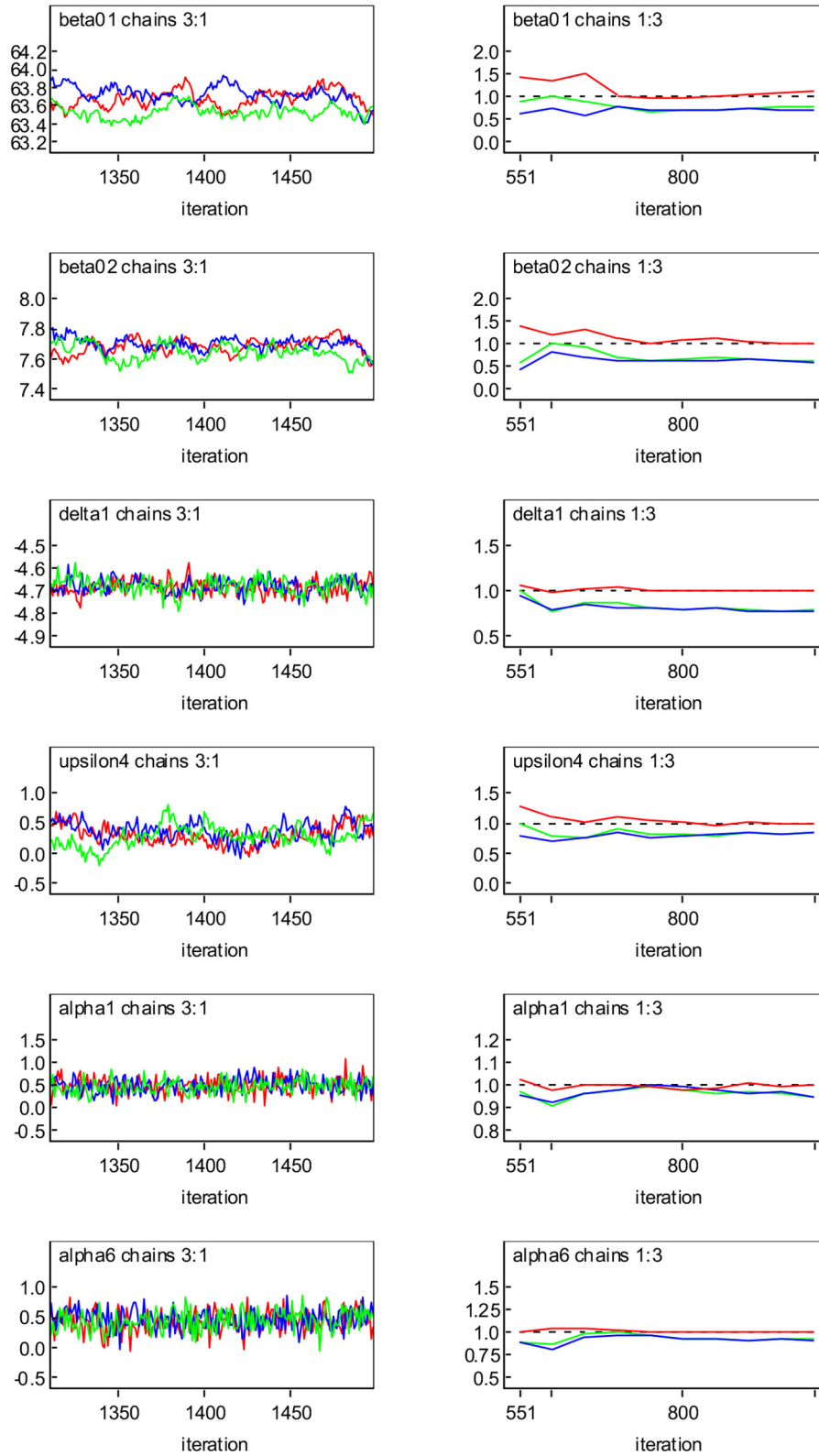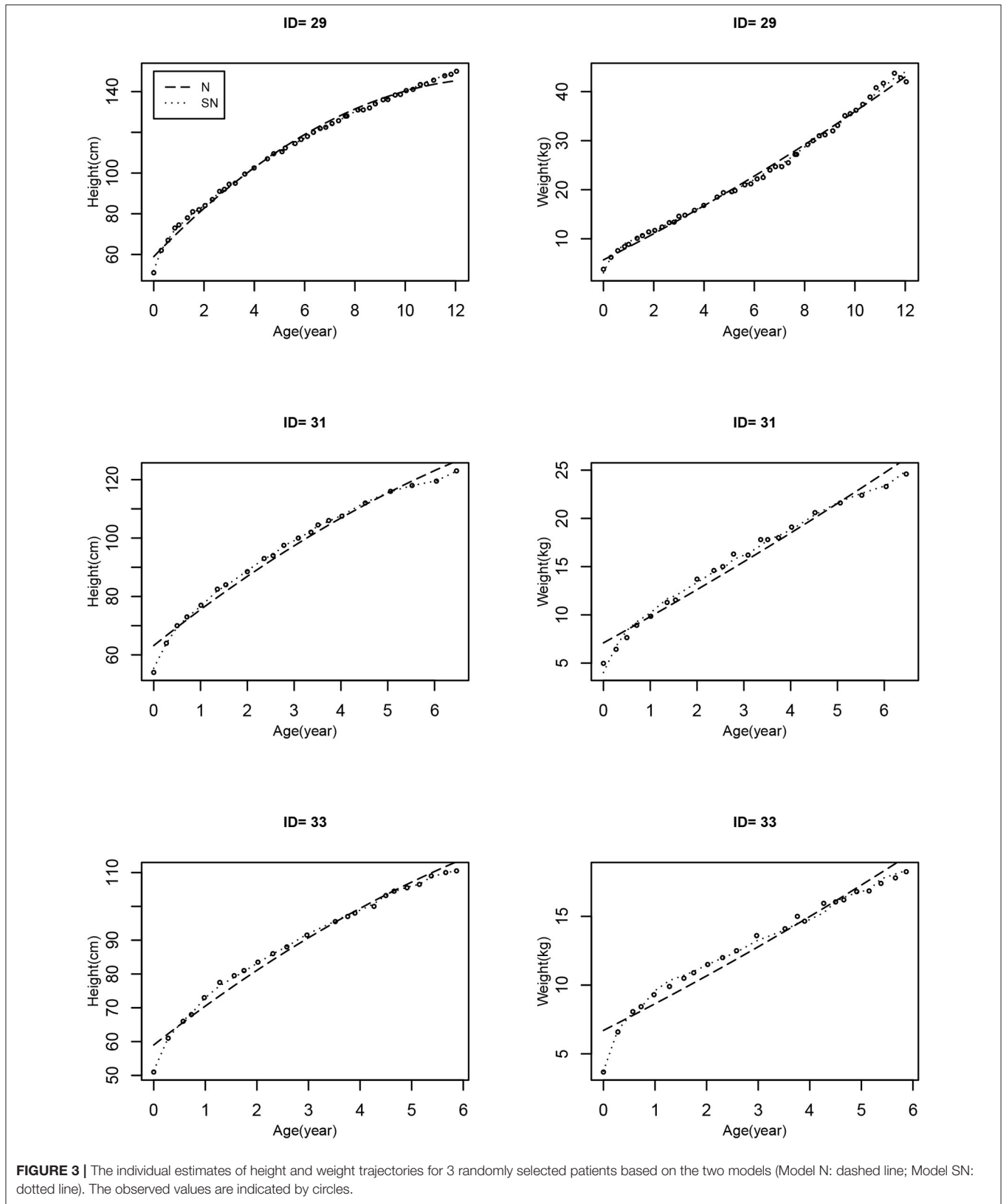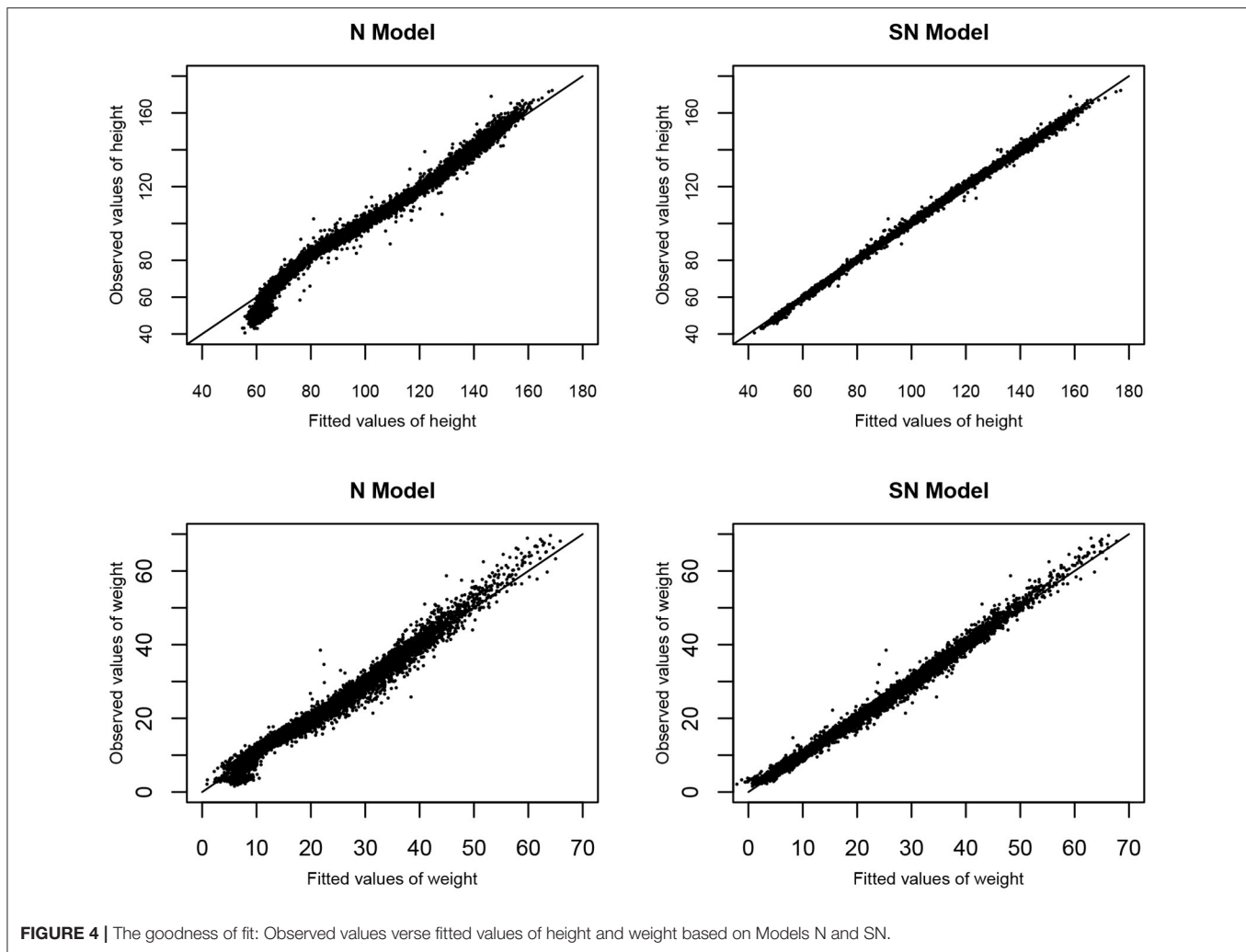
FIGURE 2 | Convergence diagnostics with three Markov chains for representative parameters based on Model SN: trace plots (left panel); Gelman-Rubin (GR) diagnostic plots (right panel), where the middle and bottom curves below the dashed horizontal line (indicated by the value one) represent the pooled posterior variance ($\hat{V}$) and average within-sample variance ($\hat{W}$), respectively, and the top curve above the dashed horizontal line represents their ratio ($\hat{R}$).

FIGURE 3 | The individual estimates of height and weight trajectories for 3 randomly selected patients based on the two models (Model N: dashed line; Model SN: dotted line). The observed values are indicated by circles.

FIGURE 4 | The goodness of fit: Observed values verse fitted values of height and weight based on Models N and SN.

Sweden and gender because of the insignificant estimates of α_3 and α_4.

In order to select the best model which fits the data adequately, we adopt a Bayesian selection criterion, known as the deviance information criterion (DIC) (Spiegelhalter et al., 2002). As we know, DIC is not intended to identify the "correct" model but is only used to find the one that fits the data best. the DIC values obtained are also summarized in **Table 1** in order to compare models under different settings. It is seen that the DIC value in Model SN is smaller than its counterpart in Models N, suggesting that Model SN produces a better fit than Model N in terms of DIC value. Thus, the further results based on Model SN are reported in detail below.

The estimated results based on Model SN in **Table 1** suggest the skewness parameters in height (-4.68) and weight (-1.82) are estimated to be significantly negative. This suggests the skewness with the heavy left tail of height and fairly left tail of weight. Thus, it may suggest that accounting for a multivariate linear mixed-effects joint modeling with the SN distribution

offers a better fit to the data which exhibit skewness and, in turn, provides more reliable parameter estimates. The estimated results of fixed-effects presented in **Table 1** based on Model SN indicate that the growth rate of height and weight with quadratic terms of age and gender may be approximated by $\hat{y}_1 = 63.63 + 11.69 \times \text{Age} - 0.37 \times \text{Age}^2 - 1.55 \times \text{Gender}$ and $\hat{y}_2 = 7.67 + 2.13 \times \text{Age} + 0.079 \times \text{Age}^2 - 0.58 \times \text{Gender}$, respectively. The quadratic of age and gender are all significant for the longitudinal sub-model of the MVJ modeling. Finally, based on the survival sub-model, the results show that the hazard ratio of the estimated association parameter υ_2 is $\exp(\upsilon_2) = 2.59$ with 95% CI being $(1.99, 3.39)$ which is statistically significant, indicating that a positive association between estimated change rate of height and risk of T1D diagnosis is found after the covariates of the country of residence, gender, HLA genotype, and FDR status are adjusted in the model. We also find that HR=$\exp(\alpha_5) = 1.40$ with 95% CI $(1.06, 1.86)$ for HLA genotype and HR=$\exp(\alpha_6) = 1.54$ with 95% CI $(1.08, 2.14)$ for FDR are significantly associated with a higher risk of T1D. However, gender is not significantly associated with the risk of T1D.

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

Parameter	Model SN			Model N		
	PM	SD	95% CI	PM	SD	95% CI
Parameter estimates in longitudinal sub-model						
β_{01}	63.63	0.13	(63.38, 63.88)	61.89	0.12	(61.66, 62.15)
β_{11}	11.69	0.05	(11.6, 11.8)	12.74	0.04	(12.66, 12.81)
β_{21}	−0.37	0.01	(−0.38, -0.36)	-0.467	0.01	(−0.473, −0.461)
β_{31}	−1.55	0.21	(−1.96, −1.15)	−1.52	0.17	(−1.87, −1.18)
β_{02}	7.67	0.07	(7.53, 7.79)	7.01	0.06	(6.89, 7.13)
β_{12}	2.13	0.03	(2.07, 2.19)	2.53	0.06	(2.47, 2.59)
β_{22}	0.079	0.01	(0.077, 0.082)	0.044	0.01	(0.041, 0.046)
β_{32}	−0.58	0.09	(−0.76, −0.40)	−0.57	0.08	(−0.73, -0.41)
δ_1	−4.68	0.03	(−4.75, −4.61)	–	–	–
δ_2	−1.82	0.03	(−1.86, −1.77)	–	–	–
σ_{11}^2	1.21	0.05	(1.12, 1.33)	9.99	0.11	(9.78, 10.21)
σ_{12}^2	−0.44	0.02	(−0.48, −0.39)	2.96	0.04	(2.87, 3.04)
σ_{22}^2	0.74	0.02	(0.74, 0.79)	2.06	0.02	(2.01, 2.10)
Parameter estimates in survival sub-model						
υ_1	−0.27	0.03	(−0.33, −0.21)	−0.35	0.04	(−0.43, −0.27)
υ_2	0.95	0.13	(0.69, 1.22)	1.26	0.18	(0.94, 1.67)
υ_3	0.42	0.09	(0.25, 0.62)	0.43	0.09	(0.26, 0.63)
υ_4	0.32	0.17	(−0.027, 0.65)	0.29	0.18	(−0.07, 0.65)
α_1	0.51	0.17	(0.18, 0.84)	0.41	0.18	(0.04, 0.41)
α_2	0.58	0.26	(0.037, 1.08)	0.51	0.18	(0.01, 0.52)
α_3	−0.15	0.18	(−0.51, 0.19)	−0.34	0.18	(-0.70, 0.03)
α_4	−0.054	0.14	(−0.33, 0.21)	−0.03	0.14	(-0.31, 0.25)
α_5	0.34	0.14	(0.061, 0.62)	0.38	0.14	(0.09, 0.66)
α_6	0.43	0.17	(0.076, 0.76)	0.45	0.18	(0.10, 0.79)
DIC	122,274			160,068		

4. SIMULATION STUDIES

In order to evaluate the performance of the introduced MVJ models and methods, the following limited simulation studies are conducted. The design of the simulated data mimics that of the diabetes data used in Section 3. Specifically, we choose the sample size $n = 500$ and assume that each individual has 32 scheduled longitudinal values. The time points of measurement are mimicked from the real application, and the true values of parameters are selected as follows. $\boldsymbol{\beta}_1^{\dagger T} = (\beta_{01}, \beta_{11}, \beta_{21}, \beta_{31})^T = (59, 11, -0.3, 0.1)^T$, $\boldsymbol{\beta}_2^{\dagger T} = (\beta_{02}, \beta_{12}, \beta_{22}, \beta_{32})^T = (6, 2, -0.5, -0.5)^T$, $\Upsilon = (\upsilon_1, \upsilon_2, \upsilon_3, \upsilon_4)^T = (-0.2, 1, 0.4, 0.3)^T$, and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6)^T = (0.7, 0.8, -0.3, -0.4, 0.3, 0.4)^T$. Longitudinal data are simulated based on Equation (6), with each model including individual random intercepts and random slopes with $\boldsymbol{b}_i = (b_{i01}, b_{i11}, b_{i02}, b_{i12})^T \sim N(\boldsymbol{0}, \text{diag}(1, 1, 1, 1))$, correlation is induced between the two longitudinal outcomes by generating the random intercepts and slopes for each outcome from the multivariate normal distribution; we simulate the model errors ϵ_{ijk} under weight sub-model with

$\Gamma(2, 0.8)$ distribution and height sub-model with $\Gamma(1, 0.5)$ which yield skewed distribution, respectively. To generate the event time data, the constant baseline hazard of 0.1 is set, and an exponential distribution with a mean of 0.1 is used to generate censoring time. The covariates in the survival model are simulated depending on variable types. For example, gender is simulated from a Bernoulli distribution with $p = 0.5$, etc. According to the settings described above, due to the heavy computational burden, we generate 50 data sets, which are fitted by Models N and SN. It is noted that the prior distributions are all close to non-informative similar to those in real data analysis. Thus, the results are expected to be somewhat robust with respect to prior distributions.

Table 2 summarizes the simulation results including the true parameter (TP) values, percent bias (defined by $100 \times \text{bias}_l/|\text{TP}_l|$) and percent mean-square-error (MSE) (defined by $100 \times \sqrt{MSE_l}/|\text{TP}_l|$) of fixed-effects $\boldsymbol{\beta}$, Υ and $\boldsymbol{\alpha}$. First, it is seen that the estimated parameters in height and weight for our two longitudinal outcomes, which means extending the univariate joint model to MVJ modeling allow us to

TABLE 2 | Summary of true parameter (TP) values, estimated parameters, Bias, and MSE for Models N and SN based on 50 simulated data sets.

Model		Model SN			Model N		
Parameter	TP	EST	Bias	MSE	EST	Bias	MSE
Parameter estimates in longitudinal sub-model							
β_{01}	59	59.61	1.04	0.64	59.61	1.04	0.64
β_{11}	11	11.81	7.37	6.21	11.85	7.74	0.60
β_{21}	−0.3	−0.30	−1.56	−0.01	−0.30	−1.60	0.01
β_{31}	0.1	−0.10	−1.32	0.01	0.09	-1.16	0.01
β_{02}	6	6.81	13.52	11.21	6.80	13.41	10.84
β_{12}	2	2.66	33.08	32.11	2.80	39.96	32.24
β_{22}	−0.5	−0.46	8.42	−1.02	−0.45	9.99	1.13
β_{32}	−0.5	−0.54	−9.60	0.73	−0.55	−10.67	0.72
δ_1		1.44	–	–	–	–	–
δ_2		1.57	–	–	–	–	–
Parameter estimates in survival sub-model							
υ_1	−0.2	−0.19	2.80	0.10	−0.20	0.11	0.08
υ_2	1	0.66	−33.60	11.54	0.67	−33.22	11.31
υ_3	0.4	0.29	−27.42	4.50	0.31	−23.54	3.52
υ_4	0.3	0.24	−19.74	1.63	0.24	−19.80	1.64
α_1	0.7	0.74	6.04	0.39	0.75	6.55	0.35
α_2	0.8	0.90	12.36	1.70	0.91	14.10	1.66
α_3	−0.3	−0.15	49.99	10.08	−0.14	52.77	8.44
α_4	−0.4	−0.20	49.99	10.08	-0.19	50.63	10.27
α_5	0.3	0.30	−0.88	0.37	0.24	−1.94	0.06
α_6	0.4	0.32	−19.28	1.54	0.32	−19.78	1.58

EST is the average of estimates, Bias and MSE are quantified by the percent bias $= 100 \times bias_l/|TP_l|$ and percent $\sqrt{MSE} = 100 \times \sqrt{MSE_l}/|TP_l|$, respectively.

incorporate more information and improves the efficiency in estimation. Second, for the multivariate linear mixed-effects sub-model, we find that, in comparison to Models SN and N, Model SN generally outperforms Model N in terms of smaller bias and MSE. For all the scenarios considered in this simulation study, it is seen that all estimated biases for β_{21}, β_{31}, and β_{22} are negative, suggesting that these parameters are underestimated, while estimated biases for β_{01}, β_{11}, β_{12}, and β_{02} are positive, indicating that these parameters are overestimated. We note that the larger bias of the growth rate of height and weight is reasonable, and this is consistent with the results from the real data analysis. The average estimates of skewness parameters $\delta_1 = 1.44$ for height and $\delta_2 = 1.57$ for weight in Model SN indicate a departure from (symmetric) normal distribution. Finally, for the estimated parameters in the survival sub-model, Model SN obviously outperforms Model N for the association parameters and baseline covariates except υ_1, υ_2, and υ_3. Some parameters in the survival sub-model are slightly overestimated and some parameters are slightly underestimated in both models. In summary, the simulation results confirm the importance of accounting for the non-normality of the data. This suggests that adopting the assumption of normal distribution may lead to inaccurate inference on fixed-effects of interest, in particular, when longitudinal data exhibit non-normal features.

5. CONCLUSION

In this study, we propose a Bayesian MVJ model with multiple longitudinal responses and survival processes. The MLME sub-model and the survival sub-model are linked through the random-effects which are served to characterize the underlying subject-specific longitudinal process (Xu, 2021). We also consider some important data features such as non-normality which may impact the discovery of the true disease diagnosis progression. Comparing with the classic frequentist's methods, a Bayesian approach is powerful when the dimension of parameters is high in the complicated MVJ modeling. Although the joint modeling for longitudinal and time-to-event data has been an active area of statistical methodological study (Huang et al., 2011; Chen et al., 2014; Huang and Chen, 2016; Zhang and Huang, 2020), this paper extends to investigate joint models for survival and multivariate longitudinal data with SN distribution, accounting for multiple data features, simultaneously. Although this article is motivated by a diabetes study, the innovations of the proposed multivariate joint models and methods may help applied researchers analyze complicated longitudinal and survival data under a wide range of applications.

The proposed MVJ model offers some advantages in comparison with traditional joint models. First, the majority of joint models focus mainly on a single longitudinal outcome

associated with the survival endpoint. However, in many clinical and observational studies, multiple longitude data are collected together and they may be significantly correlated. The MVJ model proposed here is able to reduce the bias and can increase the efficiency in parameter estimation. These interesting findings have important clinical indications. Our results suggest that there is a positive association between rates of growth and risk of T1D (i.e., there is an increased risk of T1D for larger height and weight at baseline). Second, due to its importance to measure height and weight appropriately when they show non-normality with heavy tails, this article considers two statistical models (Models N and SN) with different scenarios. We find that Model SN is favorable to model N. In model SN, the estimates of skewness parameters δ_1 and δ_2 are statistically significant for height and weight, indicating that the skewness with a heavy tail exists in height and weight measurements. Therefore, the MVJ model with the SN distribution provides more efficient and accurate estimates of parameters, and, thus, serves as a better alternative to the normal distribution-based model which is widely adopted.

We apply the Bayesian MVJ modeling approach to analyze the diabetes data set in this article. The results demonstrate the use of the MVJ model to investigate how the patterns of longitudinal height and weight trajectories are associated with the risk of T1D. Furthermore, our results indicate that it is of importance to consider the MVJ model with the skewed distribution in order to obtain more accurate and less biased parameter estimates in the presence of non-normal features in longitudinal height and weight data. Although this article is motivated by a diabetes study, the fundamental concepts of the proposed Bayesian modeling method should have generally broader applications in practice whenever the two different sources of dependence among longitudinal data over time and between longitudinal and survival variables are presented, and the relevant technical specifications are met (Chen et al., 2021). Our models may have the potential to extend to more complicated models. For example, (i) the missing data mechanism may be considered by introducing a non-ignorable missing data model for longitudinal measurements

(Huang, 2016; Huang and Chen, 2016). (ii) although a single-type event time is investigated only in this article, the developed MVJ model may be extended to accommodate competing risks survival data in the presence of multiple "failure" types of events (Elashoff et al., 2008; Hu et al., 2009). Although these interesting topics are beyond the focus of this paper, they are warranted in our research pipeline under investigation.

DATA AVAILABILITY STATEMENT

The data are available upon request to the authors.

AUTHOR CONTRIBUTIONS

YH: research concept and design. YH, JC, and LX: collection and/or assembly of data and statistical analysis. YH, JC, LX, and N-ST: data analysis and interpretation, writing the article, critical revision of the article, and final approval of article. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdata.2022.812725/full#supplementary-material

REFERENCES

Albert, P. S., and Shih, J. H. (2010). An approach for jointly modeling multivariate longitudinal measurements and discrete time-to-event data. *Ann. Appl. Stat.* 4, 1517–1532. doi: 10.1214/10-AOAS339

Arellano-Valle, R. B., and Genton, M. G. (2005). On fundamental skew distributions. *J. Multivariate Anal.* 96, 93–116. doi: 10.1016/j.jmva.2004.10.002

Azzalini, A., and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *J. R. Stat. Soc. B* 65, 367–389. doi: 10.1111/1467-9868.00391

Baxter, J., Vehik, K., Johnson, S., Lernmark, B., Roth, R., Simell, T., et al. (2012). Differences in recruitment and early retention among ethnic minority participants in a large pediatric cohort: the teddy study. *Contemp Clin. Trials* 33, 633–640. doi: 10.1016/j.cct.2012.03.009

Brown, E. R., and Ibrahim, J. G. (2003). A Bayesian semiparametric joint hierarchical model for longitudinal and survival data. *Biometrics* 59, 221–228. doi: 10.1111/1541-0420.00028

Brown, E. R., Ibrahim, J. G., and DeGruttola, V. (2005). A flexible b-spline model for multiple longitudinal biomarkers and survival. *Biometrics* 61, 64–73. doi: 10.1111/j.0006-341X.2005.030929.x

Chen, J., Huang, Y., and Tang, N.-S. (2021). Bayesian change-point joint models for multivariate longitudinal and time-to-event data. *Stat. Biopharmaceut. Res.* 20, 94. doi: 10.1080/19466315.2020.1837234

Chen, Q., May, R. C., Ibrahim, J. G., Chu, H., and Cole, S. R. (2014). Joint modeling of longitudinal and survival data with missing and left-censored time-varying covariates. *Stat. Med.* 33, 4560–4576. doi: 10.1002/sim.6242

Chen, Y., and Wang, Y. (2017). Variable selection for joint models of multivariate longitudinal measurements and event time data. *Stat. Med.* 36, 3820–3829. doi: 10.1002/sim.7391

Chi, Y.-Y., and Ibrahim, J. G. (2006). Joint models for multivariate longitudinal and multivariate survival data. *Biometrics.* 62, 432–445. doi: 10.1111/j.1541-0420.2005.00448.x

Clayton, D. G. (1991). A monte carlo method for bayesian inference in frailty models. *Biometrics* 46, 467–485. doi: 10.2307/2532139

Cox, D. R. (1972). Regression models and life-tables. *J. R. Stat. Soc. B* 34, 187–202. doi: 10.1111/j.2517-6161.1972.tb00899.x

Elashoff, R. M., Li, G., and Li, N. (2008). A joint model for longitudinal measurements and survival data in the presence of multiple failure types. *Biometrics* 64, 762–771. doi: 10.1111/j.1541-0420.2007.00952.x

Elding Larsson, H., Vehik, K., Gesualdo, P., Akolkar, B., Hagopian, W., Krischer, J., et al. (2014). Children followed in the teddy study are diagnosed with type 1 diabetes at an early stage of disease. *Pediatr. Diabetes* 15, 118–126. doi: 10.1111/pedi.12066

Farcomeni, A., and Viviani, S. (2015). Longitudinal quantile regression in the presence of informative dropout through longitudinal-survival joint modeling. *Stat. Med.* 34, 1199–1213. doi: 10.1002/sim.6393

Fieuws, S., and Verbeke, G. (2006). Pairwise fitting of mixed models for the joint modeling of multivariate longitudinal profiles. *Biometrics* 62, 424–431. doi: 10.1111/j.1541-0420.2006.00507.x

Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7, 457–472. doi: 10.1214/ss/1177011136

Hagopian, W. A., Erlich, H., Lernmark, Å., Rewers, M., Ziegler, A. G., Simell, O., et al. (2011). The environmental determinants of diabetes in the young: genetic criteria and international diabetes risk screening of 421,000 infants. *Pediatr. Diabetes* 12, 733–743. doi: 10.1111/j.1399-5448.2011.00774.x

Henderson, R., Diggle, P., and Dobson, A. (2000). Joint modelling of longitudinal measurements and event time data. *Biostatistics* 1, 465–480. doi: 10.1093/biostatistics/1.4.465

Hu, W., Li, G., and Li, N. (2009). A Bayesian approach to joint analysis of longitudinal measurements and competing risks failure time data. *Stat. Med.* 28, 1601–1619. doi: 10.1002/sim.3562

Huang, Y. (2016). Quantile regression-based bayesian semiparametric mixed-effects models for longitudinal data with non-normal, missing and mismeasured covariate. *J. Stat. Comput. Simul.* 86, 1183–1202. doi: 10.1080/00949655.2015.1057732

Huang, Y., and Chen, J. (2016). Bayesian quantile regression-based nonlinear mixed-effects joint models for time-to-event and longitudinal data with multiple features. *Stat. Med.* 35, 5666–5685. doi: 10.1002/sim.7092

Huang, Y., and Dagne, G. (2011). A Bayesian approach to joint mixed-effects models with a skew-normal distribution and measurement errors in covariates. *Biometrics* 67, 260–269. doi: 10.1111/j.1541-0420.2010.01425.x

Huang, Y., Dagne, G., and Wu, L. (2011). Bayesian inference on joint models of HIV dynamics for time-to-event and longitudinal data with skewness and covariate measurement errors. *Stat. Med.* 30, 2930–2946. doi: 10.1002/sim.4321

Kim, S., and Albert, P. S. (2016). A class of joint models for multivariate longitudinal measurements and a binary event. *Biometrics* 72, 917–925. doi: 10.1111/biom.12463

Kiviniemi, M., Hermann, R., Nurmi, J., Ziegler, A. G., Knip, M., Simell, O., et al. (2007). A high-throughput population screening system for the estimation of genetic risk for type 1 diabetes: an application for the teddy (the environmental determinants of diabetes in the young) study. *Diabetes Technol. Therapeut.* 9, 460–472. doi: 10.1089/dia.2007.0229

Larsson, H. E., Hansson, G., Carlsson, A., Cederwall, E., Jonsson, B., Jönsson, B., et al. (2008). Children developing type 1 diabetes before 6 years of age have increased linear growth independent of HLA genotypes. *Diabetologia* 51, 1623. doi: 10.1007/s00125-008-1074-0

Lernmark, B., Johnson, S., Vehik, K., Smith, L., Ballard, L., et al. (2011). Enrollment experiences in a pediatric longitudinal observational study: the environmental determinants of diabetes in the young (teddy) study. *Contemp Clin. Trials* 32, 517–523. doi: 10.1016/j.cct.2011.03.009

Lin, H., McCulloch, C. E., and Mayne, S. T. (2002). Maximum likelihood estimation in the joint analysis of time-to-event and multiple longitudinal variables. *Stat. Med.* 21, 2369–2382. doi: 10.1002/sim.1179

Lunn, D. J., Thomas, A., Best, N., and Spiegelhalter, D. (2000). Winbugs-a bayesian modelling framework: concepts, structure, and extensibility. *Stat. Comput.* 10, 325–337. doi: 10.1023/A:1008929526011

Proudfoot, J., Faig, W., Natarajan, L., and Xu, R. (2018). A joint marginal-conditional model for multivariate longitudinal data. *Stat. Med.* 37, 813–828. doi: 10.1002/sim.7552

Rizopoulos, D., et al. (2010). JM: An R package for the joint modelling of longitudinal and time-to-event data. *J. Stat. Softw.* 35, 1–33. doi: 10.18637/jss.v035.i09

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

Rizopoulos, D. (2012). *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R*. Boca Raton, FL: CRC Press.

Rizopoulos, D., and Ghosh, P. (2011). A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. *Stat. Med.* 30, 1366–1380. doi: 10.1002/sim.4205

Sahu, S. K., Dey, D. K., and Branco, M. D. (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *Can. J. Stat.* 31, 129–150. doi: 10.2307/3316064

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B* 64, 583–639. doi: 10.1111/1467-9868.00353

Tang, A.-M., Tang, N.-S., and Zhu, H. (2017a). Influence analysis for skew-normal semiparametric joint models of multivariate longitudinal and multivariate survival data. *Stat. Med.* 36, 1476–1490. doi: 10.1002/sim.7211

Tang, A.-M., Zhao, X., and Tang, N.-S. (2017b). Bayesian variable selection and estimation in semiparametric joint models of multivariate longitudinal and survival data. *Biom. J.* 59, 57–78. doi: 10.1002/bimj.201500070

TEDDY Study Group. (2007). The environmental determinants of diabetes in the young (teddy) study: study design. *Pediatr. Diabetes* 8, 286–298. doi: 10.1111/j.1399-5448.2007.00269.x

Tsiatis, A. A., and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: an overview. *Stat. Sin.* 14:809–834.

Wu, L., Liu, W., and Hu, X. (2010). Joint inference on hiv viral dynamics and immune suppression in presence of measurement errors. *Biometrics* 66, 327–335. doi: 10.1111/j.1541-0420.2009.01308.x

Xu, L. (2021). *Bayesian Multivariate Joint Modeling for Skewed-longitudinal and Time-to-eveant Data* (Ph.D. Dissertation).

Zhang, H., and Huang, Y. (2020). Quantile regression-based Bayesian joint modeling analysis of longitudinal-survival data, with application to an AIDS cohort study. *Lifetime Data Anal.* 26, 339–368. doi: 10.1007/s10985-019-09478-w

Zhang, H., and Huang, Y. (2021). 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, 295–316. doi: 10.1080/10543406.2020.1852248