

## Research Article

# Mixed-Effects Tobit Joint Models for Longitudinal Data with Skewness, Detection Limits, and Measurement Errors

**Getachew A. Dagne and Yangxin Huang**

*Department of Epidemiology & Biostatistics, College of Public Health, University of South Florida, Tampa, FL 33612, USA*

Correspondence should be addressed to Getachew A. Dagne, [gdagne@health.usf.edu](mailto:gdagne@health.usf.edu)

Received 29 May 2011; Accepted 13 August 2011

Academic Editor: Lang Wu

Copyright © 2012 G. A. Dagne and Y. Huang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Complex longitudinal data are commonly analyzed using nonlinear mixed-effects (NLME) models with a normal distribution. However, a departure from normality may lead to invalid inference and unreasonable parameter estimates. Some covariates may be measured with substantial errors, and the response observations may also be subjected to left-censoring due to a detection limit. Inferential procedures can be complicated dramatically when such data with asymmetric characteristics, left censoring, and measurement errors are analyzed. There is relatively little work concerning all of the three features simultaneously. In this paper, we jointly investigate a skew- $t$  NLME Tobit model for response (with left censoring) process and a skew- $t$  nonparametric mixed-effects model for covariate (with measurement errors) process under a Bayesian framework. A real data example is used to illustrate the proposed methods.

## 1. Introduction

Modeling of longitudinal data is an active area of biostatistics and statistics research that has received a lot of attention in the recent years. Various statistical modeling and analysis methods have been suggested in the literature for analyzing such data with complex features (Higgins et al. [1], Liu and Wu [2], Wulfsohn and Tsiatis [3], and Wu [4]). However, there is a relatively little work done on simultaneously accounting for skewness, left censoring due to a detection limit (for example, a threshold below which viral loads are not quantifiable) and covariate measurement errors, which are inherent features of longitudinal data. This paper proposes a joint skew- $t$  NLME Tobit model for a response and measurement errors in covariate by simultaneously accounting for left-censoring and skewness. Thus, the proposed model addresses three important features of longitudinal data such as viral load in an AIDS study.

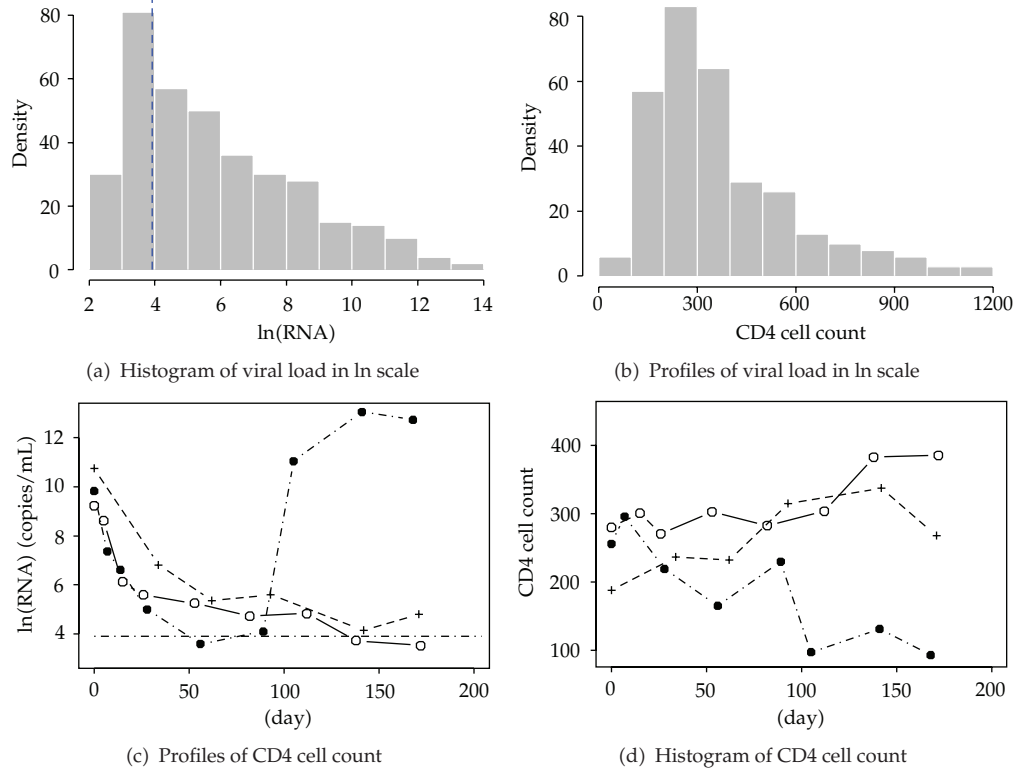
Firstly, our model relaxes the normality assumption for random errors and random-effects by using flexible skew-normal and skew- $t$  distributions. It has been documented in the literature that the normality assumption lacks robustness against extreme values, obscures important features of between- and within-subject variations, and leads to biased or misleading results (Huang and Dagne [5], Verbeke and Lesaffre [6], and Sahu et al. [7]). Specially, nonnormal characteristics such as skewness with heavy tails appear very often in virologic responses. For example, Figures 1(a) and 1(b) displays the histograms of repeated viral load (in  $\ln$  scale) and CD4 cell count measurements for 44 subjects enrolled in an AIDS clinical study (Acosta et al. [8]). For this data set, which is analyzed in this paper, both viral load (even after  $\ln$ -transformation) and CD4 cell count are highly skewed, and thus a normality assumption may be violated.

Secondly, an outcome of a longitudinal study may be subject to a detection limit because of low sensitivity of current standard assays (Perelson et al. [9]). For example, for a longitudinal AIDS study, designed to collect data on every individual at each assessment, the response (viral load) measurements may be subject to left censoring due to a detection limit of quantification. Figures 1(c) and 1(d) shows the measurements of viral load and CD4 cell count for three randomly selected patients in the study. We can see that for some patients their viral loads are below detection limit (BDL), which is 50 (in copies/mL). When observations fall below the BDL, a common practice is to impute the censored values by either the detection limit or half of the detection limit (Wu [4], Ding and Wu [10], and Davidian and Giltinan [11]). Such *ad hoc* methods may produce biased results (Hughes [12]). In this paper, instead of arbitrarily imputing the observations below detection limit, we impute them using fully Bayesian predictive distributions based on a Tobit model (Tobin [13]), which is discussed in Section 2.

Thirdly, another feature of a longitudinal data set is the existence of time-varying covariates which suffer from random measurement errors. This is usually the case in a longitudinal AIDS study where CD4 cell counts are often measured with substantial measurement errors. Thus, any statistical inference without considering measurement errors in covariates may result in biased results (Liu and Wu [2], Wu [4], and Huang and Dagne [5]). In this paper, we jointly model measurement errors in covariate process along with the response process. The distributional assumption for the covariate model is a skew- $t$  distribution which is relatively robust against potential extreme values and heavy tails.

Our research was motivated by the AIDS clinical trial considered by Acosta et al. [8]. In this study, 44 HIV-1-infected patients were treated with a potent antiretroviral regimen. RNA viral load was measured in copies/mL at study days 0, 7, 14, 28, 56, 84, 112, 140, and 168 of followup. Covariates such as CD4 cell counts were also measured throughout the study on similar scheme. In this study, the viral load detectable limit is 50 copies/mL, and there are 107 out of 357 (30 percent) of all viral load measurements that are below the detection limit. Previous studies show that change in viral load may be associated with change in CD4 cell counts. It is important to study the patterns of virological response to treatment in order to make clinical decisions and provide individualized treatments. Since viral load measurements appear to be skewed and censored, and in addition CD4 cell counts are typically measured with substantial errors and skewness, statistical analyses must take all these factors into account.

For longitudinal data, it is not clear how asymmetric nature, left censoring due to BDL, and covariate measurement error may interact and simultaneously influence inferential procedures. It is the objective of this paper to investigate the effects on inference when all of the three typical features exist in the longitudinal data. To achieve our objective,



**Figure 1:** The histograms (a,b) of viral load measured from RNA levels (in natural ln scale) and standardized CD4 cell count in plasma for 44 patients in an AIDS clinical trial study. Profiles (c,d) of viral load (response) in ln scale and CD4 cell count (covariate) for three randomly selected patients. The vertical and horizontal lines in (a) and (c) are below the detectable level of viral load ( $3.91 = \ln(50)$ ).

we employ a fairly general framework to accommodate a large class of problems with various features. Accordingly, we explore a flexible class of skew-elliptical (SE) distributions (see the Appendix for details) which include skew-normal (SN) and skew- $t$  (ST) distributions as special cases for accounting skewness and heavy tails of longitudinal data, extend the Tobit model (Tobin [13]) to treat all left-censored observations as missing values, and investigate nonparametric mixed effects model for covariate measured with error under the framework of joint models. Because the SN distribution is a special case of the ST distribution when the degrees of freedom approach infinity, for the completeness and convenient presentation, we chose ST distributions to develop NLME Tobit joint models (i.e., the ST distribution is assumed for within-subject random errors and between-subject random effects). The skewness in both within-subject random errors and random-effects distributions may jointly contribute to the skewness of response and covariate variables in a longitudinal study, which makes the assumption of normality unrealistic.

The remaining of the paper is structured as follows. In Section 2, we present the joint models with ST distribution and associated Bayesian modeling approach in general forms so that they can be applicable to other scientific fields. In Section 3, we discuss specific joint models for HIV response process with left censoring and CD4 covariate process with measurement error that are used to illustrate the proposed methods using the data set

described above and report the analysis results. Finally, the paper concludes with some discussions in Section 4.

## 2. Joint Models and Bayesian Inferential Methods

### 2.1. Skew- $t$ Mixed-Effects Tobit Joint Models

In this section, we present the models and methods in general forms so that our methods may be applicable to other areas of research. An approach we present in this paper treats censored values as realizations of a latent (unobserved) continuous variable that has been left-censored. This idea was popularized by Tobin ([13]) and the resulting model is commonly referred to as the Tobit model. Denote the number of subjects by  $n$  and the number of measurements on the  $i$ th subject by  $n_i$ . Let  $y_{ij} = y_i(t_{ij})$  and  $z_{ij} = z_i(t_{ij})$  be observed response and covariate for individual  $i$  at time  $t_{ij}$  ( $i = 1, 2, \dots, n$ ;  $j = 1, 2, \dots, n_i$ ) and  $q_{ij}$  denote the latent response variable that would be measured if the assay did not have a lower detectable limit  $\rho$ . In our case the Tobit model can be formulated as

$$y_{ij} = \begin{cases} q_{ij} & \text{if } q_{ij} > \rho, \\ \text{missing} & \text{if } q_{ij} \leq \rho, \end{cases} \quad (2.1)$$

where  $\rho$  is a nonstochastic BDL, which in our example below is equivalent to  $\ln(50)$ . Note that the value of  $y_{ij}$  is missing when it is less than or equal to  $\rho$ .

For the response process with left-censoring, we consider the following NLME model with an ST distribution which incorporates possibly mismeasured time-varying covariates

$$\begin{aligned} y_{ij} &= g(t_{ij}, \mathbf{x}_{ij}, \boldsymbol{\beta}_{ij}) + e_{ij}, & \mathbf{e}_i &\text{ iid } \sim ST_{n_i, \nu_e}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i}, \boldsymbol{\Delta}(\boldsymbol{\delta}_{e_i})), \\ \boldsymbol{\beta}_{ij} &= \mathbf{d}(z_{ij}^*, \boldsymbol{\beta}, \mathbf{b}_i), & \mathbf{b}_i &\text{ iid } \sim ST_{s_3, \nu_b}(\mathbf{0}, \boldsymbol{\Sigma}_b, \boldsymbol{\Delta}(\boldsymbol{\delta}_b)), \end{aligned} \quad (2.2)$$

where  $\mathbf{x}_{ij}$  is an  $s_1 \times 1$  design vector,  $g(\cdot)$  is a linear or nonlinear known function,  $\mathbf{d}(\cdot)$  is an  $s_1$ -dimensional vector-valued linear function,  $\boldsymbol{\beta}_{ij}$  is an  $s_1 \times 1$  individual-specific time-dependent parameter vector,  $\boldsymbol{\beta}$  is an  $s_2 \times 1$  population parameter vector ( $s_2 \geq s_1$ ); in the model (2.2), we assume that the individual-specific parameters  $\boldsymbol{\beta}_{ij}$  depend on the true (but unobservable) covariate  $z_{ij}^*$  rather than the observed covariate  $z_{ij}$ , which may be measured with errors, and we discuss a covariate model (2.3) below.

It is noticed that we assume that an  $s_3 \times 1$  vector of random effects  $\mathbf{b}_i = (b_{i1}, \dots, b_{is_3})^T$  ( $s_3 \leq s_1$ ) follows a multivariate ST distribution with the unrestricted covariance matrix  $\boldsymbol{\Sigma}_b$ , the  $s_3 \times s_3$  skewness diagonal matrix  $\boldsymbol{\Delta}(\boldsymbol{\delta}_b) = \text{diag}(\delta_1^b, \dots, \delta_{s_3}^b)$ , and the degree of freedom  $\nu_b$ ; the model random error  $\mathbf{e}_i = (e_{i1}, \dots, e_{in_i})^T$  follows a multivariate ST distribution with the unknown scale parameter  $\sigma^2$ , the degree of freedom  $\nu_e$ , and the  $n_i \times n_i$  skewness diagonal matrix  $\boldsymbol{\Delta}(\boldsymbol{\delta}_{e_i}) = \text{diag}(\delta_{e_{i1}}, \dots, \delta_{e_{in_i}})$ , where the  $n_i \times 1$  skewness parameter vector  $\boldsymbol{\delta}_{e_i} = (\delta_{e_{i1}}, \dots, \delta_{e_{in_i}})^T$ . In particular, if  $\delta_{e_{i1}} = \dots = \delta_{e_{in_i}} \hat{=} \delta_e$ , then  $\boldsymbol{\Delta}(\boldsymbol{\delta}_{e_i}) = \delta_e \mathbf{I}_{n_i}$  and  $\boldsymbol{\delta}_{e_i} = \delta_e \mathbf{1}_{n_i}$  with  $\mathbf{1}_{n_i} = (1, \dots, 1)^T$ ; this indicates that we are interested in skewness of overall data set and is the case to be used in real data analysis in Section 3.

Covariate models have been investigated extensively in the literature (Higgins et al. [1], Liu and Wu [2], Wu [4], and Carroll et al. [14]). However, those models used the

normality assumption for random measurement errors. As we pointed out earlier, this assumption lacks robustness against departures from normality and may also lead to misleading results. In this paper, we extend the covariate models by assuming an ST distribution for the random errors. We adopt a flexible empirical nonparametric mixed-effects model with an ST to quantify the covariate process as follows:

$$z_{ij} = w(t_{ij}) + h_i(t_{ij}) + \epsilon_{ij} \left( \equiv z_{ij}^* + \epsilon_{ij} \right) \quad \epsilon_i \text{ iid} \sim ST_{n_i, \nu_\epsilon} \left( \mathbf{0}, \tau^2 \mathbf{I}_{n_i}, \mathbf{\Delta}(\boldsymbol{\delta}_{\epsilon_i}) \right), \quad (2.3)$$

where  $w(t_{ij})$  and  $h_i(t_{ij})$  are unknown nonparametric smooth fixed-effects and random effects functions, respectively, and  $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \dots, \epsilon_{in_i})^T$  follows a multivariate ST distribution with degrees of freedom  $\nu_\epsilon$ , the unknown scale parameter  $\tau^2$ , and the  $n_i \times n_i$  skewness diagonal matrix  $\mathbf{\Delta}(\boldsymbol{\delta}_{\epsilon_i}) = \text{diag}(\delta_{\epsilon_{i1}}, \dots, \delta_{\epsilon_{in_i}})$  with  $n_i \times 1$  skewness parameter vector  $\boldsymbol{\delta}_{\epsilon_i} = (\delta_{\epsilon_{i1}}, \dots, \delta_{\epsilon_{in_i}})^T$ . In particular, if  $\delta_{\epsilon_{i1}} = \dots = \delta_{\epsilon_{in_i}} \hat{=} \delta_\epsilon$ , then  $\mathbf{\Delta}(\boldsymbol{\delta}_{\epsilon_i}) = \delta_\epsilon \mathbf{I}_{n_i}$  and  $\boldsymbol{\delta}_{\epsilon_i} = \delta_\epsilon \mathbf{1}_{n_i}$ .  $z_{ij}^* = w(t_{ij}) + h_i(t_{ij})$  are the true but unobservable covariate values at time  $t_{ij}$ . The fixed smooth function  $w(t)$  represents population average of the covariate process, while the random smooth function  $h_i(t)$  is introduced to incorporate the large interindividual variation in the covariate process. We assume that  $h_i(t)$  is the realization of a zero-mean stochastic process.

Nonparametric mixed-effects model (2.3) is more flexible than parametric mixed-effects models. To fit model (2.3), we apply a regression spline method to  $w(t)$  and  $h_i(t)$ . The working principle is briefly described as follows and more details can be found in the literature (Davidian and Giltinan [11] and Wu and Zhang [15]). The main idea of regression spline is to approximate  $w(t)$  and  $h_i(t)$  by using a linear combination of spline basis functions. For instance,  $w(t)$  and  $h_i(t)$  can be approximated by a linear combination of basis functions  $\boldsymbol{\Psi}_p(t) = \{\psi_0(t), \psi_1(t), \dots, \psi_{p-1}(t)\}^T$  and  $\boldsymbol{\Phi}_q(t) = \{\phi_0(t), \phi_1(t), \dots, \phi_{q-1}(t)\}^T$ , respectively. That is,

$$w(t) \approx w_p(t) = \sum_{l=0}^{p-1} \alpha_l \psi_l(t) = \boldsymbol{\Psi}_p(t)^T \boldsymbol{\alpha}, \quad h_i(t) \approx h_{iq}(t) = \sum_{l=0}^{q-1} a_{il} \phi_l(t) = \boldsymbol{\Phi}_q(t)^T \mathbf{a}_i, \quad (2.4)$$

where  $\boldsymbol{\alpha} = (\alpha_0, \dots, \alpha_{p-1})^T$  is a  $p \times 1$  vector of fixed-effects and  $\mathbf{a}_i = (a_{i0}, \dots, a_{i,q-1})^T$  ( $q \leq p$ ) is a  $q \times 1$  vector of random-effects with  $\mathbf{a}_i \text{ iid} \sim ST_{q, \nu_a}(\mathbf{0}, \boldsymbol{\Sigma}_a, \mathbf{\Delta}(\boldsymbol{\delta}_a))$  with the unrestricted covariance matrix  $\boldsymbol{\Sigma}_a$ , the skewness diagonal matrix  $\mathbf{\Delta}(\boldsymbol{\delta}_a) = \text{diag}(\delta_1^a, \dots, \delta_q^a)$ , and the degrees of freedom  $\nu_a$ . Based on the assumption of  $h_i(t)$ , we can regard  $\mathbf{a}_i$  as iid realizations of a zero-mean random vector. For our model, we consider natural cubic spline bases with the percentile-based knots. To select an optimal degree of regression spline and numbers of knots, that is, optimal sizes of  $p$  and  $q$ , the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) is often applied (Davidian and Giltinan [11] and Wu and Zhang [15]). Replacing  $w(t)$  and  $h_i(t)$  by their approximations  $w_p(t)$  and  $h_{iq}(t)$ , we can approximate model (2.3) by the following linear mixed-effects (LME) model:

$$z_{ij} \approx \boldsymbol{\Psi}_p(t_{ij})^T \boldsymbol{\alpha} + \boldsymbol{\Phi}_q(t_{ij})^T \mathbf{a}_i + \epsilon_{ij} \approx z_{ij}^* + \epsilon_{ij}, \quad \epsilon_i \text{ iid} \sim ST_{n_i, \nu_\epsilon} \left( \mathbf{0}, \tau^2 \mathbf{I}_{n_i}, \mathbf{\Delta}(\boldsymbol{\delta}_{\epsilon_i}) \right). \quad (2.5)$$

## 2.2. Simultaneous Bayesian Inference

In a longitudinal study, such as the AIDS study described previously, the longitudinal response and covariate processes are usually connected physically or biologically. Statistical

inference based on the commonly used two-step method may be undesirable since it fails to take the covariate estimation into account (Higgins et al. [1]). Although a simultaneous inference method based on a joint likelihood for the covariate and response data may be favorable, the computation associated with the joint likelihood inference in joint models of longitudinal data can be extremely intensive and may lead to convergence problems and in some cases it can even be computationally infeasible (Liu and Wu [2] and Wu [4]). Here we propose a simultaneous Bayesian inference method based on MCMC procedure for longitudinal data of response with left censoring and covariate with measurement error. The Bayesian joint modeling approach may pave a way to alleviate the computational burdens and to overcome convergence problems.

We assume that  $\mathbf{a}_i$ ,  $\mathbf{b}_i$ ,  $\boldsymbol{\epsilon}_i$ , and  $\mathbf{e}_i$  are mutually independent of each other. Following Sahu et al. [7] and properties of ST distribution, in order to specify the models (2.5) and (2.2) for MCMC computation, it can be shown that by introducing four random variable vectors  $\mathbf{w}_{\boldsymbol{\epsilon}_i} = (w_{\epsilon_{i1}}, \dots, w_{\epsilon_{im_i}})^T$ ,  $\mathbf{w}_{\mathbf{e}_i} = (w_{e_{i1}}, \dots, w_{e_{im_i}})^T$ ,  $\mathbf{w}_{\mathbf{b}_i} = (w_{b_{i1}}, \dots, w_{b_{is_3}})^T$  and  $\mathbf{w}_{\mathbf{a}_i} = (w_{a_{i1}}, \dots, w_{a_{iq}})^T$  and four random variables  $\xi_{\boldsymbol{\epsilon}_i}$ ,  $\xi_{\mathbf{e}_i}$ ,  $\xi_{\mathbf{b}_i}$ , and  $\xi_{\mathbf{a}_i}$  ( $i = 1, \dots, n$ ) based on the stochastic representation for the ST distribution (see the Appendix for details),  $z_{ij}$  and  $y_{ij}$  can be hierarchically formulated as

$$\begin{aligned}
y_{ij} \mid \mathbf{b}_i, w_{\boldsymbol{\epsilon}_i}, \xi_{\boldsymbol{\epsilon}_i}; \boldsymbol{\beta}, \sigma^2, \delta_{\boldsymbol{\epsilon}_i} &\sim N\left(g\left(t_{ij}, \mathbf{x}_{ij}, \mathbf{d}\left(z_{ij}^*, \boldsymbol{\beta}, \mathbf{b}_i\right)\right) + \delta_{\boldsymbol{\epsilon}_i} w_{\boldsymbol{\epsilon}_i}, \xi_{\boldsymbol{\epsilon}_i}^{-1} \sigma^2\right), \\
w_{\boldsymbol{\epsilon}_i} &\sim N(0, 1)I(w_{\boldsymbol{\epsilon}_i} > 0), \quad \xi_{\boldsymbol{\epsilon}_i} \mid \nu_{\boldsymbol{\epsilon}} \sim G\left(\frac{\nu_{\boldsymbol{\epsilon}}}{2}, \frac{\nu_{\boldsymbol{\epsilon}}}{2}\right), \\
\mathbf{b}_i \mid \mathbf{w}_{\mathbf{b}_i}, \xi_{\mathbf{b}_i}; \boldsymbol{\Sigma}_b, \boldsymbol{\delta}_b &\sim N_{s_3}\left(\boldsymbol{\Delta}(\boldsymbol{\delta}_b) \mathbf{w}_{\mathbf{b}_i}, \xi_{\mathbf{b}_i}^{-1} \boldsymbol{\Sigma}_b\right), \\
\mathbf{w}_{\mathbf{b}_i} &\sim N_{s_3}(\mathbf{0}, \mathbf{I}_{s_3})I(\mathbf{w}_{\mathbf{b}_i} > \mathbf{0}), \quad \xi_{\mathbf{b}_i} \mid \nu_b \sim G\left(\frac{\nu_b}{2}, \frac{\nu_b}{2}\right), \\
z_{ij} \mid \mathbf{a}_i, w_{\boldsymbol{\epsilon}_i}, \xi_{\boldsymbol{\epsilon}_i}; \boldsymbol{\alpha}, \tau^2, \delta_{\boldsymbol{\epsilon}_i} &\sim N\left(z_{ij}^* + \delta_{\boldsymbol{\epsilon}_i} w_{\boldsymbol{\epsilon}_i}, \xi_{\boldsymbol{\epsilon}_i}^{-1} \tau^2\right), \\
w_{\boldsymbol{\epsilon}_i} &\sim N(0, 1)I(w_{\boldsymbol{\epsilon}_i} > 0), \quad \xi_{\boldsymbol{\epsilon}_i} \mid \nu_{\boldsymbol{\epsilon}} \sim G\left(\frac{\nu_{\boldsymbol{\epsilon}}}{2}, \frac{\nu_{\boldsymbol{\epsilon}}}{2}\right), \\
\mathbf{a}_i \mid \mathbf{w}_{\mathbf{a}_i}, \xi_{\mathbf{a}_i}; \boldsymbol{\Sigma}_a, \boldsymbol{\delta}_a &\sim N_q\left(\boldsymbol{\Delta}(\boldsymbol{\delta}_a) \mathbf{w}_{\mathbf{a}_i}, \xi_{\mathbf{a}_i}^{-1} \boldsymbol{\Sigma}_a\right), \\
\mathbf{w}_{\mathbf{a}_i} &\sim N_q(\mathbf{0}, \mathbf{I}_q)I(\mathbf{w}_{\mathbf{a}_i} > \mathbf{0}), \quad \xi_{\mathbf{a}_i} \mid \nu_a \sim G\left(\frac{\nu_a}{2}, \frac{\nu_a}{2}\right),
\end{aligned} \tag{2.6}$$

where  $G(\cdot)$  is a gamma distribution,  $I(w_{\boldsymbol{\epsilon}_i} > 0)$  is an indicator function, and  $w_{\boldsymbol{\epsilon}_i} \sim N(0, 1)$  truncated in the space  $w_{\boldsymbol{\epsilon}_i} > 0$  (standard half-normal distribution);  $w_{\boldsymbol{\epsilon}_i}$ ,  $\mathbf{w}_{\mathbf{a}_i}$ , and  $\mathbf{w}_{\mathbf{b}_i}$  can be defined similarly.  $z_{ij}^*$  is viewed as the true but unobservable covariate values at time  $t_{ij}$ . It is noted that, as discussed in the Appendix, the hierarchical model with the ST distribution (2.6) can be reduced to the following three special cases: (i) a model with skew-normal (SN) distribution as  $\nu_{\boldsymbol{\epsilon}}, \nu_{\mathbf{e}}, \nu_b, \nu_a \rightarrow \infty$  and  $\xi_{\boldsymbol{\epsilon}_i}, \xi_{\mathbf{e}_i}, \xi_{\mathbf{b}_i}$  and  $\xi_{\mathbf{a}_i} \rightarrow 1$  with probability 1 (i.e., the four corresponding distributional specifications are omitted in (2.6)); (ii) a model with standard  $t$ -distribution as  $\delta_{\boldsymbol{\epsilon}_i} = \delta_{\mathbf{e}_i} = 0$ ,  $\boldsymbol{\delta}_b = \boldsymbol{\delta}_a = \mathbf{0}$ , and thus the four distributional specifications of  $w_{\boldsymbol{\epsilon}_i}$ ,  $w_{\mathbf{e}_i}$ ,  $\mathbf{w}_{\mathbf{a}_i}$ , and  $\mathbf{w}_{\mathbf{b}_i}$  are omitted in (2.6); (iii) a model with standard normal distribution as  $\nu_{\boldsymbol{\epsilon}}, \nu_{\mathbf{e}}, \nu_a, \nu_b \rightarrow \infty$  and  $\delta_{\boldsymbol{\epsilon}_i} = \delta_{\mathbf{e}_i} = 0$  and  $\boldsymbol{\delta}_b = \boldsymbol{\delta}_a = \mathbf{0}$ ; in this case, the eight corresponding distributional specifications are omitted in (2.6).



Let  $\theta = \{\alpha, \beta, \tau^2, \sigma^2, \Sigma_a, \Sigma_b, \nu_e, \nu_e, \nu_a, \nu_b, \delta_a, \delta_b, \delta_{e_i}, \delta_{e_i}; i = 1, \dots, n\}$  be the collection of unknown parameters in models (2.2) and (2.5). To complete the Bayesian formulation, we need to specify prior distributions for unknown parameters in the models (2.2) and (2.5) as follows:

$$\begin{aligned} \alpha &\sim N_p(\alpha_0, \Lambda_1), & \tau^2 &\sim \text{IG}(\omega_1, \omega_2), & \Sigma_a &\sim \text{IW}(\Omega_1, \rho_1), & \delta_{e_i} &\sim N_{n_i}(\mathbf{0}, \Gamma_1), \\ \beta &\sim N_{s_2}(\beta_0, \Lambda_2), & \sigma^2 &\sim \text{IG}(\omega_3, \omega_4), & \Sigma_b &\sim \text{IW}(\Omega_2, \rho_2), & \delta_{e_i} &\sim N_{n_i}(\mathbf{0}, \Gamma_2), \\ \nu_e &\sim G(\nu_{e0}, \nu_{e1})I(\nu_e > 3), & \nu_e &\sim G(\nu_{e0}, \nu_{e1})I(\nu_e > 3), & \nu_a &\sim G(\nu_{a0}, \nu_{a1})I(\nu_a > 3), \\ \nu_b &\sim G(\nu_{b0}, \nu_{b1})I(\nu_b > 3), & \delta_a &\sim N_q(\mathbf{0}, \Gamma_3), & \delta_b &\sim N_{s_3}(\mathbf{0}, \Gamma_4), \end{aligned} \quad (2.7)$$

where the mutually independent Inverse Gamma (IG), Normal (N), Gamma (G), and Inverse Wishart (IW) prior distributions are chosen to facilitate computations (Pinheiro and Bates [16]). The hyperparameter matrices  $\Lambda_1, \Lambda_2, \Omega_1, \Omega_2, \Gamma_1, \Gamma_2, \Gamma_3$ , and  $\Gamma_4$  can be assumed to be diagonal for convenient implementation.

Let  $f(\cdot | \cdot), F(\cdot | \cdot)$  and  $\pi(\cdot)$  denote a probability density function (pdf), cumulative density function (cdf), and prior density function, respectively. Conditional on the random variables and some unknown parameters, a detectable measurement  $y_{ij}$  contributes  $f(y_{ij} | \mathbf{b}_i, w_{e_{ij}}, u_{e_i})$ , whereas a nondetectable measurement contributes  $F(\rho | \mathbf{b}_i, w_{e_{ij}}, u_{e_i}) \equiv P(y_{ij} < \rho | \mathbf{b}_i, w_{e_{ij}}, u_{e_i})$  in the likelihood. We assume that  $\alpha, \beta, \tau^2, \sigma^2, \Sigma_a, \Sigma_b, \nu_e, \nu_e, \delta_{e_i}, \delta_{e_i} (i = 1, \dots, n)$  are independent of each other, that is,  $\pi(\theta) = \pi(\alpha)\pi(\beta)\pi(\tau^2)\pi(\sigma^2)\pi(\Sigma_a)\pi(\Sigma_b)\pi(\nu_e)\pi(\nu_e)\pi(\nu_a)\pi(\nu_b)\pi(\delta_a)\pi(\delta_b)\prod_i \pi(\delta_{e_i})\pi(\delta_{e_i})$ . After we specify the models for the observed data and the prior distributions for the unknown model parameters, we can make statistical inference for the parameters based on their posterior distributions under the Bayesian framework. Letting  $\mathbf{y}_i = (y_{i1}, \dots, y_{in_i})^T$  and  $\mathbf{z}_i = (z_{i1}, \dots, z_{in_i})^T$ , the joint posterior density of  $\theta$  based on the observed data can be given by

$$f(\theta | \text{data}) \propto \left\{ \prod_{i=1}^n \int \int L_{y_i} L_{z_i} L_{\mathbf{a}_i} L_{\mathbf{b}_i} d\mathbf{a}_i d\mathbf{b}_i \right\} \pi(\theta), \quad (2.8)$$

where  $L_{y_i} = \prod_{j=1}^{n_i} f(y_{ij} | \mathbf{b}_i, w_{e_{ij}}, \xi_{e_i})^{1-c_{ij}} F(\rho | \mathbf{b}_i, w_{e_{ij}}, \xi_{e_i})^{c_{ij}} f(w_{e_{ij}} | w_{e_{ij}} > 0) f(\xi_{e_i})$  is the likelihood for the observed response data,  $c_{ij}$  is the censoring indicator such that  $y_{ij}$  is observed if  $c_{ij} = 0$ , and  $y_{ij}$  is left-censored if  $c_{ij} = 1$ , that is,  $y_{ij} = q_{ij}$  if  $c_{ij} = 0$ , and  $y_{ij}$  is treated as missing if  $c_{ij} = 1$ , and  $L_{z_i} = \prod_{j=1}^{n_i} f(z_{ij} | \mathbf{a}_i, w_{e_{ij}}, \xi_{e_i}) f(w_{e_{ij}} | w_{e_{ij}} > 0) f(\xi_{e_i})$  is the likelihood for the observed covariate data  $\{\mathbf{z}_i, i = 1, \dots, n\}$ ,  $L_{\mathbf{b}_i} = f(\mathbf{b}_i | \mathbf{w}_{b_i}, \xi_{b_i}) f(\mathbf{w}_{b_i} | \mathbf{w}_{b_i} > 0) f(\xi_{b_i})$ , and  $L_{\mathbf{a}_i} = f(\mathbf{a}_i | \mathbf{w}_{a_i}, \xi_{a_i}) f(\mathbf{w}_{a_i} | \mathbf{w}_{a_i} > 0) f(\xi_{a_i})$ .

In general, the integrals in (2.8) are of high dimension and do not have closed form solutions. Therefore, it is prohibitive to directly calculate the posterior distribution of  $\theta$  based on the observed data. As an alternative, MCMC procedures can be used to sample based on (2.8) using the Gibbs sampler along with the Metropolis-Hasting (M-H) algorithm. An important advantage of the above representations based on the hierarchical models (2.6) and (2.7) is that they can be very easily implemented using the freely available WinBUGS software (Lunn et al. [17]) and that the computational effort is equivalent to the one necessary to fit the normal version of the model. Note that when using WinBUGS to implement our modeling approach, it is not necessary to explicitly specify the full conditional distributions. Thus we omit those here to save space.

### 3. Data Analysis

#### 3.1. Specification of Models

We now analyze the data set described in Section 1 based on the proposed method. Among the 44 eligible patients, the number of viral load measurements for each patient varies from 4 to 9 measurements. As is evident from Figures 1(c) and 1(d), the interpatient variations in viral load appear to be large and these variations appear to change over time. Previous studies suggest that the interpatient variation in viral load may be partially explained by time-varying CD4 cell count (Wu [4] and Huang et al. [18]).

Models for covariate processes are needed in order to incorporate measurement errors in covariates. CD4 cell counts often have nonnegligible measurement errors, and ignoring these errors can lead to severely misleading results in a statistical inference (Carroll et al. [14]). In A5055 study, roughly 10 per cent of the CD4 measurement times are inconsistent with the viral load measurement times. Consequently, CD4 measurements may be missed at viral load measurement times mainly due to a different CD4 measurement scheme as designed in the study (e.g., CD4 measurements were missed at day 7 as displayed in Figures 1(c) and 1(d)). There seem to be no particular patterns for the missingness. Thus we assume that the missing data in CD4 are missing at random (MAR) in the sense of Rubin [19], so that the missing data mechanism can be ignored in the analysis. With CD4 measures collected over time from the AIDS study, we may model the CD4 process to partially address the measurement errors (Wu [4]). However, the CD4 trajectories are often complicated, and there is no well-established model for the CD4 process. We, thus, model the CD4 process empirically using a nonparametric mixed-effects model, which is flexible and works well for complex longitudinal data. We use linear combinations of natural cubic splines with percentile-based knots to approximate  $w(t)$  and  $h_i(t)$ . Following the study in (Liu and Wu [2]), we set  $\psi_0(t) = \phi_0(t) = 1$  and take the same natural cubic splines in the approximations (2.4) with  $q \leq p$  (in order to limit the dimension of random-effects). The values of  $p$  and  $q$  are determined based on the AIC/BIC criteria. The AIC/BIC values are evaluated for various models with  $(p, q) = \{(1, 1), (2, 1), (2, 2), (3, 1), (3, 2), (3, 3)\}$  which was found that the model with  $(p, q) = (3, 3)$  has the smallest AIC/BIC values being 703.6/744.4. We thus adopted the following ST nonparametric mixed-effects CD4 covariate model:

$$z_{ij} = (\alpha_0 + a_{i0}) + (\alpha_1 + a_{i1})\psi_1(t_{ij}) + (\alpha_2 + a_{i2})\psi_2(t_{ij}) + \epsilon_{ij} \left( \equiv z_{ij}^* + \epsilon_{ij} \right), \quad (3.1)$$

where  $z_{ij}$  is the observed CD4 value at time  $t_{ij}$ ,  $\psi_1(\cdot)$  and  $\psi_2(\cdot)$  are two basis functions given in Section 2.1 and taking the same natural cubic splines for  $\phi(\cdot)$ ,  $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2)^T$  is a vector of population parameters (fixed-effects),  $\mathbf{a}_i = (a_{i0}, a_{i1}, a_{i2})^T$  is a vector of random-effects, and  $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \dots, \epsilon_{in_i})^T \sim ST_{n_i, \nu_\epsilon}(\mathbf{0}, \tau^2 \mathbf{I}_{n_i}, \delta_\epsilon \mathbf{I}_{n_i})$ . In addition, in order to avoid too small or large estimates which may be unstable, we standardize the time-varying covariate CD4 cell counts (each CD4 value is subtracted by mean 375.46 and divided by standard deviation 228.57) and rescale the original time (in days) so that the time scale is between 0 and 1.

For the initial stage of viral decay after treatment, a biologically reasonable viral load model can be formulated by the uniexponential form (Ho et al. [20]),  $V(t) = V(0) \exp(-\lambda t)$ , where  $V(t)$  is the total virus at time  $t$  and  $\lambda$  is the rate of change in viral load. To model the complete viral load trajectory, one possible extension of the model given above is to allow  $\lambda$  to vary over time. A simple determinant for time-varying  $\lambda$  is the linear function  $\lambda(t) = a + bt$ .



For HIV viral dynamic models, it is typical to take ln-transformation of the viral load in order to stabilize the variance and to speed up estimation algorithm (Ding and Wu [10]). After ln-transformation of  $V(t)$ , substituting  $\lambda$  by the linear function  $\lambda(t) = a + bt$ , we obtain the following quadratic linear mixed-effects model:

$$y_{ij} = \beta_{i0} + \beta_{ij1}t_{ij} + \beta_{ij2}t_{ij}^2 + e_{ij}, \quad (3.2)$$

where  $y_{ij} = \ln(V_i(t_{ij}))$ , parameter  $\beta_{i0}$  represents the initial viral load in ln scale, and parameters  $\beta_{ij1}$  and  $\beta_{ij2}$  incorporate change in viral decay rate over time, with  $\lambda_{ij} \equiv -(\beta_{ij1} + \beta_{ij2}t_{ij})$  being the time-varying exponential decay rate.  $\mathbf{e}_i = (e_{i1}, \dots, e_{in_i})^T \sim \text{ST}_{n_i, \nu_e}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i}, \delta_e \mathbf{I}_{n_i})$ ;  $\boldsymbol{\beta}_{ij} = (\beta_{ij0}, \beta_{ij1}, \beta_{ij2})^T$  is a vector of individual parameters for the  $i$ th subject at time  $t_{ij}$ .

Since CD4 cell counts are measured with errors, we assume that the individual-specific and time-varying parameters  $\boldsymbol{\beta}_{ij}$  are related to the summary of true CD4 values  $z_{ij}^*$ , which may be interpreted as the “regularized” CD4 covariate value. As discussed by Wu [21], to determine whether CD4 values influence the dynamic parameters  $\boldsymbol{\beta}_{ij}$ , AIC/BIC criteria are used again as guidance (Pinheiro and Bates [16]) to find the following model

$$\beta_{i0} = \beta_1 + b_{i1}, \quad \beta_{ij1} = \beta_2 + \beta_3 z_{ij}^* + b_{i2}, \quad \beta_{ij2} = \beta_4 + \beta_5 z_{ij}^* + b_{i3}, \quad (3.3)$$

where  $\mathbf{b}_i = (b_{i1}, \dots, b_{i3})^T$  is individual random-effect, and  $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_5)^T$  is a vector of population parameters. The model (3.3) indicates that the current (regularized) CD4 values  $z_{ij}^*$  rather than the past (observed) CD4 values  $z_{ij}$  are most predictive of the change in viral load at time  $t_{ij}$ . One possible explanation is that, since CD4 measurements for each individual are often sparse, the current CD4 value may be the best summary of immediate past CD4 values, while the early CD4 values may not be very predictive of the current change in viral load.

### 3.2. Model Implementation

In this section, we analyze the AIDS data set described in Section 1 to illustrate the proposed joint modeling method (denoted by JM) based on the joint models (3.2) in conjunction with the covariate model (3.1) and the corresponding specifications of prior distributions. As shown in Figures 1(a) and 1(b), the histograms of viral load in ln scale and CD4 cell count clearly indicate their asymmetric nature and it seems logical to fit the joint model with a skew distribution to the data set. Along with this consideration, the following statistical models with different distributions of both model errors and random-effects for both the response model (3.2) and the covariate model (3.1) are employed to compare their performance.

- (i) **SN Model:**  $\mathbf{e}_i$ ,  $\boldsymbol{\epsilon}_i$ ,  $\mathbf{b}_i$ , and  $\mathbf{a}_i$  follow an SN-distribution.
- (ii) **ST Model:**  $\mathbf{e}_i$ ,  $\boldsymbol{\epsilon}_i$ ,  $\mathbf{b}_i$ , and  $\mathbf{a}_i$  follow an ST-distribution.
- (iii) **N Model:**  $\mathbf{e}_i$ ,  $\boldsymbol{\epsilon}_i$ ,  $\mathbf{b}_i$ , and  $\mathbf{a}_i$  follow a normal ( $N$ ) distribution.

We investigate the following *three scenarios*. First, since a normal distribution is a special case of an SN distribution when skewness parameter is zero, while the ST distribution reduces to the SN distribution when the degree of freedom approaches infinity, we investigate how an asymmetric (SN or ST) distribution contributes to modeling results and parameter estimation in comparison with a symmetric (normal) distribution. Second, we estimate the

model parameters by using the “naive” method (denoted by NV), which ignores measurement errors in CD4, and missing responses are imputed by the half (i.e.,  $\ln(25)$ ) of the BDL. That is, the “naive” method only uses the observed CD4 values  $z_{ij}$  rather than true (unobservable) CD4 values  $z_{ij}^*$  in the response model (3.2) and the missing data in the Tobit model (2.1) is imputed by  $\ln(25)$ . We use it as a comparison to the JM proposed in Section 2. This comparison attempts to investigate how the measurement errors in CD4 and missing data in viral load together contribute to modeling results. Third, when covariates are measured with errors, a common approach is the so-called two-step (TS) method (Higgins et al. [1]): the first step estimates the “true” covariate values based on the covariate model (3.1); at the second step the covariate in the response model (2.6) is substituted by the estimate from the first step. Thus we use the two-step (TS) method to assess the performance of the JM method.

The progress in the Bayesian posterior computation due to MCMC procedures has made it possible to fit increasingly complex statistical models (Lunn et al. [17] and Huang et al. [18]). To choose the best model among candidate models, it has become more important to develop efficient model selection criteria. A recent publication by Spiegelhalter et al. [22] suggested a generalization of AIC called deviance information criterion (DIC). Since the structure of DIC allows for an automatic computation in WinBUGS, we use DIC to compare the models in this paper. As with other model selection criteria, we caution that DIC is not intended for identification of the “correct” model, but rather merely as a method of comparing a collection of alternative formulations. In our models with different distribution specifications for model errors, DIC can be used to find out how assumption of a skew-normal distribution contributes to virologic response in comparison with that of a normal distribution and how the proposed joint modeling approach influences parameter estimation compared with the “naive” method and imputation method.

To carry out the Bayesian inference, we need to specify the values of the hyperparameters in the prior distributions. In the Bayesian approach, we only need to specify the priors at the population level. The values of the hyperparameters were mostly chosen from previous studies in the literature (Liu and Wu [2], Huang and Dagne [5], Sahu et al. [7], Wu [21], and among others). We take weakly informative prior distribution for the parameters in the models. In particular, (i) fixed-effects were taken to be independent normal distribution  $N(0, 100)$  for each component of the population parameter vectors  $\alpha$  and  $\beta$ . (ii) For the scale parameters  $\tau^2$  and  $\sigma^2$ , we assume a limiting noninformative inverse gamma prior distribution,  $IG(0.01, 0.01)$  so that the distribution has mean 1 and variance 100. (iii) The priors for the variance-covariance matrices of the random-effects  $\Sigma_a$  and  $\Sigma_b$  are taken to be inverse Wishart distributions  $IW(\Omega_1, \rho_1)$  and  $IW(\Omega_2, \rho_2)$  with covariance matrices  $\Omega_1 = \Omega_2 = \text{diag}(0.01, 0.01, 0.01)$  and degrees of freedom  $\rho_1 = \rho_2 = 5$ , respectively. (iv) The degrees of freedom parameters  $\nu_e, \nu_e, \nu_a,$  and  $\nu_b$  follow a truncated gamma distribution with two hyperparameter values being 1 and 0.1, respectively. (v) For each of the skewness parameters  $\delta_e, \delta_e, \delta_k^a,$  and  $\delta_k^b$  ( $k = 1, 2, 3$ ), we choose independent normal distribution  $N(0, 100)$ , where we assume that  $\delta_{e_i} = \delta_e \mathbf{1}_{n_i}$  and  $\delta_{e_i} = \delta_e \mathbf{1}_{n_i}$  to indicate that we are interested in skewness of overall viral load data and overall CD4 cell count data. The MCMC sampler was implemented using WinBUGS software, and the program codes are available from authors on request. The convergence of MCMC implementation was assessed using standard tools (such as trace plots which are not shown here to save space) within WinBUGS, and convergence was achieved after initial 50,000 burn-in iterations. After convergence diagnostics was done, one long chain of 200,000 iterations, retaining every 20th, was run to obtain 10,000 samples for Bayesian inference. Next, we report analysis results of the three scenarios proposed above.

### 3.3. Comparison of Joint Modeling Results

The population posterior mean (PM), the corresponding standard deviation (SD), and 95% credible interval for fixed-effects parameters based on the three models (SN, ST, and  $N$ ) for JM method are presented in the upper part of Table 1. The significant findings are presented as follows. (i) For the response model (3.2), where the most substantively interesting parameters are  $(\beta_2, \beta_3, \beta_4, \beta_5)$ , the estimates of  $\beta_2$  and  $\beta_4$ , the linear coefficient and quadratic coefficient of time, respectively, under the three models, are significant since the 95% credible intervals do not contain zero. Among the coefficients of the true CD4 covariate  $(\beta_3, \beta_5)$  in model (3.3), the posterior means of  $\beta_5$  are significantly different from zero for all the three models under JM method. Moreover, the posterior mean values for  $\beta_5$  are quite different between models SN ( $-4.76$ ), ST ( $-6.31$ ), and  $N$  ( $-6.26$ ), implying that the posterior means may be substantially biased if model distribution ignores skewness. We will see later that SN gives better fit than either ST or  $N$ . In addition, for the scale parameter  $\sigma^2$ , the posterior mean value (2.63) in  $N$  model is much larger than that of any other corresponding posterior means in SN and ST models. (ii) For parameter estimates of the CD4 covariate model (3.1), the posterior means of intercept  $\alpha_0$  and coefficient  $\alpha_1$  based on SN and ST models are significant, while the posterior mean of  $\alpha_2$  turns out to be nonsignificant under all the three models. For the scale parameter  $\tau^2$  of the covariate model, the posterior mean value (0.13) is the largest under  $N$  model. This is expected since the model based on ordinary normal distribution does not account for skewness and heaviness in tails for the type of data analyzed here.

To assess the goodness-of-fit of the proposed JM method, the diagnosis plots for the SN, ST, and  $N$  models comparing the residuals and the fitted values (Figures 2(a)–2(c)) and the observed values versus the fitted values (Figures 2(d)–2(f)). The distribution of the residuals for SN model looks tighter than those for either ST model or  $N$  model, showing a better fit. Similar results are observed by looking at the plots in Figures 2(d)–2(f). The plot for SN model has most of the points close the line showing a strong agreement between the observed and the fitted values. Clearly, it can be seen from the plots that  $N$  model, which ignores skewness, does not fit the data very well as compared to either SN model or ST model. Note that the horizontal line designates the below detection limit (BDL), which is at  $\ln(50)$ . The recorded observations less than BDL are not accurate and, therefore, have not been used in the analysis, but instead they were treated as missing and predicted values are obtained. These predicted values are plotted against the recorded observations below detection limit as shown in the lower-row plots. In general, from the model fitting results, both SN and ST models provide a reasonably good fit to the observed data even though SN model is slightly better than ST model.

In order to further investigate whether SN model under JM method can provide better fit to the data than ST model, the DIC values are obtained and found to be 863.0 for SN model and 985.6 for ST model. The DIC value for SN model is smaller than that of ST model, confirming that SN model is better than ST model in fitting the proposed joint model. As mentioned before, it is hard sometimes to tell which model is “correct” but which one fits data better. The model which fits the data better may be more appealing in order to describe the mechanism of HIV infection and CD4 changing process. Thus, based on the DIC criterion, the results indicate that SN model is relatively better than either ST model or  $N$  model. These findings are consistent with those displayed in the goodness-of-fit in Figure 2 indicating that SN model outperforms both ST model and  $N$  model. In summary, our results suggest that it is very important to assume an SN distribution for the response Tobit model and the CD4 covariate model in order to achieve reliable results, in particular if the data exhibit skewness,

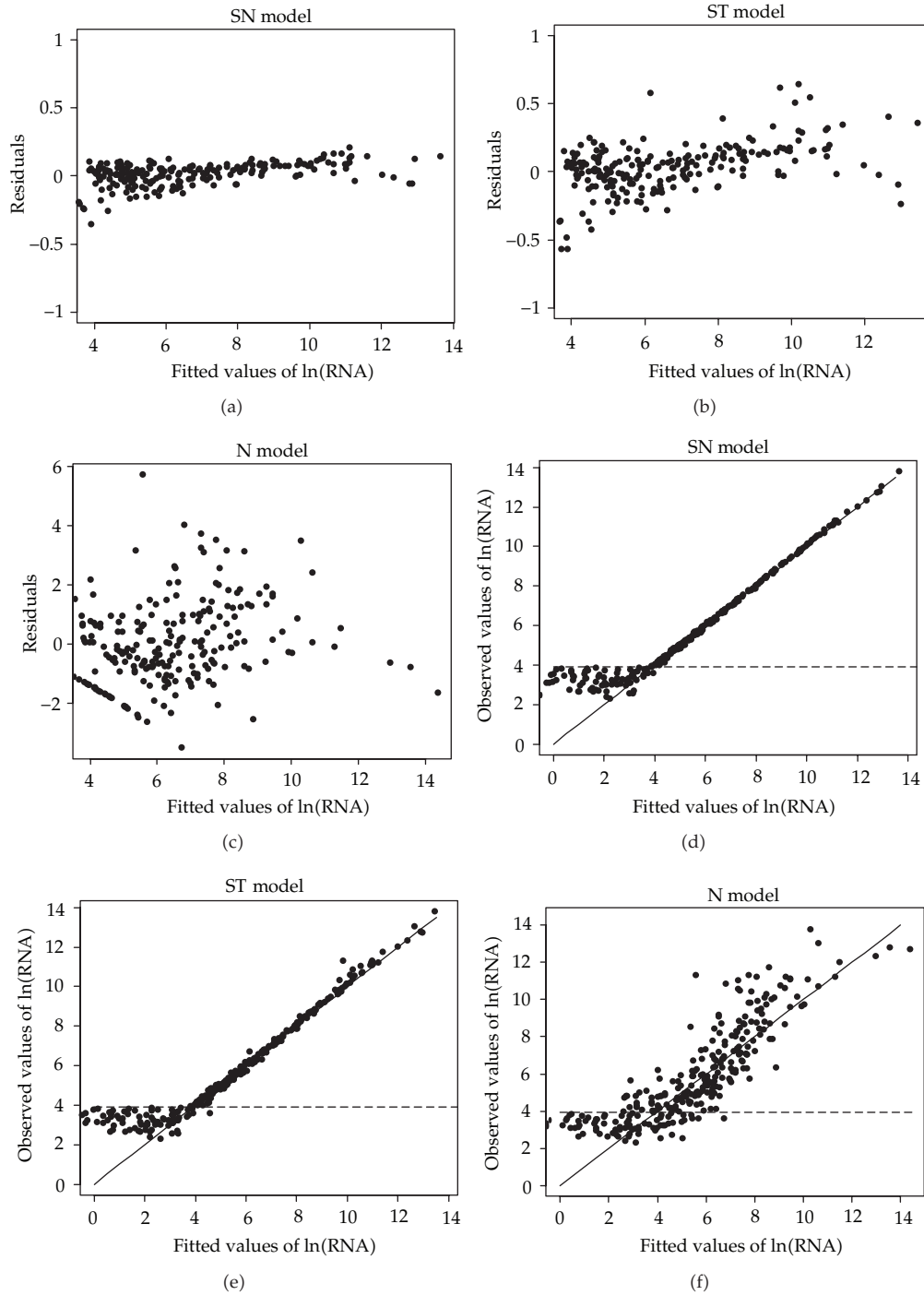
**Table 1:** A summary of the estimated posterior mean (PM) of population (fixed-effects) and scale parameters, the corresponding standard deviation (SD) and lower limit ( $L_{CI}$ ) and upper limit ( $U_{CI}$ ) of 95% equal-tail credible interval (CI) as well as DIC values based on the joint modeling (JM), the naive (NV), and the two-step (TS) methods.

Method	Model		$\alpha_0$	$\alpha_1$	$\alpha_2$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\tau^2$	$\sigma^2$	DIC	
JM	SN	PM	-0.95	0.15	-0.23	5.62	-14.6	-2.34	11.7	-4.76	0.07	0.14	863.0	
		$L_{CI}$	-1.58	0.06	-15.2	4.17	-22.1	-5.14	4.52	-9.92	0.04	0.01		
		$U_{CI}$	-0.01	0.90	14.8	7.59	-8.14	1.44	21.7	-0.62	0.11	0.64		
	SD	0.47	0.37	7.63	0.96	3.98	1.65	5.25	2.34	0.02	0.18			
	ST	PM	-0.94	0.34	-0.31	5.84	-12.0	-1.20	8.12	-6.31	0.04	0.21		985.6
		$L_{CI}$	-1.41	0.18	-14.1	4.15	-16.5	-5.72	2.20	-12.6	0.02	0.01		
		$U_{CI}$	-0.06	0.88	13.4	8.02	-7.72	2.72	19.2	-1.41	0.05	0.86		
	N	SD	0.35	0.26	7.09	1.10	2.22	2.22	4.14	2.77	0.01	0.26		
		PM	-0.21	0.45	-2.87	7.74	-15.4	-0.80	13.6	-6.26	0.13	2.63		1242.3
$L_{CI}$		-0.46	0.22	-15.9	7.20	-18.3	-4.16	9.97	-11.7	0.11	2.06			
$U_{CI}$	0.04	0.68	9.90	8.29	-12.6	2.53	17.2	-1.43	0.16	3.35				
NV	SN	SD	0.13	0.12	6.54	0.28	1.48	1.73	1.85	2.61	0.01	0.33		
		PM	—	—	—	5.03	-11.1	0.58	6.83	-2.10	—	0.10	1083.5	
		$L_{CI}$	—	—	—	3.82	-13.6	-0.94	4.52	-4.18	—	0.01		
		$U_{CI}$	—	—	—	6.59	-8.73	2.08	9.18	0.07	—	0.35		
SD	—	—	—	0.75	1.23	0.77	1.19	1.04	—	0.09				
TS	SN	PM	-0.99	0.19	2.71	5.91	-14.4	-1.24	8.47	-5.90	0.09	0.14	1023.8	
		$L_{CI}$	-1.58	-0.43	-12.1	4.12	-22.1	-5.01	1.83	-10.6	0.05	0.01		
		$U_{CI}$	0.07	0.90	17.1	7.72	-8.50	2.16	21.2	-0.80	0.14	0.65		
		SD	0.42	0.36	7.54	1.05	3.88	1.79	5.14	2.52	0.02	0.18		

but not heaviness in the tails. Along with these observations, next we provide detailed fitting results and interpretations based on the SN Model.

### 3.4. Estimation Results Based on SN Model

For comparison, we used the “naive” (NV) method to estimate the model parameters presented in the lower part of Table 1 where the raw (observed) CD4 values  $z_{ij}$  rather than the true (unobserved) CD4 values  $z_{ij}^*$  are substituted in the response model (3.3). It can be seen that there are important differences in the posterior means for the parameters  $\beta_3$  and  $\beta_5$ , which are coefficients of CD4 covariate. These posterior means are  $\hat{\beta}_3 = 0.58$  and  $\hat{\beta}_5 = -2.10$  for the NV method, and  $\hat{\beta}_3 = -2.34$  and  $\hat{\beta}_5 = -4.76$  for the JM method. The NV method may produce biased posterior means and may substantially overestimate the covariate CD4 effect. The estimated standard deviations (SD) for the CD4 effect ( $\beta_3$  and  $\beta_5$ ) using the JM method are 1.65 and 2.34, which are approximately twice as large as those (0.77 and 1.04) using the NV method, respectively, probably because the JM method incorporates the variation from fitting the CD4 process. The differences of the NV estimates and the JM estimates suggest that the estimated parameters may be substantially biased if measurement errors in CD4 covariate are ignored. We also obtained DIC value of 1083.5 for the NV method, while the DIC value for the JM method is 863.0. We can see from the estimated DIC values that the JM approach provides



**Figure 2:** The goodness-of-fit. (a–c): Residuals versus fitted values of  $\ln(\text{RNA})$  under skew-normal (SN), skew-t (ST), and normal (N) models based on the JM method; the values below detection limit ( $\ln(50)$ ) are not included in the plots since there are no corresponding residuals but only predicted values. (d–f): Observed values versus fitted values of  $\ln(\text{RNA})$  under SN, ST, and N models, where the horizontal line at  $\ln(50)$  represents the detection limit.

**Table 2:** A summary of the estimated posterior mean (PM) of skewness and degree of freedom parameters, the corresponding standard deviation (SD), and lower limit ( $L_{CI}$ ) and upper limit ( $U_{CI}$ ) of 95% equal-tail credible interval (CI) based on the joint modeling (JM), the naive (NV), and the two-step (TS) methods.

Method	Model		$\delta_\epsilon$	$\delta_e$	$\delta_1^a$	$\delta_2^a$	$\delta_3^a$	$\delta_1^b$	$\delta_2^b$	$\delta_3^b$	$\nu_\epsilon$	$\nu_e$	$\nu_a$	$\nu_b$	
JM	SN	PM	0.41	2.34	0.58	0.34	0.26	0.52	-1.81	2.58	—	—	—	—	
		$L_{CI}$	0.25	1.93	-0.62	-0.58	-16.1	-1.90	-10.1	-10.7	—	—	—	—	
		$U_{CI}$	0.54	2.73	1.41	1.09	17.0	2.31	7.12	11.0	—	—	—	—	
	ST	SD	0.07	0.20	0.60	0.47	8.57	1.18	5.30	6.93	—	—	—	—	
		PM	0.05	2.26	0.87	0.04	-0.56	0.32	-4.70	6.45	3.32	10.2	14.0	14.4	
		$L_{CI}$	-0.14	1.59	-0.35	-0.60	-16.5	-2.23	-10.1	-8.23	3.01	3.07	3.52	3.52	
			$U_{CI}$	0.25	2.70	1.54	0.65	14.7	2.37	1.91	12.3	4.18	35.2	41.1	41.9
			SD	0.11	0.33	0.50	0.32	8.41	1.31	2.93	4.88	0.32	8.98	10.3	10.3
	NV	SN	PM	—	2.24	—	—	—	0.80	0.15	5.53	—	—	—	—
$L_{CI}$			—	1.95	—	—	—	-1.05	-1.71	3.62	—	—	—	—	
$U_{CI}$			—	2.55	—	—	—	2.25	2.30	7.74	—	—	—	—	
			SD	—	0.15	—	—	—	0.92	1.00	1.06	—	—	—	—
TS		SN	PM	0.16	2.44	0.89	0.28	3.06	0.04	-0.94	5.18	—	—	—	—
			$L_{CI}$	-0.39	2.07	-0.48	-0.58	-11.7	-2.20	-8.53	-12.4	—	—	—	—
	$U_{CI}$		0.51	2.79	1.55	1.04	21.0	2.23	7.49	12.2	—	—	—	—	
		SD	0.29	0.18	0.50	0.45	8.30	1.31	4.75	6.67	—	—	—	—	

a better fit to the data in comparison with the NV method. Thus it is important to take the measurement errors into account when covariates are measured with errors.

Comparing the JM method against the two-step (TS) method from the lower part of Table 1, we can see that the TS estimates and the JM estimates are somewhat different. In particular, there are important differences in the posterior means for the parameters  $\beta_4$  and  $\beta_5$  which is directly associated CD4 covariate. For the parameter  $\beta_5$ , the posterior means are -4.76 (95% CI = (-9.92, -0.62)) and -5.90 (95% CI = (-10.60, -0.80)) for the JM and TS methods, respectively. The TS method slightly underestimates the effect of CD4 covariate.

The estimated results based on the JM method for SN model in Table 2 presents the estimated skewness parameters, and the only significant skewness parameters are those for the response model errors and CD4 covariate model errors, but not random-effects. These are  $\hat{\delta}_e = 2.34$  (95% CI = (1.93, 2.73)) and  $\hat{\delta}_\epsilon = 0.41$  (95% CI = (0.25, 0.54)) for viral load and CD4 cell count, respectively. They are significantly positive confirming the right-skewed viral load and CD4 cell count as was depicted in Figure 1. Thus, the results suggest that accounting for significant skewness, when the data exhibit skewness, provides a better model fit to the data and gives more accurate estimates to the parameters.

In summary, the results indicate that the SN model under the JM method is a better suited model for viral loads and CD4 covariate with measurement errors. Looking now at the estimated population initial stage of viral decay after treatment bases on the JM method, we get  $\hat{\lambda}(t) = -(-14.6 - 2.34z^*(t) + 11.7t - 4.76z^*(t)t)$ , where  $z^*(t)$  is the standardized true CD4 value at time  $t$  which may be interpreted as the “regularized” covariate value. Thus, the population viral load process may be approximated by  $\hat{V}(t) = \exp[5.62 - \hat{\lambda}(t)t]$ . Since the viral decay rate ( $\lambda(t)$ ) is significantly associated with the true CD4 values (due to statistically significant estimate of  $\beta_5$ ), this suggests that the viral load change  $V(t)$  may be significantly associated with the true CD4 process. Note that, although the true association described



above may be complicated, the simple approximation considered here may provide a reasonable guidance and suggest a further research.

#### 4. Discussion

Attempts to jointly fit the viral load data and CD4 cell counts with measurement errors are compromised by left censoring in viral load response due to detection limits. We addressed this problem using Bayesian nonlinear mixed-effects Tobit models with skew distributions. The models were fitted based on the assumption that the viral dynamic model (2.2) continues to hold for those unobserved left-censored viral loads. This assumption may be reasonable since the dynamic model considered here is a natural extension of a biologically justified model (Ding and Wu [10]). Even though left censoring effects are the focus of this paper, right-censoring (ceiling) effects can also be dealt with in very similar ways. It is therefore important for researchers to pay attention to censoring effects in a longitudinal data analysis, and Bayesian Tobit models with skew distributions make best use of both censored and uncensored data information.

Our results suggest that both ST (skew- $t$ ) and SN (skew-normal) models show superiority to the  $N$  (normal) model. Our results also indicate that the JM method outperformed the NV and TS methods in the sense that it produces more accurate parameter estimates. The JM method is quite general and so can be applied to other application areas, allowing accurate inferences of parameters while adjusting for skewness, left-censoring, and measurement errors. In short, skew distributions show potentials to gain efficiency and accuracy in estimating certain parameters when the normality assumption does not hold in the data.

The proposed NLME Tobit joint model with skew distributions can be easily fitted using MCMC procedure by using the WinBUGS package that is available publicly and has a computational cost similar to the normal version of the model due to the features of its hierarchically stochastic representations. Implementation via MCMC makes it straightforward to compare the proposed models and methods with various scenarios for real data analysis in comparison with symmetric distributions and asymmetric distributions for model errors. This makes our approach quite powerful and also accessible to practicing statisticians in the fields. In order to examine the sensitivity of parameter estimates to the prior distributions and initial values, we also conducted a limited sensitivity analysis using different values of hyperparameters of prior distributions and different initial values (data not shown). The results of the sensitivity analysis showed that the estimated dynamic parameters were not sensitive to changes of both priors and initial values. Thus, the final results are reasonable and robust, and the conclusions of our analysis remain unchanged (see Huang et al. [18] for more details).

The methods of this paper may be extended to accommodate various subpopulations of patients whose viral decay trajectories after treatment may differ. In addition, the purpose of this paper is to demonstrate the proposed models and methods with various scenarios for real data analysis for comparing asymmetric distributions for model errors to a symmetric distribution, although a limited simulation study might have been conducted to evaluate our results from different model specifications and the corresponding methods. However, since this paper investigated many different scenarios-based models and methods with real data analysis, the complex natures considered, especially skew distributions involved, will pose some challenges for such a simulation study which requires additional efforts, and it is beyond the purpose of this paper. We are currently investigating these related problems and will report the findings in the near future.

## Appendix

### A. Multivariate Skew- $t$ and Skew Normal Distributions

Different versions of the multivariate skew-elliptical (SE) distributions have been proposed and used in the literature (Sahu et al. [7], Azzalini and Capitanio [23], Jara et al. [24], Arellano-Valle et al. [25], and among others). We adopt a class of multivariate SE distributions proposed by Sahu et al. [7], which is obtained by using transformation and conditioning and contains multivariate skew- $t$  (ST) and skew-normal (SN) distributions as special cases. A  $k$ -dimensional random vector  $\mathbf{Y}$  follows a  $k$ -variate SE distribution if its probability density function (pdf) is given by

$$f(\mathbf{y} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}(\boldsymbol{\delta}); m_{\nu}^{(k)}) = 2^k f(\mathbf{y} \mid \boldsymbol{\mu}, \mathbf{A}; m_{\nu}^{(k)}) P(\mathbf{V} > \mathbf{0}), \quad (\text{A.1})$$

where  $\mathbf{A} = \boldsymbol{\Sigma} + \boldsymbol{\Delta}^2(\boldsymbol{\delta})$ ,  $\boldsymbol{\mu}$  is a location parameter vector,  $\boldsymbol{\Sigma}$  is a covariance matrix,  $\boldsymbol{\Delta}(\boldsymbol{\delta})$  is a skewness diagonal matrix with the skewness parameter vector  $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_k)^T$ ;  $\mathbf{V}$  follows the elliptical distribution  $El(\boldsymbol{\Delta}(\boldsymbol{\delta})\mathbf{A}^{-1}(\mathbf{y} - \boldsymbol{\mu}), \mathbf{I}_k - \boldsymbol{\Delta}(\boldsymbol{\delta})\mathbf{A}^{-1}\boldsymbol{\Delta}(\boldsymbol{\delta}); m_{\nu}^{(k)})$  and the density generator function  $m_{\nu}^{(k)}(u) = (\Gamma(k/2)/\pi^{k/2})(m_{\nu}(u)/\int_0^{\infty} r^{k/2-1}m_{\nu}(u)dr)$ , with  $m_{\nu}(u)$  being a function such that  $\int_0^{\infty} r^{k/2-1}m_{\nu}(u)dr$  exists. The function  $m_{\nu}(u)$  provides the kernel of the original elliptical density and may depend on the parameter  $\nu$ . We denote this SE distribution by  $SE(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}(\boldsymbol{\delta}); m^{(k)})$ . Two examples of  $m_{\nu}(u)$ , leading to important special cases used throughout the paper, are  $m_{\nu}(u) = \exp(-u/2)$  and  $m_{\nu}(u) = (u/\nu)^{-(\nu+k)/2}$ , where  $\nu > 0$ . These two expressions lead to the multivariate ST and SN distributions, respectively. In the latter case,  $\nu$  corresponds to the degree of freedom parameter.

Since the SN distribution is a special case of the ST distribution when the degree of freedom approaches infinity, for completeness, this section is started by discussing the multivariate ST distribution that will be used in defining the ST joint models considered in this paper. For detailed discussions on properties and differences among various versions of ST and SN distributions, see the references above. We consider a multivariate ST distribution introduced by Sahu et al. [7], which is suitable for a Bayesian inference since it is built using conditional method and is defined below.

An  $k$ -dimensional random vector  $\mathbf{Y}$  follows an  $k$ -variate ST distribution if its pdf is given by

$$f(\mathbf{y} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}(\boldsymbol{\delta}), \nu) = 2^k t_{k,\nu}(\mathbf{y} \mid \boldsymbol{\mu}, \mathbf{A}) P(\mathbf{V} > \mathbf{0}). \quad (\text{A.2})$$

We denote the  $k$ -variate  $t$  distribution with parameters  $\boldsymbol{\mu}, \mathbf{A}$  and degrees of freedom  $\nu$  by  $t_{k,\nu}(\boldsymbol{\mu}, \mathbf{A})$  and the corresponding pdf by  $t_{k,\nu}(\mathbf{y} \mid \boldsymbol{\mu}, \mathbf{A})$  henceforth,  $\mathbf{V}$  follows the  $t$  distribution  $t_{k,\nu+k}$ . We denote this distribution by  $ST_{k,\nu}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}(\boldsymbol{\delta}))$ . In particular, when  $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}_k$  and  $\boldsymbol{\Delta}(\boldsymbol{\delta}) = \delta\mathbf{I}_k$ , (A.2) simplifies to

$$\begin{aligned} f(\mathbf{y} \mid \boldsymbol{\mu}, \sigma^2, \delta, \nu) &= 2^k (\sigma^2 + \delta^2)^{-k/2} \frac{\Gamma((\nu + m)/2)}{\Gamma(\nu/2)(\nu\pi)^{k/2}} \left\{ 1 + \frac{(\mathbf{y} - \boldsymbol{\mu})^T (\mathbf{y} - \boldsymbol{\mu})}{\nu(\sigma^2 + \delta^2)} \right\}^{-(\nu+k)/2} \\ &\times T_{k,\nu+k} \left[ \left\{ \frac{\nu + (\sigma^2 + \delta^2)^{-1} (\mathbf{y} - \boldsymbol{\mu})^T (\mathbf{y} - \boldsymbol{\mu})}{\nu + k} \right\}^{-1/2} \frac{\delta(\mathbf{y} - \boldsymbol{\mu})}{\sigma\sqrt{\sigma^2 + \delta^2}} \right], \end{aligned} \quad (\text{A.3})$$

where  $T_{k,\nu+k}(\cdot)$  denotes the cumulative distribution function (cdf) of  $t_{k,\nu+k}(\mathbf{0}, \mathbf{I}_k)$ . However, unlike in the SN distribution to be discussed below, the ST density cannot be written as the product of univariate ST densities. Here  $\mathbf{Y}$  are dependent but uncorrelated.

The mean and covariance matrix of the ST distribution  $ST_{k,\nu}(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k, \boldsymbol{\Delta}(\boldsymbol{\delta}))$  are given by

$$\begin{aligned} E(\mathbf{Y}) &= \boldsymbol{\mu} + \left(\frac{\nu}{\pi}\right)^{1/2} \frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)} \boldsymbol{\delta}, \\ \text{cov}(\mathbf{Y}) &= \left[\sigma^2 \mathbf{I}_k + \boldsymbol{\Delta}^2(\boldsymbol{\delta})\right] \frac{\nu}{\nu-2} - \frac{\nu}{\pi} \left[\frac{\Gamma\{(\nu-1)/2\}}{\Gamma(\nu/2)}\right]^2 \boldsymbol{\Delta}^2(\boldsymbol{\delta}). \end{aligned} \quad (\text{A.4})$$

According to Lemma 1 of Azzalini and Capitanio [23], if  $\mathbf{Y}$  follows  $ST_{k,\nu}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}(\boldsymbol{\delta}))$ , it can be represented by

$$\mathbf{Y} = \boldsymbol{\mu} + \xi^{-1/2} \mathbf{X}, \quad (\text{A.5})$$

where  $\xi$  follows a Gamma distribution  $\Gamma(\nu/2, \nu/2)$ , which is independent of  $\mathbf{X}$ , and  $\mathbf{X}$  follows a  $k$ -dimensional skew-normal (SN) distribution, denoted by  $SN_k(\mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}(\boldsymbol{\delta}))$ . It follows from (A.5) that  $\mathbf{Y} \mid \xi \sim SN_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}/\xi, \boldsymbol{\Delta}(\boldsymbol{\delta}))$ . By Proposition 1 of Arellano-Valle et al. [25], the SN distribution of  $\mathbf{Y}$  conditional on  $\xi$  has a convenient stochastic representation as follows:

$$\mathbf{Y} = \boldsymbol{\mu} + \boldsymbol{\Delta}(\boldsymbol{\delta}) |X_0| + \xi^{-1/2} \boldsymbol{\Sigma}^{1/2} X_1, \quad (\text{A.6})$$

where  $X_0$  and  $X_1$  are two independent  $N_k(\mathbf{0}, \mathbf{I}_k)$  random vectors. Note that the expression (A.6) provides a convenience device for random number generation and for implementation purpose. Let  $\mathbf{w} = |X_0|$ ; then  $\mathbf{w}$  follows a  $k$ -dimensional standard normal distribution  $N_k(\mathbf{0}, \mathbf{I}_m)$  truncated in the space  $\mathbf{w} > \mathbf{0}$  (i.e., the standard half-normal distribution). Thus, following Sahu et al. [7], a hierarchical representation of (A.6) is given by

$$\mathbf{Y} \mid \mathbf{w}, \xi \sim N_k\left(\boldsymbol{\mu} + \boldsymbol{\Delta}(\boldsymbol{\delta}) \mathbf{w}, \xi^{-1} \boldsymbol{\Sigma}\right), \quad \mathbf{w} \sim N_k(\mathbf{0}, \mathbf{I}_k) \mathbf{I}(\mathbf{w} > \mathbf{0}), \quad \xi \sim G\left(\frac{\nu}{2}, \frac{\nu}{2}\right), \quad (\text{A.7})$$

where  $G(\cdot)$  is a gamma distribution. Note that the ST distribution presented in (A.7) can be reduced to the following three special cases: (i) as  $\nu \rightarrow \infty$  and  $\xi \rightarrow 1$  with probability 1 (i.e., the last distributional specification is omitted), then the hierarchical expression (A.7) becomes an SN distribution  $SN_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}(\boldsymbol{\delta}))$ ; (ii) as  $\boldsymbol{\Delta}(\boldsymbol{\delta}) = \mathbf{0}$ , then the hierarchical expression (A.7) is a standard multivariate  $t$ -distribution; (iii) as  $\nu \rightarrow \infty$ ,  $\xi \rightarrow 1$  with probability 1, and  $\boldsymbol{\Delta}(\boldsymbol{\delta}) = \mathbf{0}$ , then the hierarchical expression (A.7) is a standard multivariate normal distribution.

Specifically, if a  $k$ -dimensional random vector  $\mathbf{Y}$  follows a  $k$ -variate SN distribution, then (A.2)–(A.4) revert to the following forms, respectively:

$$f(\mathbf{y} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}(\boldsymbol{\delta})) = 2^k |\mathbf{A}|^{-1/2} \phi_k \left\{ \mathbf{A}^{-1/2} (\mathbf{y} - \boldsymbol{\mu}) \right\} P(\mathbf{V} > \mathbf{0}), \quad (\text{A.8})$$

where  $\mathbf{V} \sim N_k\{\boldsymbol{\Delta}(\boldsymbol{\delta}) \mathbf{A}^{-1} (\mathbf{y} - \boldsymbol{\mu}), \mathbf{I}_k - \boldsymbol{\Delta}(\boldsymbol{\delta}) \mathbf{A}^{-1} \boldsymbol{\Delta}(\boldsymbol{\delta})\}$ , and  $\phi_k(\cdot)$  is the pdf of  $N_k(\mathbf{0}, \mathbf{I}_k)$ . We denote the above distribution by  $SN_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}(\boldsymbol{\delta}))$ . An appealing feature of (A.8) is that it

gives independent marginal when  $\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2)$ . The pdf (A.8) thus simplifies to

$$f(\mathbf{y} | \boldsymbol{\mu}, \Sigma, \Delta(\boldsymbol{\delta})) = \prod_{i=1}^k \left[ \frac{2}{\sqrt{\sigma_i^2 + \delta_i^2}} \phi \left\{ \frac{y_i - \mu_i}{\sqrt{\sigma_i^2 + \delta_i^2}} \right\} \Phi \left\{ \frac{\delta_i}{\sigma_i} \frac{y_i - \mu_i}{\sqrt{\sigma_i^2 + \delta_i^2}} \right\} \right], \quad (\text{A.9})$$

where  $\phi(\cdot)$  and  $\Phi(\cdot)$  are the pdf and cdf of the standard normal distribution, respectively. The mean and covariance matrix are given by  $E(\mathbf{Y}) = \boldsymbol{\mu} + \sqrt{2/\pi} \boldsymbol{\delta}$ ,  $\text{cov}(\mathbf{Y}) = \Sigma + (1 - 2/\pi) \Delta(\boldsymbol{\delta})^2$ . It is noted that when  $\boldsymbol{\delta} = \mathbf{0}$ , the SN distribution reduces to usual normal distribution.

## Acknowledgments

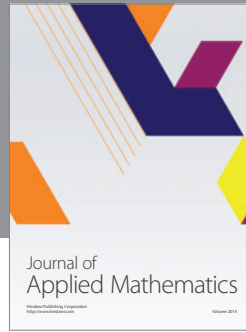
The authors are grateful to the Guest Editor and three reviewers for their insightful comments and suggestions that led to a marked improvement of the paper. They gratefully acknowledge A5055 study investigators for allowing them to use the clinical data from their study. This research was partially supported by NIAID/NIH Grant R03 AI080338 and MSP/NSA Grant H98230-09-1-0053 to Y. Huang.

## References

- [1] M. Higgins, M. Davidian, and D. M. Giltinan, "A two-step approach to measurement error in time-dependent covariates in nonlinear mixed-effects models, with application to IGF-I pharmacokinetics," *Journal of the American Statistical Association*, vol. 92, pp. 436–448, 1997.
- [2] W. Liu and L. Wu, "Simultaneous inference for semiparametric nonlinear mixed-effects models with covariate measurement errors and missing responses," *Biometrics*, vol. 63, no. 2, pp. 342–350, 2007.
- [3] M. S. Wulfsohn and A. A. Tsiatis, "A joint model for survival and longitudinal data measured with error," *Biometrics*, vol. 53, no. 1, pp. 330–339, 1997.
- [4] L. Wu, "A joint model for nonlinear mixed-effects models with censoring and covariants measured with error, with application to AIDS studies," *Journal of the American Statistical Association*, vol. 97, no. 460, pp. 955–964, 2002.
- [5] Y. Huang and G. Dagne, "A Bayesian approach to joint mixed-effects models with a skew-normal distribution and measurement errors in covariates," *Biometrics*, vol. 67, no. 1, pp. 260–269, 2011.
- [6] G. Verbeke and E. Lesaffre, "A linear mixed-effects model with heterogeneity in the random-effects population," *Journal of the American Statistical Association*, vol. 91, no. 433, pp. 217–221, 1996.
- [7] S. K. Sahu, D. K. Dey, and M. D. Branco, "A new class of multivariate skew distributions with applications to Bayesian regression models," *The Canadian Journal of Statistics*, vol. 31, no. 2, pp. 129–150, 2003.
- [8] E. P. Acosta, H. Wu, S. M. Hammer et al., "Comparison of two indinavir/ritonavir regimens in the treatment of HIV-infected individuals," *Journal of Acquired Immune Deficiency Syndromes*, vol. 37, no. 3, pp. 1358–1366, 2004.
- [9] A. S. Perelson, P. Essunger, Y. Cao et al., "Decay characteristics of HIV-1-infected compartments during combination therapy," *Nature*, vol. 387, no. 6629, pp. 188–191, 1997.
- [10] A. A. Ding and H. Wu, "Relationships between antiviral treatment effects and biphasic viral decay rates in modeling HIV dynamics," *Mathematical Biosciences*, vol. 160, no. 1, pp. 63–82, 1999.
- [11] M. Davidian and D. M. Giltinan, *Nonlinear Models for Repeated Measurement Data*, Chapman & Hall, London, UK, 1995.
- [12] J. P. Hughes, "Mixed effects models with censored data with application to HIV RNA levels," *Biometrics*, vol. 55, no. 2, pp. 625–629, 1999.
- [13] J. Tobin, "Estimation of relationships for limited dependent variables," *Econometrica*, vol. 26, pp. 24–36, 1958.
- [14] R. J. Carroll, D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu, *Measurement Error in Nonlinear Models: A Modern Perspective*, Chapman & Hall, Boca Raton, Fla, USA, 2nd edition, 2006.

- [15] H. Wu and J. T. Zhang, "The study of long-term HIV dynamics using semi-parametric non-linear mixed-effects models," *Statistics in Medicine*, vol. 21, no. 23, pp. 3655–3675, 2002.
- [16] J. C. Pinheiro and M. D. Bates, *Mixed-Effects Models in S and S-PLUS*, Springer, New York, NY, USA, 2000.
- [17] D. J. Lunn, A. Thomas, N. Best, and D. Spiegelhalter, "WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility," *Statistics and Computing*, vol. 10, no. 4, pp. 325–337, 2000.
- [18] Y. Huang, D. Liu, and H. Wu, "Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system," *Biometrics*, vol. 62, no. 2, pp. 413–423, 2006.
- [19] D. B. Rubin, "Inference and missing data," *Biometrika*, vol. 63, no. 3, pp. 581–592, 1976.
- [20] D. D. Ho, A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz, "Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection," *Nature*, vol. 373, no. 6510, pp. 123–126, 1995.
- [21] L. Wu, "Simultaneous inference for longitudinal data with detection limits and covariates measured with errors, with application to AIDS studies," *Statistics in Medicine*, vol. 23, no. 11, pp. 1715–1731, 2004.
- [22] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. Van der Linde, "Bayesian measures of model complexity and fit," *Journal of the Royal Statistical Society. Series B*, vol. 64, no. 4, pp. 583–639, 2002.
- [23] A. Azzalini and A. Capitanio, "Distributions generated by perturbation of symmetry with emphasis on a multivariate skew  $t$ -distribution," *Journal of the Royal Statistical Society. Series B*, vol. 65, no. 2, pp. 367–389, 2003.
- [24] A. Jara, F. Quintana, and E. San Martín, "Linear mixed models with skew-elliptical distributions: a Bayesian approach," *Computational Statistics & Data Analysis*, vol. 52, no. 11, pp. 5033–5045, 2008.
- [25] R. B. Arellano-Valle, H. Bolfarine, and V. H. Lachos, "Bayesian inference for skew-normal linear mixed models," *Journal of Applied Statistics*, vol. 34, no. 5-6, pp. 663–682, 2007.





# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

