Statistics and Its Interface Volume 8 (2015) 203–215

Quantile regression for censored mixed-effects models with applications to HIV studies*

Victor H. Lachos, Ming-Hui Chen,
Carlos A. Abanto-Valle†, and Caio L. N. Azevedo

HIV RNA viral load measures are often subjected to some upper and lower detection limits depending on the quantification assays. Hence, the responses are either left or right censored. Linear/nonlinear mixed-effects models, with slight modifications to accommodate censoring, are routinely used to analyze this type of data. Usually, the inference procedures are based on normality (or elliptical distribution) assumptions for the random terms. However, those analyses might not provide robust inference when the distribution assumptions are questionable. In this paper, we discuss a fully Bayesian quantile regression inference using Markov Chain Monte Carlo (MCMC) methods for longitudinal data models with random effects and censored responses. Compared to the conventional mean regression approach, quantile regression can characterize the entire conditional distribution of the outcome variable, and is more robust to outliers and misspecification of the error distribution. Under the assumption that the error term follows an asymmetric Laplace distribution, we develop a hierarchical Bayesian model and obtain the posterior distribution of unknown parameters at the pth level, with the median regression ($p = 0.5$) as a special case. The proposed procedures are illustrated with two HIV AIDS studies on viral loads that were initially analyzed using the typical normal (censored) mean regression mixed-effects models, as well as a simulation study.

Keywords and phrases: Censored regression model, HIV viral load, Quantile regression, Asymmetric Laplace distribution, Gibbs sampling.

1. INTRODUCTION

Studies of HIV viral dynamics are the centerpiece of AIDS research. Such studies often consider repeated/longitudinal measures over the period of treatment, which are routinely analyzed using linear/nonlinear mixed effects models (LME/NLME) to assess rates of changes in HIV-1 RNA level or viral load [17, 18]. Viral load measures the amount of an actively replicating virus and its reduction is frequently used as a primary endpoint in clinical trials of anti-retroviral (ARV) therapy. However, depending upon the diagnostic assays used, its measurement may be subjected to some upper and lower detection limits (hence, left or right censored), below or above which they are not quantifiable. The proportion of censored data in these studies may not be trivial [4]. The crude/ad hoc methods viz. substituting threshold value or some arbitrary point such as mid-point between zero and cut-off for detection [16] might lead to biased estimates of fixed effects and variance components [18].

Our motivating datasets in this study consist of (i) the HIV-1 viral load after unstructured treatment interruption or UTI [14] and (ii) the setpoint for acutely infected subjects from the AIEDRP program [16]. The former has about 7% observations below the detection-limits (left-censored), whereas the latter has about 22% lying above the limits of assay quantifications (right-censored). As alternatives to crude imputation methods in the context of mean regression, [4] proposed a likelihood-based Monte Carlo EM algorithm (MCEM) for LME with censored responses (LMEC). [15] proposed a hybrid EM using a more efficient Hughes' algorithm, extending it to NLME with censored data (NLMEC). Recently, [16] proposed an exact EM algorithm for LMEC/NLMEC, which uses closed-form expressions at the E-step, as opposed to Monte Carlo simulations. In the framework of LMEC/NLMEC, the random effects and the within-subject errors are routinely assumed to follow a normal distribution for mathematical convenience. However, such assumption may not be always reasonable since they are vulnerable to the presence of atypical observations. To deal with the problem of atypical observations in the context of heavy–tailed LMEC/NLMEC, [9] advocated the use of the normal/independent (NI) class of distributions [11] and adopted a Bayesian framework to carry out posterior inference. More recently, [13] proposed a robust parametric modeling of LMEC/NLMEC based on the multivariate-t distribution so that the t-LMEC/t-NLMEC is defined and a fully likelihood based approach is carried out, including the implementation of an exact conditional EM (ECM) algorithm for maximum likelihood (ML) esti-

*The research of V. H. Lachos was supported by Grant 305054/2011-2 from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-Brazil) and by Grant 2011/17400-6 from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP-Brazil). M.-H. Chen's research was partially supported by US National Institutes of Health (NIH) grants #GM 70335 and #CA 74015. C. A. Abanto-Valle is deeply indebted to CNPq-Brazil, via grant 481159/2013-4 and Fundação de Amparo à Pesquisa do Estado de Rio de Janeiro (FAPERJ) via grant E-26/110.359/2014.
†Corresponding author.

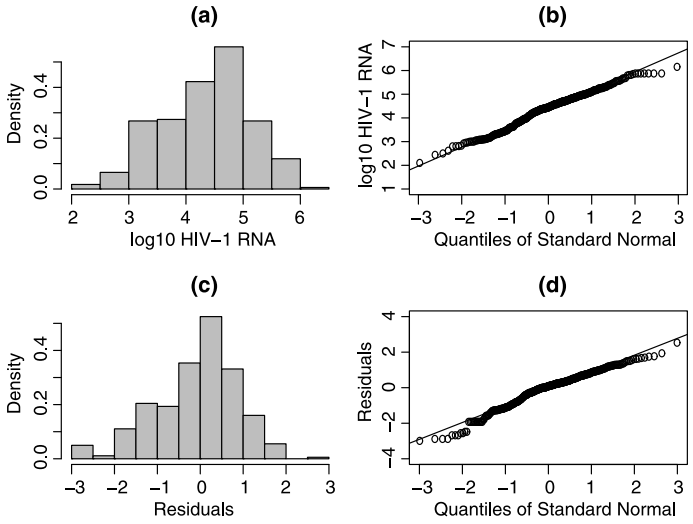

Figure 1. UTI data: density histogram and corresponding Q–Q plots for raw HIV viral load measures (in log10 scale; Panels a and b), and model residuals (Panels c and d), respectively, after fitting an Normal LMEC model using R package lmec.

mation. Note however that the majority of these methods focus on mean regression which is not a good measure of centrality when the conditional distribution of the response variable is skewed or multimodal, and therefore the mean regression estimator may be inadequate to make inferences about the shapes of these distributions. In contrast to the mean regression model, quantile regression (QR) belongs to a robust model family, which can provide an overall assessment of the covariate effects at different quantiles of the outcome [5]. Unlike conventional 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. In addition, QR does not impose any distribution assumption on the error, except requiring that the error has a zero conditional quantile.

An additional complication in the analysis of the HIV data is that viral-load measurements are often highly right-skewed with heavy right (or left) tail, and even log-transformations on the responses do not render normality or symmetry. These characteristics further complicate analysis of mixed-effects models, since both the random error (within-subject) and random effects (between-subject) might contribute to the "shift from symmetry". For example, panels (a) and (b) in Figure 1 display the density histogram and associated Q–Q plots for (repeated and noncensored) viral-load measurements (in the log10 scale) from the above study, which reveals some degree of left skewness in the response and panels ((c) and (d)) show the residuals, which all were obtained after fitting a NLMEC model to the UTI data using the R package *lmec*() [16]. These plots reveal left-skewed nature of the responses and the slightly symmetric behavior for the random errors. To the best of our knowledge, there are no studies on QR from a Bayesian perspective for LMEC/NLMEC. Thus, in this paper, we propose a QR model for LMEC/NLMEC based on the asymmetric Laplace distribution (ALD). The hierarchical representation of the ALD facilitates the convenient implementation of an efficient Gibbs algorithm with known generating distributions. In the Bayesian paradigm, the estimation and inference based on the proposed model can be easily implemented using the Markov chain Monte Carlo (MCMC) procedure.

The rest of the paper proceeds as follows. Section 2 introduces the connection between QR and ALD as well as outlines the main results related to ALD. In Section 3, the QR-LMEC model and related Gibbs sampling algorithm to estimate all of the model parameters is presented. In Section 4, the extension to the QR-NLMEC model is discussed. The advantage of the proposed methodology is illustrated through the analysis of two case studies of HIV viral load in Section 5. Section 6 presents a simulation study to compare the performance of our methods with mean regression-based methods. Section 7 concludes with a short discussion of issues that arise in our study and some possible directions for the future research.

2. PRELIMINARIES

Let y_i, be a response variable and $\mathbf{x}_i$ a $k \times 1$ vector of covariates for the ith subject for $i = 1, \ldots, n$. Let $Q_p(\mathbf{x}_i)$ denote the pth $(0 < p < 1)$ quantile regression function of y_i given $\mathbf{x}_i$. Suppose that the relationship between $Q_p(\mathbf{x}_i)$ and $\mathbf{x}_i$ can be modeled as $Q_p(\mathbf{x}_i) = \mathbf{x}_i^\top \boldsymbol{\beta}_p$, where $\boldsymbol{\beta}_p$ is a vector of unknown parameters of interest. Then, we consider the quantile regression model given by

$$y_i = \mathbf{x}_i^\top \boldsymbol{\beta}_p + \epsilon_i, \ i = 1, \ldots, n,$$

where ϵ_i is the error term whose distribution (with density, say, $f_p(.)$) is restricted to have the pth quantile equal to zero, that is, $\int_{-\infty}^{0} f_p(\epsilon_i) d\epsilon_i = p$.

The error density $f_p(.)$ is often left unspecified in the classical literature. Thus, quantile regression estimation for $\boldsymbol{\beta}_p$ proceeds by minimizing

$$\widehat{\boldsymbol{\beta}}_p = \arg \min_{\beta \in R^k} \sum_{i=1}^{n} \rho_p(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}_p), \quad (1)$$

where $\rho_p(.)$ is the so called check (or loss) function defined by $\rho_p(u) = u(p - \mathbb{I}\{u < 0\})$ and $\mathbb{I}\{.\}$ denotes the usual indicator function. Then $\widehat{\boldsymbol{\beta}}_p$ is called the pth quantile regression estimate. Note that the case where $p = 0.5$, corresponds to median regression. As the check function is not differentiable at zero, we cannot derive explicit solutions to the minimization problem. Therefore, linear programming methods are commonly used to obtain quantile regression estimates for $\boldsymbol{\beta}_p$. A connection between the minimization of

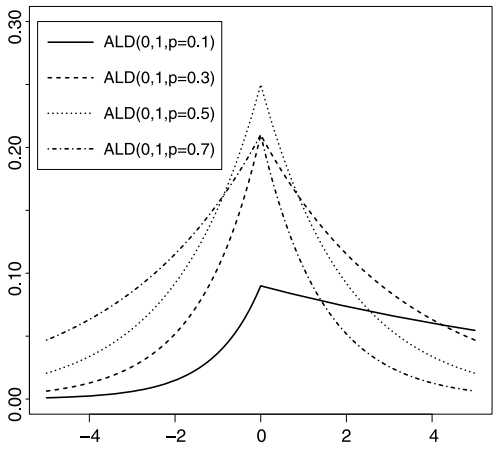

Figure 2. Standard asymmetric Laplace density (ALD).

the sum in Equation (1) and the maximum-likelihood theory is provided by the ALD. This skewed distribution was discussed in [6] and [9] among others. A random variable Y is distributed as an ALD with location parameter μ, scale parameter $\sigma > 0$ and skewness parameter $p \in (0,1)$, denoted by $ALD(\mu, \sigma, p)$, if its probability density function (pdf) is given by

$$f(y|\mu,\sigma,p) = \frac{p(1-p)}{\sigma} \exp\{-\rho_p(\frac{y-\mu}{\sigma})\}. \tag{2}$$

Set $\mu = \mu_i = \mathbf{x}_i^\top \boldsymbol{\beta}$ and write $\mathbf{y} = (y_1, \ldots, y_n)^\top$. Assuming that $y_i \sim ALD(\mu_i, \sigma, p)$, then the likelihood for n independent observations is

$$L(\boldsymbol{\beta}, \sigma|\mathbf{y}) = \frac{p^n(1-p)^n}{\sigma^n} \exp\{-\sum_{i=1}^{n} \rho_p(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}_p}{\sigma})\}. \tag{3}$$

Note that if we consider σ as a nuisance parameter, then the maximization of the likelihood in (3) with respect to the parameter $\boldsymbol{\beta}_p$ is equivalent to the minimization of the objective function in Equation (1).

In quantile regression, it is often of interest to compare slope coefficients for different quantiles. Then ALD can deal with the case when slope coefficients might be different for different quantile levels. In the Bayesian model using ALD, we impose the assumption $y \sim ALD(\mu, \sigma, p)$, which implies that the different quantiles of y conditional on x has the same slope. However, we only compute the pth quantile of y if $y \sim AL(\mu, \sigma, p)$ and for different p, we actually use a different model. Thus as long as $Q_p(\mathbf{x}_i) = \mathbf{x}_i^\top \boldsymbol{\beta}_p$, the likelihood is consistent in the sense that the maximum likelihood estimator (MLE) will converge to the true $\boldsymbol{\beta}_p$ in Equation (1). Thus, when using ALD in Bayesian analysis, we still can obtain consistent estimation of the quantile function and the slope coefficients that might be different for different p.

Figure 2 shows how the skewness of the ALD changes with altering the value of p. For example, when $p = 0.1$, almost all the mass of the ALD is situated in the right tail. In the case where $p = 0.5$, both tails of the ALD have equal mass and the distribution then reduces to a standard double exponential distribution. In contrast to the normal distribution with a quadratic term in the exponent, the ALD is linear in the exponent. This results in a more peaked mode for the ALD together with thicker tails. On the other hand, the normal distribution has heavier shoulders compared to the ALD.

To develop the Gibbs sampling algorithm in our development, we utilize a mixture representation based on exponential and normal distributions, which is found in [7] and is summarized as follows:

Lemma 1. *Let $Y \sim AL(\mu, \sigma, p)$, $Z \sim N(0,1)$ independent of $V \sim \exp(\sigma)$. Then*

$$Y \stackrel{d}{=} \mu + \vartheta_p V + \tau_p \sqrt{\sigma V} Z,$$

where $\vartheta_p = \frac{1-2p}{p(1-p)}$ and $\tau_p^2 = \frac{2}{p(1-p)}$, $\exp(\sigma)$ represents the exponential distribution with mean $1/\sigma$ and $\stackrel{d}{=}$ denotes the equality in distribution.

The result given in Lemma 1 yields a further hierarchical representation of Y in the following:

$$Y|V = v \sim N(\mu + \vartheta_p v, \tau_p^2 \sigma v), \tag{4}$$

$$V \sim \exp(\sigma). \tag{5}$$

It follows that the conditional distribution of V given Y is given by

$$V|Y \sim GIG(\frac{1}{2}, \delta, \gamma), \tag{6}$$

where $\delta = \frac{|y-\mu|}{\tau_p \sqrt{\sigma}}$ and $\gamma = \sqrt{\frac{1}{\sigma}(2 + \frac{\vartheta_p^2}{\tau_p^2})}$. In (6), $GIG(\nu, a, b)$ is a generalized inverse Gaussian distribution with pdf and moments, respectively, given by

$$f(x|\nu,a,b) = \frac{(b/a)^\nu}{2K_\nu(ab)} x^{\nu-1} \exp(-\frac{1}{2}(a^2 x^{-1} + b^2 x))$$

$$E[X^k] = \left(\frac{a}{b}\right)^k \frac{K_{\nu+k}(ab)}{K_\nu(ab)}, \quad k \in \mathbb{R},$$

where $x > 0$, $\nu \in \mathbb{R}$ $a, b > 0$ and $K_\nu(.)$ is a modified Bessel function of the third kind. See [1] for details.

3. QR LINEAR MIXED EFFECTS WITH CENSORED RESPONSES

We consider the following general LME model

$$y_{ij} = \mathbf{x}_{ij}^\top \boldsymbol{\beta} + \mathbf{z}_{ij} \mathbf{b}_i + \epsilon_{ij}, \; i = 1, \ldots, n, \; j = 1, \ldots, n_i, \tag{7}$$

where y_{ij} is the jth measurement of a continuous random variable on the ith subject, the $\mathbf{x}_{ij}^\top$'s are row vectors of a

know design matrix of dimension $n_i \times k$ corresponding to the fixed effects, $\boldsymbol{\beta}$ is a $k \times 1$ vector of population-averaged regression coefficients called fixed effects, and $\mathbf{z}_{ij}$ is a $q \times 1$ design matrix corresponding to the $q \times 1$ vector of random effects $\mathbf{b}_i$.

We define the LME quantile function of the response y_{ij} as

$$Q_p(y_{ij}|\mathbf{x}_{ij}, \mathbf{b}_i) = \mathbf{x}_{ij}^\top \boldsymbol{\beta}_p + \mathbf{z}_{ij}\mathbf{b}_i. \tag{8}$$

We assume that y_{ij}, conditionally on $\mathbf{b}_i$, for $i = 1, \ldots, n$, $j = 1, \ldots, n_i$ are independently distributed according to the ALD:

$$\begin{aligned} f(y_{ij}|\boldsymbol{\beta}_p, \mathbf{b}_i, \sigma) &= \frac{p(1-p)}{\sigma} \\ &\times \exp\left\{-\rho_p\left(\frac{y_{ij} - \mathbf{x}_{ij}^\top\boldsymbol{\beta}_p - \mathbf{z}_{ij}\mathbf{b}_i}{\sigma}\right)\right\}. \end{aligned} \tag{9}$$

In addition, we assume that the $\mathbf{b}_i$'s are distributed as $\mathbf{b}_i \overset{\text{iid}}{\sim} \mathrm{N}_q(0, \boldsymbol{D})$, where the dispersion matrix $\mathbf{D} = \mathbf{D}(\boldsymbol{\alpha})$ depends on unknown and reduced parameters $\boldsymbol{\alpha}$. In the present formulation, we consider the case where the response Y_{ij} is not fully observed for all i, j [16]. The observed data for the i-th subject is $(\mathbf{Q}_i, \mathbf{C}_i)$, where $\mathbf{Q}_i$ represents the vector of uncensored reading or censoring level, and $\mathbf{C}_i$ the vector of censoring indicators, such that

$$\begin{aligned} y_{ij} &\leq Q_{ij} \quad \text{if } C_{ij} = 1, \\ y_{ij} &= Q_{ij} \quad \text{if } C_{ij} = 0. \end{aligned} \tag{10}$$

For simplicity, we assume that the data are left-censored and thus the quantile regression censored linear mixed effect model (QR-LMEC) is defined. The extensions to arbitrary censoring are immediate. For normal LMEC, an EM algorithm was proposed by [4], with computational improvements considered in [15] and [16].

3.1 Prior and posterior distributions

Let $\mathbf{y}_i = (y_{i1}, \ldots, y_{in_i})^\top$, $\mathbf{X}_i = (x_{i1}, \ldots, x_{in_i})^\top$, $\mathbf{Z}_i = (z_{i1}, \ldots, z_{in_i})$, $\mathbf{V}_i = (V_{i1}, \ldots, V_{in_i})^\top$, $i = 1, \ldots, n$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\alpha}^\top, \sigma)^\top$. A key feature of this model is that, from Lemma 1, it can be formulated in a flexible hierarchical representation as follows:

$$\begin{aligned} &\mathbf{y}_i|\mathbf{b}_i, \mathbf{C}_i, \mathbf{Q}_i, \mathbf{V}_i = \mathbf{v}_i, \boldsymbol{\theta} \\ &\overset{\text{ind}}{\sim} TN_{n_i}(\mathbf{X}_i\boldsymbol{\beta}_p + \mathbf{Z}_i\mathbf{b}_i + \vartheta_p\mathbf{v}_i, \tau^2\sigma\boldsymbol{\Omega}_i; \mathbb{A}_i), \end{aligned} \tag{11}$$

$$V_{ij}|\sigma \overset{\text{iid}}{\sim} \exp(\sigma),\ i = 1, \ldots, n,\ j = 1, \ldots, n_i, \tag{12}$$

$$\mathbf{b}_i|\boldsymbol{\alpha} \overset{\text{ind}}{\sim} N(\mathbf{0}, \boldsymbol{D}), \tag{13}$$

where the observed data for the i-th subject is $(\mathbf{Q}_i, \mathbf{C}_i)$, for $i = 1, \ldots, n$; ϑ_p and τ^2 are as in Lemma 1; $\boldsymbol{\Omega}_i = Diag(\mathbf{v}_i) = Diag(v_{i1}, \ldots, v_{in_i})$, $TN_{n_i}(.; \mathbb{A})$ denotes the truncated normal distribution on the interval $\mathbb{A}_i = A_{i1} \times \ldots, \times A_{in_i}$, with A_{ij} as the interval $(-\infty, \infty)$ if $C_{ij} = 0$ and $(-\infty, Q_{ij}]$ if $C_{ij} = 1$. Specifically, a k-dimensional vector $\mathbf{X} \sim TN_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathbb{A})$ if its density is given by $\mathrm{TN}_k(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathbb{A}) = \frac{\phi_k(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma})}{\prod_{r=1}^k \int_{-\infty}^{a_r} \phi_k(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma})d\mathbf{x}}\mathbb{I}_{\{\mathbb{A}\}}(\mathbf{x})$, where the notation $\prod_{r=1}^k \int_{-\infty}^{a_r} = \int_{-\infty}^{a_1} \ldots \int_{-\infty}^{a_r}$ stand for the abbreviation of multiple integrals and $\phi_k(.; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the pdf of the k-variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariate matrix $\boldsymbol{\Sigma}$ ($N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$).

Let $\mathbf{y} = (\mathbf{y}_1^\top, \ldots, \mathbf{y}_n^\top)^\top$, $\mathbf{b} = (\mathbf{b}_1^\top, \ldots, \mathbf{b}_n^\top)^\top$, $\mathbf{u} = (u_1, \ldots, u_n)^\top$, $\mathbf{t} = (t_1, \ldots, t_n)^\top$, $\mathbf{Q} = vec(\mathbf{Q}_1, \ldots, \mathbf{Q}_n)$ and $\mathbf{C} = vec(\mathbf{C}_1, \ldots, \mathbf{C}_n)$. It follows that the complete likelihood function associated with $(\mathbf{y}, \mathbf{b}, \mathbf{Q}, \mathbf{C}, \mathbf{v})$ is given by

$$\begin{aligned} &L(\boldsymbol{\theta}|\mathbf{y}, \mathbf{b}, \mathbf{Q}, \mathbf{C}, \mathbf{v}) \\ &\propto \prod_{i=1}^n \Big[TN_{n_i}(\mathbf{y}_i|\mathbf{X}_i\boldsymbol{\beta}_p + \mathbf{Z}_i\mathbf{b}_i + \vartheta_p\mathbf{v}_i, \tau^2\sigma\boldsymbol{\Omega}_i; \mathbb{A}_i) \\ &\times \phi_q(\mathbf{b}_i; \mathbf{0}, \boldsymbol{D}) \times \frac{1}{\sigma^{n_i}} \prod_{i=1}^{n_i} \exp(-\frac{v_{ij}}{\sigma})\Big]. \end{aligned} \tag{14}$$

In order to carry out Bayesian inference, we need to specify prior distributions for all the unknown parameters $\boldsymbol{\theta} = (\boldsymbol{\beta}_p^\top, \sigma^2, \boldsymbol{\alpha}^\top)^\top$. A popular choice to ensure posterior propriety in a LME is to specify proper (but diffuse) conditionally conjugate priors [3, 20]. Following [10], we take

$$\begin{aligned} \boldsymbol{\beta}_p &\sim N_p(\boldsymbol{\beta}_0, \mathbf{S}_\beta), \\ \sigma &\sim IGamma(q_0, \lambda_0), \\ \boldsymbol{D} &\sim IWish_q(\boldsymbol{\Lambda}_0^{-1}, \nu_0), \end{aligned}$$

where $IGamma(a, b)$ denotes an inverse gamma distribution with mean $b/(a-1)$, $a > 1$, and $IWish_q(\mathbf{M}^{-1}, \nu_0)$ denotes an inverse Wishart distribution with mean $\mathbf{M}^{-1}/(\nu_0 - q - 1)$, $\nu_0 > q + 1$, where $\mathbf{M}$ is a $q \times q$ known positive definite matrix. Assuming the elements of the parameter vector to be independent, the joint prior distribution of all the unknown parameters is given by

$$\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta}_p)\pi(\sigma)\pi(\boldsymbol{D}). \tag{15}$$

Combining the likelihood function (14) and the prior distribution, the joint posterior distribution of all model parameters is then given by

$$\begin{aligned} &\pi(\boldsymbol{\beta}_p, \sigma^2, \boldsymbol{D}, \mathbf{v}, \mathbf{y}|\mathbf{Q}, \mathbf{C}) \\ &\propto \prod_{i=1}^n \Big[TN_{n_i}(\mathbf{y}_i|\mathbf{X}_i\boldsymbol{\beta}_p + \mathbf{Z}_i\mathbf{b}_i + \vartheta_p\mathbf{v}_i, \tau^2\sigma\boldsymbol{\Omega}_i; \mathbb{A}_i) \\ &\times \phi_q(\mathbf{b}_i; \mathbf{0}, \boldsymbol{D}) \times \frac{1}{\sigma^{n_i}} \prod_{j=1}^{n_i} \exp(-\frac{v_{ij}}{\sigma})\Big]\pi(\boldsymbol{\theta}). \end{aligned} \tag{16}$$

Our Bayesian model allows for a straightforward implementation of a Gibbs sampler via the hierarchical representation given in (12)–(14). To proceed, it is necessary to obtain the

full conditional distribution of one variable given the values of all the other variables – $(\mathbf{C}_i, \mathbf{Q}_i)$ included. We have the following expressions:

1. $\mathbf{y}_i|\mathbf{b}_i, \mathbf{v}_i, \mathbf{C}_i, \mathbf{Q}_i, \boldsymbol{\theta} \sim f(\mathbf{y}_i|\mathbf{b}_i, \mathbf{v}_i, \mathbf{C}_i, \mathbf{Q}_i, \boldsymbol{\theta})$. Thus, conditional on $(\mathbf{b}_i, \mathbf{v}_i)$, $\mathbf{y}_i$ is a vector of independent observations, whose distributions are truncated normal, each with untruncated variance $\tau^2\sigma v_{ij}$ and untruncated mean $\mathbf{x}_{ij}^\top\boldsymbol{\beta}_p + \mathbf{z}_{ij}\mathbf{b}_i + \vartheta_p v_{ij}$ on the interval $y_{ij} \leq Q_{ij}$, i.e. $TN_1(\mathbf{x}_{ij}^\top\boldsymbol{\beta}_p + \mathbf{z}_{ij}\mathbf{b}_i + \vartheta_p v_{ij}, \tau^2\sigma v_{ij}; (-\infty, Q_{ij}))$.
2. $\mathbf{b}_i|\mathbf{y}_i, \mathbf{v}_i, \mathbf{C}_i, \mathbf{Q}_i, \boldsymbol{\theta} \equiv \mathbf{b}_i|\mathbf{y}_i, \mathbf{v}_i, \boldsymbol{\theta} \sim f(\mathbf{b}_i|\mathbf{y}_i, \mathbf{v}_i, \boldsymbol{\theta})$. This distribution is multivariate normal with mean $\widehat{\mathbf{b}}_i = \boldsymbol{\Lambda}_i(\mathbf{Z}_i^\top\boldsymbol{\Sigma}_{vi}^{-1}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_p - \vartheta_p\mathbf{v}_i))$ and variance $\boldsymbol{\Lambda}_i$, with $\boldsymbol{\Lambda}_i = (\boldsymbol{D}^{-1} + \mathbf{Z}_i^\top\boldsymbol{\Sigma}_{vi}^{-1}\mathbf{Z}_i)^{-1}$ and $\boldsymbol{\Sigma}_{vi} = \tau^2\sigma\boldsymbol{\Omega}_i$. Note that the entire vector $\mathbf{y}_i$ is used for sampling from $\mathbf{b}_i$.
3. $V_{ij}|y_{ij}, \mathbf{b}_i, C_{ij}, Q_{ij}, \boldsymbol{\theta} \equiv \pi(v_{ij}|y_{ij}, \mathbf{b}_i, \boldsymbol{\theta}) \propto v_{ij}^{1/2}\exp\{-\frac{1}{2}\times(\frac{A_{ij}^2}{\tau^2\sigma}v_{ij}^{-1} + (\frac{\vartheta_p^2}{\tau^2\sigma} + \frac{2}{\sigma})v_{ij})$, where $A_{ij} = y_{ij} - \mathbf{x}_{ij}^\top\boldsymbol{\beta}_p - \mathbf{z}_{ij}\mathbf{b}_i$, i.e., $V_{ij}|y_{ij}, \mathbf{b}_i, C_{ij}, Q_{ij}, \boldsymbol{\theta} \sim GIG(\frac{1}{2}, \sqrt{\frac{A_{ij}^2}{\tau^2\sigma}}, \sqrt{\frac{\vartheta_p^2}{\tau^2\sigma} + \frac{1}{\sigma}})$, $i = 1, \ldots, n,\ j = 1, \ldots, n_i$, where $GIG(\nu, a, b)$ is the generalized inverse Gaussian defined in Section 2.
4. Now, by observing that $\boldsymbol{\theta}_1|\mathbf{y}, \mathbf{C}, \mathbf{Q}, \mathbf{b}_i, \mathbf{v}_i, \boldsymbol{\theta}_{(-\boldsymbol{\theta}_1)}$ and $\boldsymbol{\theta}_1|\mathbf{y}, \mathbf{b}_i, \mathbf{v}_i, \boldsymbol{\theta}_{(-\boldsymbol{\theta}_1)}$ are identical, we have

$$\begin{aligned}
\boldsymbol{\beta}_p|\mathbf{y}, \mathbf{v}, \mathbf{b}, \boldsymbol{\theta}_{(-\boldsymbol{\beta}_p)} &\sim N\left(\mathbf{A}_\beta\boldsymbol{\mu}_\beta,\ \mathbf{A}_\beta\right), \\
\sigma|\mathbf{y}, \mathbf{v}, \mathbf{b}, \boldsymbol{\theta}_{(-\sigma^2)} &\sim IGamma(q_0 + \frac{3N}{2}, \lambda_0 + s), \\
\boldsymbol{D}|\mathbf{y}, \mathbf{v}, \mathbf{b}, \boldsymbol{\theta}_{(-\boldsymbol{\alpha})} &\sim IWish_q(\boldsymbol{\Lambda}^{-1}, \nu_0 + n),
\end{aligned}$$

where $\boldsymbol{\mu}_\beta = (\mathbf{S}_\beta^{-1}\boldsymbol{\beta}_0 + \sum_{i=1}^n \mathbf{X}_i^\top\boldsymbol{\Sigma}_{vi}^{-1}(\mathbf{y}_i - \mathbf{Z}_i\mathbf{b}_i - \vartheta_p\mathbf{v}_i))$, $\mathbf{A}_\beta = (\mathbf{S}_\beta^{-1} + \sum_{i=1}^n \mathbf{X}_i^\top\boldsymbol{\Sigma}_{vi}^{-1}\mathbf{X}_i)^{-1}$, $N = \sum_{i=1}^n n_i$, $s = \sum_{i=1}^n[\frac{1}{2\tau^2}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_p - \mathbf{Z}_i\mathbf{b}_i)^\top\boldsymbol{\Omega}_i^{-1}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_p - \mathbf{Z}_i\mathbf{b}_i) + \sum_{i=1}^{ni} v_{ij}]$, $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_0 + \sum_{i=1}^n \mathbf{b}_i\mathbf{b}_i^\top$.

Note that all the full conditional distributions have closed forms and hence can be easily implemented, particularly using the popular Bayesian software WinBUGS.

4. THE NONLINEAR CASE

4.1 Model specification

Extending the notation defines in the previous section and ignoring censoring, we first propose the following general mixed-effects model. Let $\mathbf{y}_i = (y_{i1}, \ldots, y_{in_i})^\top$ denote the (continuous) response vector for subject i and also let $\eta = (\eta(\mathbf{x}_{i1}, \boldsymbol{\phi}_i), \ldots, \eta(\mathbf{x}_{in_i}, \boldsymbol{\phi}_i))^\top$ be a nonlinear vector-valued differentiable function of the random parameter $\boldsymbol{\phi}_i$ of dimension r and a vector (or matrix) of covariates $\mathbf{x}_i$. The NLME can then be expressed as

$$\mathbf{y}_i = \eta(\boldsymbol{\phi}_i, \mathbf{x}_i) + \boldsymbol{\epsilon}_i,\ \ \boldsymbol{\phi}_i = \mathrm{A}_i\boldsymbol{\beta} + \mathbf{B}_i\mathbf{b}_i, \tag{17}$$

where A_i and $\mathbf{B}_i$ are known design matrices of dimensions $r \times k$ and $r \times q$, respectively, possibly depending on some covariates, $\boldsymbol{\beta}$ is the $(k \times 1)$ vector of fixed effects, and $\mathbf{b}_i$ is the $(q \times 1)$ vector of random effects. In mean regression, it is common to assume that, $\mathbf{b}_i \overset{ind}{\sim} N_q(0, \mathbf{D})$ and $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{in_i})^\top \overset{ind.}{\sim} N_{n_i}(\mathbf{0}, \sigma^2\mathbf{I}_{n_i})$ [see, 9]. Here, we define the NLME quantile function of the response y_{ij} as

$$Q_p(y_{ij}|\mathbf{x}_{ij}, \mathbf{b}_i) = \eta(\phi_i, \mathbf{x}_{ij}) = \eta(\mathrm{A}_i\boldsymbol{\beta}_p + \mathbf{B}_i\mathbf{b}_i, \mathbf{x}_{ij}). \tag{18}$$

We assume that, given $\mathbf{b}_i$, y_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, n_i$ are independent distributed according to the ALD, i.e.,

$$\begin{aligned}
f(y_{ij}|\boldsymbol{\beta}_p, \mathbf{b}_i, \sigma) &= \frac{p(1-p)}{\sigma} \\
&\times\ \mathrm{e}^{-\rho_p\left(\frac{y_{ij} - \eta(\mathrm{A}_i\boldsymbol{\beta}_p + \mathbf{B}_i\mathbf{b}_i, \mathbf{x}_{ij})}{\sigma}\right)},
\end{aligned} \tag{19}$$

and in addition, we assume that $\mathbf{b}_i$ is distributed as $\mathbf{b}_i \overset{\text{iid}}{\sim} \mathrm{N}_q(0, \boldsymbol{D})$, where the dispersion matrix $\mathbf{D} = \mathbf{D}(\boldsymbol{\alpha})$ depends on unknown and reduced parameters $\boldsymbol{\alpha}$ and hence the quantile regression nonlinear mixed effects model is defined (QR-NLME).

For QR-NLME with complete responses, the marginal distribution is given by

$$f(\mathbf{y}|\boldsymbol{\theta}) = \prod_{i=1}^n \int_{\mathbb{R}^q} \left[\prod_{j=1}^{n_i} f(y_{ij}|\boldsymbol{\beta}_p, \mathbf{b}_i, \sigma)\right] \phi_q(\mathbf{b}_i; 0, \mathbf{D}) d\mathbf{b}_i,$$

which generally does not have a closed form expression because the model function is not linear in the random effects.

Now assuming left-censoring, such that the observed data for the i-th subject are $(\mathbf{Q}_i, \mathbf{C}_i)$, the individual observations within cluster i follows (10), the QR-NLMEC is defined. Using the same notation as in Section 3.1 and Lemma 1, we have the following hierarchical representation for the QR-NLMEC:

$$\begin{aligned}
&\mathbf{y}_i|\mathbf{b}_i, \mathbf{C}_i, \mathbf{Q}_i, \mathbf{V}_i, = \mathbf{v}_i, \boldsymbol{\theta} \\
&\quad \overset{\text{ind}}{\sim} TN_{n_i}(\eta(\mathrm{A}_i\boldsymbol{\beta}_p + \mathbf{B}_i\mathbf{b}_i, \mathbf{x}_i) + \vartheta_p\mathbf{v}_i, \tau^2\sigma\boldsymbol{\Omega}_i; \mathbb{A}_i),
\end{aligned} \tag{20}$$

$$V_{ij}|\sigma \overset{\text{iid}}{\sim} \exp(\sigma),\ i = 1, \ldots, n,\ j = 1, \ldots, n_i, \tag{21}$$

$$\mathbf{b}_i|\boldsymbol{\alpha} \overset{\text{ind}}{\sim} N_q(\mathbf{0}, \boldsymbol{D}). \tag{22}$$

4.2 Prior and posterior specifications

Under the same prior specifications as discussed in Subsection 3.1, the full conditional distributions for QR-NLMEC models are as follows:

$$\begin{aligned}
&y_{ij}|\mathbf{b}_i, \mathbf{v}_i, \mathbf{C}_i, \mathbf{Q}_i, \boldsymbol{\theta} \\
&\sim\ TN_1(\eta(\mathrm{A}_i\boldsymbol{\beta}_p + \mathbf{B}_i\mathbf{b}_i, \mathbf{x}_{ij}) + \vartheta_p v_{ij}, \tau^2\sigma v_{ij}; (-\infty, Q_{ij}))
\end{aligned}$$

Table 1. Posterior parameter estimates for the UTI data

	mean regression			$p = 0.5$ (median regression)		
Parameter	Mean	SD	95% CI	Mean	sd	95% CI
β_1	3.6501	0.1308	[3.3983; 3.9095]	3.8749	0.1225	[3.6305; 4.1088]
β_2	4.1765	0.1336	[3.9142; 4.4417]	4.2105	0.1158	[3.9750; 4.4310]
β_3	4.2464	0.1357	[3.9768; 4.5051]	4.2621	0.1136	[4.0332; 4.4799]
β_4	4.3637	0.1365	[4.1043; 4.6268]	4.4238	0.1181	[4.1831; 4.6478]
β_5	4.5697	0.1429	[4.2903; 4.8450]	4.5465	0.1189	[4.3063; 4.7703]
β_6	4.5881	0.1517	[4.2927; 4.8822]	4.5417	0.1222	[4.2949; 4.7764]
β_7	4.6957	0.1709	[4.35951; 5.0321]	4.7042	0.1382	[4.4280; 4.9723]
β_8	4.8079	0.2065	[4.4086; 5.2108]	4.7793	0.1636	[4.4582; 5.1007]
σ	0.3339	0.0304	[0.2798; 0.3969]	0.1851	0.0110	[0.1646; 0.2081]
α	0.7864	0.1526	[0.5375; 1.1534]	0.8070	0.1530	[0.5546; 1.1491]

$$\begin{aligned}
\mathbf{b}_i|\mathbf{y}_i,\mathbf{v}_i,\boldsymbol{\theta} &\propto \phi_{n_i}\left(\mathbf{y}_i;\eta(\mathrm{A}_i\boldsymbol{\beta}_p+\mathbf{B}_i\mathbf{b}_i,\mathbf{x}_i),\sigma^2\mathbf{I}_{n_i}\right)\\
&\quad\times\phi_q\left(\mathbf{b}_i;\mathbf{0},\boldsymbol{D}\right);\\
V_{ij}|y_{ij},\mathbf{b}_i,C_{ij},Q_{ij},\boldsymbol{\theta} &\sim GIG(\frac{1}{2},\sqrt{\frac{A_{ij}^2}{\tau^2\sigma}},\sqrt{\frac{\vartheta_p^2}{\tau^2\sigma}+\frac{1}{\sigma}}),\\
&\quad i=1,\ldots,n,\ j=1,\ldots,n_i;\\
\boldsymbol{D}|\mathbf{y},\mathbf{b},\mathbf{v},\boldsymbol{\theta}_{(-\boldsymbol{\alpha})} &\sim IWish_q\left(\boldsymbol{\Lambda}^{-1},\nu_0+n\right);\\
\boldsymbol{\beta}_p|\mathbf{y},\mathbf{b},\mathbf{v},\boldsymbol{\theta}_{(-\boldsymbol{\beta})} &\sim N_p\left(\mathbf{A}_\beta\boldsymbol{\mu}_\beta,\ \mathbf{A}_\beta\right);\\
\sigma^2|\mathbf{y},\mathbf{b},\mathbf{u},\boldsymbol{\theta}_{(-\sigma^2)} &\sim IGamma\left(\frac{3N}{2}+q_0,\ \lambda_0+s\right),
\end{aligned}$$

where $\mathbf{A}_\beta = (\mathbf{S}_\beta^{-1}+\sum_{i=1}^n\mathbf{A}_i^\top(\mathbf{B}_i\boldsymbol{D}\mathbf{B}_i^\top)^{-1}\mathbf{A}_i)^{-1}$, $\boldsymbol{\mu}_\beta = (\mathbf{S}_\beta^{-1}\boldsymbol{\beta}_0+\sum_{i=1}^n\mathbf{A}_i^\top(\mathbf{B}_i\boldsymbol{D}\mathbf{B}_i^\top)^{-1}\boldsymbol{\phi}_i)$, $N=\sum_{i=1}^n n_i$, $s=\sum_{i=1}^n[\frac{1}{2\tau^2}(\mathbf{y}_i-\eta(\boldsymbol{\phi}_i,\mathbf{x}_i))^\top\boldsymbol{\Omega}_i^{-1}(\mathbf{y}_i-\eta(\boldsymbol{\phi}_i,\mathbf{x}_i))+\sum_{i=1}^{ni}v_{ij}]$, $\boldsymbol{\Lambda}=\boldsymbol{\Lambda}_0^{-1}+\sum_{i=1}^n\mathbf{b}_i\mathbf{b}_i^\top$, and $A_{ij}=y_{ij}-\eta(\mathrm{A}_i\boldsymbol{\beta}_p+\mathbf{B}_i\mathbf{b}_i,\mathbf{x}_{ij})$. Note that sampling from the full conditional distribution for $\mathbf{b}_i$ requires the using of some auxiliary algorihtm, for example Metropolis-Hastings steps.

5. APPLICATIONS

We apply the proposed methods to the two HIV data sets previously analyzed using mean regression LMEC models.

5.1 UTI data

The HIV UTI data were from a study of 72 perinatally HIV-infected children [14, 16]. The dataset is available in the R package *lmec*. Primarily due to treatment fatigue, an unstructure treatment interruption (UTI) is common in this population. Suboptimal adherence can lead to ARV resistance and diminished treatment options in the future. The subjects in the study had taken ARV therapy for at least 6 months before UTI, and the medication was discontinued for more than 3 months. The HIV viral load from the closest time points at 0, 1, 3, 6, 9, 12, 18, 24 months after UTI were studied. The numbers of observations from baseline (month 0) to month 24 are 71, 62, 58, 57, 43, 34, 24, and 13, respectively. Out of 362 observations, 26 (7%) observations were below the detection limits (50 or 400 copies/mL) and were left-censored at these values. The individual profiles of viral load at different followup times after UTI are plotted in Figure 2. Following [16], we consider a profile LME model with random intercepts b_i's as

$$y_{ij} = b_i + \beta_j + \epsilon_{ij}, \tag{23}$$

where y_{ij} is the $\log_{10}$ HIV RNA for subject i at time t_j, $t_1=0$, $t_2=1$, $t_3=3$, $t_4=6$, $t_5=9$, $t_6=12$, $t_7=18$, $t_8=24$. [16] analyzed the same dataset by fitting a N-LMEC from a frequentist perspective, but from Figure 1, it is clear that inference under the normality assumption can be questionable. In our analysis, we assume a QR-LMEC as defined in (8)–(10). The priors are specified as $\beta_j \sim N_1(\mathbf{0},10^3)$, $j=1,\ldots,8$, $\sigma\sim IGamma(0.1,0.1)$, and $\sigma_b^2=\alpha\sim IGamma(0.1,0.1)$. We generated two parallel independent MCMC runs of size 100,000 with widely dispersed initial values, where the first 20,000 iterations (burn-in samples) were discarded, for computing posterior estimates. To eliminate potential problems due to autocorrelations, we considered a spacing of size 40. The convergence of the MCMC chains was monitored using trace plots, auto-correlation (ACF) plots and Gelman-Rubin $\hat{R}$ diagnostics. Following [2], we carried out sensitivity analysis on the specification of the inverse-gamma priors on the variance components and found that the results are fairly robust under different choices of the priors. The posterior summaries of the parameters do not present a noticeable difference and do not impair the results given in Table 1.

In Table 1, we report the posterior means, standard deviations (SDs) and 95% credible intervals (CI) of the model parameters from the popular mean regression (N-LMEC) and the QR-LMEC for $p=0.5$ (i.e., median regression). Note that the posterior estimates of $\beta_1-\beta_8$ (the slope parameters corresponding to the time points) for the QR-LMEC models are quite close (to first decimal place) to those from N-LMEC. The 95% posterior CIs to $\boldsymbol{\beta}$ are tighter (and also the standard deviations) than those in the mean regression model, indicating that

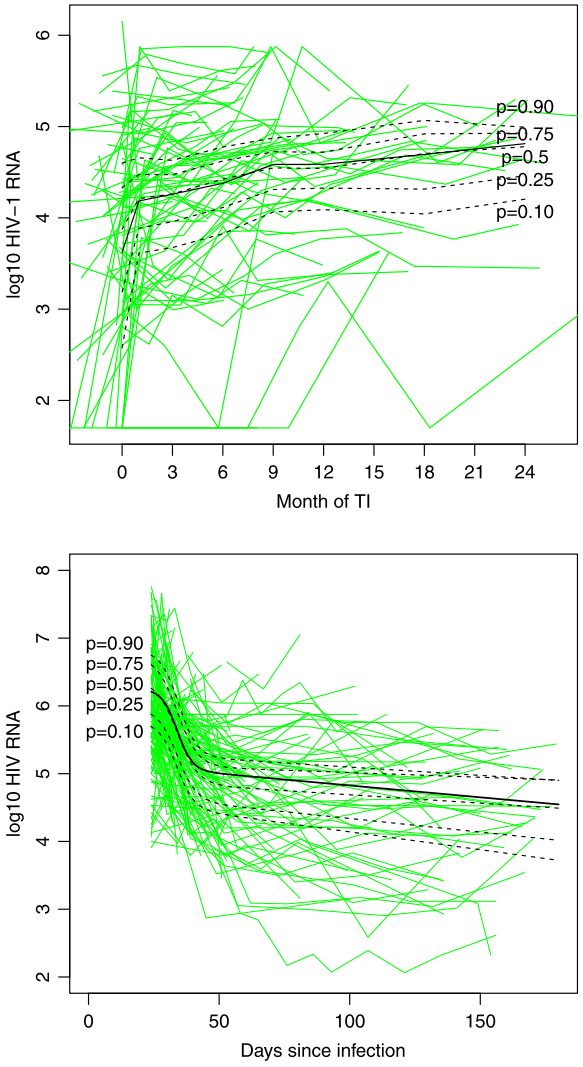

Figure 3. Individual profiles and overall mean (in log10 scale) at different quantiles for HIV viral load at different follow-up times for (left panel) UTI Data and (right panel) AIEDRP data.

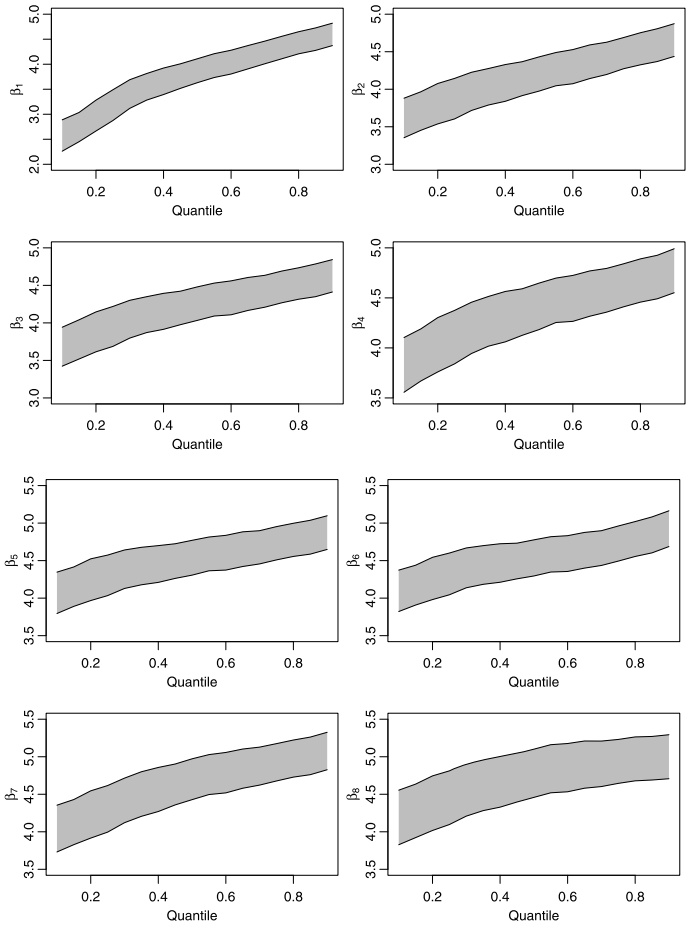

Figure 4. UTI data: Posterior means and 95% credible intervals for various values of p.

the median regression seems to produce more precise estimates.

As in [16], our dropout (censored) model does not bias the inference regarding the mean of β_j. The median and mean viral load β_j's increase gradually throughout 24 months for all the models. For the N-LMEC, it increases from 3.65 at the time of UTI to 4.80 at 24 months whereas in the median regression it increases from 3.87 to 4.77.

To obtain a more complete picture of the effects, series of QR models over the grid $p = \{0.1, 0.15, \ldots, 0.9\}$ are estimated. Figure 4 gives a graphical summary of this analysis. The solid lines are the $Q_{0.025}$ percentile and the $Q_{0.975}$ percentile obtained from the marginal posterior distribution of the different parameters. Thus, the shaded area depicted the 95% credible band from the marginal posterior distribution. From Figure 4, we can observe some interesting evidences which cannot be detected by mean regression. For example, the effects of most variables become stronger for the higher conditional quantiles, indicating that the viral load at different time points are positively correlated with the quantiles. This finding can be also appreciated in Figure 3, where the overall mean (in log10 scale) at different quantiles for HIV viral load at different follow-up times are depicted.

5.2 AIEDRP data

The second AIDS case study is from the AIEDRP program, a large multi-center observational study of subjects with acute and early HIV infection. We consider 320 untreated individuals with acute HIV infection; for more details see [16]. Of the 830 recorded observations, 185 (22%) were above the limit of assay quantification, hence right-censored. So, we consider a right-censored version of (10) and accommodate it within our NLME. Following [16], we choose a five-parameter NLME model (inverted S-shaped curve) as follows:

Table 2. AIEDRP data: Posterior estimates from censored N-LMEC (mean regression) and QR–NLMEC with $p = 0.5$

	mean regression			$p = 0.5$ (median regression)		
Parameter	Mean	sd	95% CI	Mean	sd	95% CI
β_1	1.5612	0.0191	[1.5209; 1.5971]	1.5709	0.0145	[1.5430; 1.5993]
β_2	0.4983	0.1840	[0.2088; 0.9263]	0.4016	0.1284	[0.1824; 0.6848]
β_3	3.5184	0.0650	[3.3548; 3.5926]	3.5294	0.0306	[3.4568; 3.5805]
β_4	1.6468	0.3066	[1.0379; 2.2489]	1.4200	0.2716	[0.8970; 1.9150]
β_5	-0.0018	0.0027	[-0.0072; 0.0036]	-0.0024	0.0023	[-0.0070; 0.0020]
σ	0.2324	0.0184	[0.1996; 0.2720]	0.1732	0.0085	[0.1571; 0.1906]
D_{11}	0.0188	0.0028	[0.0141; 0.0249]	0.0186	0.0026	[0.0140; 0.0245]
D_{12}	0.0004	0.0003	[-0.0001; 0.0011]	0.0004	0.0003	[-0.0001; 0.0011]
D_{22}	0.0003	0.0001	[0.0002; 0.0005]	0.0003	0.0001	[0.0002; 0.0004]

$$y_{ij} = \alpha_{1i} + \frac{\alpha_2}{(1 + \exp((t_{ij} - \alpha_3)/\alpha_4))} + \alpha_{5i}(t_{ij} - 50) + \epsilon_{ij},$$

where y_{ij} is the $\log_{10}$ HIV RNA for subject i at time t_{ij}. The choice of an appropriate non-linear model is difficult to assess for any HIV data, but the above model was considered in [16] primarily because the residual plots did not exhibit any serial auto-correlations, and the model fit seems adequate. The parameters α_{1i} and α_2 are the setpoint value and the decrease from the maximum HIV RNA. In the absence of treatment (following acute infection), the HIV RNA varies around a set-point which may differ among individuals, hence the setpoint is chosen to be subject-specific. The location parameter α_3 indicates the time point at which half of the change in HIV RNA is attained, α_4 is a scale parameter modeling the rate of decline, and α_{5i} allows for increasing HIV RNA trajectory after day 50. The smooth (mean) curve for the observed data in Figure 3 (right Panel) agrees with the postulated shape of the HIV RNA trajectory for this study. To force the parameters to be positive, we re-parameterize as follows: $\beta_{1i} = \log(\alpha_{1i}) = \beta_1 + b_{1i}$; $\beta_k = \log(\alpha_k), k = 2, 3, 4$ and $\alpha_{5i} = \beta_5 + b_{2i}$. Within the Bayesian framework, we use the Normal mean regression (N-NLMEC) considered by [16] and the QR-NLMEC with $p = 0.5$, where (b_{1i}, b_{2i}) are assumed to be an i.i.d. multivariate Normal distribution with the unrestricted scale matrix $\boldsymbol{D}$. The MCMC scheme was similar to the previous application for the UTI data, as well as the procedures described in Section 3. We further consider $\boldsymbol{D} \sim IWish_2(\mathbf{T}^{-1}, 2)$ with $\boldsymbol{T} = \text{Diag}(0.01, 0.01)$.

Table 2 gives the estimates for the different parameters in the QR-NLMEC for $p = 0.5$ (median case) and the N-NLMEC (mean regression). From Table 2, we observe that the estimates of the slope parameters β_2 and β_4 for the median regression model are somewhat different than the mean regression model and the standard errors of the QR–NLMEC are smaller, indicating that the median regression seem to produce more precise estimates. Residual plots in our analysis (omitted for brevity) revealed no serial correlations.

As in the linear case, to obtain a more complete picture of the effects, a series of QR models over the grid

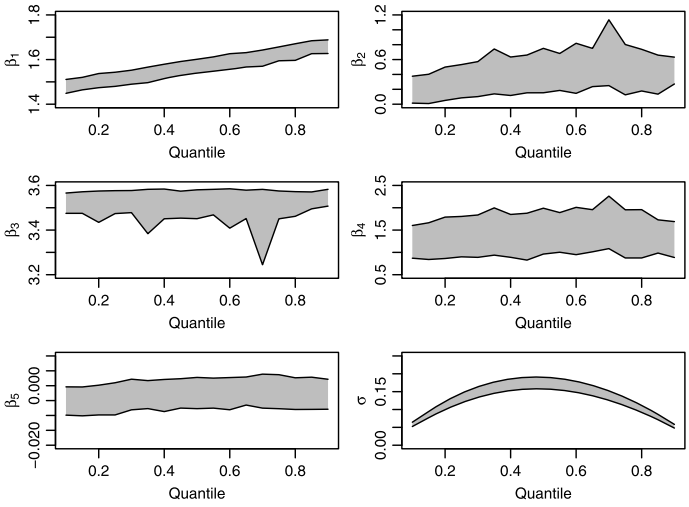

Figure 5. AIEDRP data: Posterior means and 95% credible intervals for various values of p.

$p = \{0.1, 0.15, \ldots, 0.9\}$ are estimated. Figure 5 gives a graphical summary of this analysis. The solid lines are the $Q_{0.025}$ percentile and the $Q_{0.975}$ percentile obtained from the marginal posterior distributions of the different parameters. Thus, the shaded area depicted the 95% credible band from the marginal posterior distribution. From Figure 5, we can see that the effect β_1 and β_2 become stronger as the value of the conditional quantile p increases, and on the other hand, β_3, β_4 and β_5 have constat effects on the HIV viral load (in log10 scale).

6. SIMULATION STUDY

In this section, we conduct a simulation study to examine the performance of our proposed methodology concerning parameter recovery. Similarly to [8] and [12] and differently from [21], our goal is to measure the accuracy of the proposed algorithm and model, in terms of parameter recovery, under unfavourable scenarios. That is, we simulate the observations using error distributions and quantile levels

Table 3. SE, Bias and RMSE for (β_0, β_1) based on R = 100 Monte Carlo replicas

Distr.	Perc. (%)	Quantile (%)	β_0			β_1		
			SE	Bias	RMSE	SE	Bias	RMSE
Normal	5	25	0.164	-0.283	0.327	0.021	-0.01	0.023
		50	0.162	0.025	0.164	0.021	-0.006	0.022
		75	0.167	0.343	0.381	0.02	-0.005	0.021
	10	25	0.161	-0.287	0.329	0.021	-0.008	0.022
		50	0.154	0.031	0.158	0.021	-0.006	0.022
		75	0.159	0.351	0.385	0.022	-0.003	0.022
	15	25	0.156	-0.268	0.310	0.023	-0.014	0.027
		50	0.156	0.049	0.163	0.024	-0.012	0.027
		75	0.17	0.371	0.408	0.023	-0.009	0.025
Student t	5	25	0.190	-0.291	0.347	0.020	-0.009	0.022
		50	0.188	0.027	0.190	0.020	-0.007	0.021
		75	0.188	0.349	0.396	0.020	-0.005	0.020
	10	25	0.179	-0.267	0.322	0.023	-0.012	0.026
		50	0.178	0.048	0.184	0.022	-0.01	0.024
		75	0.18	0.365	0.407	0.024	-0.008	0.025
	15	25	0.166	-0.274	0.320	0.025	-0.017	0.030
		50	0.164	0.047	0.171	0.022	-0.015	0.027
		75	0.171	0.369	0.406	0.023	-0.013	0.027
χ^2	5	25	0.169	-0.280	0.327	0.02	-0.007	0.022
		50	0.163	0.030	0.165	0.02	-0.005	0.021
		75	0.159	0.339	0.375	0.02	-0.004	0.021
	10	25	0.166	-0.299	0.342	0.023	-0.016	0.028
		50	0.157	0.018	0.158	0.021	-0.013	0.025
		75	0.169	0.331	0.372	0.021	-0.011	0.023
	15	25	0.168	-0.244	0.296	0.021	-0.014	0.025
		50	0.162	0.084	0.183	0.02	-0.012	0.023
		75	0.167	0.417	0.449	0.021	-0.010	0.023
Mixture	5	25	0.156	-0.309	0.346	0.022	-0.009	0.024
		50	0.151	0.008	0.151	0.021	-0.007	0.022
		75	0.161	0.326	0.364	0.022	-0.004	0.023
	10	25	0.164	-0.297	0.339	0.024	-0.007	0.025
		50	0.158	0.018	0.159	0.023	-0.005	0.024
		75	0.164	0.340	0.377	0.023	-0.002	0.024
	15	25	0.141	-0.277	0.310	0.021	-0.008	0.023
		50	0.138	0.035	0.143	0.021	-0.006	0.022
		75	0.146	0.354	0.382	0.022	-0.005	0.023

different from the ones that we consider in the model estimation. Since the estimates under these adverse scenarios are shown to be relatively accurate, we expect to observe even more accurate results when the simulated and estimated model are the same. Also, for the sake of simplicity, we decided not to perform a simulation study for a nonlinear model.

We considered the following regression model:

$$y_{ij} = -2.83 - 0.18x_{1ij} + 0.50x_{2ij} + b_{1i}z_{1ij} + b_{2i}z_{2ij} + \epsilon_{ij}$$
$$i = 1, 2, ..., 50,\ j = 1, ..., 6.$$

where $(b_{1i}, b_{2i}) \overset{i.i.d.}{\sim} N_2[\binom{0}{0}, \left(\begin{smallmatrix} 0.49 & 0.01 \\ 0.01 & 0.02 \end{smallmatrix}\right)]$ and $\xi \overset{i.i.d.}{\sim} 0.15D(\nu)$, with $D(\nu)$ being a suitable distribution (as we will explain ahead). We examined different scenarios produced by crossing the levels of two factors: the percentual of censored response (PCR) and the error distribution (ED), which corresponds to the term $D(\nu)$). For PCR we considered (5%, 10%, 15%) and for ED we considered $N(0,1), t_{(4)}, \chi^2_{(4)}, 0.5N(2, 0.36) + 0.5N(-2, 0.36)$ namely, henceforth, Normal, Student t, χ^2 and mixture. Therefore, we have a total of 12 scenarios. For each of these scenarios, we generated R = 100 replicas (responses) according to model (24) and we estimated the model parameters, considering the quantiles 0.25, 0.50 and 0.75, by using the MCMC algorithm presented in Subsection 3.1. The following priors were specified: $\beta_i \overset{i.i.d}{\sim} N(0{,}100)$, $\sigma^{-1} \sim U(0{,}100)$ and $\boldsymbol{D} \sim$ Wishart$(\boldsymbol{\Omega}, 2)$, where $\boldsymbol{\Omega} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. For the four scenarios, we computed the standard error (SE), the bias and the square root of the mean square error (RMSE), for each parameter over the 100 replicas. They are defined as:

Table 4. SE, Bias and RMSE for (β_2, σ) based on R = 100 Monte Carlo replicas

Distr.	Perc. (%)	Quantile (%)	β_2			σ		
			SE	Bias	RMSE	SE	Bias	RMSE
Normal	5	25	0.210	0.006	0.21	0.015	-0.009	0.017
		50	0.216	0.013	0.216	0.019	0.035	0.039
		75	0.224	0.008	0.224	0.015	-0.008	0.017
	10	25	0.231	0.005	0.231	0.014	-0.010	0.017
		50	0.235	0.002	0.235	0.018	0.035	0.040
		75	0.243	0.005	0.243	0.015	-0.007	0.017
	15	25	0.227	0.004	0.227	0.014	-0.012	0.018
		50	0.236	-0.001	0.236	0.018	0.034	0.039
		75	0.246	-0.017	0.246	0.014	-0.008	0.016
Student t	5	250	0.246	-0.011	0.247	0.014	-0.009	0.016
		50	0.247	-0.009	0.247	0.017	0.036	0.040
		75	0.250	-0.017	0.251	0.013	-0.008	0.015
	10	25	0.258	-0.012	0.259	0.013	-0.012	0.018
		50	0.260	-0.012	0.260	0.017	0.032	0.036
		75	0.272	-0.016	0.273	0.014	-0.01	0.017
	15	25	0.245	-0.013	0.246	0.016	-0.011	0.019
		50	0.235	-0.014	0.235	0.022	0.036	0.042
		75	0.240	-0.018	0.241	0.019	-0.007	0.020
χ^2	5	25	0.226	-0.005	0.226	0.013	-0.012	0.018
		50	0.228	-0.003	0.228	0.017	0.031	0.036
		75	0.236	0.004	0.236	0.013	-0.011	0.017
	10	25	0.221	-0.001	0.221	0.013	-0.009	0.016
		50	0.211	0.008	0.211	0.018	0.037	0.042
		75	0.229	0.026	0.23	0.015	-0.006	0.016
	15	25	0.237	-0.030	0.239	0.014	-0.011	0.018
		50	0.223	-0.045	0.227	0.019	0.036	0.041
		75	0.235	-0.06	0.243	0.015	-0.006	0.017
Mixture	5	25	0.203	-0.004	0.203	0.014	-0.009	0.017
		50	0.201	-0.004	0.201	0.019	0.036	0.041
		75	0.218	-0.009	0.218	0.016	-0.007	0.018
	10	25	0.219	0.010	0.219	0.012	-0.009	0.015
		50	0.218	0.019	0.219	0.015	0.036	0.039
		75	0.225	0.017	0.226	0.013	-0.007	0.014
	15	25	0.213	-0.027	0.215	0.012	-0.011	0.016
		50	0.205	-0.010	0.205	0.016	0.036	0.039
		75	0.218	¡0.001	0.218	0.013	-0.006	0.015

$$SE(\gamma) = \sqrt{\frac{1}{99}\sum_{i=1}^{100}\left(\widehat{\gamma}_i - \overline{\widehat{\gamma}}\right)^2}\,;\; Bias(\gamma) = (\overline{\widehat{\gamma}} - \gamma)\,;$$

$$RMSE(\gamma) = \sqrt{SE(\gamma)^2 + Bias(\gamma)^2}\,;\; \overline{\widehat{\gamma}} = \frac{1}{100}\sum_{i=1}^{100}\widehat{\gamma}_i$$

γ denotes each of (β_0, β_1, β_2, σ^2, D_{11}, D_{12}, and D_{22}), $\boldsymbol{D} = \left(\begin{smallmatrix} D_{11} & D_{12} \\ D_{12} & D_{22} \end{smallmatrix}\right)$, $\widehat{\gamma}_i$ is the estimate (the posterior expectation) obtained in replica i and γ is the respective true value. The results are summarized in Tables 3 to 6. It can be seen that the most accurate results are obtained when the error distribution used to simulate the responses matches the distribution used to obtain the Bayesian estimates. Also, the higher the PCR is the less precise the estimates are. In addition, the results when the median is the quantile of interest are more accurate compared to the results related to the other quantiles. All these results agree with our expectations.

Figure 6 presents the box-plots of the estimates obtained in each one of the 100 replicas related to β_0. The horizontal line corresponds to the true value. The labels along the x-axis indicate the distribution and percentual of censored response. For example, N10, indicates that the normal distribution was considered for simulating the error distribution with 10% of censored response, and so on. Again, we can see that the more accurate results are obtained for the normal distribution with 5% of censored response when the median is the quantile of interest. Also, we can see that the estimates, when the median is the quantile of interest, are approximately unbi-

Table 5. SE, Bias and RMSE for (D_{11}, D_{12}) based on R = 100 Monte Carlo replicas

Distr.	Perc. (%)	Quantile (%)	D_{11}			D_{12}		
			SE	Bias	RMSE	SE	Bias	RMSE
		50	0.149	0.062	0.161	0.016	-0.011	0.019
		75	0.156	0.096	0.184	0.017	-0.013	0.021
	15	25	0.166	0.109	0.198	0.019	-0.013	0.023
		50	0.145	0.044	0.152	0.018	-0.010	0.021
		75	0.163	0.071	0.177	0.020	-0.012	0.024
Student t	5	25	0.147	0.092	0.173	0.017	-0.012	0.021
		50	0.130	0.055	0.141	0.017	-0.011	0.020
		75	0.136	0.092	0.164	0.017	-0.012	0.021
	10	25	0.152	0.077	0.170	0.016	-0.013	0.020
		50	0.124	0.021	0.126	0.014	-0.011	0.018
		75	0.155	0.058	0.166	0.016	-0.013	0.021
	15	25	0.145	0.079	0.165	0.019	-0.014	0.024
		50	0.120	0.025	0.122	0.017	-0.012	0.021
		75	0.152	0.072	0.169	0.018	-0.015	0.023
		50	0.125	0.040	0.131	0.013	-0.012	0.017
		75	0.198	0.104	0.223	0.021	-0.017	0.027
	15	25	0.146	0.079	0.166	0.019	-0.015	0.024
		50	0.127	0.028	0.130	0.018	-0.013	0.022
		75	0.134	0.074	0.153	0.019	-0.016	0.024
Mixture	5	25	0.133	0.068	0.149	0.017	-0.013	0.022
		50	0.125	0.025	0.128	0.016	-0.011	0.020
		75	0.159	0.072	0.175	0.018	-0.013	0.022
	10	25	0.146	0.095	0.174	0.017	-0.010	0.020
		50	0.134	0.045	0.141	0.016	-0.008	0.018
		75	0.161	0.081	0.180	0.017	-0.011	0.020
	15	25	0.139	0.087	0.164	0.016	-0.013	0.021
		50	0.127	0.037	0.132	0.016	-0.010	0.019
		75	0.159	0.072	0.175	0.018	-0.012	0.021

ased, whereas the parameter is underestimated and overestimated when the quantiles of interest are 0.25 and 0.75, respectively. For the other parameters, in general, the results were quite similar. Therefore, those plots were not presented.

7. CONCLUSIONS

In this paper, we have considered Bayesian quantile regression for censored mixed effects models with the likelihood function based on the asymmetric Laplace distribution. The use of the asymmetric Laplace distribution makes it easy to implement the Bayesian inference based on the posterior distributions of parameters of interest via Gibbs sampling. We apply our methodology to a recent AIDS study (freely downloadable from R) to illustrate how the procedure developed can be used to obtain robust parameter estimates when the distribution assumptions are questionable. Depending on assay quantifications, censoring can be both left or right. Our application is based on right-censoring, consideration for left-censoring is immediate and follows from (10) by reversing the role of y_{ij} and Q_{ij}. We believe that this paper provides a first attempt to incorporate censoring in the context of Quantile regression mixed-effects models (QR-LMEC/NLMEC) and thus, our method provides an improvement over the one of [16], who considered analysis of these datasets using normal LMEC/NLMEC models. The models can be fitted using standard available software packages, such as R and WinBUGS (code available upon request) and hence can be easily accessible to practitioners in the field.

The models developed here do not consider skewness in the random effects because typically in HIV-AIDS studies, the responses (censored viral load) is log transformed to achieve a "close to normality" shape. Recently, [10] adopted a Markov chain Monte Carlo approach to draw Bayesian inferences in Linear mixed models with multivariate skew-normal (SNI) distributions for both random effects and error terms. Therefore, it would be a worthwhile task to investigate the applicability of a likelihood based treatment in the context of QR-LMEC/NLMEC models with SNI distributions. Incorporating measurement error models [18] within our robust framework for related HIV viral load covariates (namely, CD4 cell counts) is also topic of our future research.

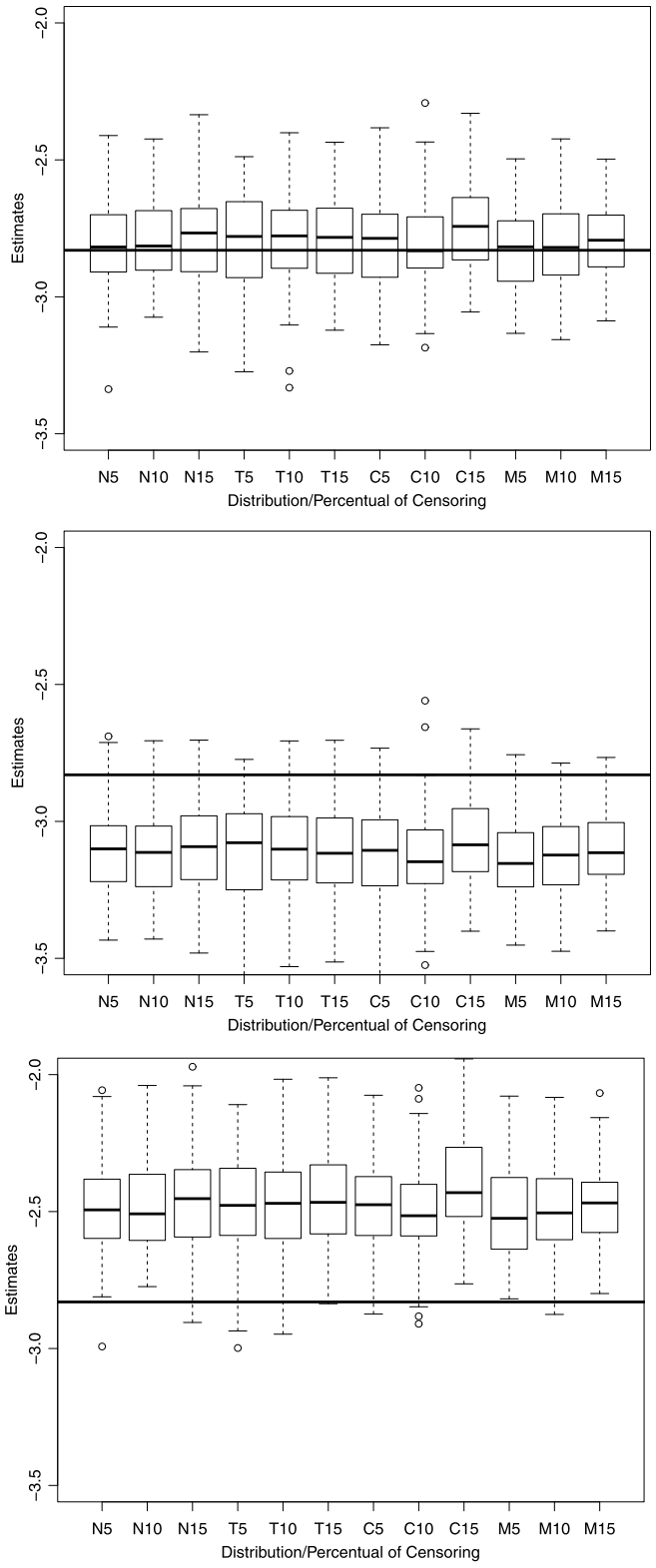

Figure 6. Box-plots of the estimates along the 100 replicas, considering the combinations of the levels of the factors, for the parameter β_0, when the (top) median is being modeling (middle) the quantile $p = 0.25$ is being modeling and (botom) when the quantile $p = 0.75$ is being modeling.

Table 6. SE, Bias and RMSE for D_{22} based on $R = 100$ Monte Carlo replicas

Distr.	Perc. (%)	Quantile (%)	D_{22}		
			SE	Bias	RMSE
Normal	5	25	0.004	0.024	0.024
		50	0.004	0.024	0.024
		75	0.004	0.024	0.024
	10	25	0.006	0.026	0.026
		50	0.005	0.025	0.026
		75	0.005	0.025	0.026
	15	25	0.005	0.028	0.028
		50	0.005	0.027	0.028
		75	0.005	0.027	0.028
Student	5	25	0.005	0.024	0.025
		50	0.005	0.024	0.024
		75	0.005	0.024	0.025
	10	25	0.005	0.025	0.026
		50	0.005	0.025	0.025
		75	0.005	0.025	0.026
	15	25	0.005	0.028	0.028
		50	0.005	0.027	0.028
		75	0.005	0.027	0.028
χ^2	5	25	0.005	0.024	0.024
		50	0.005	0.024	0.024
		75	0.005	0.024	0.025
	10	25	0.005	0.025	0.026
		50	0.004	0.025	0.025
		75	0.005	0.026	0.026
	15	25	0.005	0.028	0.028
		50	0.005	0.027	0.028
		75	0.005	0.027	0.028
Mixture	5	25	0.005	0.025	0.025
		50	0.005	0.024	0.025
		75	0.004	0.025	0.025
	10	25	0.004	0.027	0.027
		50	0.005	0.027	0.027
		75	0.005	0.027	0.027
	15	25	0.006	0.027	0.028
		50	0.006	0.027	0.027
		75	0.006	0.027	0.028

ACKNOWLEDGEMENTS

The authors are grateful to the Editor, Associate Editor and an anonymous referee for numerous suggestions that significantly improved the paper.

Received 7 November 2013

REFERENCES

[1] Barndorff-Nielsen, O. E. and Shephard, N. [2001], "Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2), 167–241. MR1841412

[2] Gelman, A., Carlin, J., and Rubin, D. [2006], *Bayesian Data Analysis*, New York, NY: Chapman & Hall/CRC.

[3] Hobert, J. and Casella, G. [1996], "The effect of improper priors on Gibbs sampling in hierarchical linear mixed models," *Journal of the American Statistical Association*, 91(436). MR1439086
[4] Hughes, J. [1999], "Mixed effects models with censored data with application to HIV RNA levels," *Biometrics*, 55(2), 625–629.
[5] Koenker, R. [2005], *Quantile Regression*, Vol. 38, Cambridge University Press. MR2268657
[6] Koenker, R. and Machado, J. [1999], "Goodness of fit and related inference processes for quantile regression," *Journal of the American Statistical Association*, 94(448), 1296–1310. MR1731491
[7] Kotz, S., Kozubowski, T. and Podgorski, K. [2001], *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*, Vol. 183, Boston: Birkhöuser. MR1935481
[8] Kozumui, H. and Kobayashi, G. [2011], "Gibbs sampling methods for Bayesian quantile regression," *Journal of Statistical Computation and Simulation*, 81(11), 1565–1578. MR2851270
[9] Lachos, V., Bandyopadhyay, D., and Dey, D. [2011], "Linear and nonlinear mixed-effects models for censored HIV viral loads using normal/independent distributions," *Biometrics*, 67(4), 1594–1604. MR2872410
[10] Lachos, V. H., Dey, D. K., and Cancho, V. G. [2009], "Robust linear mixed models with skew-normal independent distributions from a Bayesian perspective," *Journal of Statistical Planning and Inference*, 139, 4098–4110. MR2558353
[11] Liu, C. [1996], "Bayesian robust multivariate linear regression with incomplete data," *Journal of the American Statistical Association*, 91, 1219–1227. MR1424619
[12] Luo, Y. L. and Tian, M. [2012], "Bayesian quantile regression for longitudinal models," *Journal of Statistical Computation and Simulation*, 82(11), 1635–1649. MR2984566
[13] Matos, L. A., Prates, M. O., Chen, M.-H., and Lachos, V. [2013], "Likelihood-based inference for mixed-effects models with censored response using the multivariate-t distribution," *Statistica Sinica*, 23, 1299–1322. MR3114716
[14] Saitoh, A., Foca, M., Viani, R., Heffernan-Vacca, S., Vaida, F., Lujan-Zilbermann, J., Emmanuel, P., Deville, J., and Spector, S. [2008], "Clinical outcomes after an unstructured treatment interruption in children and adolescents with perinatally acquired HIV infection," *Pediatrics*, 121(3), e513–521.
[15] Vaida, F., Fitzgerald, A., and DeGruttola, V. [2007], "Efficient hybrid EM for linear and nonlinear mixed effects models with censored response," *Computational Statistics & Data Analysis*, 51(12), 5718–5730. MR2407672
[16] Vaida, F. and Liu, L. [2009], "Fast implementation for normal mixed effects models with censored response," *Journal of Computational and Graphical Statistics*, 18(4), 797–817. MR2750442
[17] Wu, H. [2005], "Statistical methods for HIV dynamic studies in AIDS clinical trials," *Statistical Methods in Medical Research*, 14(2), 171. MR2135921
[18] Wu, L. [2010], *Mixed Effects Models for Complex Data*, Boca Raton, FL: Chapman & Hall/CRC. MR2598844
[19] Yu, K. and Moyeed, R. [2001], "Bayesian quantile regression," *Statistics & Probability Letters*, 54(4), 437–447. MR1861390
[20] Zhao, Y., Staudenmayer, J., Coull, B., and Wand, M. [2006], "General design Bayesian generalized linear mixed models," *Statistical Science*, 21(1), 35–51. MR2275966
[21] Zhou, Y.-H., xin Ni, Z., and Li, Y. [2014], "Quantile regression via the EM algorithm," *Communications in Statistics: Simulation and Computation*, 43(10), 2162–2172. MR3223660

Victor H. Lachos
Department of Statistics
Campinas States University
Rua Sergio Buarque de Holanda, 651
Cidade Universitária-Barão Geraldo
CEP: 13083-859
Campinas, SP
Brazil
E-mail address: hlachos@ime.unicamp.br

Ming-Hui Chen
Department of Statistics
University of Connecticut
215 Glenbrook Rd, U-4120
Storrs, CT 06269
USA
E-mail address: ming-hui.chen@uconn.edu

Carlos A. Abanto-Valle
Departament of Statistics
Federal University of Rio de Janeiro
Caixa Postal 68530
CEP: 21945-970
Rio de Janeiro
Brazil
E-mail address: cabantovalle@im.ufrj.br

Caio L. N. Azevedo
Department of Statistics
Campinas States University
Rua Sergio Buarque de Holanda, 651
Cidade Universitária-Barão Geraldo
CEP: 13083-859
Campinas, SP
Brazil
E-mail address: cnaber@ime.unicamp.br