Comput Stat (2017) 32:179–196
DOI 10.1007/s00180-016-0671-5

ORIGINAL PAPER

Bayesian inference on partially linear mixed-effects joint models for longitudinal data with multiple features

Yangxin Huang[1] · **Tao Lu**[2]

Received: 12 August 2014 / Accepted: 2 July 2016 / Published online: 8 July 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract The relationship between viral load and CD4 cell count is one of the interesting questions in AIDS research. Statistical models are powerful tools for clarifying this important problem. Partially linear mixed-effects (PLME) model which accounts for the unknown function of time effect is one of the important models for this purpose. Meanwhile, the mixed-effects modeling approach is suitable for the longitudinal data analysis. However, the complex process of data collection in clinical trials has made it impossible to rely on one particular model to address the issues. Asymmetric distribution, measurement error and left censoring are features commonly arisen in longitudinal studies. It is crucial to take into account these features in the modeling process to achieve reliable estimation and valid conclusion. In this article, we establish a joint model that accounts for all these features in the framework of PLME models. A Bayesian inferential procedure is proposed to estimate parameters in the joint model. A real data example is analyzed to demonstrate the proposed modeling approach for inference and the results are reported by comparing various scenarios-based models.

Keywords AIDS clinical trial · Bayesian analysis · PLME models · Skew-*t* distribution · Viral load and CD4 cell count

✉ Yangxin Huang
yhuang@health.usf.edu

[1] Department of Epidemiology and Biostatistics, College of Public Health, University of South Florida, Tampa, FL 33612, USA

[2] Department of Epidemiology and Biostatistics, State University of New York, Albany, NY 12144, USA

1 Introduction

In AIDS clinical trial, it is interesting to study the relationship between viral load (plasma HIV RNA copies per milliliter) and CD4 cell count. The goal of the highly active antiretroviral therapy (HAART) is to suppress the viral load at very low level and to improve the patient's immune system (quantified by CD4 cell counts/mm^3). It is well known that the CD4 cells play an important role in eradicating and/or maintaining the suppressed viral load. The better understanding of the interaction between viral load and CD4 cells may help clinicians design treatment regimes in a more efficient way.

Partially linear model (PLM), a special case of semiparametric model, has received much attention in the statistical community (Speckman 1988; Härdle et al. 2000; Liang and Ren 2005). PLM, as a combination of both parametric and nonparametric components, enjoys advantages from both sides. The interpretation of the linear component is made easier by the parametric formation while the nonparametric component allows flexibility of modeling complex nonlinear relationships between the response variable and independent variables. Liang et al. (2004) firstly applied the PLM to model the relationship between viral load and CD4 cell counts in presence of a nonlinear relationship between viral load and treatment time. To account for the longitudinal nature of clinical data, the PLM has been extended to the generalized linear mixed-effects model framework (Liang and Ren 2005). In these models, the nonparametric component was usually approximated by the local linear smoother, such as local polynomial kernel, before the optimization method was applied to obtain parameter estimates. Nevertheless, one major drawback, as many other existing statistical models share, is that the random errors are assumed to be normally and symmetrically distributed, which lacks robustness against departure from normality and outliers. The statistical inference based on normal assumption may lead to misleading results (Ho and Lin 2010; Huang and Dagne 2011; Huang et al. 2012; Sahu et al. 2003; Lachos et al. 2011).

Some data features are worthy of attention when one models the relationship between viral load and CD4 cell counts. In practice, when the quantity of viral load falls below a certain threshold, referred to as limit of detection (LOD), it is not detectable by the device because of the low sensitivity of the current standard assay. Thus, for computational convenience, data below LOD are usually imputed by the LOD or half of the LOD. However, it is to be suspected that such manipulations may impact the statistical inference eventually. Further, CD4 cell counts are generally measured with much noise. It has been noted that ignoring measurement error in CD4 cell counts incurs significant bias on parameter estimation (Liang and Ren 2005).

The simultaneous inference for longitudinal data with features of asymmetry, LOD and covariate measurement errors is of interest recently (Huang and Dagne 2011; Huang et al. 2012). Most of the work were focused on nonlinear mixed-effects models in the presence of multiple data features (Huang and Dagne 2011; Huang et al. 2011). However, it is not clear how asymmetry, LOD and measurement error in covariates may interact and simultaneously impact the statistical procedure under the framework of partially linear mixed-effects (PLME) joint model. Statistical inferences complicate dramatically when all of these issues are present. In this article, we develop a PLME

joint model with skew distributions to investigate the effects on inference when all of these data features are present.

We adopt a skew-t (ST) distribution (Azzalini and Genton 2008; Azzalini and Capitanio 2003; Ho and Lin 2010; Huang and Dagne 2011, 2012; Jara et al. 2008; Sahu et al. 2003) to develop joint models for longitudinal data with features of asymmetry, LOD and measurement error in covariates. The performance based on an ST distribution is compared with that based on a normal distribution. We employ a fully Bayesian approach to investigate the ST-PLME joint models. It is noted that the ST distribution reduces to the standard normal distribution when its degrees of freedom approach infinity and its skewness parameter is zero. Therefore, the joint model built on ST distributions can be easily reverted to other distributions such as normal distributions. In this article, to account for many data features in longitudinal studies, we developed a joint model with following components: (1) PLME model with ST distribution for response process with considering left censoring due to LOD; (2) linear mixed-effects (LME) model with ST distribution for covariate process. It is noted that the LOD data for viral load is in the similar sense to the left-censored data in survival analysis. Toward this end, to account for the LOD in viral load, the joint likelihood that takes into consideration of the censoring mechanism is adopted for Bayesian analysis. We then present a flexible and robust approach to model the complicated process under the Bayesian framework. The Bayesian approach avoids the high dimensional integration as well as complicated approximations usually adopted by a frequentist approach, thus it is more convenient for drawing statistical inferences.

The rest of the article is organized as follows. In Sect. 2, the clinical data that motivate this research are briefly described and the joint statistical models that account for asymmetric distribution and LOD in response as well as measurement error in covariate are introduced. The Bayesian inferential approach that estimates the parameters in the joint model is presented in Sect. 3. In Sect. 4, the proposed models and inferential method are applied to an AIDS clinical data and analysis results are presented. The article is concluded with a discussion in Sect. 5.

2 Motivating data and joint model setup

2.1 Motivating data

The data that motivate this study are from an HIV/AIDS clinical study (ACTG 398) (Hammer et al. 2002). This study is a randomized, double-blind, placebo-controlled, with an extension to more than 48 week study comparing four-drug class regimens for 481 patients with virologic failure. The plasma HIV-1 RNA (viral load) was designed to be repeatedly measured in copies per milliliter at weeks 0, 2, 4, 8, 16 and every 8 weeks until the last patient completed 48 weeks on study. The number of viral load measurements for each individual varies from 2 to 13. Out of total 481 patients, the 379 patients who had more than two observations were included in the data analysis. CD4 cell counts were also measured throughout the study on a similar scheme. A $\log_{10}$ transformation of viral load and standardized CD4 cell counts were used in the analysis in order to reduce the variation of the measurements and, in turn, to stabilize

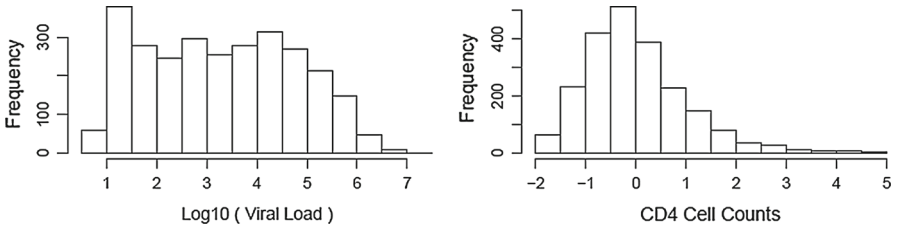

Fig. 1 Histogram of $\log_{10}$ viral load (*left*) and standardized CD4 cell counts (*right*) of 379 patients from ACTG 398 clinical trial

the estimation algorithm and to speed up convergence of the algorithm. In addition, to avoid too small or too large estimates which may be unstable, we rescaled the original time (in days) so that the time scale is between 0 and 1.

Figure 1 shows the distribution of repeated viral load ($\log_{10}$ scale) and standardized CD4 cell count measurements for 379 patients in an ACTG 398 clinical trial study. It is obvious that, both viral load (response) and CD4 cell counts (covariate) are highly skewed. Thus, the normality assumption is not quite realistic for this dataset. As an alternative, an asymmetric distribution such as skew-t (ST) distribution (Sahu et al. 2003; Arellano-Valle and Genton 2005; Azzalini and Capitanio 2003; Azzalini and Genton 2008; Ho and Lin 2010; Jara et al. 2008) is more suitable than a symmetric normal distribution for modeling skewed-longitudinal data. Another challenge for analyzing longitudinal data in AIDS clinical studies is that CD4 cell counts are often measured with substantial errors. Additionally, due to technological constraints, the viral load are usually not measurable when the quantity falls below LOD. In ACTG 398 study, 26 % of viral load measurements were below LOD which is $\log_{10}(50)$. Therefore, it is critical to take into account all these data features when the relationship between viral load and CD4 cell counts is modeled in practice.

2.2 Covariate measurement error model with ST distribution

In AIDS clinical studies, the important biomarker, CD4 cell count, is usually measured with substantial errors and distribution of this variable is highly skewed. To relax the normality assumption for the measurement errors in CD4 cell counts, we extend a covariate (CD4) mixed-effects model by assuming an ST distribution for the measurement error. Denote the number of subjects by n and the number of measurements on the ith subject by n_i; the observed CD4 cell counts for individual i at time t_{ij} ($i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, n_i$) as z_{ij}; the vector $\boldsymbol{z}_i = (z_{i1}, \ldots, z_{in_i})^T$. $\boldsymbol{U}_i$ and $\boldsymbol{V}_i$ are $n_i \times l$ design vectors, where l is the number of factors (see model (9) for a special case), $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_l)^T$ and $\boldsymbol{a}_i = (a_{1i}, \ldots, a_{li})^T$ are unknown population (fixed-effects) and individual-specific (random-effects) parameter vectors, respectively. The mixed-effects covariate model that assumes an ST distribution for the measurement error is as follows:

$$\boldsymbol{z}_i = \boldsymbol{U}_i\boldsymbol{\alpha} + \boldsymbol{V}_i\boldsymbol{a}_i + \boldsymbol{\epsilon}_i \ \left(\equiv \boldsymbol{z}_i^* + \boldsymbol{\epsilon}_i\right),$$

$$\boldsymbol{\epsilon}_i \overset{\text{iid}}{\sim} ST_{n_i,\nu_1}\left(-J(\nu_1)\delta_\epsilon \mathbf{1}_{n_i},\, \sigma_1^2 \boldsymbol{I}_{n_i},\, \delta_\epsilon \boldsymbol{I}_{n_i}\right), \tag{1}$$

where $\boldsymbol{z}_i^* = (z_{i1}^*, \ldots, z_{in_i}^*)^T$ and $\boldsymbol{z}_i^* = \boldsymbol{U}_i\boldsymbol{\alpha} + \boldsymbol{V}_i\boldsymbol{a}_i$ may be viewed as the true (but unobservable) covariate values. The measurement error $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{in_i})^T$ follows a multivariate ST distribution with degrees of freedom ν_1, unknown scale parameter σ_1^2 and skewness parameter δ_ϵ, $J(\nu_1) = (\nu_1/\pi)^{1/2}[\Gamma((\nu_1 - 1)/2)/\Gamma(\nu_1/2)]$, and $\mathbf{1}_{n_i} = (1, \ldots, 1)^T$. We assume that $\boldsymbol{a}_i \overset{\text{iid}}{\sim} N_l(0, \boldsymbol{\Sigma}_a)$, where $\boldsymbol{\Sigma}_a$ is an unrestricted covariance matrix for random-effects. Model (1) may be interpreted as an ST covariate measurement error model that accounts for data skewness and incorporates the correlation of the repeated measurements on each subject.

2.3 Partially linear mixed-effects model with ST distribution

Denote y_{ij} as the viral load for the ith subject at time t_{ij} ($i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n_i$). As mentioned previously, when the viral load in patients is below LOD, the device is unable to detect it. This type of data is typically left censored. Denote the observed value of y_{ij} as (q_{ij}, c_{ij}) where c_{ij} is the censoring indicator so that y_{ij} is observed, i.e. $y_{ij} = q_{ij}$ when $c_{ij} = 0$. When $c_{ij} = 1$, y_{ij} is unobserved, i.e. $y_{ij} \leq d$ where d is the LOD. Denote the observed data from an AIDS clinical study as $\mathfrak{R} = \{(\boldsymbol{z}_i, \boldsymbol{q}_i, \boldsymbol{c}_i), i = 1, \ldots, n\}$ where $\boldsymbol{q}_i = (q_{i1}, \ldots, q_{in_i})^T$ and $\boldsymbol{c}_i = (c_{i1}, \ldots, c_{in_i})^T$. $\boldsymbol{z}_i$ is defined accordingly in Sect. 2.2. Further, let $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{in_i})^T$, $\boldsymbol{t}_i = (t_{i1}, \ldots, t_{in_i})^T$, $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \ldots, \varepsilon_{in_i})^T$. To study the relationship between the viral load response and CD4 cell counts during the course of clinical trial, we consider the following PLME model:

$$\begin{aligned} \boldsymbol{y}_i &= \beta_i \boldsymbol{z}_i^* + \boldsymbol{g}_i(t_i) + \boldsymbol{\varepsilon}_i, \\ \beta_i &= \beta + b_i,\ \boldsymbol{g}_i(t_i) = \boldsymbol{g}(t_i) + \boldsymbol{h}_i(t_i), \end{aligned} \tag{2}$$

where β_i is the individual coefficient that quantifies the relationship between viral load and actual CD4 cell counts for individual i; β is the population coefficient (fixed-effects) and b_i is the random-effects variable that quantifies the departure from the population for individual i. We assume b_i is normally distributed with mean 0 and variance σ_b^2. Both $\boldsymbol{g}(\cdot)$ and $\boldsymbol{h}_i(\cdot)$ are unknown smoothing functions. $\boldsymbol{g}(\cdot)$ stands for the population smoothing curve while $\boldsymbol{h}_i(\cdot)$ represents the random-effects. Altogether, $\boldsymbol{g}_i(\cdot)$ is the smoothing curve for individual i. The error term $\boldsymbol{\varepsilon}_i(\cdot)$ and random smoothing function $\boldsymbol{g}_i(\cdot)$ are zero mean stochastic processes and are independent from each other. b_i is independent of both $\boldsymbol{\varepsilon}_i(\cdot)$ and $\boldsymbol{g}_i(\cdot)$. Note that in model (2), we replaced the observed CD4 cell counts ($\boldsymbol{z}_i$) by the unobserved but actual values ($\boldsymbol{z}_i^*$) due to the measurement error in CD4 cell counts. In this way, the covariate measurement error model (1) is incorporated into the response model (2) when the parameter estimation is carried out.

To fit model (2), we adopt a regression spline method for $\boldsymbol{g}(\cdot)$ and $\boldsymbol{h}_i(\cdot)$. The main idea of regression splines is to approximate $\boldsymbol{g}(\cdot)$ and $\boldsymbol{h}_i(\cdot)$ with a linear combination of spline basis functions $\boldsymbol{\Theta}_p(t) = (\theta_0(t), \ldots, \theta_p(t))^T$ and $\boldsymbol{\Phi}_r(t) = (\phi_0(t), \ldots, \phi_r(t))^T$,

respectively. Therefore, we have

$$\begin{aligned} \boldsymbol{g}_p(t) &\approx \sum_{k=0}^{p} \xi_k \theta_k(t) = \boldsymbol{\Theta}_p(t)^T \boldsymbol{\xi}_p \\ \boldsymbol{h}_{i,r}(t) &\approx \sum_{k=0}^{r} \chi_{ki} \phi_k(t) = \boldsymbol{\Phi}_r(t)^T \boldsymbol{\chi}_{ri} \end{aligned} \tag{3}$$

where $\boldsymbol{\xi}_p = (\xi_0, \ldots, \xi_p)^T$ is a (p + 1)-dimensional vector of fixed-effects, $\boldsymbol{\chi}_{ri} = (\chi_{0i}, \ldots, \chi_{ri})^T$ is a (r + 1)-dimensional vector of random-effects; the dimensional numbers p and r can be determined by AIC or BIC. Based on the assumption of $\boldsymbol{h}_i(\cdot)$, we can regard $\boldsymbol{\chi}_{ri}$ as i.i.d. realizations of a zero-mean random vector. Denote $\boldsymbol{\Theta}_{pi} = (\boldsymbol{\Theta}_p(t_{i1}), \ldots, \boldsymbol{\Theta}_p(t_{in_i}))^T$ and $\boldsymbol{\Phi}_{ri} = (\boldsymbol{\Phi}_r(t_{i1}), \ldots, \boldsymbol{\Phi}_r(t_{in_i}))^T$. Plugging (3) into (2), we have

$$\boldsymbol{y}_i = \beta z_i^* + b_i z_i^* + \boldsymbol{\Theta}_{pi} \boldsymbol{\xi}_p + \boldsymbol{\Phi}_{ri} \boldsymbol{\chi}_{ri} + \boldsymbol{\varepsilon}_i \tag{4}$$

Let $\boldsymbol{X}_i = (z_i^*, \boldsymbol{\Theta}_{pi})$, $\boldsymbol{Z}_i = (z_i^*, \boldsymbol{\Phi}_{ri})$, $\boldsymbol{\zeta} = (\beta, \boldsymbol{\xi}_p^T)^T$ and $\boldsymbol{\varpi}_i = (b_i, \boldsymbol{\chi}_{ri}^T)^T$. We can write (4) as

$$\boldsymbol{y}_i = \boldsymbol{X}_i \boldsymbol{\zeta} + \boldsymbol{Z}_i \boldsymbol{\varpi}_i + \boldsymbol{\varepsilon}_i \tag{5}$$

which is a standard LME model if we assume $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$ are the fixed-effects and random-effects design matrices, respectively; $\boldsymbol{\zeta}$ and $\boldsymbol{\varpi}_i$ are the fixed-effects and random-effects parameter vectors, respectively; $\boldsymbol{\varpi}_i \sim N(0, \boldsymbol{\Sigma}_{\varpi})$ and $\boldsymbol{\varepsilon}_i \sim N(0, \boldsymbol{R}_i)$. However, in practice, the viral load $\boldsymbol{y}_i$ are most likely not normally distributed. In this case, we may assume $\boldsymbol{\varepsilon}_i \sim ST_{n_i,\nu_2}\left(-J(\nu_2)\delta_e \mathbf{1}_{n_i}, \sigma_2^2 \boldsymbol{I}_{n_i}, \delta_e \boldsymbol{I}_{n_i}\right)$, which follows a multivariate ST distribution with degrees of freedom ν_2, unknown scale parameter σ_2^2 and skewness parameter δ_e, where $\mathbf{1}_{n_i} = (1, \ldots, 1)^T$ and $J(\nu_2) = (\nu_2/\pi)^{1/2}[\Gamma((\nu_2 - 1)/2)/\Gamma(\nu_2/2)]$. Note that $-J(\nu_2)\delta_e \mathbf{1}_{n_i}$ is set here in order to have a zero mean vector for the ST distribution.

3 Bayesian analysis for the PLME joint model

The simultaneous parameter estimation based on the joint likelihood for the covariate and response longitudinal data associated with the joint models (1) and (5) is often computationally infeasible and may lead to convergence problems, sometimes it is even computationally prohibitive (Wu 2002; Liu and Wu 2007). We propose a fully Bayesian approach to estimate the parameters in models (1) and (5) simultaneously. Markov chain Monte Carlo (MCMC) methods enable us to sample the posterior distribution for each parameter and make inference subsequently.

Assume that $\boldsymbol{a}_i$, $\boldsymbol{\varpi}_i$, $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\epsilon}_i$ are mutually independent. Following the properties of ST distributions (Sahu et al. 2003), it can be shown by introducing two $n_i \times 1$ random vectors $\boldsymbol{w}_{e_i}$ and $\boldsymbol{w}_{\epsilon_i}$ based on the stochastic representation for the ST distribution (see

Huang and Dagne 2012; Sahu et al. 2003 in detail) that $\boldsymbol{z}_i$ and $\boldsymbol{y}_i$ can be hierarchically formulated as follows.

$$\begin{aligned}
&\boldsymbol{z}_i|\boldsymbol{a}_i, \boldsymbol{w}_{\epsilon_i} \sim t_{n_i,\nu_1+n_i}\left(\boldsymbol{z}_i^* + \delta_\epsilon[\boldsymbol{w}_{\epsilon_i} - J(\nu_1+n_i)\mathbf{1}_{n_i}], \omega_{1i}\sigma_1^2\boldsymbol{I}_{n_i}\right),\\
&\boldsymbol{y}_i|\boldsymbol{z}_i, \boldsymbol{a}_i, \boldsymbol{\varpi}_i, \boldsymbol{w}_{e_i} \sim t_{n_i,\nu_2+n_i}\left(\boldsymbol{X}_i\boldsymbol{\zeta} + \boldsymbol{Z}_i\boldsymbol{\varpi} + \delta_e[\boldsymbol{w}_{e_i} - J(\nu_2+n_i)\mathbf{1}_{n_i}],\ \omega_{2i}\sigma_2^2\boldsymbol{I}_{n_i}\right),\\
&\boldsymbol{a}_i \sim N_a(\mathbf{0}, \boldsymbol{\Sigma}_a), \quad \boldsymbol{\varpi}_i \sim N_{\varpi}(\mathbf{0}, \boldsymbol{\Sigma}_{\varpi}),\\
&\boldsymbol{w}_{\epsilon_i} \sim t_{n_i,\nu_1}(\mathbf{0}, \boldsymbol{I}_{n_i})I(\boldsymbol{w}_{\epsilon_i} > \mathbf{0}), \quad \boldsymbol{w}_{e_i} \sim t_{n_i,\nu_2}(\mathbf{0}, \boldsymbol{I}_{n_i})I(\boldsymbol{w}_{e_i} > \mathbf{0}),
\end{aligned} \tag{6}$$

where $\omega_{1i} = (\nu_1 + \boldsymbol{w}_{\epsilon_i}^T\boldsymbol{w}_{\epsilon_i})/(\nu_1+n_i)$, $\omega_{2i} = (\nu_2 + \boldsymbol{w}_{e_i}^T\boldsymbol{w}_{e_i})/(\nu_2+n_i)$, $t_{n_i,\nu}(\boldsymbol{\mu}, \boldsymbol{A})$ denote the n_i-variate t distribution with parameters $\boldsymbol{\mu}$, $\boldsymbol{A}$ and degrees of freedom ν, $I(\boldsymbol{w} > 0)$ is an indicator function and $\boldsymbol{w} = |\boldsymbol{\varsigma}|$ with $\boldsymbol{\varsigma} \sim t_{n_i,\nu}(0, \boldsymbol{I}_{n_i})$. The above representations of the hierarchical models can be easily implemented in the freely available WinBUGS software (Lunn et al. 2000) and enable that the computational effort for the model with an ST distribution is almost equivalent to that for the model with a symmetric distribution.

Under Bayesian framework, we next need to specify prior distributions for all of these parameters as follows.

$$\begin{aligned}
&\boldsymbol{\alpha} \sim N_\alpha(\boldsymbol{\tau}_1, \boldsymbol{\Lambda}_1), \quad \sigma_1^2 \sim IG(\omega_1, \omega_2), \quad \boldsymbol{\Sigma}_a \sim IW(\boldsymbol{\Omega}_1, \rho_1), \quad \delta_\epsilon \sim N(0, \gamma_1),\\
&\boldsymbol{\zeta} \sim N_\zeta(\boldsymbol{\tau}_2, \boldsymbol{\Lambda}_2), \quad \sigma_2^2 \sim IG(\omega_3, \omega_4), \quad \boldsymbol{\Sigma}_{\varpi} \sim IW(\boldsymbol{\Omega}_2, \rho_2), \quad \delta_e \sim N(0, \gamma_2),\\
&\nu_1 \sim Exp(\nu_{10})I(\nu_1 > 3), \quad \nu_2 \sim Exp(\nu_{20})I(\nu_2 > 3)
\end{aligned} \tag{7}$$

where the mutually independent Inverse Gamma (IG), Normal (N), Exponential (Exp) and Inverse Wishart (IW) prior distributions are chosen to facilitate computations. The super-parameter matrices $\boldsymbol{\Lambda}_1$, $\boldsymbol{\Lambda}_2$, $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_2$ can be assumed to be diagonal for convenient implementation. The exponential priors for ν are truncated to lie above 3 to make variance of ST distribution well-defined.

Let $f(\cdot|\cdot)$, $F(\cdot|\cdot)$ and $\pi(\cdot)$ be the conditional density function, cumulative density function (cdf) and prior density function, respectively. Let $\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\zeta}, \sigma_1^2, \sigma_2^2, \boldsymbol{\Sigma}_a, \boldsymbol{\Sigma}_{\varpi}, \nu_1, \nu_2, \delta_\epsilon, \delta_e\}$ be the collection of unknown population parameters in models (1) and (5). If we assume that $\boldsymbol{\alpha}$, $\boldsymbol{\zeta}$, σ_1^2, σ_2^2, $\boldsymbol{\Sigma}_a$, $\boldsymbol{\Sigma}_{\varpi}$, ν_1, ν_2, δ_ϵ and δ_e are independent of each other, then $\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\alpha})\pi(\boldsymbol{\zeta})\pi(\sigma_1^2)\pi(\sigma_2^2)\pi(\boldsymbol{\Sigma}_a)\pi(\boldsymbol{\Sigma}_{\varpi})\pi(\nu_1)\pi(\nu_2)\pi(\delta_\epsilon)\pi(\delta_e)$. We further need to specify the models for the observed data and the prior distributions for the unknown model parameters. Subsequently, based on the posterior distributions, we can make statistical inference for the unknown parameters. When viral load response is observable (i.e. $y_{ij} > d$), the contribution of the detectable measurement y_{ij} to the joint likelihood is $f(y_{ij}|\boldsymbol{z}_i, \boldsymbol{a}_i, \boldsymbol{\varpi}_i, \boldsymbol{w}_{e_i})$. Whereas, in the case of an undetectable measurement y_{ij} (i.e. $y_{ij} \le d$), the contribution to the joint likelihood is $\Pr(y_{ij} < d|\boldsymbol{z}_i, \boldsymbol{a}_i, \boldsymbol{\varpi}_i, \boldsymbol{w}_{e_i}) = F(d|\boldsymbol{z}_i, \boldsymbol{a}_i, \boldsymbol{\varpi}_i, \boldsymbol{w}_{e_i})$. As a result, the joint posterior density of

$\boldsymbol{\theta}$ based on the observed data $\Re$ is

$$f(\boldsymbol{\theta}|\Re) \propto \left\{ \prod_{i=1}^{n} \int \int \prod_{j=1}^{n_j} f(y_{ij}|\boldsymbol{z}_i, \boldsymbol{a}_i, \boldsymbol{\varpi}_i, \boldsymbol{w}_{e_i})^{1-c_{ij}} F(d|\boldsymbol{z}_i, \boldsymbol{a}_i, \boldsymbol{\varpi}_i, \boldsymbol{w}_{e_i})^{c_{ij}} \right.$$
$$\left. \times f(\boldsymbol{z}_i|\boldsymbol{a}_i, \boldsymbol{w}_{\epsilon_i}) f(\boldsymbol{w}_{e_i}|\boldsymbol{w}_{e_i} > \boldsymbol{0}) f(\boldsymbol{w}_{\epsilon_i}|\boldsymbol{w}_{\epsilon_i} > \boldsymbol{0}) f(\boldsymbol{a}_i) f(\boldsymbol{\varpi}_i) d\boldsymbol{a}_i d\boldsymbol{\varpi}_i \right\} \pi(\boldsymbol{\theta}) \quad (8)$$

where $\prod_{j=1}^{n_j} f(y_{ij}|\boldsymbol{z}_i, \boldsymbol{a}_i, \boldsymbol{\varpi}_i, \boldsymbol{w}_{e_i})^{1-c_{ij}} F(d|\boldsymbol{z}_i, \boldsymbol{a}_i, \boldsymbol{\varpi}_i, \boldsymbol{w}_{e_i})^{c_{ij}} f(\boldsymbol{w}_{e_i}|\boldsymbol{w}_{e_i} > \boldsymbol{0})$ is the contribution to the full likelihood by individual i after accounting for the LOD measurement for viral load response; $f(\boldsymbol{z}_i|\boldsymbol{a}_i, \boldsymbol{w}_{\epsilon_i}) f(\boldsymbol{w}_{\epsilon_i}|\boldsymbol{w}_{\epsilon_i} > \boldsymbol{0})$ is the contributed likelihood from the same individual on CD4 cell counts.

In general, the integrals in (8) are of high dimension and do not have closed form. Numerical approximations to the integrals may not be sufficiently accurate. Therefore, it is prohibitive to directly calculate the posterior distribution of $\boldsymbol{\theta}$ based on the observed data. As an alternative, the MCMC procedure can be used to sample from the posterior distributions based on (8), using the Gibbs sampler along with the Metropolis-Hastings (M-H) algorithm.

4 Analysis of an AIDS clinical data

4.1 Specification of joint models

Section 2.1 has briefly described the dataset that motivated this research. Hammer et al. (2002) has more detailed discussion on this study. As we have discussed previously, the viral load in log scale and CD4 cell counts are highly skewed (Fig. 1). It is therefore critical to consider the skewed distribution, such as ST, in the model. Furthermore, the covariate CD4 cell counts are measured with error and there are substantial observations below LOD in viral load data. The joint model with the associated inferential method accounts for multiple features in this dataset. It is our belief that the proposed model will perform better than other models that only consider few of these factors. Toward this end, we compare four statistical models with different structure and specification of random errors for both response model and covariate measurement error model.

- *Model I* A "naive" model with the independent multivariate normal distribution of random errors for the response model (5) and not accounting for covariate measurement error.
- *Model II* A model with the independent multivariate ST distribution of random errors for response model (5) but not accounting for covariate measurement error.
- *Model III* A joint model with the independent multivariate ST distribution of random errors for both the covariate model (1) and response model (5), but data feature of LOD in response is not accounted for.
- *Model IV* A joint model with the independent multivariate ST distribution of random errors for both the covariate model (1) and response model (5) with considering left censoring due to LOD in response.

Note that in Models I and II, z_i^* in (5) is replaced by the measured CD4 cell counts (z_i). In Models I through III, the data below LOD are imputed by the half of LOD (25 copies/ml in ACTG 398 study). Thus, the contribution to the joint likelihood by viral load response, i.e. $\prod_{j=1}^{n_j} f(y_{ij}|\boldsymbol{z}_i, \boldsymbol{a}_i, \boldsymbol{\varpi}_i, \boldsymbol{w}_{e_i})^{1-c_{ij}} F(d|\boldsymbol{z}_i, \boldsymbol{a}_i, \boldsymbol{\varpi}_i, \boldsymbol{w}_{e_i})^{c_{ij}} f(\boldsymbol{w}_{e_i}|\boldsymbol{w}_{e_i} > \boldsymbol{0})$, is substituted by $\prod_{j=1}^{n_j} f(y_{ij}|\boldsymbol{z}_i, \boldsymbol{a}_i, \boldsymbol{\varpi}_i, \boldsymbol{w}_{e_i}) f(\boldsymbol{w}_{e_i}|\boldsymbol{w}_{e_i} > \boldsymbol{0})$ in (8). We will perform the following comparisons: by comparing Models I and II, we test whether the asymmetric ST distribution for model error will improve model fitting as compared to normal distribution; by comparing Models II and III, we assess whether measurement error in CD4 cell counts is necessary to account for; by comparing Models III and IV, we evaluate whether the mechanism accounting for data below LOD in viral load contributes to the model fitting.

The measurements on CD4 cell counts usually contain non-negligible errors. Ignoring CD4 measurement errors leads to severely misleading results (Carroll et al. 2006). Based on the CD4 measurements collected over time during an AIDS study, we can model the CD4 process to partially address the measurement errors (Wu 2002). Nevertheless, there are no well established models for the CD4 process. We model the CD4 process empirically using an LME model which is flexible and works well for complex longitudinal data. It is noted that, in the absence of theoretical rationale, we may pick up a low order polynomial models by employing a standard model selection technique. In our case, AIC and BIC are used to determine the best covariate measurement error model (1) with a quadratic function. Thus, we model the CD4 process as

$$z_{ij} = (\alpha_1 + a_{1i}) + (\alpha_2 + a_{2i})t_{ij} + (\alpha_3 + a_{3i})t_{ij}^2 + \epsilon_{ij} \tag{9}$$

where the true CD4 cell counts $z_{ij}^* = (\alpha_1 + a_{1i}) + (\alpha_2 + a_{2i})t_{ij} + (\alpha_3 + a_{3i})t_{ij}^2$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)^T$ is the population (fixed-effects) parameter vector, $\boldsymbol{a}_i = (a_{1i}, a_{2i}, a_{3i})^T$ is an individual-specific random-effects parameter vector which is assumed to follow normal distribution centered at zero with variance-covariance $\boldsymbol{\Sigma}_a$. In addition, we assume that ϵ_{ij} independently follows ST distribution $ST_{\nu_1}\left(-J(\nu_1)\delta_\epsilon, \sigma_1^2, \delta_\epsilon\right)$, where ν_1 is the degrees of freedom, σ_1^2 is the unknown scale parameter and δ_ϵ is the skewness parameter.

To approximate the smoothing functions of t_i in model (5), we use the same basis functions for $\boldsymbol{g}(\cdot)$ and $\boldsymbol{h}_i(\cdot)$. We adopt the natural cubic spline bases with the percentile-based knots. The smoothing parameters p and r in (3) are determined by BIC. For ACTG 398 dataset, the optimal smoothing parameters are both set at 3.

In Bayesian analysis, we need to specify the values for the hyper-parameters in the prior distributions (7). We take weakly informative prior distribution for the parameters in the joint models. In particular, (1) fixed-effects were taken to be independent normal distributions $N(0, 100)$ for each component of the population parameter vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\zeta}$. (2) For the scale parameters σ_1^2 and σ_2^2 we assume a limiting non-informative inverse gamma prior distribution, $IG(0.01, 0.01)$ so that the distribution has mean 1 and variance 100. (3) The priors for the variance-covariance matrices of the random-effects $\boldsymbol{\Sigma}_a$ and $\boldsymbol{\Sigma}_{\varpi}$ are taken to be inverse Wishart distributions $IW(\boldsymbol{\Omega}_1, \rho_1)$ and $IW(\boldsymbol{\Omega}_2, \rho_2)$ with covariance matrices $\boldsymbol{\Omega}_1 = diag(0.01, 0.01, 0.01)$, $\boldsymbol{\Omega}_2 =$

$diag(0.01, 0.01, 0.01, 0.01)$ and $\rho_1 = \rho_2 = 4$, respectively. (4) For each of the skewness parameters δ_e and δ_ϵ, we choose independent normal distribution $N(0, 100)$. (5) The degrees of freedom parameters ν_1 and ν_2 follow truncated exponential distribution with $\nu_{10} = \nu_{20} = 0.5$.

The MCMC sampler is implemented using WinBUGS software (Lunn et al. 2000). The MCMC scheme for drawing samples from the posterior distributions of all parameters in the joint models is obtained by iterating between the Gibbs sampler and the M-H algorithm. After the final MCMC samples are collected, we are able to draw statistical inference for the unknown parameters. Specifically, we are interested in the posterior means and quantiles. When the MCMC procedure is applied to the actual clinical data, convergence of the generated samples is assessed using standard tools within WinBUGS software such as trace plots and Gelman–Rubin diagnostics (Gelman and Rubin 1992). When convergence was achieved, for each of three chains, after an initial number of 50,000 burn-in iterations, every 50th MCMC sample is retained from the next 50,000. Thus, we obtain 3000 samples of the targeted posterior distributions of the unknown parameters for statistical inference. The computational burden for fitting our models via MCMC procedure was reasonable. For example, to fit Model IV, it took about 9 h on a Window PC with Intel Core i7-2600 CPU @ 6.80 GHz and 16 GB RAM.

The dynamic Gelman–Rubin diagnostics for representative parameters based on Model IV is given in Fig. 2. Three curves are given: the middle and bottom curves below the dashed horizontal line (indicating value one) represent the pooled posterior variance and average within-sample variance, respectively, and the top curve above the dashed horizontal line represents their ratio. We can see that the ratio tends to 1, and both pooled posterior variance and average within-sample variance settle down as the number of iterations increases, indicating that the algorithm has reached convergence.

4.2 Comparison of modeling results

We first check the model fitting diagnostic plots (Fig. 3). By comparing the fitted values against observed viral load values from different models, we see that Model IV fits the data best, followed by Models III, II and I, respectively. This finding suggests that each of the three factors: asymmetric distribution, covariate measurement error, left censoring mechanism due to LOD in response, contributes synergistically to the model fitting. Taking into account each of them in the modeling process improves modeling fitting results. It is also obvious based on Q–Q plots that Models II, III and IV that assume an ST distribution for the random errors have fewer outliers than Model I which assumes a normal distribution. To compare individual fitting results by each model, we pick four patients and overlay their fitting results from four models in Fig. 4. It can be seen that Model I does not fit well each individual; Model II improves model fitting for each individual greatly but misses a few important observations; Model III fits the data very well except for below LOD response data; Model IV captures both below and above LOD data very well, and thus provides the best individual fitting among the four models.

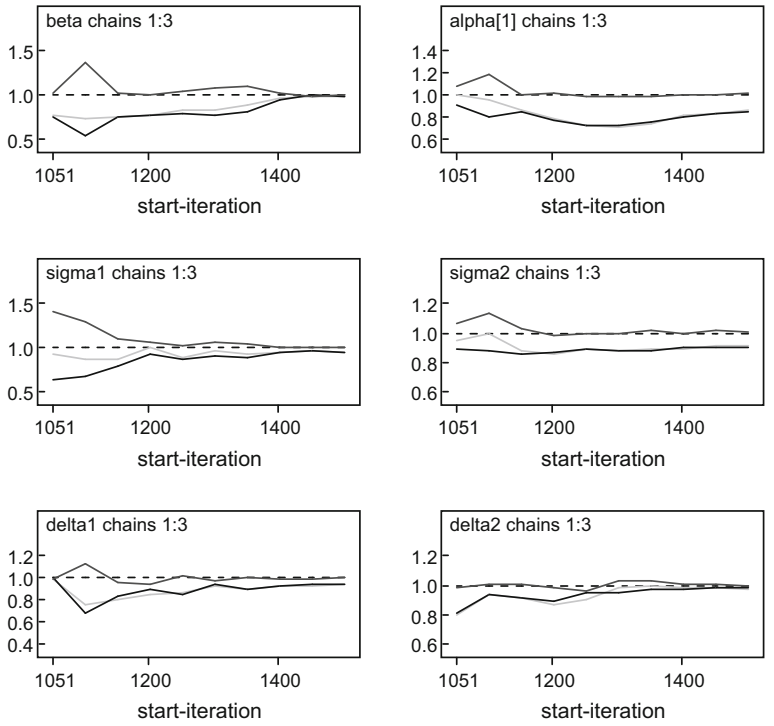

Fig. 2 Gelman–Rubin diagnostic plot based on Model IV (the joint model considering skewness, covariate measurement error and *left-censored* response) for representative parameters from three Markov chains. The *middle* and *bottom curves* below the *dashed line* which indicates value one, stand for the pooled posterior variance and average within-sample variance separately, and the *top curve* represents their ratio

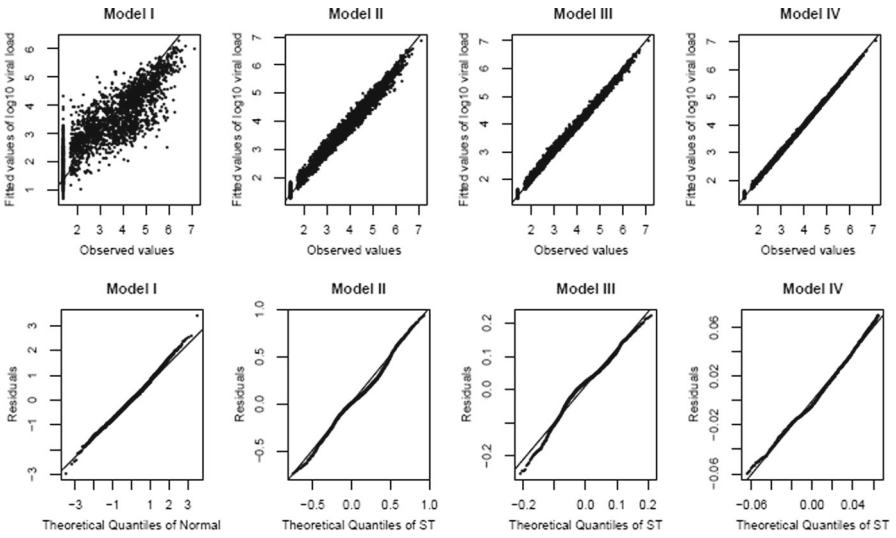

Fig. 3 Diagnostics of model fitting for the four models discussed in Sect. 4.1. *Top panel* fitted values versus observed values of $\log_{10}$ viral load; *bottom panel* Q–Q plots of residuals for the four models

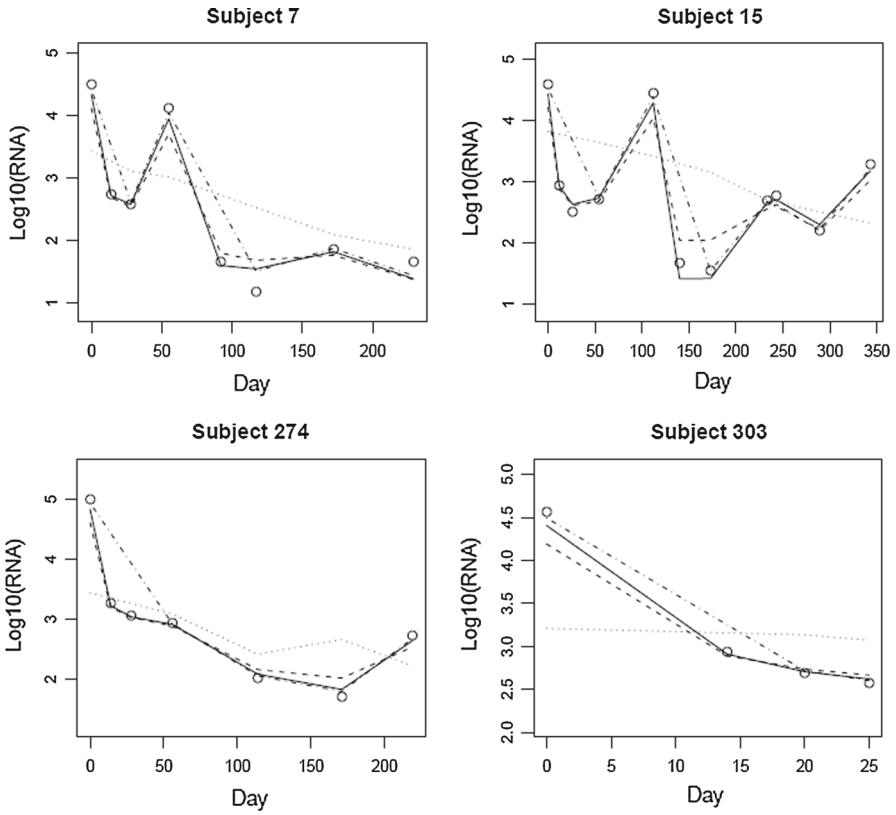

Fig. 4 Individual model fitting from the four models discussed in Sect. 4.1. Model I: *dotted lines*; Model II: *dash-dotted lines*; Model III: *dashed lines*; Model IV: *solid lines*. Viral load measurements are represented by circles. Note that number of days considerably differ by subject and range from 25 days for Subject 303–350 days for Subject 15

The population posterior means (PM), standard deviation (SD) and 95 % credible interval (CI) for fixed-effects parameters based on Models I to IV along with the scale, skewness and degrees of freedom parameters are summarized in Table 1. Based on these results, we have the following findings: (1) the parameter estimates are all significant (95 % CIs do not include zero); (2) for the key population parameter β that quantifies the relationship between viral load and CD4 cell counts in (2), the posterior mean is gradually increasing from -0.54 to -0.39 with negative values for Model I through Model IV, indicating that the relationship becomes weaker when multiple data features are accounted for in the joint model; (3) the scale parameter σ_2^2 that quantifies the dispersion of random errors in the response model (5) is shrinking from 0.76 in Model I, over 0.16 in Model II, 0.12 in Model III to 0.07 in Model IV. Further, the other two factors, covariate measurement error and left censoring in viral load response, both contribute to the attenuation of scale parameter; (4) the estimates of the skewness parameters (δ_e and δ_ϵ) of Models II, III and IV are all significantly positive, confirming that the distribution of the original data is skewed

even after transformation of viral load (see Fig. 1). Thus, it is recommended to include a skewness parameter in the model; (5) the incorporation of measurement error in Models III and IV impacts other parameters' estimation (by comparison of the values for $\beta, \sigma_2^2, \delta_e$ and ν_2 between Model II and Models III and IV in Table 1); (6) for parameter estimates of CD4 covariate model (9), the estimates of the linear coefficient α_2 based on Models III and IV are slightly different, but significantly positive, whereas the estimates of α_1 and α_3 are significantly negative. This finding suggests that there is a positive linear relationship between CD4 cell counts and measurement time; (7) there are not quite significant differences in the parameter estimates ($\alpha_1, \alpha_2, \alpha_3, \sigma_1^2, \delta_\epsilon$ and ν_1) as well as SD and 95 % CI for the covariate measurement error model between Models III and IV. This comparison suggests that incorporating the left censoring mechanism due to LOD in response does not influence the parameter estimation in covariate measurement error model (9). Nevertheless, the estimated parameters in the response model (5) are significantly impacted.

We further compute the Deviance Information Criterion (DIC) (Spiegelhalter et al. 2002) and the expected predictive deviance (EPD) for each model. EPD is defined as $\text{EPD} = E\left\{\sum_{i,j}(y_{rep,ij} - y_{obs,ij})^2\right\}$ where the predictive value $y_{rep,ij}$ is a replicate of the observed $y_{obs,ij}$ and the expectation is taken over the posterior distribution of the model parameters $\boldsymbol{\theta}$ (Gelman et al. 2003). The best model has the least discrepancy between observed values and predictive values. According to Table 1, Model IV has the least DIC and EPD among all. The order of the preferred model based on these two criteria is Models IV, III, II and I, respectively. The results are consistent with the goodness of fit diagnostics shown in Fig. 3. This finding suggests that all three components we considered here contribute to the model fitting. Therefore, it is critical to take into account asymmetric distribution, covariate measurement error and left censoring mechanism due to LOD in response.

Next, we look at the population estimate of the smooth function $\boldsymbol{g}(t)$ in (2). Figure 5 presents an estimating curve of $\boldsymbol{g}(t)$ against time along with 95 % CI for the four models. We observe that Models I and II show a similar shape of the curves. Estimated curves in Models III and IV are similar to each other, but differ from ones in Models I and II. The curves from all four models start at the similar point (around 3.5) but diverge afterwards. The curves from Models I and II begin to decrease from day 0 almost linearly but go upwards after the mid-point (around day 280) and recover to the value at day 0 eventually. On the other hand, the curves from Models III and IV remain steady after the treatment initiates until around day 250 when they turn to pick up gradually. We summarize the findings as follows: (1) the ST distribution for random error does not show an obvious change on the shape of the population curve, in comparison with the normal distribution (i.e., comparison of curves between Models I and II); (2) the addition of measurement error into the joint model change the shape of the population curves drastically (i.e. comparison of curves between Models I/II and Model III); (3) the incorporation of data below LOD into the joint model does not change the shape of the population curve dramatically (i.e., comparison of curves between Model III and Model IV). The 95 % CI is narrower at the beginning of the treatment but gradually becomes wider at later stage. This is understandable by the

Table 1 The estimated posterior mean (PM) of population (fixed-effects) parameters, as well as the corresponding standard deviation (SD) and lower limit (L_{CI}) and upper limit (U_{CI}) of 95 % equal-tail credible interval (CI) for the four models described in Sect. 4.1

Model	β	α_1	α_2	α_3	σ_1^2	σ_2^2	δ_ϵ	δ_e	ν_1	ν_2	DIC	EPD
I												
PM	−0.54	–	–	–	–	0.76	–	–	–	–	7059	1.52
L_{CI}	−0.63	–	–	–	–	0.71	–	–	–	–		
U_{CI}	−0.44	–	–	–	–	0.82	–	–	–	–		
SD	0.05	–	–	–	–	0.03	–	–	–	–		
II												
PM	−0.52	–	–	–	–	0.16	–	0.89	–	8.9	5963	0.54
L_{CI}	−0.68	–	–	–	–	0.002	–	0.41	–	3.0		
U_{CI}	−0.4	–	–	–	–	0.49	–	1.37	–	18.5		
SD	0.07	–	–	–	–	0.21	–	0.23	–	4.6		
III												
PM	−0.46	−0.17	0.62	−0.54	0.03	0.12	0.23	0.76	3.3	4.2	3869	0.47
L_{CI}	−0.58	−0.26	0.38	−0.98	0.02	0.08	0.18	0.61	3.0	3.1		
U_{CI}	−0.36	−0.08	0.9	−0.22	0.04	0.16	0.28	0.93	3.9	5.8		
SD	0.06	0.05	0.13	0.19	0.004	0.02	0.03	0.08	0.2	0.7		
IV												
PM	−0.39	−0.17	0.6	−0.49	0.03	0.07	0.23	0.88	3.4	6.5	2998	0.2
L_{CI}	−0.52	−0.26	0.37	−0.84	0.02	0.03	0.18	0.70	3.0	4.3		
U_{CI}	−0.28	−0.08	0.84	−0.17	0.04	0.11	0.28	1.02	4.0	9.7		
SD	0.07	0.05	0.12	0.17	0.004	0.02	0.03	0.08	0.3	1.4		

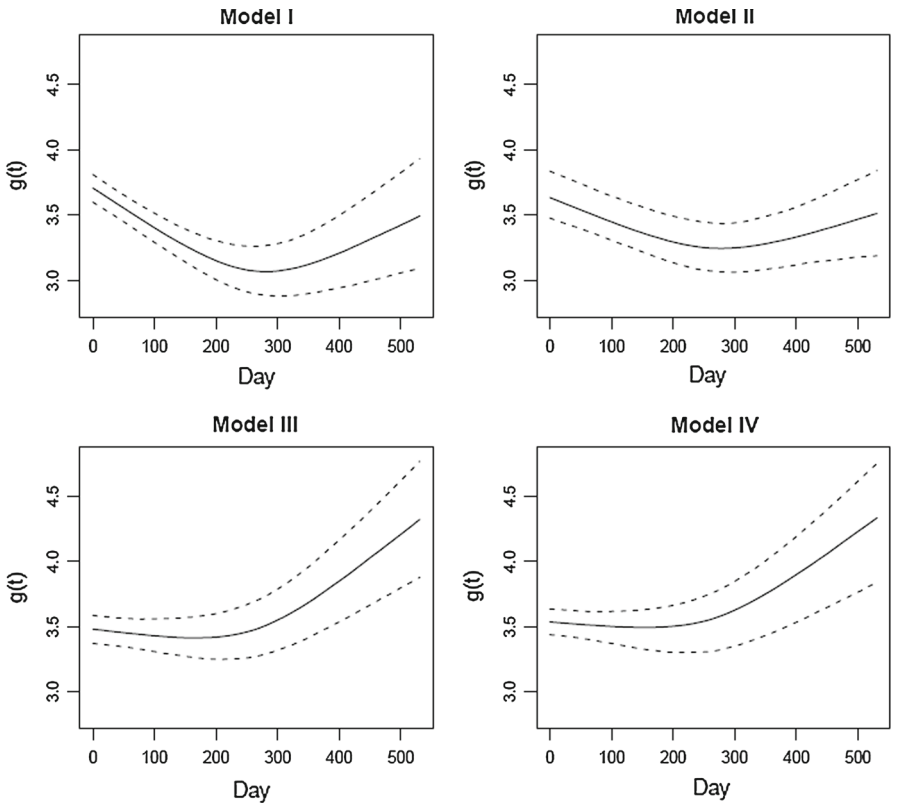

Fig. 5 The population estimating curve of $\boldsymbol{g}(t)$ based on the four models discussed in Sect. 4.1. The estimates (*solid curve*) along with the 95 % credible intervals (*dashed curves*) are presented

fact that there is more dropout at the later stage than during the earlier period so that more uncertainty evolves alongside the parameter estimates.

One advantage of the proposed joint model is that by employing the PLME model (2), we can estimate not only the population curve $\boldsymbol{g}(t)$ but also the individual curve $\boldsymbol{g}_i(t)$ which is a combination of the population curve $\boldsymbol{g}(t)$ and individual departure $\boldsymbol{h}_i(t)$. We note that the individual curve $\boldsymbol{g}_i(t)$ may not follow the pattern of the population curve if the variation between subjects is large. Since Model IV is the best model based on model fitting diagnostics and model selection criteria, for ease of demonstration, we present the individual curves for the four exemplary patients and these results illustrate the principal advantage of the proposed model and developed methods in which the estimates can be obtained for both population and individuals. For comparison purpose, the corresponding population estimate is also plotted in Fig. 6. We observe that the population and individual estimates are different not only in magnitude but also in patterns of change. Some of the individuals, such as subject 24, follow a similar pattern as the population curve. However, other patients show quite different curves from the population curve. For examples, subject 274 starts from a higher point (4 as compared to 3.5 for population estimate at day 0) and then keeps decreasing

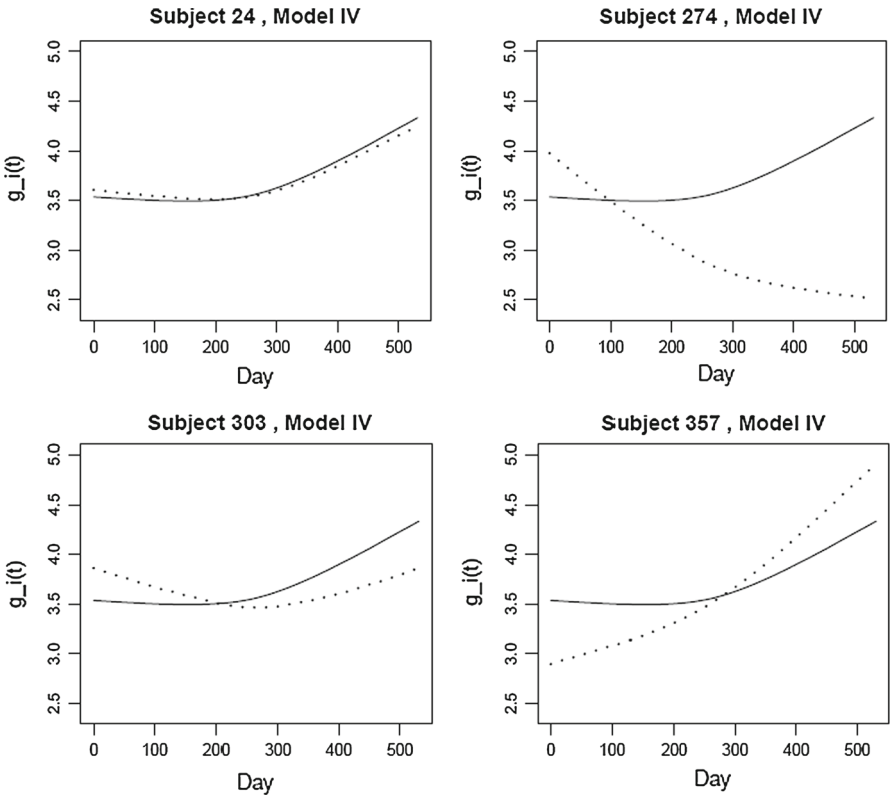

Fig. 6 Four exemplary individual estimating curves ($\boldsymbol{g}_i(t)$, *dotted curve*) based on the best selected Model IV. For comparison, the population curve (*solid curve*) is presented as well

all the way to the end of the trial; on the contrast, subject 357 starts from a lower point (3 as compared to 3.5 for population estimate at day 0), but quickly picks up and exceeds the population curve eventually; interestingly, subject 303 starts higher, followed by a concave shape and ends up lower than the population curve. Given this large between-subject variation, the estimated trajectories of the individual curves are critical for individualized treatment management and care for AIDS patients.

5 Discussion

To study the complex relationship between viral load and CD4 cell count biomarkers at both population and individual levels, we adopted a PLME joint model which is a special case of semiparametric models. However, it is critical to consider a number of data features that may potentially impact the discovery of this relationship. Toward this end, we took into account three data features that commonly arise in practice under the framework of a PLME joint model. We employed a Bayesian approach to obtain the estimates and credible intervals for interesting parameters in the joint model. To investigate whether each data feature influences the modeling results, we compared the

naive model with three other models that account for various data features. As a result, we found that the model that accounts for all three data features performed much better than those which only account for certain aspects of the data. The linear coefficient (β) that quantifies the relationship between viral load and CD4 cell counts is significantly negative in all models, suggesting that there is strong negative relationship between viral load and CD4 cell counts. This finding is consistent with biological mechanism in which CD4 cells are one of major sources to clear out HIV virus. The comparison of the nonparametric component estimates in the PLME joint model reveals different shape of curves when various data features were accounted for. Since the joint model offers a unique opportunity of both population and individual estimates, as an example, we presented several patterns at individual level based on the best Model IV. We observed that the individual curve may behave similarly to the population one, but on the other hand, they may be completely different from the population estimates.

In the Bayesian analysis, it is critical to perform sensitivity analysis to see if the posterior estimates change significantly when priors are different. Toward this end, we carried out sensitivity analysis by employing a few sets of different values for the hyper-parameters in (7) and re-run the MCMC sampling scheme. We observe that the conclusions are similar to those presented in the article. Thus, we are confident that the obtained results are robust against hyper-parameter values. In the ST-PLME model (5), we adopted the regression spline basis to represent the unknown smoothing function. There are a lot of alternative ways for approximating the unknown function, such as local polynomial kernel and smoothing splines. It is interesting to compare the modeling results based on various nonparametric methods.

To remove noises and extract the actual information for CD4 cell counts, we adopted the low order polynomial model (9) to approximate the covariate process. Note that the approximation is empirical and may not be true for the unknown CD4 path. Therefore, the fitted CD4 values based on such a model are not the 'true' ones but rather 'regularized' CD4 values. Such an operation provides a way to alleviate measurement errors in observed CD4 cell counts. Two final points related to this longitudinal study should be briefly noted. (1) Although the data set is an unbalanced panel data, it will presumably not make a difference in terms of the analysis conducted in this paper. (2) Although this article is motivated by an AIDS clinical study, the basic concepts of the developed PLME joint models and method have generally broader applications such as cancer and infectious disease studies whenever the relevant technical specifications are met.

Acknowledgements The authors gratefully acknowledge the Editor, Associate Editor and two anonymous referees for their insightful comments and detailed suggestions that led to a marked improvement of the article.

References

Arellano-Valle R, Genton M (2005) On fundamental skew distributions. J Multivar Anal 96:93–116

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

Azzalini A, Genton M (2008) Robust likelihood methods based on the skew-t and related distributions. Int Stat Rev 76:106–129

Carroll R, Ruppert D, Stefanski L, Crainiceanu C (2006) Measurement error in nonlinear models: a modern perspective. Chapman and Hall, London

Gelman A, Rubin D (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7:457–511

Gelman A, Carlin J, Stern H, Rubin D (2003) Bayesian data analysis. Chapman and Hall, London

Hammer SM, Vaida F, Bennett KK, Holohan MK, Sheiner L, Eron JJ, Wheat LJ, Mitsuyasu RT, Gulick RM, Valentine FT, Aberg JA, Rogers MD, Karol CN, Saah AJ, Lewis RH, Bessen LJ, Brosgart C, DeGruttola V, Mellors JW (2002) AIDS clinical trials group 398 study team. Dual vs single protease inhibitor therapy following antiretroviral treatment failure: a randomized trial. JAMA 288:169–180

Härdle W, Liang H, Gao J (2000) Partially linear models. Springer Physica-Verlag, Heidelberg

Ho H, Lin T (2010) Robust linear mixed models using the skew-t distribution with application to schizophrenia data. Biom J 52:449–469

Huang Y, 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

Huang Y, Dagne G (2012) Bayesian semiparametric nonlinear mixed-effects joint models for data with skewness, missing responses and measurement errors in covariates. Biometrics 68:943–953

Huang Y, Dagne G, 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

Huang Y, Dagne G, Wu L (2012) Mixed-effects joint models with skew-normal distribution for HIV dynamic response with missing and mismeasured time-varying covariate. Int J Biostat 8(1):34

Jara A, Quintana F, Martin E (2008) Linear mixed models with skew-elliptical distributions: a Bayesian approach. Comput Stat Data Anal 52:5033–5045

Lachos V, Bandyopadhyay D, Dey D (2011) Linear and nonlinear mixed-effects models for censored HIV viral loads using normal/independent distributions. Biometrics 67:1594–1604

Liang H, Ren H (2005) Generalized partially linear measurement error models. J Comput Graph Stat 14:237–250

Liang H, Wang S, Robins J, Carroll R (2004) Estimation in partially linear models with missing covariates. J Am Stat Assoc 99:357–367

Liu W, Wu L (2007) Simultaneous inference for semiparametric nonlinear mixed-effects models withco variate measurement errors and missing responses. Biometrics 63:342–350

Lunn D, Thomas A, Best N, Spiegelhalter D (2000) WinBUGS-a Bayesian modelling framework: concepts, structure, an extensibility. Stat Comput 10:325–337

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

Speckman P (1988) Kernel smoothing in partial linear models. J R Stat Soc Series B 50:413–436

Spiegelhalter D, Best N, Carlin B, Van der Linde A (2002) Bayesian measures of model complexity and fit. J R Stat Soc Series B 64:583–639

Wu L (2002) A joint model for nonlinear mixed-effects models with censoring and covariates measured with error. J Am Stat Assoc 97:955–964