Bayesian Approach to Nonlinear Mixed-Effects Quantile Regression Models for Longitudinal Data with Non-normality and Left-censoring

Yangxin Huang[1*], Jiaqing Chen[2] and Xiaosun Lu[3]

[1]School of Mathematics and Computer, Wuhan Textile University, Wuhan 430073, P.R.China and College of Public Health, University of South Florida, Tampa, Florida 33612, USA
[2]Department of Statistics, Wuhan University of Technology, Wuhan, Hubei, 430070, P.R.China
[3]Department of Biostatistics, Medpace Inc., Cincinnati, OH 45227, USA
Email: hyx123@hotmail.com

Abstract In longitudinal studies, measurements of the same individuals are taken repeatedly through time, but it often happens that some collected data are observed with the following issues. (i) Often, the primary goal is to characterize the change in response over time. Compared with conventional mean regression, quantile regression (QR) can characterize the entire conditional distribution of the outcome variable, and may be more robust to outliers and mis-specification of error distribution. (ii) longitudinal outcomes may suffer from a serious departure of normality in which normality assumption may cause lack of robustness and subsequently lead to invalid inference; and (iii) the response observations may be subject to left-censoring due to a limit of detection. Inferential procedures will become very complicated when one analyzes data with these features together. In this article, Bayesian modeling approach to nonlinear mixed-effects quantile regression models for longitudinal data is developed to study simultaneous impact of multiple data features (non-normality, left-censoring, non-linearity, outliers and heavy-tails). Simulation studies are conducted to assess the performance of the proposed models and methods. A real data example is analyzed to demonstrate the proposed methodology through comparing potential models with different distribution specifications of random-effects.

Keywords: Asymmetric Laplace distribution, Bayesian inference, left-censoring, nonlinear mixed-effects quantile regression, skew-normal distribution.

1 Introduction

Longitudinal data analysis has attracted a considerable research interest and a large number of statistical modeling and analysis methods have been suggested for analyzing such data with various features. Researchers may often confront the task of developing inference from data where longitudinal outcomes of interest may exhibit departure from normality, appear non-central location due to outliers and/or heavy-tails, and suffer from measure with left-censoring due to a limit of detection (LOD).

The data set that motivated this research is from an AIDS clinical trial group (ACTG398) study [1], which is a randomized, double-blind, placebo-controlled, with an extension to more than 48 weeks of study of saquinavir, indinavir, or nelfinavir added as second protease inhibitor to the 4-drug class regimen in patients with virologic failure defined by receiving saquinavir, nelfinavir, indinavir, or ritonavir. This study consists of 481 HIV-1 infected patients. The plasma HIV-1 RNA (viral load) is repeatedly quantified at weeks 0, 2, 4, 8, 16, and every 8 weeks until the last patient on study. The number of viral load measurements for each individual varies from 2 to 13. Out of total 481 patients, 379 patients who had more than 2 measurements were included in data analysis. A $\log_{10}$ transformation of viral load was used in the analysis in order to reduce the variation of the measurements and, in turn, to stabilize the estimation algorithm and to speed up convergence of the algorithm. The CD4 cell counts were also measured throughout study on a similar scheme. About 21% of viral load observations were measured below LOD, which is $\log_{10}(50) = 1.699$.

As we know, the change in viral load is an important indicator of HIV disease progression after an application of antiretroviral treatment (ART) [1,2,3]. The viral load measurements are usually observed

with features of non-normality, non-linearity, non-central location and LOD as shown in Figure 1. Modeling such data appears to be challenging due to the following issues of inherent features of longitudinal data. First, in the literature, most studies of longitudinal modeling assume that model error and/or random-effects in mixed-effects models follow normal distribution due to mathematical tractability and computational convenience [3,4,5,6,7,8]. This requires the variables to be "symmetrically" distributed. A violation of this assumption could lead to biased or misleading results [9,10,11,12]. In fact, observed viral load data are often far from being "symmetric" and asymmetric patterns of observations usually occur; it can be seen from Figure 1(a) that, for this data set to be analyzed in this article, the viral load response (even after $\log_{10}$ transformation) appears a skewed feature. Alternatively, an asymmetric distribution such as skew-normal (SN) distribution [9,12,13,14] should be more appropriate than a symmetric (normal) distribution to model such data set. Second, the outcome of a longitudinal study may be subject to LOD because of the low sensitivity of current standard assays. For example, in ACTG398 study, which was designed to collect data on every individual at each assessment, the response (viral load) measurements are subject to left censoring due to an LOD (50 copies/mL) for which 21% of viral load observations were measured below LOD. Thus, the proportion of data censored may not be trivial and failure to account for censoring due to LOD in the analysis may result in significant biases [5]. *Ad hoc* procedures are sometimes used to adjust for the censoring. For instance, when observations fall below LOD, a common practice is to impute the censored values by either the LOD or some arbitrary value, such as LOD/2 [3,7]. In this paper, instead of arbitrarily imputing the observations below LOD, we use fully Bayesian predictive distributions by incorporating a Tobit model [15] to treat the observations below LOD as left-censoring. That is, Tobit model is introduced to treat the inaccurate measures below LOD as missing values of a latent variable. Third, the majority of longitudinal modeling methods in the literature [9,10,11,5,6,3,7,4,8,16] and among others focuses mainly on mean regression to concentrate on the average effect of response variable conditional on covariates, which may lead to non-robust parameter estimates if obvious heavy tails exist as shown in Figure 1(b). Quantile regression (QR) models have become increasingly popular since the seminal work of [17]. In contrast to the mean regression model, QR model belongs to a robust model family, which can give an overall assessment of the covariate effects at different quantiles of the outcome [18]. Unlike conventional mean regression models, which address solely the conditional mean or the central effects of the covariates, QR models quantify the entire conditional distribution of the outcome variable, and are more robust to misspecification of error distribution [18,19,20] since a set of quantile often provides more complete description of the response distribution than the mean and, thus, leads to a robust inference by allowing estimation of parameters at different quantiles of conditional distribution and conveys a more comprehensive understanding of data. In particular, we can model the lower or higher quantiles of the outcome to provide a natural assessment of covariate effects specific for those regression quantiles. Therefore, QR may offer a practically important alternative to classic mean regression [19,21]. Although various linear mixed-effects models-based QR has been widely considered via different methods [21,20,22,23,24,25,26,27,28,29], these studies considered modeling methods either based on linear QR or under complete data setting. However, in most longitudinal studies, response trajectories may exhibit nonlinear patterns (see Figure 1(c)) and data observed are often far from "perfe" or "complete" (for example, left-censoring due to LOD displayed in Figure 1(c)). To the best of our knowledge, very limited research has been conducted for longitudinal studies based on nonlinear QR to simultaneously account for data with non-normality, non-linearity, left-censoring due to LOD and non-central location due to outliers and/or heavy-tails. It is not clear how these data features may interact and simultaneously influence inferential procedure based on nonlinear QR. The goal of this article is to propose nonlinear mixed-effects quantile regression (NLMEQR) models to simultaneously account for response with those features in the longitudinal data under the framework of Bayesian inferential approach. It is noted that, in the NLMEQR models, the following three components are incorporated to capture those data features discussed above: (i) nonlinear QR-based models using the asymmetric Laplace distribution (ALD) [30], which are robust with respect to outliers/heavy-tails and provide a natural way to deal with the Bayesian QR problem as discussed by [31]; (ii) Tobit model [15], which treats the observations below LOD as left-censoring; and (iii) an SN distribution [12] for random-effects (latent random variable), which offers an appealing robust alternative to the usual symmetric normal distribution and is equipped with flexibility in capturing the effects of skewness (non-normality). We use an SN distribution for random-effects to develop the NLMEQR modeling approach, as the normal distribution is a special case of the SN distribution when

skewness is zero. In what follows, we consider multivariate SN distribution introduced by [12], which is suitable for Bayesian inference [9,16,11,12]. The rest of this paper is organized as follows. Notation and model setup are introduced, and associated inferential methods that simultaneously account for response with non-normality, left-censoring, outliers and/or heavy-tails are discussed in Section 2. In Section 3, we discuss the specific NLMEQR models for HIV dynamics, demonstrate to apply the proposed methodologies to AIDS data set and report the analysis results. Section 4 conducts simulation studies aimed to evaluate the performance of the proposed models and method. Finally, general concluding discussion is included in Section 5.
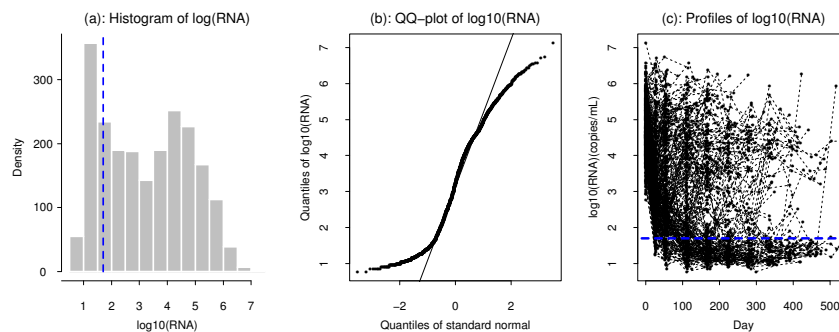

Figure 1. The histogram, Q-Q plot and trajectories of viral load measured from RNA levels in plasma (in $\log_{10}$ scale) for 379 subjects in an AIDS clinical trial study. The horizontal and vertical dashed-lines at $1.699 = \log_{10}(50)$ are LOD.

2   Quantile Regression-based Models and Bayesian Inferential Procedure

In this section, we present NLMEQR models and associated inferential method in a general form, illustrating that our models and method may be applicable in other applications. Specific models in application to AIDS study will be depicted in Section 3.1.

2.1   Nonlinear Quantile Regression for Independent Data

Let y_i and $\boldsymbol{x}_i$ denote the outcome of interest and the corresponding covariate vector for subject i $(i = 1, \ldots, n)$, where y_i is independent scalar observations of a continuous random variable with common cumulative distribution function (cdf) $F_{y_i}(\cdot)$. The τth nonlinear QR model for the response y_i given $\boldsymbol{x}_i$ takes the form of

$$Q_{y_i}(\tau|\boldsymbol{x}_i) = g(\boldsymbol{x}_i, \boldsymbol{\beta}), \tag{1}$$

where $Q_{y_i}(\tau|\boldsymbol{x}_i)(0 < \tau < 1)$ denotes the τ-quantile of the conditional distribution of $y_i|\boldsymbol{x}_i$, and $g(\cdot)$ is a nonlinear known function. The nonlinear regression coefficient vector $\boldsymbol{\beta}$ is estimated by minimizing

$$\sum_{i=1}^{n} \rho_\tau \left(y_i - g(\boldsymbol{x}_i, \boldsymbol{\beta}) \right), \tag{2}$$

where $\rho_\tau(\cdot)$ is the check function defined by $\rho_\tau(u) = u(\tau - I(u < 0))$ and $I(\cdot)$ denotes the indictor function. In order to highlight the τ-distributional dependency, the parameter vector $\boldsymbol{\beta}$ should be indexed by τ (i.e., $\boldsymbol{\beta}(\tau)$). For sake of simplicity, however, we will omit this notation in the remainder of the paper. The check function is closely related to the ALD; see [19,31,32,33] in detail. The density function of an ALD, denoted by $\mathrm{ALD}(\mu, \sigma, \tau)$, is given by

$$f(y_i|\mu, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp\left\{ -\rho_\tau \left(\frac{y_i - \mu}{\sigma} \right) \right\}, \tag{3}$$

where $\mu = g(\boldsymbol{x}_i, \boldsymbol{\beta})$. Considering σ is a nuisance parameter, it can be easily shown that the minimization of equation (2) with respect to the parameter $\boldsymbol{\beta}$ is exactly equivalent to the maximization of a likelihood function of y_i by assuming y_i from an $\mathrm{ALD}(\mu, \sigma, \tau)$.

The relationship between the check function and ALD can be used to reformulate the QR method in the likelihood framework. By utilizing this property, under independent data setting, [31] developed Bayesian QR, [19] and [33] studied the Bayesian estimation procedure for the Tobit QR model with censored data, and [32] proposed a likelihood-based goodness-of-fit test for QR. More recently, QR-based linear mixed-effects models have been considered via different methods for longitudinal data [20,21,23,24,25,26,27,28,29].

2.2 Quantile Regression-based Models for Longitudinal Data

We present NLMEQR models in a general form in this section. Denote the number of subjects by n and the number of measurements on the ith subject by n_i. Let y_{ij} be the value of response for the individual i at time t_{ij} $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, n_i)$. In order to introduce Tobit model to deal with observations below LOD in our modeling framework, denote the observed value y_{ij} by (q_{ij}, c_{ij}), where c_{ij} is the censoring indicator and q_{ij} is the latent response variable. The latent q_{ij} is observed, as y_{ij}, if and only if $y_{ij} > d$ (a known constant LOD). When q_{ij} is observed we have $c_{ij} = 0$. Otherwise we have $c_{ij} = 1$ and $q_{ij} \le d$ must be estimated. We consider time-varying covariate vectors $\boldsymbol{x}_{ij}$ and $\boldsymbol{z}_{ij}$ for the individual i at time t_{ij} that may share common covariates. Thus, the observed data are denoted by $\Re = \{(q_{ij}, c_{ij}, \boldsymbol{x}_{ij}, \boldsymbol{z}_{ij}), i = 1, \ldots, n; j = 1, 2, \ldots, n_i\}$.

For the response process with left-censoring, we consider the following NLMEQR model

$$
\begin{aligned}
q_{ij} &= g(t_{ij}, \boldsymbol{x}_{ij}, \boldsymbol{\beta}_{ij}) + e_{ij}, \quad Q_{e_{ij}}(\tau | \boldsymbol{x}_{ij}, \boldsymbol{z}_{ij}, \boldsymbol{b}_i) = 0, \\
\boldsymbol{\beta}_{ij} &= \boldsymbol{d}(\boldsymbol{z}_{ij}, \boldsymbol{\beta}, \boldsymbol{b}_i), \quad \boldsymbol{b}_i \overset{\text{iid}}{\sim} SN_{s_3}\left(-\sqrt{2/\pi}\boldsymbol{\delta}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}\right),
\end{aligned}
\tag{4}
$$

where $g(\cdot)$ is a nonlinear known function, $\boldsymbol{d}(\cdot)$ is an s_1-dimensional vector-valued linear function, $\boldsymbol{\beta}_{ij}$ is an $s_1 \times 1$ individual-specific parameter vector, $\boldsymbol{\beta}$ is an $s_2 \times 1$ population parameter vector $(s_2 \ge s_1)$; the error term e_{ij} is assumed to be independently distributed as $\mathrm{ALD}(0, \sigma, \tau)$ whose distribution is restricted to have the τth quantile equal to zero with the scale parameter $\sigma(>0)$, the $s_3 \times 1$ random-effects $\boldsymbol{b}_i$ $(s_3 \le s_1)$ follows a multivariate SN distribution with the covariance matrix $\boldsymbol{\Sigma}$, the $s_3 \times s_3$ skewness diagonal matrix $\boldsymbol{\Delta} = \mathrm{diag}(\delta_1, \ldots, \delta_{s_3})$ and $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_{s_3})^T$; see the specific model (9) for detailed descriptions of determining s_1, s_2 and s_3 where $s_1 = s_3 = 4$ and $s_2 = 6$. In order to have a zero mean vector for the random-effects $\boldsymbol{b}_i$, we assume the location parameter $\boldsymbol{\mu} = -\sqrt{2/\pi}\boldsymbol{\delta}$; see [11] and [12] in detail. The proposed NLMEQR model (4) along with the Tobit model provides a generalization and extension of commonly adopted QR models for longitudinal data in the literature [20,23,24,25,21,26,27,28,29] by the following facts that (i) linear QR-based models are extended to nonlinear QR formulation; (ii) Tobit model is introduced to deal with the inaccurate observations below LOD as left-censoring (or missing values) which are predicted using fully Bayesian predictive distributions; and (iii) random-effects are assumed to follow an SN distribution which offers a robust alternative to the usual symmetric normal distribution, a special case of the SN distribution when $\boldsymbol{\Delta} = \boldsymbol{0}$.

2.3 Bayesian Inferential Approach

Although the likelihood inferential methods based on the proposed NLMEQR models for the longitudinal response data with non-linearity, non-normality, left-censoring, outliers and/or heavy-tails may be favorable, the computation associated with the likelihood inferential approach, such as Markov chain expectation-maximization (MCEM) algorithm and bootstrap methods to estimate the parameters and associated variance-covariances, can be extremely intensive and may lead to convergence problems and in some cases it can even be computationally infeasible due to complex data features [5,6,7]. Instead, we proceed with our estimation using Bayesian inferential method based on MCMC procedure, which automatically yields the posterior variance estimates based on the posterior samples of the model parameters.

We assume that $\boldsymbol{e}_i = (e_{i1}, \ldots, e_{in_i})^T$ and $\boldsymbol{b}_i$ are independent of each other. In order to specify model (4) for MCMC computation, it can be shown that by introducing the random vector $\boldsymbol{w}_i = (w_{i1}, \ldots, w_{is_3})^T$

based on the stochastic representation for the SN distribution [9,12] and random variable v_{ij} based on the a hierarchical mixture of exponential and normal distributions [19,34], y_{ij} with random-effects $\boldsymbol{b}_i$ in the presence of left-censoring due to LOD can be hierarchically formulated as

$$
\begin{aligned}
q_{ij}|\boldsymbol{b}_i, v_{ij} &\sim N\left(g(t_{ij}, \boldsymbol{x}_{ij}, \boldsymbol{d}(\boldsymbol{z}_{ij}, \boldsymbol{\beta}, \boldsymbol{b}_i)) + \vartheta_1 v_{ij}, \vartheta_2 \sigma v_{ij}\right), \\
\boldsymbol{b}_i|\boldsymbol{w}_i &\sim N_{s_3}\left(\boldsymbol{\Delta}[\boldsymbol{w}_i - \sqrt{2/\pi}\boldsymbol{1}_{s_3}], \boldsymbol{\Sigma}\right), \\
v_{ij} &\sim Exp(\tfrac{1}{\sigma}), \quad \boldsymbol{w}_i \sim N_{s_3}(\boldsymbol{0}, \boldsymbol{I}_{s_3}) I(\boldsymbol{w}_i > \boldsymbol{0}),
\end{aligned}
\tag{5}
$$

where $\boldsymbol{1}_{s_3} = (1, \ldots, 1)^T$, $\vartheta_1 = (1 - 2\tau)/[\tau(1 - \tau)]$ and $\vartheta_2 = 2/[\tau(1 - \tau)]$, $I(\boldsymbol{w}_i > \boldsymbol{0})$ is an indicator function and $\boldsymbol{w}_i \sim N_{s_3}(\boldsymbol{0}, \boldsymbol{I}_{s_3})$ truncated in the space $\boldsymbol{w}_i > \boldsymbol{0}$ (i.e., standard half-normal distribution).

Let $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \sigma, \boldsymbol{\delta}, \boldsymbol{\Sigma}\}$ be the collection of unknown population parameters in model (4). To complete the Bayesian formulation, we need to specify prior distributions for all the unknown parameters as follows.

$$
\boldsymbol{\beta} \sim N_{s_2}(\boldsymbol{\beta}_0, \boldsymbol{\Omega}_1), \ \sigma \sim IG(\omega_1, \omega_2), \ \boldsymbol{\Sigma} \sim IW(\boldsymbol{\Omega}_2, \eta), \ \boldsymbol{\delta} \sim N_{s_3}(\boldsymbol{0}, \boldsymbol{\Omega}_3),
\tag{6}
$$

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

Let $f(\cdot|\cdot)$, $F(\cdot|\cdot)$ and $h(\cdot)$ 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 $q_{ij} = y_{ij}$ contributes $f(q_{ij}|\boldsymbol{b}_i, v_{ij})$, whereas a non-detectable measurement contributes $F(d|\boldsymbol{b}_i, v_{ij}) \equiv P(q_{ij} < d|\boldsymbol{b}_i, v_{ij})$ in the likelihood. Assume that the elements of $\boldsymbol{\theta}$ are independent of each other; that is, $h(\boldsymbol{\theta}) = h(\boldsymbol{\beta})h(\sigma)h(\boldsymbol{\delta})h(\boldsymbol{\Sigma})$. 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 Bayesian framework. The joint posterior density of $\boldsymbol{\theta}$ based on the observed data $\Re = \{(q_{ij}, c_{ij}, \boldsymbol{x}_{ij}, \boldsymbol{z}_{ij}), i = 1, \ldots, n; j = 1, 2, \ldots, n_i\}$ can be given by

$$
f(\boldsymbol{\theta}|\Re) \propto \left\{ \prod_{i=1}^{n} \int \prod_{j=1}^{n_i} f(q_{ij}|\boldsymbol{b}_i, v_{ij})^{1-c_{ij}} F(d|\boldsymbol{b}_i, v_{ij})^{c_{ij}} f(\boldsymbol{b}_i|\boldsymbol{w}_i) f(v_{ij}) f(\boldsymbol{w}_i|\boldsymbol{w}_i > \boldsymbol{0}) d\boldsymbol{b}_i \right\} h(\boldsymbol{\theta}),
\tag{7}
$$

where $L_i = \prod_{j=1}^{n_i} f(q_{ij}|\boldsymbol{b}_i, v_{ij})^{1-c_{ij}} F(d|\boldsymbol{b}_i, v_{ij})^{c_{ij}}$ is the likelihood for the observed data $\{(q_{ij}, c_{ij}), i = 1, \ldots, n; j = 1, 2, \ldots, n_i\}$ and other quantities.

In general, the integrals in (7) are of high dimension and do not have a closed form. Analytic approximations to the integrals may not be sufficiently accurate. Therefore, it is prohibitive to directly calculate the posterior distribution of $\boldsymbol{\theta}$ based on the observed data. As an alternative, the MCMC procedure can be used to sample population parameters $\boldsymbol{\theta}$, and random-effects $\boldsymbol{b}_i$ ($i = 1, ..., n$) from conditional posterior distributions based on (7), using the Gibbs sampler along with the Metropolis-Hastings (M-H) algorithm. This process is repeated in iterations of MCMC procedure until convergence is reached. When the MCMC procedure was applied to the actual clinical data, convergence of the generated samples was assessed using standard tools within WinBUGS software such as trace plots and Gelman-Rubin (GR) diagnostics[35]. An important advantage of the hierarchical representation (5) above, which facilitates us to express the proposed NLMEQR models as a commonly used normal model structure, is that they are easily implemented using the freely available WinBUGS software [36] interacted with a function called bugs in a package R2WinBUGS of R. Another advantage is that when WinBUGS software is used to implement our modeling approach, it is not necessary to explicitly specify the full conditional posterior distributions for parameters to be estimated. Although their derivations are straightforward by working the complete joint posterior (7), some cumbersome algebra will be involved. We, thus, omit those here to save space.

3 Analysis of AIDS Data

3.1 Specification of Models for HIV Dynamics

Viral dynamic models can be formulated through a system of ordinary differential equations (ODE) [2,3,37,38]. Under some reasonable assumptions and simplifications, a useful approximation of ODE

solution, which can be used to capture viral load response, has been proposed as follows.

$$y(t) = \log_{10}(e^{p_1 - \lambda_1 t} + e^{p_2 - \lambda_2 t}), \tag{8}$$

where $y(t)$ is the $\log_{10}$ scaled plasma HIV-1 RNA levels at time t. λ_1 and λ_2 are called the first- and second-phase viral decay rates, which may represent the minimum turnover rate of productively infected cells and that of latently or long-lived infected cells, respectively [2]. The parameters p_1 and p_2 are macro-parameters, and $e^{p_1} + e^{p_2}$ is the baseline viral load at time $t = 0$. It is generally assumed that $\lambda_1 > \lambda_2$, which assures that the model is identifiable and is appropriate for empirical studies [3]. It was noted that equation (8) can be only applied to the early segment or longer term of the viral load response with decreasing trajectory patterns [3,39]. However, we noted in Figure 1(c) that for some patients viral loads increase at the later period, indicating that the second-phase viral decay rate λ_2 may vary over time and negative values of the second-phase decay rate may correspond to viral increase and lead to viral rebound [3]; it suggests that variation in the dynamic parameters, particularly λ_2, may be partially associated with time-varying covariates such as repeated CD4 cell counts. Thus, it may not be reasonable to assume that second-phase viral decay rate λ_2 is a constant when viral load is rebounded at the later stage during long-term treatment. To model the long-term HIV dynamics with viral rebound, a natural extension of equation (8) is to assume that the second-phase viral decay rate λ_2 changes over time, which may be a function of time-varying covariate such as CD4 cell count to capture the complicated viral load trajectories including viral rebound. Thus, we consider the following NLME model for viral load response.

$$\begin{aligned}
y_{ij} &= \log_{10}(e^{p_{i1} - \lambda_{i1} t_{ij}} + e^{p_{i2} - \lambda_{ij2} t_{ij}}) + e_{ij} \\
p_{i1} &= \beta_1 + b_{i1}, \quad \lambda_{i1} = \beta_2 + \beta_3 z_{i0} + b_{i2}, \\
p_{i2} &= \beta_4 + b_{i3}, \quad \lambda_{ij2} = \beta_5 + \beta_6 z_{ij} + b_{i4},
\end{aligned} \tag{9}$$

where y_{ij} is the $\log_{10}$-transformation of the viral load for the ith subject at time t_{ij} $(i = 1, 2, ..., n,$ $j = 1, 2, ..., n_i)$ and the $\log_{10}$-transformation of the viral load is taken here to stabilize the variation of measurement errors and to speed up estimation algorithm, e_{ij} represents within-individual random error which follows $\mathrm{ALD}(0, \sigma, \tau)$, z_{i0} and z_{ij} are the baseline CD4 and repeated (time-varying) CD4 value for the ith subject at time t_{ij}, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_6)^T$ and $\boldsymbol{\beta}_{ij} = (p_{i1}, p_{i2}, \lambda_{i1}, \lambda_{ij2})^T$ are population parameters and individual-specific parameters for the ith subject at time t_{ij}, respectively, $\boldsymbol{b}_i = (b_{i1}, \ldots, b_{i4})^T$ is individual random-effects. Note that in AIDS studies, CD4 cell count, viral load, or both may be treated as responses, depending on research objectives. CD4 cell count is more often used as diagnostic indictor for patient AIDS progression, while viral load is often used as a primary endpoint to quantify treatment effect. For follow-up trials, CD4 cell count is viewed as a covariate to help predict virologic responses more efficiently, especially, in the HIV dynamic studies discussed here.

3.2   Model Implementation

As discussed previously, the viral loads (in $\log_{10}$ scale) described in Section 1 exhibited the asymmetric feature (see Figure 1(a)). Thus, it seems plausible to fit the data using the NLMEQR model by considering random-effects with an SN distribution. Toward this end, the following two statistical models with specifying different distributions for random-effects based on model (9) were employed to compare their performance.

  – **Model SN:** e_{ij} and $\boldsymbol{b}_i$ follow ALD and SN distribution, respectively.
  – **Model N:** e_{ij} and $\boldsymbol{b}_i$ follow ALD and normal distribution, respectively.

    We consider the following two scenarios. First, since a normal distribution is a special case of an SN distribution when skewness parameter is zero, we investigate how an asymmetric SN 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 NM), where viral load values below LOD are imputed by the half of the LOD (i.e., $\log_{10}(25)$). We use it as a comparison to the proposed method (denoted by TM) in Section 2 where viral load values below LOD are handled by Tobit model. This comparison attempts to investigate how the left-censored data in viral load influence modeling results.

To carry out the Bayesian inference, we took weakly-informative prior distributions for the parameters. In particular, (i) fixed-effects are taken to be independent normal distribution $N(0, 100)$ for each element of the population parameter vector $\boldsymbol{\beta}$; (ii) we assume a noninformative inverse Gamma prior distribution $IG(0.01, 0.01)$, which has mean 1 and variance 100, for scale parameter σ; (iii) the prior for the variance-covariance matrix of the random-effects $\boldsymbol{\Sigma}$ is taken to be inverse Wishart distribution $IW(\boldsymbol{\Omega}_2, \eta)$, with the diagonal matrix $\boldsymbol{\Omega}_2 = \mathrm{diag}(0.01, 0.01, 0.01, 0.01)$ and $\eta = 5$; (iv) for each element of skewness parameter $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_4)^T$, we choose independent normal distributions $N(0, 100)$.

The MCMC sampler was implemented using WinBUGS software [36] interacted with R2WinBUGS of R, and the program code is available from authors upon request. Figure 2 shows the trace plots, autocorrelation and dynamic version of GR diagnostic plots based on Model SN for the representative parameters β_3, β_5 and β_6. We observe from trace plots (top panel) that the lines of three different chains mix or cross in trace, implying that convergence is ensured. For the plots GR diagnostics (middle panel) three curves are given: the middle and bottom curves below the dashed horizontal line (indicating the value one) represent the pooled posterior variance ($\hat{V}$) and average within-sample variance ($\hat{W}$), respectively, and the top curve represents their ratio ($\hat{R}$). It is seen that $\hat{R}$ is generally expected to be higher than one at the initial stage of the algorithms, but $\hat{R}$ tends to 1, and $\hat{V}$ and $\hat{W}$ stabilize as the number of iterations increases, indicating that the algorithm has approached convergence. We further monitor convergence using autocorrelation plots (bottom panel) where autocorrelations are very low with a lag being 50, implying that convergence is obtained. Along with the diagnostic convergence, we proposed that, after an initial number of 100,000 burn-in iterations of three chains of length 150,000, every 50th MCMC sample was retained from the next 50,000 for each chain. Thus, we obtained a total of 3,000 samples of targeted posterior distributions of the unknown parameters for statistical inference. Note that in comparison with maximum likelihood methods, the Bayesian approach via MCMC procedure caused less convergence problems and the computational burden is reasonable. For example, to fit Model SN, it took about 3.5 hours on a Windows workstation with Intel Core 2 Quad CPU @ 2.83GHz and RAM of 16.0GB. The Bayesian modeling approach in conjunction with the NLMEQR model with different scenarios was applied to fit the AIDS data that was described in Section 1 and to estimate parameters. In what follows, we report the results based on the two scenarios highlighted above.

In order to examine the dependence of parameter estimates on the prior distributions and initial values, we carried out a sensitivity analysis. As an example, we implemented the MCMC sampling scheme and monitor several independent MCMC runs, starting from different initial values. Those runs exhibited similar and stable behavior (data not shown here). We summarize the sensitivity analysis results as follows: (i) The estimated parameters were not sensitive to both the priors and/or the initial values, and the results were reasonable and robust. (ii) When different priors and/or different initials were used, the results were similar to those presented in this paper.

3.3   Data Analysis Results

Bayesian modeling approach in conjunction with the NLMEQR model at the three quantiles of $\tau = 0.25, 0.50$ and 0.75 was used to fit the viral load data with non-normality and left-censoring due to LOD. Table 1 presents the population posterior mean (PM), the corresponding 95% credible interval (CI) for fixed-effects parameters based on the two methods (TM and NM). The following findings are obtained for the estimated parameters.

For both Models N and SN with the three quantiles of $\tau = 0.25, 0.50$ and 0.75, it is seen that as τ increases, β_1 and β_4, which are parameters related to baseline viral load, increase, while β_2 and β_3, which are parameters associated with the first-phase viral decay rate, decrease; For fixed-effects β_5 and β_6, which are parameters associated with the second-phase viral decay rate, the estimates of β_5 and β_6 are smallest and largest, respectively, at the quantile of $\tau = 0.50$. The estimated results of element parameters in $\boldsymbol{\beta}$ are quite different for different quantiles which appear to change from positive to negative values.
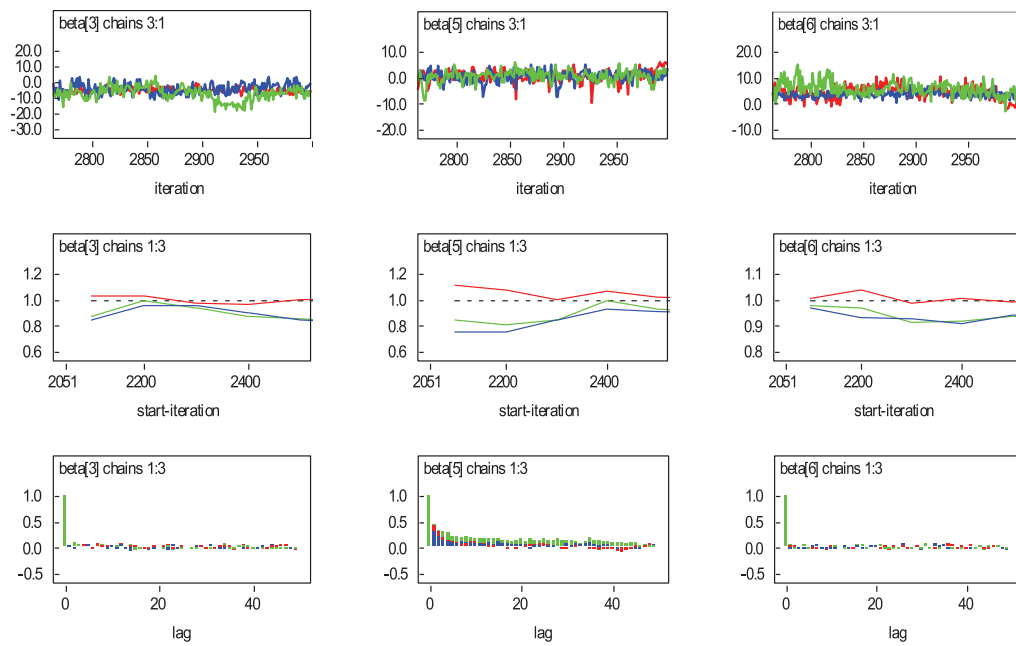
Figure 2. Convergence diagnostics with three Markov chains as obtained from the WinBUGS software for representative parameters based on Model SN with $\tau = 0.5$ quantile: (a) Trace plots (top panel). (b) Gelman-Rubin (GR) diagnostic plots (middle panel), where the middle (green) and bottom (blue) curves below the dashed horizontal line (indicating the value one) represent the pooled posterior variance ($\hat{V}$) and average within-sample variance ($\hat{W}$), respectively, and the top (red) curve above the dashed horizontal line represents their ratio ($\hat{R}$). (c) Autocorrelation plots (bottom panel).

Table 1. Summary of estimated posterior mean (PM) of population (fixed-effects) parameters, the corresponding lower limit (L_{CI}) and upper limit (U_{CI}) of 95% equal-tail credible interval (CI) as well as DIC values.

Method	Model		β_1	β_2	β_3	β_4	β_5	β_6	δ_1	δ_2	δ_3	δ_4	DIC
		$\tau = 0.25$											
TM	N	PM	-2.54	29.3	-5.48	-19.6	14.5	-2.92	-	-	-	-	12732.3
		L_{CI}	-3.85	13.5	-13.8	-30.7	-6.68	-14.2	-	-	-	-	
		U_{CI}	-0.79	51.7	-0.26	-11.0	31.6	7.98	-	-	-	-	
TM	SN	PM	-2.14	40.7	-8.21	-18.6	6.35	2.62	0.11	-3.56	-7.95	4.43	12455.3
		L_{CI}	-3.73	5.75	-16.9	-25.7	-6.18	-6.44	-1.94	-10.1	-29.8	1.24	
		U_{CI}	-0.50	79.4	-0.64	-11.7	31.8	9.77	2.14	-1.09	6.16	15.6	
NM	SN	PM	-2.14	17.1	-3.57	-4.52	0.45	1.48	-0.11	-1.69	0.03	1.36	14782.9
		L_{CI}	-13.9	3.80	-14.4	-11.6	-6.91	-0.85	-3.14	-7.11	-2.00	0.87	
		U_{CI}	-0.03	32.8	4.60	-1.70	6.03	3.75	4.90	-0.39	1.75	3.57	
		$\tau = 0.50$											
TM	N	PM	8.83	7.48	-6.85	-6.76	0.97	2.16	-	-	-	-	11397.9
		L_{CI}	7.85	-8.49	-16.0	-13.9	-12.7	-11.7	-	-	-	-	
		U_{CI}	9.76	23.1	0.96	2.06	20.5	11.6	-	-	-	-	
TM	SN	PM	9.20	31.2	-8.87	-9.29	-5.73	4.26	0.06	-2.69	-7.95	2.01	11349.0
		L_{CI}	8.19	-10.7	-19.9	-20.3	-21.7	-0.01	-1.85	-10.0	-22.8	1.42	
		U_{CI}	10.5	72.1	0.72	-2.16	4.96	10.2	1.66	-1.22	5.29	8.24	
NM	SN	PM	9.16	20.3	-3.93	6.29	0.42	1.64	-0.15	3.98	0.01	-0.18	13641.6
		L_{CI}	6.70	7.92	-13.4	3.36	-4.51	-0.02	-2.69	-7.30	-1.82	-2.67	
		U_{CI}	10.4	35.9	3.86	8.29	4.50	3.12	1.38	10.9	1.53	2.47	
		$\tau = 0.75$											
TM	N	PM	19.6	7.19	-6.93	-1.18	9.40	-5.50	-	-	-	-	12093.8
		L_{CI}	18.7	-6.25	-14.9	-16.9	-8.53	-14.8	-	-	-	-	
		U_{CI}	20.8	26.7	-0.66	11.7	21.5	6.80	-	-	-	-	
TM	SN	PM	19.9	4.44	-9.23	-4.69	1.97	1.55	-0.01	-1.31	-8.23	1.14	12086.7
		L_{CI}	19.0	-5.76	-18.4	-16.9	-11.2	-12.7	-1.64	-11.9	-19.3	1.09	
		U_{CI}	20.9	14.7	-0.28	4.62	19.2	12.4	1.47	-0.96	2.46	14.6	
NM	SN	PM	12.9	9.69	-4.25	5.84	-1.74	-0.50	0.37	6.24	-2.66	4.04	13656.6
		L_{CI}	-13.0	-1.28	-20.5	-20.3	-17.3	-12.1	-4.24	-2.92	-10.2	-5.08	
		U_{CI}	20.4	27.8	2.27	19.5	6.62	6.67	5.86	22.5	2.09	18.8	

In comparison of Models N and SN at the same quantile of τ, the estimated results indicate that β_1, β_2 and β_6 in Model N are smaller than their counterparts in Model SN except for $\tau = 0.75$ where β_2 (7.19) in Model N is larger than that (4.44) in Model SN; in contrast, β_3, β_4 and β_5 in Model N are larger than their counterparts in Model SN.

Regarding the skewness parameters δ_k ($k = 1, \ldots, 4$) for random-effects with SN distribution, their estimates vary according to quantiles. It is seen that as τ increases, δ_2 (associated with the first-phase viral decay rate) increases and δ_4 (associated with the second-phase viral decay rate) decreases. In particular, for the three quantiles of $\tau = 0.25, 0.50, 0.75$, the estimates of δ_2 are significantly negative with quite large skewness values (-3.56, -2.69 and -1.31) and the estimates of δ_4 are significantly positive with relatively large skewness values (4.43, 2.01 and 1.14). It might suggest that the first-phase decay rate appears a left-tail with increased shape and the second-phase viral decay rate seems to be a right-tail with decreased shape during the each period of process. The skewness parameters δ_1 and δ_3 associated with the macro-parameters p_1 and p_2 (which reflect the baseline viral load) are not significant and less interest in this study.

To select the best model that fits the data adequately, a Bayesian selection criterion, known as deviance information criterion (DIC) suggested by [40], is used. As with other model selection criteria, we caution that DIC is not intended for identification of the "correct" model, only which one fits the data better. In order to investigate whether SN model can provide better fit to the data than N model, the DIC values obtained are summarized in Table 1. We found that for the same quantile of τ the DIC values for Model SN are smaller than their counterparts for Model N. In addition, the DIC values are the smallest for the quantile of $\tau = 0.50$, while the DIC values are the largest for the quantile of $\tau = 0.25$. Therefore, based on the DIC, the results indicate that the models-based median regression ($\tau = 0.50$) is the best fitting model. In summary, our results suggest that it is very important to assume an SN distribution for the random-effects in order to achieve more reliable results, in particular if the data exhibit non-normality and/or heaviness in the tails.

We further investigate how left-censoring due to LOD handled by employing Tobit model in viral load data contribute to modeling results. Based on Model SN we compare the proposed method (denoted by TM) with the 'naive' method (NM) where the observed viral load values below LOD are imputed by the half of LOD (i.e., $\log_{10}(25)$). It can be seen from Table 1 that there are important differences in the estimates for the most parameters. The differences of estimates between TM and NM, due to whether the observed viral load values below LOD are treated as left-censoring or not, indicate that the observed viral load values below LOD replaced by half of LOD in the analysis might result in biased estimates. We also obtained DIC values of 12455.3, 11349.0 and 12086.7 for the three quantiles of $\tau = 0.25, 0.50, 0.75$, respectively, using TM which are smaller than the corresponding DIC values (14782.9,13641.6 and 13656.6) using NM, suggesting that TM provides a better fit to the data in comparison with NM. Thus, it is important to take left-censoring due to LOD in viral load observations into account when collected data are 'inaccurately' measured.

An advantage of the QR-based NLME models in this application allows us to examine the viral dynamics and treatment effects at different quantiles of the viral load observations in spite of one practical issue that the tuning parameters need to be adjusted for each quantile. In contrast, the mean regression-based NLME models [9,10,16,11,7,39] only quantify the viral dynamics and treatment effects at the average of the viral load observations, and thus cannot detect such quantile differences.

4   Simulation Studies

To assess the performance of the proposed NLMEQR models and method, as an illustration, we conducted the following limited simulation studies due to extensive computation involved. We mimicked the setting of ACTG398 study by taking the sample size n =300, and assuming that each subject had 13 scheduled longitudinal measurements. The measurement time points used in the simulation are similar to those in the real data analysis. We generated 200 datasets from specified model (9) according to the additional specifications described below. The true parameter values are taken by $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_6)^T = (30, 100, 4, 15, 5, 5)^T$. The time-varying CD4 covariate z_{ij} is simulated from equation $z_{ij} = (-2.5 + a_{i1}) + (2.0 + a_{i2})t_{ij} + (-1.2 + a_{i3})t_{ij}^2$, with $(a_{i1}, a_{i2}, a_{i3})^T \sim N(\mathbf{0}, \text{diag}(0.1, 0.1, 0.1))$. We simulated the model error e_{ij} from a chi-square distribution with two degrees of freedom, and for the random-effects $\boldsymbol{b}_i$, we

generated b_{ik} $(k = 1, \ldots, 4)$ from $\Gamma(2, 1)$ distribution, yielding a skewed distribution with the mean and variance being 2. Under these specifications, data generated from the model may exhibit highly skewed feature. To study the effect of the level of censoring on the posterior estimates, we consider two scenarios with approximate censoring proportions of 10% and 20%, respectively. Note that the prior distributions considered are all close to non-informative as similarly treated in real data analysis. Thus, we expect the results to be somewhat robust with respect to prior distributions.

For evaluating the objective use of the criteria, the models preferred by DIC were recorded. For example, in the MCMC sampling result, none of DIC selected the normally distributed specification for any of the 200 data sets, demonstrating the ability of selection method to detect an obvious departure from symmetry and suggesting strong evidence of skewness.

Table 2 summarizes simulation results which include the average estimates (EST) of fixed-effects $\boldsymbol{\beta}$ and skewness parameter $\boldsymbol{\delta}$ as well as associating the percent bias (defined by $100 \times \mathrm{bias}_l / \beta_l$) and percent mean-square-error (MSE), defined by $100 \times \sqrt{MSE_l} / |\beta_l|$ $(l = 1, \ldots, 6)$. We found that the proposed Model SN provides more accurate results compared with Model N in terms of parameter estimates as well as both bias and MSE. As expected, for smaller censoring rate (i.e., 10%), both Models SN and N provide more accurate parameter estimates and yield smaller biases and MSEs at different quantile levels. For all scenarios considered here, it is of interest to see that all estimated biases for β_1, β_2, β_3 and β_4 are negative, indicating that these parameters are underestimated, while estimated biases for β_5 and β_6, which are associated with the second-phase decay rate, are positive, suggesting that β_5 and β_6 are overestimated. For the skewness parameters δ_2 and δ_4 of random-effects, δ_2 associated with the

Table 2. Summary of estimated parameters as well as Bias and MSE for Models N and SN based on 100 simulated data sets under model error with χ_2^2 distribution and random-effects with $\Gamma(2, 1)$. EST is average of estimates, Bias and MSE are quantified by percent bias $= 100 \times \mathrm{bias}_l / \beta_l$ and percent $\sqrt{MSE} = 100 \times \sqrt{MSE_l} / |\beta_l|$, respectively.

		Model N						Model SN					
		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.75$		$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.75$	
		10%	20%	10%	20%	10%	20%	10%	20%	10%	20%	10%	20%
$\beta_1 = 30$	EST	27.15	25.96	27.07	25.96	26.66	25.59	28.55	27.06	28.27	27.67	28.63	27.08
	Bias	-9.51	-13.47	-9.78	-13.47	-11.12	-14.71	-4.83	-9.81	-5.78	-7.78	-4.56	-9.75
	MSE	3.19	6.14	3.46	6.14	4.16	7.14	2.96	5.27	4.34	7.72	3.99	6.91
$\beta_2 = 100$	EST	96.82	96.39	97.04	96.39	96.81	94.82	101.69	97.57	101.49	99.04	99.81	97.12
	Bias	-3.18	-3.61	-2.96	-3.61	-3.19	-5.18	1.69	-2.43	1.49	-0.96	-0.19	-2.88
	MSE	5.03	8.37	4.16	8.37	4.38	9.41	4.06	6.18	4.16	5.91	4.18	8.08
$\beta_3 = 4$	EST	2.87	2.57	4.04	2.86	2.56	3.01	2.87	2.94	2.79	3.00	2.80	2.80
	Bias	-28.18	-35.79	-29.36	-35.79	-28.41	-36.10	-24.69	-28.15	-26.42	-30.25	-25.10	-30.10
	MSE	49.54	58.24	50.17	58.24	53.28	62.16	48.55	57.74	48.23	56.34	52.12	61.65
$\beta_4 = 15$	EST	14.12	13.92	14.20	13.92	14.07	13.98	14.55	14.08	14.47	14.34	14.13	14.38
	Bias	-5.84	-7.19	-5.33	-7.19	-6.17	-6.81	-2.97	-6.14	-3.53	-4.37	-5.77	-4.13
	MSE	17.27	25.36	17.27	25.36	19.18	29.12	15.86	23.70	17.26	22.20	18.22	26.21
$\beta_5 = 5$	EST	7.97	8.21	7.87	8.21	8.11	8.24	6.36	6.74	6.46	6.67	6.52	6.76
	Bias	59.33	64.27	57.41	64.27	62.17	64.82	27.18	34.76	29.11	33.41	30.44	35.21
	MSE	44.58	55.32	41.42	55.32	44.47	57.18	40.73	51.39	40.41	52.53	42.77	56.09
$\beta_6 = 5$	EST	6.21	6.42	6.26	6.42	6.31	6.67	5.99	6.29	6.03	6.18	5.91	6.24
	Bias	24.14	28.32	25.16	28.32	26.24	33.38	19.86	25.83	20.61	23.61	18.11	24.88
	MSE	42.95	48.19	40.69	48.19	39.26	53.23	40.61	45.94	39.96	44.68	36.99	53.17
δ_2	EST	-	-	-	-	-	-	-18.11	-21.81	-16.92	-19.54	-18.06	-20.89
δ_4	EST	-	-	-	-	-	-	4.43	3.14	5.11	4.15	4.59	6.82

first-phase decay rate is significantly negative with very high skewness values and δ_4 associated with the second-phase decay rate is significantly positive with quite high skewness values in various scenarios. It might suggest that the first-phase decay rate appears a heavy left-tail with increased shape and the second-phase decay rate seems to be a fair right-tail with decreased shape during the each period of process. This finding is consistent with that reported in real data analysis. In summary, our simulation

results may suggest that it is important to account for the NLMEQR model by assuming random-effects with SN distribution, in particular, if the data exhibit skewness.

5 Concluding Discussion

This paper developed a general Bayesian modeling approach for NLMEQR models by assuming the ALD for error term for longitudinal data with left-censoring due to LOD, skewness, non-linearity, outlier and heavy-tail characteristics in which the random-effects follow an SN distribution that may be preferred over a normal distribution. By utilizing the relationship between the QR check function and the ALD, we cast the QR problem into the usual likelihood framework. This Bayesian NLMEQR model is analogous to the traditional Bayesian mixed-effects models for the mean regression, but allows the estimation at different quantiles of the conditional distribution, which provides a more robust estimator, offers a more comprehensive understanding of the data and facilitates a better understanding of the entire conditional distribution of the outcome given the covariate variables. The modeling methods are relatively straightforward to implement and, unlike frequentist approaches to quantile regression, do not rely on estimation or approximation of the asymptotic variances of the estimated parameters.

We believe that, to the best of our knowledge, this is the first attempt to work on such general distributional structure in NLMEQR models for longitudinal data with multiple features. Our proposed method is quite general and so can be used to other applications. This kind of modeling approach by assuming the random-effects with an SN distribution is important in many biostatistical application areas, allowing accurate inference of parameters while adjusting for the data with skewness. The SN distribution for random-effects is shown to provide an alternative to normal (symmetric) distribution that is often assumed in statistical models. The results indicate that, assuming the random-effects with SN distribution, there is potential to gain efficiency and accuracy in estimating certain parameters. The models considered in this paper can be easily fitted using MCMC procedure. Moreover, the proposed modeling approach is fitted using the WinBUGS package that is available publicly. This makes our approach quite powerful and accessible to applied statisticians in the fields.

A common concern with Bayesian methods is their dependence on various aspects of the nonlinear modeling process. Possible sources of uncertainty include the prior distributions and the initial values. The basic tool for investigating model uncertainty is the sensitivity analysis. That is, we simply make reasonable modifications to the assumptions in question, recompute the posterior quantities of interest, and see whether they have changed in a way that significantly affects the resulting interpretations or conclusions. If the results are robust against the suspected assumptions, we can report the results with confidence and our conclusions will be solid. However, if the results are sensitive to the assumptions, we choose to communicate the sensitivity results and interpret the results with caution [41].

Although QR originally does not require any distributional assumption, our approach assumes an ALD for the model error to achieve the right quantiles. Our simulation study has shown that the finite sample performance of the proposed methods is reasonably well even when the model error and random-effects distributions are away from the ALD and SN distribution, respectively. The mixture of stochastic representation utilized in this paper allows us to express a QR model as a commonly used normal model structure. Therefore, our approach can be further extended to more complicated models such as joint models for longitudinal-survival data [4,42]. Another future direction is to develop Bayesian QR for nonparametric or semiparametric models with random-effects for longitudinal data [6,16]. These interesting topics are left for future research.

References

1. Hammer S.M., Vaida F., Bennett K.K. et al., "Dual vs single protease inhibitor therapy following antiretroviral treatment failure: a randomized trial," *The Journal of the American Medical Association*, vol. 288, pp. 169-180, 2002.
2. Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., Markowitz, M., Ho, D.D., "Decay characteristics of HIV-1-infected compartments during combination therapy," *Nature*, vol. 387, pp. 188–191, 1997.
3. Wu, H., Ding, A.A., "Population HIV-1 dynamics in vivo: Applicable models and inferential tools for virological data from AIDS clinical trials," *Biometrics*, vol. 55, pp. 410–418, 1999.
4. Wu, L., Liu, W., Hu, X.J., "Joint inference on HIV viral dynamics and immune suppression in presence of measurement errors," *Biometrics*, vol. 66, pp. 327–335, 2010.
5. Hughes J.P., "Mixed effects models with censored data with applications to HIV RNA levels," *Biometrics*, vol. 55, pp. 625–629, 1999.
6. Liu, W., Wu, L., "Simultaneous inference for semiparametric nonlinear mixed-effects models with covariate measurement errors and missing responses," *Biometrics*, vol. 63, pp. 342–350, 2007.
7. Wu, L., "A joint model for nonlinear mixed-effects models with censoring and covariates measured with error, with application to AIDS studies," *Journal of the American Statistical Association*, vol. 97, pp. 955–964, 2002.
8. Yi, G.Y., Liu, W., Wu L., "Simultaneous inference and bias analysis for longitudinal data with covariate measurement error and missing responses," *Biometrics*, vol. 67, pp. 67–75, 2011.
9. Arellano-Valle R.B., Bolfarine H. and Lachos V.H., "Bayesian inference for skew-normal linear mixed models," *Journal of Applied Statistics*, vol. 34, pp. 663-682.
10. Ho HJ, Lin TI. "Robust linear mixed models using the skew-t distribution with application to schizophrenia data," *Biometrical Journal*, vol. 52, pp. 449–469, 2010.
11. A. Jara, F. Quintana and E.S. Martin, "Linear mixed models with skew-elliptical distributions: A Bayesian approach," *Computational Statistics and Data Analysis* , vol. 52, pp. 5033-5045, 2008.
12. 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, pp. 129-150, 2003.
13. Arellano-Valle R.B. and Genton, M., "On fundamental skew distributions," *Journal of Multivariate Analysis*, vol. 96, pp. 93-116, 2005.
14. Azzalini A. and Capitanio A., "Statistical applications of the multivariate skew normal distribution," *Journal of Royal Statistical Society, Series B*, vol. 67, pp. 579-602, 1999.
15. Tobin, J., "Estimation of relationships for limited dependent variables," *Econometrica*, vol. 26, pp. 24–36, 1958.
16. Huang Y., Dagne G.A., "Bayesian semiparametric nonlinear mixed-effects joint models for data with skewness, missing responses and measurement errors in covariates," *Biometrics*, vol. 68(3), pp. 943–953, 2012.
17. Koenker, R., Bassett, G., "Regression quantiles," *Econometrica*, vol. 46, pp. 33–50, 1978.
18. Koenker, R., *Quantile Regression*. Cambridge University Press, New York, 2005.
19. Kozumi, H., Kobayashi G., "Gibbs sampling methods for Bayesian quantile regression," *Journal of Statistical Computation and Simulation*, vol. 81(11), pp. 1565–1578, 2011.
20. Yuan, Y., Yin, G., "Bayesian quantile regression for longitudinal studies with nonignorable missing data," *Biometrics*, vol. 66, pp. 105–114, 2010.
21. Koenker, R., "Quantile regression for longitudinal data," *Journal of Multivariate Analysis*, vol. 91(1), pp. 74–89, 2004.
22. Alhamzawi R. and Yu K., "Bayesian Lasso-mixed quantile regression," *Journal of Statistical Computation and Simulation*, vol. 84(4), pp. 868-880, 2014.
23. Farcomeni A., "Quantile regression for longitudinal data based on latent Markov subjectspecific parameters," *Journal Statistics and Computing*, vol. 22, pp. 141–152, 2012.
24. Geraci M. and Bottai M., "Quantile regression for longitudinal data using the asymmetric Laplace distribution," *Biostatistics*, vol. 8, pp. 140–154, 2007.
25. Kim, M.O., Yang, Y., "Semiparametric approach to a random effects quantile regression model," *Journal of the American Statistical Association*, vol. 106(496), pp. 1405–1417, 2011.
26. Lipsitz, S.R., Fitzmaurice, G.M., Molenberghs, G., Zhao, L.P., "Quantile regression methods for longitudinal data with drop-outs: Application to CD4 cell counts of patients infected with the human immunodeficiency virus," *Journal of the Royal Statistical Society, Series C*, vol. 46, pp. 463–476, 1997.
27. Liu, Y., Bottai, M., "Mixed-effects models for conditional quantiles with longitudinal data," *The International Journal of Biostatistics*, vol. 5(1), Article 28, 2009.
28. Luo, Y., Lain, H., Tian, M., "Bayesain quantile regression for longitudinal data models," *Journal of Statistical Computation and Simulation*, vol. 82(11), pp. 1635–1649, 2012.

29. Wang, H.J., Fygenson, M., "Inference for censored quantile regression models in longitudinal studies," *The Annals of Statistics*, vol. 37(2), pp. 756–781, 2009.

30. Yu, K., Zhang, J.,"A three-parametric asymmetric Laplace distribution and its extension," *Communication in Statistics–Theory and Methods*, vol. 34, pp. 1867–1879, 2005.

31. Yu, K., Moyeed, R.A., "Bayesian quantile regression," *Statistics and Probability Letters*, vol. 54, pp. 437–447, 2001.

32. Koenker, R., Machado, J., "Goodness of fit and related inference process for quantile regression," *Journal of the American Statistical Association*, vol. 94, pp. 1296–1310, 1999.

33. Yu, K., Stander, J., "Bayesian analysis of a Tobit quantile regression model," *Journal of Econometrics*, vol. 137, pp. 260–276, 2007.

34. Kotz, S., Kozabowski, T.J., Podgorski K., *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*, Birkhauser, Boston, 2001.

35. Gelman, A. and Rubin, D., "Inference from iterative simulation using multiple sequences," *Statistical Science*, vol. 7, pp. 457-511, 1992.

36. Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., "WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility," *Statistics and Computing*, vol. 10, pp. 325–337, 2000.

37. Nowak, MA., May, RM., *Virus Dynamics: Mathematical Principles of Immunology and Virology.* Oxford University Press, Oxford, 2000.

38. Huang, Y., Liu, D., Wu, H., "Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system," *Biometrics*, vol. 62, pp. 413–423, 2006.

39. Huang Y., Chen, J., Yan C., "Mixed-effects joint models with skew-normal distribution for HIV dynamic response with missing and mismeasured time-varying covariate," *International Journal of Biostatistics*, vol. 8(1), Article 34, 1–30, 2012.

40. Spiegelhalter, D. J., Best, N. G., Carlin, B. P., Van der Linde, A., "Bayesian measures of model complexity and fit," *Journal of the Royal Statistical Society, Series B*, vol. 64, pp. 583–639, 2002.

41. Gelman, A. Carlin, J.B., Stern, H.S., Rubin. D.B., *Bayesian Data Analysis.* Chapman and Hall, London, 2003.

42. Huang Y, Dagne GA, Wu L., "Bayesian inference on joint models of HIV dynamics for time-to-event and longitudinal data with skewness and covariate measurement errors," *Statistics in Medicine*, vol. 30, pp. 2930–2946, 2011.