Published in final edited form as:

Stat Methods Med Res. 2019 ; 28(10-11): 3415–3436. doi:10.1177/0962280218801147.

Bayesian multivariate skew meta-regression models for individual patient data

Joseph G Ibrahim¹, Sungduk Kim², Ming-Hui Chen³, Arvind K Shah⁴, Jianxin Lin⁴

¹Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

²Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA

³Department of Statistics, University of Connecticut, Storrs, CT, USA

⁴Clinical Biostatistics, Merck Research Laboratories, Rahway, NJ, USA

Abstract

We examine a class of multivariate meta-regression models in the presence of Individual Patient Data (IPD). The methodology is well motivated from several studies of cholesterol lowering drugs where the goal is to jointly analyze the multivariate outcomes, Low Density Lipoprotein Cholesterol (LDL-C), High Density Lipoprotein Cholesterol (HDL-C), and Triglycerides (TG). These three continuous outcome measures are correlated and shed much light on a subject's lipid status. One of the main goals in lipid research is the joint analysis of these three outcome measures in a meta-regression setting. Since these outcome measures are not typically multivariate normal, one must consider classes of distributions that allow for skewness in one or more of the outcomes. In this paper, we consider a new general class of multivariate skew distributions for multivariate meta-regression and examine their theoretical properties. Using these distributions, we construct a Bayesian model for the meta-data and develop an efficient Markov chain Monte Carlo (MCMC) computational scheme for carrying out the computations. In addition, we develop a multivariate L measure for model comparison, Bayesian residuals for model assessment, and a Bayesian procedure for detecting outlying trials. The proposed multivariate L measure, Bayesian residuals, and Bayesian outlying trial detection procedure are particularly suitable and computationally attractive in the multivariate meta-regression setting. A detailed case study demonstrating the usefulness of the proposed methodology is carried out in an IPD multivariate meta-regression setting using 26 pivotal Merck clinical trials that compare statins (cholesterol lowering drugs) in combination with ezetimibe and statins alone on treatment-naïve patients and those continuing on statins at baseline.

Corresponding author: Joseph G. Ibrahim, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA, ibrahim@bios.unc.edu.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Supplemental material

Supplementary materials for Bayesian Multivariate Skew Meta-Regression Models for Individual Patient Data (SMVMeta2017SMMRSupplemental.pdf). This supplement consists of the posterior and full conditional distributions and additional tables of posterior estimates under the Normal, t , and Skew t models with a diagonal for the cholesterol data.

Keywords

Bayesian inference; heterogeneity; multivariate L measure; multi-dimensional random effects; multiple trials; outlying trials

1 Introduction

An estimated 12 million Americans are taking cholesterol-lowering drugs and experts are recommending many millions more should be taking them. Cholesterol levels are on the rise and many medical and government organizations have recommended aggressive lipid-lowering goals in patients with established coronary heart disease or those at high risk of developing cardiovascular diseases. It is now well known that Low Density Lipoprotein Cholesterol, abbreviated LDL-C, (i.e., “bad” cholesterol) is a causative factor for coronary heart disease and is the main target for cholesterol and lipid lowering therapy. This fact has been established in many studies around the world.

In lipid research, three common outcome measures are LDL-C, High Density Lipoprotein Cholesterol (HDL-C), and Triglycerides (TG). These three continuous outcomes measures are correlated and shed much light on a subject’s lipid status. Thus, one of the main goals in lipid research is the joint analysis of these three outcome measures. Research in Kim et al.¹ has shown that these outcome measures are typically not jointly normally distributed and the joint distribution of these outcome variables is in fact skewed. Thus, an analysis assuming a multivariate normal distribution may not be valid. In addition to accommodating skewness, another challenge for analyzing such data in meta-regression is that one must model the between-trial heterogeneity as well as the within-trial covariance structure for the multivariate response vector.

Meta-regression of individual patient data (IPD) is a useful and effective statistical tool for synthesizing evidence across studies and improving investigation and explanation of heterogeneity. Examples of applications of meta-regression of IPD in different fields exist in literature and can be found in publications by Edwards et al.², Gorman et al.³, and Pereira et al.⁴. Some of the notable methodological work in this area is done by Simmonds and Higgins⁵, Ritz et al.⁶, and Riley et al.⁷. The literature on multivariate skew distributions for meta-regression is essentially non-existent.

Skew models have been used quite extensively in the literature for univariate and multivariate response data in non meta-analysis settings^{8–16}. Such models have proved to be quite powerful for modeling skew distributions. In this paper, we develop multivariate meta-analysis methods for IPD based on jointly modeling the three outcome variables (LDL-C, HDL-C, TG). Since the three outcome variables cannot be assumed to be jointly multivariate normal, we consider a general multivariate skew meta-regression model that accommodates for skewness in each of the outcome variables and allows for random effects and covariates for modeling between trial heterogeneity, as well as inducing a covariance structure on the multivariate response. The model is quite general and flexible in that (i) the model is a multivariate random effects model that captures several sources of between-study variation involving treatments, while simultaneously accommodating both patient-level and study-

level covariates; (ii) it allows for different skewness parameters for different outcome variables; (iii) similar to Adcock¹³ and Chang and Zimmerman¹⁶, our model allows the correlations of multivariate outcome variables to depend on skewness parameters in addition to the covariance matrix of these outcome variables; and (iv) it can also account for different heavy-tailed distributions for the error terms via the specification of a mixing distribution. The proposed model includes multivariate normal meta-regression models, multivariate skew normal meta-regression models; and multivariate skew t meta-regression models as special cases. Compared to Kim et al.¹ who used a Box-Cox transformation on the response variables, our proposed multivariate skew t meta-regression model is more attractive since the point estimates are more directly comparable and easier to interpret than those under a Box-Cox transformation model and the proposed model requires much less computational time. Estimation of the parameters requires a sophisticated and computationally intensive Markov chain Monte Carlo (MCMC) sampling algorithm. To the best of our knowledge, multivariate skew regression models have not been used in the meta-analysis setting. In addition to developing a novel model for multivariate meta-regression with IPD, we also develop a multivariate analog of the L measure for model comparison proposed in Ibrahim and Laud¹⁷ and Laud and Ibrahim,¹⁸ Bayesian residuals for assessment of overall fit, and a new Bayesian procedure for detecting outlying trials based on the posterior distributions of the mixing variables in the proposed multivariate t meta-regression model. The multivariate L measure is shown to have an attractive decomposition and can be efficiently computed via MCMC methods. Bayesian residuals are easy to compute and provide an alternative to the L measure for model assessment. Our proposed Bayesian procedure for detecting outlying trials is analogous to the approach discussed in Johnson¹⁹ for checking the second level of a hierarchical model. The proposed procedure is quite attractive as it allows for different sample sizes across trials and is empirically shown to be quite promising in both simulated and real data.

The rest of this paper is organized as follows. In Section 2, we present a detailed description of the cholesterol meta-data that has motivated the proposed model. In Section 3, we present a general development of the multivariate skew meta-regression model and discuss several of its attractive properties. We also give a general discussion of the random effects structure, its implications, and the induced covariance structures of the proposed model. In Section 4, we lay out, in detail, the computational development for sampling from the posterior distribution along with the full conditionals needed for MCMC sampling based on the proposed model, provide a detailed development of the multivariate L measure, and examine its decomposition and theoretical properties. In this section, we also develop Bayesian residuals and a method for detecting outlying trials. The empirical performance of the proposed Bayesian diagnostic procedure is investigated in Section 5.1. In Section 5.2, a simulation study is conducted to empirically compare the proposed model to the Box-Cox transformation model¹. In Section 6, we carry out a detailed analysis of the cholesterol meta-data. We conclude the article with a brief discussion in Section 7.

2 The Cholesterol Data

The IPD is based on 26 Merck-sponsored double-blind, randomized, active or placebo-controlled clinical trials on adult patients with primary hypercholesterolemia. The primary

goal of these clinical trials was to evaluate the LDL-C lowering effects of ezetimibe (which works in the digestive tract) in combination with statin (which works in the liver) in comparison with statin alone on treatment-naïve patients at baseline (on a first-line therapy) or patients who underwent washout of previous lipid-modifying therapy at baseline (on a second-line therapy). In our analyses, different statins and their doses are combined to form the ‘statin’ and ‘statin + ezetimibe’ treatment groups. Ezetimibe (EZE) is available at only one dose of 10 mg, and the statins used in these trials included simvastatin, atorvastatin, lovastatin, rosuvastatin, pravastatin, and fluvastatin. The covariates include treatment (trt: 0 = statin and 1 = statin + EZE), baseline LDL-C (bl_ldlc), baseline HDL-C (bl_hdlc), baseline TG (bl_trig), age, race (White (reference), Black, Hispanic, and Other), gender (Female: 0 = male (reference), 1 = 1 female), diabetes (DM: 0 = No, 1 = Yes), coronary heart disease (CHD: 0 = No, 1 = Yes), body mass index (BMI), statin potency (low (reference), med (potency2), and high (potency3)), and trial duration (duration) (6 – 12 weeks). Statin potency simply describes the chemical/medicinal strength of the statin and it was categorized into three potency classes (low, medium and high) considered in this paper. In this analysis, we include only the patients whose outcome variables and covariates were available. There were a total of 21,694 patients in total, of which 510 patients had missing values in either outcome variables or covariates including baseline LDL-C, baseline HDL-C, baseline TG, CHD, statin potency, and BMI. There were no missing values for age, race, gender, DM, and trial duration. BMI had the most missing values (336, 1.55%).

The IPD considered in our analyses are a subset of the meta-data published in Leiter et al.²⁰. The citations of the primary papers that were published in clinical journals for the 26 trials considered here can be also found in Leiter et al.²⁰ Leiter et al.²⁰ carried out a meta-analysis based on the pooled data. Kim et al.¹ provided a detailed summary of the covariates for these 26 clinical trials in Tables I and II. From these tables, we can see a considerable amount of heterogeneity in the covariates across the trials. Specifically, the ranges of the within-trial means of the continuous covariates are (89.2, 186.0), (43.1, 55.3), (127.0, 199.5), (52.3, 71.2), and (27.2, 33.6) for baseline LDL-C, baseline HDL-C, baseline TG, age, and BMI, respectively. We also see drastically different proportions of the categorical covariates across the trials. For example, almost every patient except for one in trial 9 had DM. Also, there was only medium statin potency in trial 13, whereas there were no low or high statin potencies in some other trials.

This descriptive summary shows that to examine the treatment effects, there is a need to adjust for these covariates. We consider three primary outcome variables including percent changes from baseline in LDL-C, HDL-C, and TG. For ease of presentation, we simply denote these three outcome variables by LDL-C, HDL-C, and TG.

3 Multivariate Skew Meta-Regression Models with Multi-Dimensional Random Effects

Consider K randomized trials, where each trial has two treatment arms (“Statin” or “Statin + EZE”), and patients in each trial were either all on statin or all not on statin prior to the trial. The sample size of the individual patient data for the k^{th} trial is n_k . Let $y_{ik} = (y_{1ik}, \dots, y_{iik})'$

denote a J -dimensional vector of the responses for the i^{th} patient in the k^{th} trial. Also let $\text{trt}_{ik} = 1$ if the i^{th} patient received “Statin + EZE” and 0 if “Statin” alone, and $\text{onstatin}_k = 1$ if patients were on statin and 0 if not on statin prior to the trial. Furthermore, let x_{ijk} denote a p_j -dimensional vector of covariates for the j^{th} response corresponding to the i^{th} patient. We propose the following multivariate random effects regression model for the meta-analysis:

$$y_{ijk} = [\gamma_{jk0} + \gamma_{jk1}\text{trt}_{ik}](1 - \text{onstatin}_k) + [\gamma_{jk2} + \gamma_{jk3}\text{trt}_{ik}]\text{onstatin}_k + \delta_j(z_{ijk} - E[z_{ijk}]) + x'_{ijk}\beta_j + \epsilon_{ijk}, \quad (1)$$

where $\beta_j = (\beta_{j1}, \dots, \beta_{jpj})'$ is the vector of fixed effects regression coefficients corresponding to the p_j covariates. In (1), we assume $z_{i1k}, \dots, z_{iJk}$ are dependent. First, we assume $z_{ijk}|\psi_{ik}$ follows an exponential distribution with mean $\frac{1}{\psi_{ik}}$. Then, we assume $\psi_{ik} \sim \text{Gamma}(\tau + 1, \tau)$

with $E[\psi_{ik}] = \frac{\tau + 1}{\tau}$, and $\text{Var}(\psi_{ik}) = \frac{\tau + 1}{\tau^2}$. Under this assumption, we have

$$E[z_{ijk}] = E[E(z_{ijk}|\psi_{ik})] = E\left[\frac{1}{\psi_{ik}}\right] = 1 \text{ and } \text{Var}(z_{ijk}) = \frac{\tau + 1}{\tau - 1} \text{ for } \tau > 1. \text{ Write}$$

$z_{ik}^\delta = (z_{i1k}^\delta, \dots, z_{iJk}^\delta)' = (\delta_1(z_{i1k} - E[z_{i1k}]), \dots, \delta_J(z_{iJk} - E[z_{iJk}]))'$ Then, the covariance matrix of z_{ik}^δ is given by

$$\begin{aligned} \text{Var}(z_{ik}^\delta) &= E[\text{Var}(z_{ik}^\delta|\omega_{ik})] + \text{Var}(E[z_{ik}^\delta|\omega_{ik}]) \\ &= \frac{\tau}{\tau - 1} \text{diag}(\delta_1^2, \dots, \delta_J^2) + \frac{1}{\tau - 1} \delta \delta'. \end{aligned} \quad (2)$$

In (1), the variable z_{ijk} is latent, which mimics a random effect. As discussed in Chen et al.⁸, a nice feature of this approach is that the outcome variable y_{ijk} has a random-effects model structure and its distribution is skewed when a skewed distribution is assumed for the latent variable z_{ijk} . Furthermore, the amount of skewness is further controlled by a skewness parameter δ_j . We follow Kim et al.¹⁴ to subtract $E[z_{ijk}]$ from z_{ijk} so that δ_j is less confounded with the intercept term. We see from (2) that (i) the correlations of the z_{ijk}^δ depend on δ , which essentially implies that the correlations of the responses, y_{ijk} , depend on the skewness parameters, and (ii) the z_{ijk} 's become independent when $\tau \rightarrow \infty$.

In (1), we further assume that

$$\epsilon_{ik} | \lambda_{ik} = (\epsilon_{i1k}, \dots, \epsilon_{iJk})' | \lambda_{ik} \sim N(0, \lambda_{ik}^{-1} \Sigma), \quad (3)$$

independently, for $i = 1, \dots, n_k$ and $k = 1, \dots, K$, where $\Sigma = (\delta_{jj})$ is a $J \times J$ unstructured covariance matrix, the λ_{ik} 's are independent, and each follows

$$\lambda_{ik} \sim \text{Gamma}(v/2, v/2) \quad (4)$$

with a probability density function $h(\lambda_{ik}) = \frac{\left(\frac{v}{2}\right)^{\frac{v}{2}}}{\Gamma\left(\frac{v}{2}\right)} \lambda_{ik}^{v/2-1} \exp\{-(v/2) \lambda_{ik}\}$ if $\lambda_{ik} > 0$ and 0 otherwise for $i = 1, \dots, n_k$ and $k=1, \dots, K$. The joint density of ϵ_{ik} is given by

$$f(\epsilon_{ik} | \Sigma) = \int \frac{|\Sigma|^{-1/2}}{(2\pi)^{J/2}} \exp\left\{-\frac{\lambda_{ik}}{2} \epsilon'_{ik} \Sigma^{-1} \epsilon_{ik}\right\} \times \lambda_{ik}^{J/2} h(\lambda_{ik}) d\lambda_{ik}, \quad (5)$$

which is a t -distribution denoted by t_v with v degrees of freedom. In addition, the covariance matrix of ϵ_{ik} is given by

$$\text{Var}(\epsilon_{ik}) = E[\text{Var}(\epsilon_{ik} | \lambda_{ik})] = \Sigma \int_0^\infty \frac{1}{\lambda_{ik}} h(\lambda_{ik}) d\lambda_{ik} = \left(\frac{v}{v-2}\right) \Sigma \text{ for } v > 2. \quad (6)$$

From (5), we see that the main role of λ_{ik} is to introduce a heavy-tailed distribution to the multivariate responses y_{ik} . The use of λ_{ik} also facilitates a convenient implementation of the Markov chain Monte Carlo sampling algorithm in Section 4.3.

Let $\gamma_{jk} = (\gamma_{jk0}, \gamma_{jk1}, \gamma_{jk2}, \gamma_{jk3})'$ so that γ_{jk} represents the vector of random effects for the j^{th} response in (1). Let $z_{ik} = (z_{i1k}, \dots, z_{iJk})$. We assume that z_{ik} , ϵ_{ik} , and γ_{jk} are independent. We further assume

$$\gamma_{jk} \sim N_4(\gamma_j, \Omega_j), \quad (7)$$

where $\gamma_{jk} = (\gamma_{jk0}, \gamma_{jk1}, \gamma_{jk2}, \gamma_{jk3})'$ denotes the vector of the overall treatment effects for the j^{th} response. To ensure model identifiability, we assume

$$\Omega_j = \begin{pmatrix} \Omega_{j00} & \Omega_{j01} & 0 & 0 \\ \Omega_{j01} & \Omega_{j11} & 0 & 0 \\ 0 & 0 & \Omega_{j22} & \Omega_{j23} \\ 0 & 0 & \Omega_{j23} & \Omega_{j33} \end{pmatrix} = \begin{pmatrix} \Omega_j^1 & 0 \\ 0 & \Omega_j^2 \end{pmatrix}. \quad (8)$$

In (8), Ω_{j00} and Ω_{j11} capture the variability in γ_{jk0} and γ_{jk1} , and Ω_{j01} captures the correlation between γ_{jk0} and γ_{jk1} among the trials in which patients were not on statin; and similarly, Ω_{j22} and Ω_{j33} capture the variability in γ_{jk2} and γ_{jk3} , and Ω_{j23} captures the correlation between γ_{jk2} and γ_{jk3} among the trials in which patients were on statin. We

emphasize here that Ω_j captures only the response-specific correlations within each trial but it does not capture the correlations across the different responses.

The multivariate meta-regression model defined in (1), (3), (4), and (7) is general and flexible, including the multivariate normal meta-regression model, the multivariate t meta-regression model, and the multivariate skew t meta-regression model as special cases. Simultaneous estimation of $(\beta_1, \dots, \beta_J, \Sigma, \gamma_1, \dots, \gamma_J, \Omega_1, \dots, \Omega_J, \delta_1, \dots, \delta_J, \nu)$ is not easy and requires a sophisticated and computationally intensive MCMC sampling algorithm.

4 Bayesian Inference

4.1 The Complete-Data Likelihood Function

Write $T_{ik} = (1 - \text{onstatin}_k, \text{trt}_{ik}(1 - \text{onstatin}_k), \text{onstatin}_k, \text{trt}_{ik}\text{onstatin}_k)'$. Let $X_{ik} = \text{diag}(x'_{ik1}, \dots, x'_{ikJ})$, $U_{ik} = \text{diag}(T'_{ik}, \dots, T'_{ik}) = (\beta'_1, \dots, \beta'_J)'$, $\gamma = (\gamma'_1, \dots, \gamma'_J)'$, $\delta = (\delta_1, \dots, \delta_J)$, $\gamma_k^R = (\gamma'_{1k}, \dots, \gamma'_{Jk})'$, and $\psi = (\psi_{11}, \dots, \psi_{n_k K})'$. Also let $y = (y'_{11}, \dots, y'_{n_k K})'$,

$U = (U_{11}, \dots, U_{n_k K})$ and $U = (U_{11}, \dots, U_{n_k K})$. Furthermore, we let $D_{\text{obs}} = (y, X, U)$ denote

observed data. Write $\Delta = \text{diag}(\delta_1, \dots, \delta_J)$, $\Omega = \text{diag}(\Omega_1, \dots, \Omega_J)$, $z = (z'_{11}, \dots, z'_{n_k K})'$,

$\gamma^R = ((\gamma_1^R)', \dots, (\gamma_K^R)')$ and $\lambda = (\lambda_{11}, \dots, \lambda_{n_k K})$. Then the complete-data likelihood function

is given by

$$\begin{aligned} L(\beta, \gamma, \Sigma, \Omega, \delta, \nu, \gamma^R, z, \psi, \lambda | D_{\text{obs}}) \\ = \prod_{k=1}^K \prod_{i=1}^{n_k} f(y_{ik} | \beta, \Sigma, \Delta, \gamma_k^R, z_{ik}, \lambda_{ik}, X_{ik}, U_{ik}) \times \frac{\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)} \lambda_{ik}^{\frac{\nu}{2}-1} \exp\left\{-\frac{\nu}{2} \lambda_{ik}\right\} \\ \times \prod_{k=1}^K \prod_{i=1}^{n_k} \left[\prod_{j=1}^J \psi_{ik} \exp(-\psi_{ik} z_{ijk}) \right] \times \frac{\tau^{\tau+1}}{\Gamma(\tau+1)} \psi_{ik}^{\tau} \exp\left\{-\tau \psi_{ik}\right\} \\ \times \prod_{k=1}^K \prod_{j=1}^J (2\pi)^{-\frac{4}{2}} \left| \Omega_j \right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} (\gamma_{jk} - \gamma_j)' \Omega_j^{-1} (\gamma_{jk} - \gamma_j)\right\} \end{aligned} \quad (9)$$

where $f(y_{ik} | \beta, \Sigma, \Delta, \gamma_k^R, z_{ik}, \lambda_{ik}, X_{ik}, U_{ik}) = \frac{|\Sigma|^{-1/2}}{(2\pi)^{J/2}} \lambda_{ik}^{J/2} \exp$ with

$$\left\{ -\frac{\lambda_{ik}}{2} (y_{ik} - X_{ik}\beta - U_{ik}\gamma_k^R - \Delta z_{ik}^*)' \Sigma^{-1} (y_{ik} - X_{ik}\beta - U_{ik}\gamma_k^R - \Delta z_{ik}^*) \right\}$$

$$z_{ik}^* = z_{ik} - E[z_{ik}]$$

4.2 Prior and Posterior

We assume that $\beta, \gamma, \Sigma, \Omega, \delta$, and ν are independent *a priori*. Thus, the joint prior for $(\beta, \gamma, \Sigma, \Omega, \delta, \nu)$ takes the form

$$\pi(\beta, \gamma, \Sigma, \Omega, \delta, \nu) = \pi(\beta)\pi(\gamma)\pi(\Sigma)\pi(\Omega)\pi(\delta)\pi(\nu). \quad (10)$$

We further assume $\beta \sim N_p(0, c_1 I_p)$, where $p = \sum_{j=1}^J p_j$, $\gamma \sim N_{4J}(0; c_2 I_{4J})$, $\delta \sim N_J(0; c_3 I_J)$, $\Sigma^{-1} \sim \text{Wishart}_J(d_0, S_0)$, $\Omega_j^{l-1} \sim \text{Wishart}_2(d_1, S_1)$, $l = 1, 2, j = 1, \dots, J$, and $\nu \sim \text{Gamma}(a_0, b_0)$. The quantities $c_1, c_2, c_3, d_0, S_0, d_1, S_1, a_0$, and b_0 are prespecified hyperparameters. Further, we have $\pi(\Sigma^{-1} | d_0, S_0) \propto |\Sigma^{-1}|^{(d_0 - J - 1)/2} \exp(-\frac{1}{2} \text{tr}(S_0^{-1} \Sigma^{-1}))$, $\pi(\Omega_j^{l-1} | d_1, S_1) \propto |\Omega_j^{l-1}|^{(d_1 - 2 - 1)/2} \exp(-\frac{1}{2} \text{tr}(S_1^{-1} \Omega_j^{l-1}))$, and $p(\nu | a_0, b_0) \propto \nu^{a_0 - 1} \exp(-b_0 \nu) 1(\nu > c_0)$. Using (9) and (10), the joint posterior distribution of all the parameters is given by

$$\begin{aligned} & \pi(\beta, \gamma, \Sigma, \Omega, \delta, \nu, \gamma^R, z, \psi, \lambda | D_{obs}) \\ & \propto L(\beta, \gamma, \Sigma, \Omega, \delta, \nu, \gamma^R, z, \psi, \lambda | D_{obs}) \times \pi(\beta, \gamma, \Sigma, \Omega, \delta, \nu). \end{aligned} \quad (11)$$

The hyperparameters of the prior in (10) were specified as $c_0 = 3$, $c_1 = 100$, $c_2 = 100$, $c_3 = 100$, $d_0 = J + 0.1$, $S_0 = 0.1$, $d_1 = 2.1$, $S_1 = 0.1$, $a_0 = 1$, and $b_0 = 0.1$ in the analysis. We chose $c_0 = 3$ to ensure the existence of $\text{Var}(\epsilon_{ik})$ in (6). The other choices of hyperparameters lead to noninformative priors. The flow diagram of the proposed model with prior specifications is given in Figure 1. Figure 1 shows the role of each set of variables, the relations between different variables within the proposed skew t meta regression models, and the hyperparameters in the prior specifications. Under the proposed model, most of the parameters except for τ , which controls the dependence among the skewness variables $z_{1k}, z_{2k}, \dots, z_{Jk}$, are identifiable under very mild conditions. For τ , we use the proposed multivariate L measure in Section 4.4 to select an “optimal” value. Due to the availability of the IPD with 26 trials, it is easy to see that the overall treatment effects γ_j and the regression coefficients β_j are identifiable as long as the design matrix is of full rank. The identifiability of the skewness parameters δ_j is discussed in Chen et al.⁸ and Kim et al.¹⁴. Due to a large number of the trials and the availability of the 26 large datasets, the variability of random effects j_k , namely, Ω_j and the dependence among these response variables, i.e., Σ , can be estimated.

4.3 Computational Development

We consider the following one-to-one transformations: $\gamma_k^{*R} = (\gamma_{1k}^{*'}, \gamma_{2k}^{*'}, \dots, \gamma_{Jk}^{*'})' = \gamma_k^R - \gamma$ for $k = 1, \dots, K$. Thus, $\gamma_{jk}^{*} = \gamma_{jk} - \gamma_j$ for $j = 1, \dots, J$ and $k = 1, \dots, K$. Write $\gamma^{*R} = ((\gamma_1^{*R})', \dots, (\gamma_K^{*R})')'$. Also, let $X_{ik}^{*} = (X_{ik}, U_{ik})$ for $i = 1, \dots, n$ and $k = 1, \dots, K$ and $\theta = (\beta', \gamma')'$. We present a detailed development of the MCMC sampling algorithm. Although the analytic evaluation of the joint posterior distribution of $(\theta, \Sigma, \Omega, \delta, \nu, \gamma^{*R}, z, \psi, \lambda)$ based on the observed data D_{obs} given in Equation (S.1) of the supplementary materials is

not possible, the proposed model allows us to develop an efficient MCMC sampling algorithm to sample from Equation (S.1). The MCMC sampling algorithm requires sampling from the following full conditional distributions in turn: (i) $[\theta, \delta, \gamma^{*R} | \Sigma, \Omega, \nu, z, \psi, \lambda, D_{\text{obs}}]$; (ii) $[z | \theta, \Sigma, \Omega, \delta, \nu, \gamma^{*R}, \psi, \lambda, D_{\text{obs}}]$; (iii) $[\psi | \theta, \Sigma, \Omega, \delta, \nu, \gamma^{*R}, z, \lambda, D_{\text{obs}}]$; (iv) $[\lambda, \nu | \theta, \Sigma, \Omega, \delta, \gamma^{*R}, z, \psi, D_{\text{obs}}]$; (v) $[\Sigma | \theta, \Omega, \delta, \nu, \gamma^{*R}, z, \psi, \lambda, D_{\text{obs}}]$; (vi) $[\Omega | \theta, \Sigma, \delta, \nu, \gamma^{*R}, z, \psi, \lambda, D_{\text{obs}}]$; For (i), we apply the collapsed Gibbs technique in Liu²¹ via the following identity:

$$\begin{aligned} & [\theta, \delta, \gamma^{*R} | \Sigma, \Omega, \nu, z, \psi, \lambda, D_{\text{obs}}] \\ &= [\gamma^{*R} | \theta, \delta, \Sigma, \Omega, \nu, z, \psi, \lambda, D_{\text{obs}}] \times [\theta, \delta | \Sigma, \Omega, \nu, z, \psi, \lambda, D_{\text{obs}}]. \end{aligned} \quad (12)$$

That is, we sample θ and δ after collapsing on γ^{*R} . For (iv), we apply the collapsed Gibbs technique of Liu²¹ through the identity

$$\begin{aligned} & [\lambda, \nu | \theta, \Sigma, \Omega, \delta, \gamma^{*R}, z, \psi, D_{\text{obs}}] \\ &= [\lambda | \theta, \Sigma, \Omega, \delta, \nu, \gamma^{*R}, z, \psi, D_{\text{obs}}] \times [\nu | \theta, \Sigma, \Omega, \delta, \gamma^{*R}, z, \psi, D_{\text{obs}}]. \end{aligned} \quad (13)$$

That is, we sample ν after collapsing out λ . The details of the above full conditional distributions are given in Section S1 of the supplementary materials.

4.4 Bayesian Model Comparison via the L Measure

Within the multivariate meta-regression framework, it is important to check whether a skew and/or heavy-tailed distribution is needed for modeling the y_{ik} 's. Also, we may need to investigate whether the J responses, $y_{11k}, \dots, y_{iJk}$, are indeed dependent. To help address these issues, we propose a useful multivariate criterion for model assessment called the *multivariate L measure*.

The multivariate L measure is motivated as follows. Consider an experiment that yields the data $y = (y_{11}, \dots, y_{n_1 1}, \dots, y_{1K}, \dots, y_{n_K K})'$, where each y_{ik} is a J -dimensional vector of

response variables, and thus y is an $n \times J$ matrix, where $n = \sum_{k=1}^K n_k$. Let

$w = (w_{11}, \dots, w_{n_1 1}, \dots, w_{1K}, \dots, w_{n_K K})'$ denote future values of a replicate experiment, where

$w_{ik} = (w_{i1k}, \dots, w_{iJk})'$ for $i = 1, \dots, n_k$ and $k = 1, \dots, K$. That is, w is a future response vector with the same sampling density as y/θ , where θ is defined in Section 3. The idea of using future responses in developing such a criterion for assessing a model or comparing several models has been well motivated in the statistical literature.^{17,18,22,23}

For a given model, we first propose the statistic

$$L_1(y, b) = E[(w - b)'(w - b)|D_{\text{obs}}] + \kappa(y - b)'(y - b), \quad (14)$$

where $k = 0$ and the expectation is taken with respect to the posterior predictive distribution of w/D_{obs} . We note here that (14) is a $J \times J$ matrix. The posterior predictive density of w/D_{obs} is given by

$$\pi(w|D_{\text{obs}}) = \int f(w|\theta)\pi(\theta|D_{\text{obs}})d\theta, \quad (15)$$

where $f(w|\theta)$ is the sampling distribution of the future matrix w , and $\pi(\theta|D_{\text{obs}})$ denotes the posterior distribution of θ . The matrix in (14) takes the form of a weighted discrepancy measure. The matrix $b = (b_{11}, \dots, b_{n_1 1}, \dots, b_{1K}, \dots, b_{n_K K})'$ is an arbitrary location matrix to be chosen and k is a nonnegative scalar that weights the discrepancy based on the future values relative to the observed response matrix y . The case $k = 0$, for example, can be interpreted as a squared error loss discrepancy measure in the future response matrix.

Expression (14) can be alternatively written as

$$L_1(y, b) = \sum_{k=1}^K \sum_{i=1}^{n_k} \{ \text{Var}(w_{ik}|D_{\text{obs}}) + (\mu_{ik} - b_{ik})(\mu_{ik} - b_{ik})' + \kappa(y_{ik} - b_{ik})(y_{ik} - b_{ik})' \}, \quad (16)$$

where $\mu_{ik} = E(w_{ik}|D_{\text{obs}})$ and $\text{Var}(w_{ik}|D_{\text{obs}})$ are the mean and variance of w_{ik} with respect to the posterior predictive distribution in (15). Thus we see that (16) has the appealing decomposition as a sum involving the predictive covariance matrices plus two “bias” terms, $(\mu_{ik} - b_{ik})(\mu_{ik} - b_{ik})'$ and $\kappa(y_{ik} - b_{ik})(y_{ik} - b_{ik})'$, where κ is a weight for the second bias component. The development of $L_1(y, b)$ serves as a useful multivariate model assessment statistic.

Since $L_1(y, b)$ is a $J \times J$ matrix, we propose to take the determinant of $L_1(y, b)$ as a model assessment statistic. That is, we take

$$L \equiv |L_1(y, b)|. \quad (17)$$

The reason why we take a determinant of $L_1(y, b)$ is that the determinant measures the size (volume) of a matrix and hence represents the best possible summary of a matrix in terms of its volume. For example, the volume of an ellipsoid is proportional to the determinant of the scale matrix defining the ellipsoid. Since the determinant is the product of the eigenvalues, it captures the largest eigenvalue as a result, and the determinant of a matrix also captures the sizes of the off-diagonal entries, which represent covariances if the matrix is a covariance matrix, which the L measure is. If we think about a matrix as representing a linear

transformation, then the determinant (technically the absolute value of the determinant) represents the “volume distortion” experienced by a region after being transformed. So for instance, the matrix $2I_2$ stretches a square of area 1 into a square with area 4, since the determinant is 4. This idea works in all dimensions. This also translates well when we get in to more general mappings in n dimensions. If the function is nice enough, we can represent it “locally” by a linear transformation (matrix). The (absolute value of the) determinant of this linear transformation gives the “local” volume distortion of the function - i.e. how much the function is stretching or compressing regions of space near a point. Thus, the determinant of a matrix is the best possible summary of its size in terms of volume distortion.

Selecting $\mathbf{b}$ as the minimizer of (17) gives us, in some sense, an appropriate weight of the bias component of the L measure. We are now led to the following theorem.

Theorem 1. *The $\mathbf{b}_{ik}$'s which minimize (17) are given by*

$$\hat{b}_{ik} = (1 - \varphi)\mu_{ik} + \varphi y_{ik}, \quad (18)$$

where $\varphi = \kappa/(\kappa + 1)$, for $i = 1, \dots, n_k$ and $k = 1, \dots, K$. Therefore the $\mathbf{b}$ which minimizes (17) is given by $\hat{\mathbf{b}} = (\hat{b}_{11}, \dots, \hat{b}_{n_1 1}, \dots, \hat{b}_{1K}, \dots, \hat{b}_{n_K K})'$.

The proof of Theorem 1 is given in Appendix A. We see that (18) has an appealing interpretation and consists of a weighted linear combination of the posterior predictive mean μ_{ik} and the observed data y_{ik} . Substitution of (18) into (16) leads to

$$L_2(\mathbf{y}) = \sum_{k=1}^K \sum_{i=1}^{n_k} \text{Var}(w_{ik}|D_{\text{obs}}) + \varphi \sum_{k=1}^K \sum_{i=1}^{n_k} (\mu_{ik} - y_{ik})(\mu_{ik} - y_{ik})'. \quad (19)$$

Clearly, $0 \leq \varphi < 1$, where $\varphi = 0$ if $\kappa = 0$, and $\varphi \rightarrow 1$ as $\kappa \rightarrow \infty$, and therefore our multivariate L measure criterion is given by

$$\hat{L} = |L_2(\mathbf{y})|. \quad (20)$$

Allowing φ to vary between zero and one gives the user a great deal of flexibility in the tradeoff between bias and variance.

To compute the L measure $\hat{L}$ in (20), we need to derive a computationally attractive form of μ_{ik} and $\text{Var}(w_{ik}|D_{\text{obs}})$. To this end, we state a useful theorem presented as follows.

Theorem 2. *Under the general multivariate skew meta-regression model specified by (1), (3), and (7), we have*

$$\mu_{ik} = E[X_{ik}\beta + (I_J \otimes T'_{ik})\gamma | D_{\text{obs}}], \quad (21)$$

where I_J is the $J \times J$ identity matrix, denotes the Kronecker product, and the expectation is taken with respect to the posterior distribution of β and given the data D_{obs} ; and

$$\begin{aligned} \text{Var}(w_{ik} | D_{\text{obs}}) &= E[\text{Var}(\epsilon_{ik}) | D_{\text{obs}}] \\ &+ E[\text{diag}(T'_{ik} \Omega_1 T_{ik}, \dots, T'_{ik} \Omega_J T_{ik}) | D_{\text{obs}}] \\ &+ \frac{\tau}{\tau - 1} \text{diag}\{\delta_1^2, \dots, \delta_J^2\} + \frac{1}{\tau - 1} \delta \delta' \end{aligned} \quad (22)$$

where $\text{Var}(\epsilon_{ik})$ is defined by (6).

4.5 Bayesian Assessment of Overall Fit and Outlying Trials

To examine the goodness-of-fit of the proposed models, we first introduce “residuals” using (1) as

$$e_{ijk} = y_{ijk} - \left\{ \left[\gamma_{ij0} + \gamma_{ij1} \text{trt}_{ik} \right] (1 - \text{onstain}_k) + \left[\gamma_{jk2} + \gamma_{jk3} \text{trt}_{ik} \right] \text{onstain}_k + \delta_j (z_{ijk} - E[z_{ijk}]) + x'_{ijk} \beta_j \right\}$$

and then define the Bayesian standardized residuals as

$$r_{ijk} = E(e_{ijk} | D_{\text{obs}}) / \sqrt{\text{Var}(e_{ijk} | D_{\text{obs}})}, \quad (23)$$

where $E(e_{ijk} | D_{\text{obs}})$ and $\text{Var}(e_{ijk} | D_{\text{obs}})$ are the posterior mean and the posterior variance of e_{ijk} , respectively, for $i = 1, \dots, n_k$, $k = 1, \dots, K$, and $j = 1, \dots, J$. We note that $E(e_{ijk} | D_{\text{obs}}) = y_{ijk} - \mu_{ijk}$, where μ_{ijk} is the j^{th} component of μ_{ik} given in (21). We further note that the Bayesian standardized residuals in (23) are different than those in Kim et al.,¹ in which the unstandardized Bayesian residuals were considered.

For the IPD, the sample sizes are typically different across trials. For the cholesterol data considered in this paper, the sample sizes range from 184 to 2829. Such drastic differences in sample sizes across trials pose a challenge for assessing outlying trials. To account for sample sizes, we develop the following new diagnostic procedure. From (4), *a priori*, the λ_{ik} 's are i.i.d. from $\text{Gamma}(\nu/2, \nu/2)$. Define

$$\lambda_k = \sum_{i=1}^{n_k} \lambda_{ik}.$$

Then, *a priori*, we have $\lambda_k \sim \text{Gamma}(n_k \nu/2, \nu/2)$. We then standardize λ_k based on its prior mean and prior variance to adjust for the size of the IPD from the k^{th} trial as

$$\lambda_k^* = (\lambda_k - n_k)/(n_k I(v/2))^{1/2}, \quad (24)$$

which implies $E[\lambda_k^*] = 0$ and $\text{Var}(\lambda_k^*) = 1$ *a priori*. Then, following Johnson,¹⁹ we can simply sample from the posterior distribution of λ_k^* to see which trial departs from the prior distribution of λ_k^* the most.

5 Simulation Studies

5.1 An Empirical Study of the Diagnostic Procedure for Detecting Outlying Trials

In order to better understand the proposed diagnostic procedure developed in Section 4.5, we simulate a meta-data set with 30 trials. Specifically, we independently generate $x_{ik1} \sim N(0, 1)$ and $x_{ik2} \sim \text{Bernoulli}(0.3)$, then generate y_{ik} based on the following meta-regression model:

$$y_{ik} = x'_{ik}\beta + \gamma_k + \epsilon_{ik}, \quad (25)$$

where $x_{ik} = (1, x_{ik1}, x_{ik2})'$ for $i = 1, \dots, n_k$, $\gamma_k \sim N(0, 1)$ for $k = 1, \dots, K = 30$, and $\beta = (\beta_0, \beta_1, \beta_2)'$ with $\beta_0 = 5$, $\beta_1 = 0.5$, and $\beta_2 = -0.3$. In addition, we independently generate $\epsilon_{ik} \sim t_5(0, 5)$ (a t -distribution with location parameter 0, scale parameter 5, and 5 degrees of freedom) for $k = 1, \dots, 27$; $\epsilon_{i,28} \sim t_1(0, 5)$ (Cauchy); $\epsilon_{i,29} \sim N(0, 5)$ (normal); and $\epsilon_{i,30} \sim t_5(0, 1)$ (a t -distribution with location parameter 0, scale parameter 1, and 5 degrees of freedom). Furthermore, we take $n_k = 500$ for $k = 1, \dots, 10$, $n_k = 1000$ for $k = 11, \dots, 20$, $n_k = 2000$ for $k = 21, \dots, 27$, $n_{28} = 500$, $n_{29} = 1000$, and $n_{30} = 2000$. We fit the model in (25) assuming that $\gamma_k \sim N(0, \Omega)$ and $\epsilon_{i,30} \sim t_5(0, \sigma^2)$, where Ω and σ^2 are unknown. The priors are specified as $\beta \sim N(0, 100I_3)$, $\Omega \sim \text{IG}(0.1, 0.1)$, where $\text{IG}(a, b)$ denotes an inverse gamma distribution with density proportional to $\Omega^{-(a+1)} \exp(-b\Omega)$, and $\sigma^2 \sim \text{IG}(0.1, 0.1)$. Under this simulation setting, the error distributions have heavier tails in trial 28 and lighter tails in trial 29 compared to those in trials 1 to 27, while the scale parameter of the error distribution in trial 30 is much smaller. Thus, trials 28, 29, and 30 are clearly outlying ones.

We generated 20,000 MCMC samples, which were taken from every fifth iteration, after a burn-in of 20,000 iterations from the posterior distribution, and then calculated the standardized λ_k^* for $k = 1, \dots, 30$. Figure 2 shows the boxplots of the λ_k^* 's from the posterior distribution for these 30 trials. From this figure, we see that for trials 1 to 27, the centers of the boxplots are close to zero and the shapes of the boxplots are quite similar. In addition, the 95% HPD intervals of the λ_k^* for trials 1 to 27 contained zero. Although the sample sizes range from 500 to 2000, the effect of different sample sizes was well adjusted in the posterior distributions of the λ_k^* 's. For the heavy-tailed distribution of $\epsilon_{i,28}$, the corresponding boxplot tends to be below zero and the 95% HPD interval of λ_{28} was $(-7.092, -3.755)$. On

the other hand, when a light-tailed distribution or a smaller scale parameter are specified for $\epsilon_{i,29}$ and $\epsilon_{i,30}$, respectively, the corresponding boxplots tend to be above zero and the 95% HPD intervals were $(-0.182, 3.561)$ for λ_{29}^* and $(7.671, 11.734)$ for λ_{30}^* . This empirical investigation helps us to understand and interpret potentially outlying trials.

5.2 A Simulation Study for Examining the Empirical Performance of the Proposed Model

In this simulation study, we simulate meta-data sets with 15 and 30 trials on the first-line therapy. Specifically, we independently generate three covariates and a treatment variable: $x_{ik1} \sim N(0, 1)$, $x_{ik2} \sim N(0, 1)$, $x_{ik3} \sim \text{Bernoulli}(0.3)$, which is standardized by mean 0.3 and SD 0.458, and $\text{trt}_{ik} \sim \text{Bernoulli}(0.5)$. We then generate y_{ijk} from the skew t model with an unstructured Σ and $\tau = 30$, which corresponds to the model, which fits the cholesterol data best, as follows

$$y_{ijk} = \gamma_{jk0} + \gamma_{jk1}\text{trt}_{ik} + \delta_j(z_{ijk} - E[z_{ijk}]) + x'_{ijk}\beta_j + \epsilon_{ijk}, \quad (26)$$

where $\mathbf{x}_{ikj} = (x_{ik1}, x_{ik2}, x_{ik3})'$, $\gamma_{jk} \sim N_2(\gamma_j, \Omega_j)$ with $\gamma_j = (\gamma_{j0}, \gamma_{j1})'$ for $i = 1, \dots, n_k$ and $k = 1, \dots, K$. Furthermore, we take $n_k = 200$ for $k = 1, \dots, 5$ and $k = 16, \dots, 20$, $n_k = 350$ for $k = 6, \dots, 10$ and $k = 21, \dots, 25$, and $n_k = 500$ for $k = 11, \dots, 15$ and $k = 26, \dots, 30$. Then 500 simulated datasets are generated. The prior distributions are those specified in Section 4.2. We use 20,000 Gibbs samples to compute all the posterior estimates, including the mean and 95% highest posterior density (HPD) intervals after a burn-in of 10,000 iterations. Furthermore, we also fit the Box-Cox model¹ to the simulated data sets, and evaluate its empirical performance.

Tables 1 – 4 present the mean of the posterior means (Mean), the standard deviation of the posterior means (SE), the mean of the posterior standard deviations (MSD), the 95% confidence interval (CI), the coverage probability (CP) of the 95% HPD intervals, and power for the simulation study. The estimates for all parameters in Tables 1 and 2 are close to the true values. From these two tables, we also see that the 95% CP's and the powers improve when K increases from 15 to 30. When fitting the data generated under the proposed model, the posterior estimates of the regression coefficients under the Box-Cox model¹ were not expected to be close to the true values, however, the signs (plus or minus) of these estimates were expected to be consistent with the true values except for those intercepts ($\gamma_{j0}, j = 1, 2, 3$). We see from Tables 3 and 4 that (i) the posterior estimates of the regression coefficients, $(\beta_{j1}, \beta_{j2}, \beta_{j3})$ and γ_{j1} for $j = 1, 2, 3$, the covariances of the random effects, $\Omega_{j0,1}$ for $j = 1, 2, 3$, and the covariances of the responses, Σ_{12} , Σ_{13} , and Σ_{23} under the Box-Cox model have the exact same signs as the true values; (ii) the powers of these estimates under the Box-Cox model improve when K increases from 15 to 30; and (iii) for $K = 15$ or $K = 30$, the posterior estimates under the Box-Cox model always had lower powers than those corresponding estimates under the proposed model given in Tables 1 and 2

6 Analysis of the Cholesterol Data

We carry out a detailed analysis of the cholesterol data discussed in Section 2. In (1), x_{ijk} consists of 14 covariates, including bl_ldlc, bl_hdlc, bl_tg, BMI, age, duration, Female, DM, CHD, potency2, potency3, black, hispanic, and other, as well as 14 interaction terms between the 14 covariates and onstatin. We note here that these interaction effects were not examined in Kim et al.¹ The outcome variables were LDL-C, HDL-C, and TG, which were defined as percent changes from baseline in LDL-C, HDL-C, and TG. We model these three outcome variables jointly via (1), (3), (4), and (7) with $J=3$ and $K=26$. For numerical stability in the posterior computations, we standardized all of the fourteen covariates, in which each covariate was subtracted from its sample mean and divided by its sample standard deviation computed from the pooled data.

We fit the multivariate meta-regression models defined by (1), (3), (4), and (7) with different structures of the covariance matrix Σ , and different distributions of λ_{ik} in (3), and different values of τ in (2). The models considered here include the normal model, the t model, and the skew t models with various values of τ . For each of these models, we specify an unstructured Σ and a diagonal Σ , i.e., $\Sigma = \text{diag}(\Sigma_{11}, \Sigma_{22}, \Sigma_{33})$ in (3). For the normal and t models, we assume that $\delta_j = 0$ with probability 1 in (1) for $j = 1, 2, 3$. For the normal models, we further assume that H is $\{1\}$ in (4), i.e., $P(\lambda_{ik} = 1) = 1$. For the skew t models, we take values of τ to be 15, 20, 25, 30, 35, 40, and ∞ , where $\tau = I$ corresponds to the model in which the z_{ijk} 's in (1) are independent. We then computed the L measures defined by (20) under these 18 models for the cholesterol data, and the results are reported in Table 5 for normal, t , and skew t with $\tau = 20, 25, 30, 35$, and ∞ and Table S1 of the Supplementary Materials for skew t with $\tau = 15$ and 40. From Tables 5 and S1, we see that, for example, when $\phi = 0.5$, the values of $10^{-4}n^{-3}|L_2(y)|$, where $n = 21,184$ is the total sample size, were 10358.71, 10279.28, 9033.06, 8984.90, 8970.14, 8974.47, 8977.12, 8979.17, and 9034.77, respectively, for the normal model, the t model, and the skew t models with $\tau = 15, 20, 25, 30, 35, 40$, and ∞ under a diagonal Σ ; and 9577.31, 9334.06, 8363.13, 8326.39, 8315.03, 8311.82, 8315.29, 8316.04, and 8363.44 for the normal model, the t model, and the skew t models with $\tau = 15, 20, 25, 30, 35, 40$, and ∞ under an unstructured Σ . Clearly, for $\phi = 0.5$, the smallest value of $10^{-4}n^{-3}|L_2(y)|$ was 8311.82, which was achieved by the skew t model with $\tau = 30$ and an unstructured Σ . 'shown in Tables 5 and S1 were consistent with those for $\phi = 0.5$ under these 18 models. These L measures indicate that (i) the models with an unstructured Σ fit the cholesterol data much better than the respective models with a diagonal Σ ; (ii) under an unstructured Σ , the t model fit the data better than the normal model; (iii) under both the diagonal Σ and unstructured Σ , the skew t models fit the data better than the symmetric normal and t models; (iv) under both the diagonal Σ and unstructured Σ , the skew t models with dependent z_{ijk} 's ($\tau < \infty$) outperformed the corresponding models with independent z_{ijk} 's ($\tau = \infty$); and (v) among the 9 models with a diagonal Σ , the skew t model with $\tau = 25$ was the best while the skew t model with $\tau = 30$ was the best among the 9 models with an unstructured Σ according to the L measure. These results are quite interesting and appealing since they showed that the three outcome variables were indeed dependent and the distributions of these variables were not symmetric for the cholesterol data as shown in Kim et al.¹

The boxplots of the Bayesian standardized residuals defined in (23) for each of the three outcome variables under the normal, t , and skew t models with an unstructured Σ are shown in Figure 3. The first quartiles, medians, and third quartiles of these residuals were -10.887 , -2.041 , and 7.85 for the outcome variable LDL-C, -12.665 , -0.992 , and 11.177 for the outcome variable HDL-C, and -12.841 , -2.955 , and 8.833 for the outcome variable TG under the normal model; -10.994 , -0.547 , 11.081 for the outcome variable LDL-C, -13.462 , -0.692 , and 12.574 for the outcome variable HDL-C, and -13.013 , -0.761 , and 14.123 for the outcome variable TG under the t model; and -0.833 , 0.200 , 0.688 for the outcome variable LDL-C, -1.540 , 0.140 , and 1.064 for the outcome variable HDL-C, and -0.755 , 0.157 , and 0.623 for the outcome variable TG under the skew t model with $\tau = 30$. From Figure 3 and the above summary of the Bayesian standardized residuals, we see that the Bayesian standardized residuals under the skew t model were much more symmetric and smaller than those under the normal and t models. Figure 3 also shows that the skew model had a great improvement in the residuals for the outcome variable TG over the normal or t models. These results were consistent with those based on the L measure, which further confirms the need of skew and heavy-tailed distributions for all three outcome variables for these data.

Figure 4 shows the boxplots of the λ_k^* 's defined in (24) under the skew t model with $\tau = 30$ and an unstructured Σ for the cholesterol data. In Figure 4, subjects in trials 1 to 13 were on the first-line therapy (i.e., onstatin = 0) while the subjects in the rest of the 13 trials were on the second-line therapy (i.e., onstatin = 1). The boxplots in Figure 4 were more heterogeneous than those in Figure 2. Most of these boxplots in Figure 4 had a median close to zero. The boxplot corresponding to trial 8 was clearly above all of the rest boxplots. Two boxplots that were below all other boxplots were from trials 14 and 25. The 95% HPD intervals of λ_8^* , λ_{14}^* , and λ_{25}^* were $(4.510, 8.334)$, $(-5.222, -1.655)$, and $(-6.060, -2.603)$, respectively. These three trials were more likely to be outlying. After we re-examined these 26 trials, we found that the 8th trial was the only first-line study including ROSUVA statin, which has the highest potency, while trial 14 was an add-on second-line study, which includes different statins (Atorvastatin, Simvastatin, Pravastatin, and Other statins) and trial 25 was the only second-line study with all Type-2 diabetic patients.

The posterior estimates, including the posterior means, posterior standard deviations (SDs), and 95% highest posterior density (HPD) intervals of the parameters under the skew t models with an unstructured Σ are reported in Tables 6 – 10 while these posterior estimates under the normal and t models with an unstructured Σ are reported in Tables S2 – S6 of the Supplementary Materials. We note here that we define a posterior estimate to be “statistically significant at a significance level of 0.05” if the corresponding 95% HPD interval does not contain 0. The results shown in Tables S2 and 6 indicate that patients on “statin + EZE” had significantly more percent changes from baseline in all three outcome variables (LDL-C, HDL-C, and TG) than those on statin alone for both the first-line and second-line therapies under all three models with an unstructured Σ . For first-line therapy, the 95% HPD intervals were $(-15.066, -12.012)$, $(-16.011, -12.745)$, and $(-15.662, -12.454)$ for the percent change from baseline in LDL-C ($\gamma_{1,1}$); $(1.323, 3.077)$, $(1.335, 3.026)$, and

(1.285, 2.870) for the percent change from baseline in HDL-C ($\gamma_{2,1}$); and (−7.853, −4.915), (−8.246, −5.180), and (−7.316, −4.369) for the percent change from baseline in TG ($\gamma_{3,1}$), respectively, under the normal, t , and skew t models with an unstructured Σ . For second-line therapy, the 95% HPD intervals were (−21.131, −15.415), (−21.975, −16.041), and (−21.586, −15.974) for the percent change from baseline in LDL-C ($\gamma_{1,3}$); (0.503, 1.679), (0.737, 1.776), and (0.726, 1.736) for the percent change from baseline in HDL-C ($\gamma_{2,3}$); and (−10.482, −7.128), (−10.338, −7.074), and (−9.213, −5.868) for the percent change from baseline in TG ($\gamma_{3,3}$), respectively, under the normal, t , and skew t models with an unstructured Σ . From Table 6, we also see that the respective 95% HPD intervals for the skewness parameters under the skew t model with $\tau = 30$ were (10.429, 11.046) for LDL-C (δ_1), (5.203, 5.924) for HDL-C (δ_2), and (19.831, 20.915) for TG (δ_3), respectively. Thus, TG has the largest skewness parameter while HDL-C has the smallest skewness parameter according to their posterior estimates. These results indicate that the distributions of all three outcome variables are skewed and also are consistent with the results based on the L measure since the skew t models yield much smaller L -measure values than the symmetric models. In addition, the 95% HPD interval of ν under the t and skew t models were (4.248, 4.559) and (4.778, 5.497), which implies that a heavy-tailed joint distribution for the three outcome variables is required in order to adequately fit the cholesterol data. The posterior estimates for these parameters under the normal, t , and skew t models with a diagonal Σ were similar and the results are given in Tables S7 – S11 of the Supplementary Materials. We note that for the diagonal Σ , the skew t model with $\tau = 25$ was the best according to the multivariate L measure. Therefore, the posterior estimates under the skew t model with $\tau = 25$ were reported in Tables S7 – S11.

In Tables 6 and S2, ρ_{12} , ρ_{13} , and ρ_{23} denote the correlations between LDL-C and HDL-C, between LDL-C and TG, and between HDL-C and TG, respectively. From these two tables, we see that there were moderate correlations among these three outcome variables. In particular, the percent change from baseline in LDL-C was positively correlated with both the percent changes from baseline in HDL-C and TG, while the percent change from baseline in HDL-C was negatively correlated with the percent change from baseline in TG. The 95% HPD intervals of ρ_{12} , ρ_{13} , and ρ_{23} were (0.061, 0.088), (0.147, 0.174), and (−0.240, −0.214) under the normal model; (0.078, 0.107), (0.150, 0.180), and (−0.238, −0.210) under the t model; and (0.120, 0.165), (0.276, 0.335), and (−0.387, −0.343) under the skew t model with $\tau = 30$. Therefore, all three correlations were significantly different from 0, which further confirms the results that the models with an unstructured Σ greatly improve the fit according to the L measure. We also note that the values of the posterior estimates of the correlations (ρ_{jj} 's) between the three outcome variables LDL-C, HDL-C, and TG were (0.074, 0.161, −0.227) under the normal model and (0.092, 0.165, −0.223) under the t model, which are consistently smaller than (0.0.143, 0.305, −0.366) under the skew t model with $\tau = 30$.

Since these interaction effects between the 14 covariates and onstatin were not examined in Kim et al. (2013), the results shown in Tables 6 - 9 are not directly comparable to those in Kim et al. (2013). In order to make these two different methods more comparable, we modified and re-ran the code for the Box-Cox transformation model in Kim et al. (2013) by including these 14 interaction terms. The posterior estimates under the Box-Cox

transformation model are given in Tables S12 and S13 of the Supplementary Materials. Although the magnitudes of the posterior estimates under these two models are quite different, the significance and the sign of the coefficient corresponding to each covariate should be comparable. From Table 6 and Table S12, we see that the posterior estimates of the correlations ρ_{12} , ρ_{13} , and ρ_{23} are similar under the skew t model with an unstructured Σ and $\tau = 30$ and the Box-Cox transformation model. The posterior estimates of the regression coefficients are similar in terms of significance, i.e., whether the 95% HPD interval contains 0 or not, for most covariates. Specifically, from Tables 3 – 5 and Table S13, we see that (i) for LDL-C, Hispanic was not significant with a 95% HPD interval of $(-0.382, 1.670)$ and onstatin $\times$ DM was significant with a 95% HPD interval of $(-2.088, -0.225)$ under the skew t model while Hispanic was significant with a 95% HPD interval of $(0.008, 0.0218)$ and onstatin $\times$ DM was not significant with a 95% HPD interval of $(-0.014, 0.0134)$ under the Box-Cox model; (ii) for HDL-C, the posterior estimates of the regression coefficients under skew t model were completely in agreement with those under the Box-Cox model in terms of significance; (iii) for TG, bl_hdlc, age, and Hispanic were not significant and onstatin $\times$ bl_tg and onstatin $\times$ duration were significant under the skew t model while bl_hdlc, age, and Hispanic were significant and onstatin $\times$ bl_tg and onstatin $\times$ duration were not significant under the Box-Cox model; and (iv) the signs of the posterior estimates of the treatment effects were consistent under these two models except for $\gamma_{1,0}$, which corresponds to an intercept. The computational time for the Box-Cox transformation model and the skewed t model with an unstructured Σ and $\tau = 30$ were 20 hours and 19 minutes and 8 hours and 20 minutes, respectively, for 10,000 iterations, by including the same covariates. Thus the skew t model is more advantageous in terms of computational time.

In all the Bayesian computations, we used 20,000 MCMC samples, which were taken from every fifth iteration, after a burn-in of 4,000 iterations for each model to compute all posterior estimates, including posterior means, posterior SDs, 95% HPD intervals, and the L measure. The convergence of the MCMC sampling algorithm was checked using several diagnostic procedures discussed in Chen et al.²⁴ The HPD intervals were computed via the Monte Carlo method developed by Chen and Shao.²⁵ Computer code was written for the FORTRAN 95 compiler, and we used IMSL subroutines with double precision accuracy. The FORTRAN code is available from the authors upon request.

7 Discussion

In this paper, we have proposed a new multivariate skew meta-regression model for modeling individual level patient meta-data. One of the innovations is that our proposed model allows the correlations among the multivariate outcomes to depend on skewness parameters. We have also developed an efficient MCMC sampling algorithm using the collapsed Gibbs technique of Liu²¹ to carry out challenging posterior computations due to the large size of the meta-data and the high-dimensions of the random effects. In addition, we have proposed the multivariate L measure for model comparison, the Bayesian residuals for model assessment, and the diagnostic procedure for detecting outlying trials. As was seen from the analysis of the (LDL-C, HDL-C, TG) data, the proposed model is quite useful and highly needed since the multivariate outcome measures have skewed distributions and are highly dependent. The proposed diagnostic procedure automatically adjusts for the size

of the IPD and is quite effective in detecting outlying trials in both the simulated and real data. In addition, we have carried out a simulation study to compare the proposed model with the Box-Cox model¹.

The proposed multivariate skew meta-regression model can be extended by adding additional hierarchical structures of the skewness parameters as well as the degrees of freedom. Under the proposed model, the multivariate responses share the common degrees of freedom in the t -distribution. It may be more desirable to relax this assumption so that each response has its own degrees of freedom. These extensions are quite challenging and extensive and, therefore, they are deferred to future work.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

Dr. Ibrahim and Dr. Chen's research was partially supported by NIH grants #GM70335 and #P01CA142538. Dr. Kim's research was supported by the Intramural Research Program of National Institutes of Health, National Cancer Institute.

Appendix. Proof of Theorem 1

We can write

$$\begin{aligned}\kappa(b_{ik} - y_{ik})(b_{ik} - y_{ik})' &= \kappa(b_{ik} - \hat{b}_{ik} + \hat{b}_{ik} - y_{ik})(b_{ik} - \hat{b}_{ik} + \hat{b}_{ik} - y_{ik})' \\ &= \kappa(b_{ik} - \hat{b}_{ik})(b_{ik} - \hat{b}_{ik})' + \kappa(\hat{b}_{ik} - y_{ik})(\hat{b}_{ik} - y_{ik})' + \kappa(b_{ik} - \hat{b}_{ik})(\hat{b}_{ik} - y_{ik})' \\ &\quad + \kappa(\hat{b}_{ik} - y_{ik})(b_{ik} - \hat{b}_{ik})'\end{aligned}$$

Now we can write $\hat{b}_{ik} - y_{ik} = (1 - \varphi)\mu_{ik} + \varphi y_{ik} - y_{ik} = (1 - \varphi)(\mu_{ik} - y_{ik})$ and

$$\begin{aligned}(b_{ik} - \mu_{ik})(b_{ik} - \mu_{ik})' &= (b_{ik} - \hat{b}_{ik} + \hat{b}_{ik} - \mu_{ik})(b_{ik} - \hat{b}_{ik} + \hat{b}_{ik} - \mu_{ik})' \\ &= (b_{ik} - \hat{b}_{ik})(b_{ik} - \hat{b}_{ik})' + (\hat{b}_{ik} - \mu_{ik})(\hat{b}_{ik} - \mu_{ik})' + (b_{ik} - \hat{b}_{ik})(\hat{b}_{ik} - \mu_{ik})' \\ &\quad + (\hat{b}_{ik} - \mu_{ik})(b_{ik} - \hat{b}_{ik})'\end{aligned}$$

Thus, we have $\hat{b}_{ik} - \mu_{ik} = (1 - \varphi)\mu_{ik} + \varphi y_{ik} - \mu_{ik} = \varphi(y_{ik} - \mu_{ik}) = -\varphi(\mu_{ik} - y_{ik})$. Since $\kappa(1 - \varphi) = \varphi$, we obtain

$$\begin{aligned}
& \kappa(b_{ik} - y_{ik})(b_{ik} - y_{ik})' + (b_{ik} - \mu_{ik})(b_{ik} - \mu_{ik})' \\
& = \kappa(b_{ik} - \hat{b}_{ik})(b_{ik} - \hat{b}_{ik})' + \kappa(1 - \varphi)^2(\mu_{ik} - y_{ik})(\mu_{ik} - y_{ik})' \\
& + \kappa(1 - \varphi)(b_{ik} - \hat{b}_{ik})(\mu_{ik} - y_{ik})' \\
& + \kappa(1 - \varphi)(\mu_{ik} - y_{ik})(b_{ik} - \hat{b}_{ik})' + (b_{ik} - \hat{b}_{ik})(b_{ik} - \hat{b}_{ik})' \\
& + \varphi^2(\mu_{ik} - y_{ik})(\mu_{ik} - y_{ik})' - \varphi(b_{ik} - \hat{b}_{ik})(\mu_{ik} - y_{ik})' \\
& + \varphi(\mu_{ik} - y_{ik})(b_{ik} - \hat{b}_{ik})' \\
& = (\kappa + 1)(b_{ik} - \hat{b}_{ik})(b_{ik} - \hat{b}_{ik})' + \varphi(\mu_{ik} - y_{ik})(\mu_{ik} - y_{ik})'.
\end{aligned}$$

Let $A = \sum_{k=1}^K \sum_{i=1}^{n_k} [\text{Var}(w_{ik}|D_{\text{obs}}) + \varphi(\mu_{ik} - y_{ik})(\mu_{ik} - y_{ik})']$ and

$B = \sum_{k=1}^K \sum_{i=1}^{n_k} (\kappa + 1)(b_{ik} - \hat{b}_{ik})(b_{ik} - \hat{b}_{ik})'$. Then, it is easy to show that A and B are symmetric and semi-positive definite, and $L \equiv [L_1(y, b)] = |A + B|$. Thus, Theorems 6.7 and 6.8 in Zhang²⁶ yield that $L \geq |A|$ with the minimum achieved at $b_{ik} = \hat{b}_{ik}$ for $i = 1, \dots, n_k$ and $k = 1, \dots, K$. In addition, if A is positive definite, then equality holds if and only if $b_{ik} = \hat{b}_{ik}$ for $i = 1, \dots, n_k$ and $k = 1, \dots, K$. This completes the proof.

References

1. Kim S, Chen MH, Ibrahim JG et al. Bayesian inference for multivariate meta-analysis Box-Cox transformation models for individual patient data with applications to evaluation of cholesterol lowering drugs. *Statistics in Medicine* 2013; 32: 3972–3990. [PubMed: 23580436]
2. Edwards JE, Moore RA and McQuay HJ. Rofecoxib for dysmenorrhoea: meta-analysis using individual patient data. *BMC Women's Health* 2004; 4: 1–7. [PubMed: 15068485]
3. Gorman JD, David-Vaudey E, Pai M et al. Lack of association of the hla-drb1 shared epitope with rheumatoid nodules: An individual patient data meta-analysis of 3,272 caucasian patients with rheumatoid arthritis. *Arthritis & Rheumatism* 2004; 50: 753–762. [PubMed: 15022316]
4. Pereira TV, Kimura L, Suwazono Y et al. Multivariate meta-analysis of the association of G-protein beta 3 gene (gnb3) haplotypes with cardiovascular phenotypes. *Molecular Biology Reports* 2014; 41: 3113–3125. [PubMed: 24477587]
5. Simmonds MC and Higgins JPT. Covariate heterogeneity in meta-analysis: Criteria for deciding between meta-regression and individual patient data. *Statistics in Medicine* 2007; 26: 2982–2999. [PubMed: 17195960]
6. Ritz J, Demidenko E and Spiegelman D. Multivariate meta-analysis for data consortia, individual patient meta-analysis, and pooling projects. *Journal of Statistical Planning and Inference* 2008; 138: 1919–1933.
7. Riley RD, Price MJ, Jackson D et al. Multivariate meta-analysis using individual participant data. *Research Synthesis Methods* 2015; 6: 157–174. [PubMed: 26099484]
8. Chen MH, Dey DK and Shao QM. A new skewed link model for dichotomous quantal response data. *Journal of the American Statistical Association* 1999; 94: 1172–1186.
9. Branco MD and Dey DK. A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis* 2001; 79: 99–113.
10. Azzalini A and Capitanio A. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t -distribution. *Journal of The Royal Statistical Society, Series B* 2003; 65: 367–389.

11. Sahu SK, Dey DK and Branco MD. A new class of multivariate skew distributions with applications to Bayesian regression models. *Canadian Journal of Statistics* 2003; 31(2): 129–150.
12. Genton MG. *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. Boca Raton, FL: Chapman & Hall/CRC, 2004.
13. Adcock CJ. Capital asset pricing for UK stocks under the multivariate skew-normal distribution. Boca Raton, FL: Chapman & Hall/CRC, 2004.
14. Kim S, Chen MH and Dey DK. Flexible generalized t -link models for binary response data. *Biometrika* 2008; 95: 93–106.
15. Arellano-Valle RB and Genton MG. Multivariate unified skew-elliptical distributions. *Chilean Journal of Statistics* 2010; 1: 17–33.
16. Chang SC and Zimmerman DL. Skew-normal antedependence models for skewed longitudinal data. *Biometrika* 2016; 103(2): 363–376. [PubMed: 27279663]
17. Ibrahim JG and Laud PW. A predictive approach to the analysis of designed experiments. *Journal of the American Statistical Association* 1994; 89: 309–319.
18. Laud PW and Ibrahim JG. Predictive model selection. *Journal of the Royal Statistical Society, Series B* 1995; 57: 247–262.
19. Johnson VE. Comment: Bayesian checking of the second levels of hierarchical models. *Statistical Sciences* 2007; 22: 353–358.
20. Leiter LA, Betteridge DJ, Farnier M et al. Lipid-altering efficacy and safety profile of combination therapy with ezetimibe/statin vs. statin monotherapy in patients with and without diabetes: an analysis of pooled data from 27 clinical trials. *Diabetes, Obesity and Metabolism* 2011; 13: 615–628.
21. Liu JS. The collapsed Gibbs sampler in Bayesian computations with applications to a gene regulation problem. *Journal of the American Statistical Association* 1994; 89: 958–966.
22. Ibrahim JG, Chen MH and Sinha D. Criterion based methods for Bayesian model assessment. *Statistica Sinica* 2001; 11: 419–443.
23. Ibrahim JG, Chen MH and Sinha D. Bayesian methods for joint modeling of longitudinal and survival data with applications to cancer vaccine studies. *Statistica Sinica* 2004; 14: 863–883.
24. Chen MH, Shao QM and Ibrahim JG. *Monte Carlo Methods in Bayesian Computation*. New York, Springer-Verlag, 2000.
25. Chen MH and Shao QM. Monte carlo estimation of Bayesian credible and hpd intervals. *Journal of Computational and Graphical Statistics* 1999; 8: 69–92.
26. Zhang F *Matrix Theory*. New York, Springer-Verlag, 1999.

Figure 1.
Flow Diagram of the Proposed Model with Prior Specifications.

Figure 2.
Boxplots of λ_k^* 's from the posterior distribution based on the simulated data with 30 trials.

Figure 3. Boxplots of the Bayesian standardized residuals for the normal model (a), the t model (b), and the skew t model with $\tau = 30$ (c) under an unstructured Σ .

Figure 4. Boxplots of the λ_k^* 's from the posterior distribution under the skew t model with $\tau = 30$ and an unstructured Σ for the cholesterol data.

Table 1.Simulation results for the skew t model with an unstructured Σ and $\tau = 30$ based on $K = 15$

Parameter	True	Mean	SE	MSD	95% CI	95% CP	Power
$\beta_{1,1}$	0.5	0.499	0.164	0.161	(0.212, 0.818)	95.0	88.4
$\beta_{1,2}$	-1.5	-1.508	0.153	0.162	(-1.823, -1.207)	95.2	100.0
$\beta_{1,3}$	1.5	1.503	0.165	0.162	(1.193, 1.846)	93.6	100.0
$\beta_{2,1}$	-1.0	-0.983	0.142	0.140	(-1.274, -0.706)	93.8	100.0
$\beta_{2,2}$	1.0	1.004	0.135	0.140	(0.737, 1.277)	95.2	100.0
$\beta_{2,3}$	-2.0	-2.003	0.137	0.140	(-2.261, -1.732)	95.4	100.0
$\beta_{3,1}$	1.5	1.505	0.229	0.223	(1.062, 1.914)	95.2	100.0
$\beta_{3,2}$	-0.5	-0.497	0.223	0.223	(-0.937, -0.076)	95.6	59.4
$\beta_{3,3}$	-1.0	-0.988	0.233	0.223	(-1.443, -0.527)	93.4	99.2
$\gamma_{1,0}$	-20.0	-19.861	1.110	1.046	(-21.844, -17.585)	92.2	100.0
$\gamma_{1,1}$	-15.0	-15.027	0.769	0.723	(-16.595, -13.611)	89.4	100.0
$\gamma_{2,0}$	10.0	10.008	0.411	0.399	(9.187, 10.785)	93.0	100.0
$\gamma_{2,1}$	2.0	1.986	0.429	0.397	(1.087, 2.778)	92.0	99.2
$\gamma_{3,0}$	1.5	1.462	0.779	0.723	(-0.076, 3.005)	89.6	50.6
$\gamma_{3,1}$	-5.0	-4.940	0.735	0.667	(-6.290, -3.523)	89.8	100.0
Σ_{11}	55.0	54.883	2.793	2.847	(49.381, 60.508)	94.6	100.0
Σ_{12}	10.0	10.056	1.333	1.352	(7.520, 12.639)	95.6	100.0
Σ_{13}	24.0	25.089	2.274	2.222	(20.516, 29.561)	93.6	100.0
Σ_{22}	60.0	60.173	3.182	3.017	(54.340, 66.812)	94.6	100.0
Σ_{23}	-30.0	-30.058	1.895	1.949	(-33.880, -26.456)	95.8	100.0
Σ_{33}	100.0	100.095	5.026	5.081	(89.977, 110.218)	94.6	100.0
$\Omega_{1,0,0}$	15.0	16.286	6.834	7.283	(5.688, 33.067)	92.6	100.0
$\Omega_{1,0,1}$	-7.5	-7.791	4.312	4.343	(-17.212, -1.082)	88.8	76.0
$\Omega_{1,1,1}$	7.0	6.645	3.564	3.566	(0.668, 14.372)	84.0	100.0
$\Omega_{2,0,0}$	2.0	1.824	1.027	1.051	(0.258, 4.019)	82.2	100.0
$\Omega_{2,0,1}$	-1.5	-1.254	0.944	0.905	(-3.389, -0.014)	78.6	35.2
$\Omega_{2,1,1}$	1.5	1.229	0.976	0.935	(0.177, 3.631)	74.4	100.0
$\Omega_{3,0,0}$	7.0	6.296	3.504	3.397	(0.732, 14.385)	81.6	100.0
$\Omega_{3,0,1}$	-5.5	-4.571	3.036	2.871	(-11.636, -0.204)	79.6	56.4
$\Omega_{3,1,1}$	5.0	3.888	2.810	2.744	(0.368, 10.883)	76.2	100.0
δ_1	10.0	10.044	0.290	0.294	(9.531, 10.602)	94.8	100.0
δ_2	5.0	4.977	0.380	0.360	(4.165, 5.644)	94.4	100.0
δ_3	15.0	15.012	0.401	0.390	(14.240, 15.768)	94.6	100.0
ν	5.0	5.028	0.293	0.299	(4.498, 5.673)	94.6	100.0

Table 2.Simulation results for the skew t model with an unstructured Σ and $\tau = 30$ based on $K = 30$

Parameter	True	Mean	SE	MSD	95% CI	95% CP	Power
$\beta_{1,1}$	0.5	0.493	0.116	0.114	(0.266, 0.727)	94.2	98.8
$\beta_{1,2}$	-1.5	-1.505	0.112	0.114	(-1.720, -1.263)	94.6	100.0
$\beta_{1,3}$	1.5	1.501	0.116	0.114	(1.266, 1.736)	94.0	100.0
$\beta_{2,1}$	-1.0	-0.991	0.096	0.099	(-1.179, -0.804)	95.0	100.0
$\beta_{2,2}$	1.0	0.999	0.095	0.099	(0.817, 1.197)	95.2	100.0
$\beta_{2,3}$	-2.0	-2.007	0.097	0.099	(-2.194, -1.809)	95.0	100.0
$\beta_{3,1}$	1.5	1.492	0.158	0.158	(1.167, 1.797)	95.2	100.0
$\beta_{3,2}$	-0.5	-0.496	0.159	0.158	(-0.813, -0.215)	95.4	87.6
$\beta_{3,3}$	-1.0	-0.992	0.158	0.158	(-1.294, -0.695)	95.6	100.0
$\gamma_{1,0}$	-20.0	-19.972	0.736	0.738	(-21.301, -18.501)	95.0	100.0
$\gamma_{1,1}$	-15.0	-15.011	0.505	0.530	(-15.974, -14.102)	95.4	100.0
$\gamma_{2,0}$	10.0	9.991	0.287	0.289	(9.421, 10.508)	93.8	100.0
$\gamma_{2,1}$	2.0	1.997	0.302	0.284	(1.483, 2.607)	93.0	100.0
$\gamma_{3,0}$	1.5	1.471	0.559	0.532	(0.327, 2.576)	91.4	78.0
$\gamma_{3,1}$	-5.0	-4.947	0.517	0.489	(-5.971, -3.980)	92.4	100.0
Σ_{11}	55.0	54.924	2.053	2.011	(51.019, 58.832)	94.2	100.0
Σ_{12}	10.0	10.021	0.907	0.954	(8.237, 11.584)	97.0	100.0
Σ_{13}	24.0	25.062	1.628	1.570	(22.007, 28.328)	94.0	100.0
Σ_{22}	60.0	60.220	2.152	2.111	(55.884, 64.673)	93.8	100.0
Σ_{23}	-30.0	-30.054	1.347	1.376	(-32.747, -27.701)	96.0	100.0
Σ_{33}	100.0	100.067	3.545	3.582	(93.209, 107.259)	95.0	100.0
$\Omega_{1,0,0}$	15.0	15.725	4.563	4.596	(8.457, 26.085)	94.2	100.0
$\Omega_{1,0,1}$	-7.5	-7.671	2.903	2.849	(-13.425, -2.858)	93.0	97.8
$\Omega_{1,1,1}$	7.0	6.933	2.480	2.394	(2.642, 12.465)	89.4	100.0
$\Omega_{2,0,0}$	2.0	1.870	0.676	0.700	(0.747, 3.305)	89.8	100.0
$\Omega_{2,0,1}$	-1.5	-1.330	0.647	0.618	(-2.720, -0.268)	85.4	77.6
$\Omega_{2,1,1}$	1.5	1.240	0.696	0.648	(0.256, 2.905)	80.8	100.0
$\Omega_{3,0,0}$	7.0	6.675	2.358	2.309	(2.809, 12.235)	88.6	100.0
$\Omega_{3,0,1}$	-5.5	-5.044	2.091	1.999	(-9.603, -1.686)	86.4	94.6
$\Omega_{3,1,1}$	5.0	4.220	2.008	1.942	(1.164, 8.641)	81.6	100.0
δ_1	10.0	10.017	0.193	0.207	(9.643, 10.386)	97.0	100.0
δ_2	5.0	4.982	0.250	0.249	(4.451, 5.461)	94.2	100.0
δ_3	15.0	15.007	0.288	0.276	(14.464, 15.564)	94.6	100.0
ν	5.0	5.020	0.214	0.209	(4.624, 5.466)	94.6	100.0

Table 3.Simulation results for the Box-Cox model with an unstructured Σ based on $K = 15$

Parameter	True	Mean	SE	MSD	95% CI	Power
$\beta_{1,1}$	0.5	0.075	0.035	0.030	(0.020, 0.160)	75.0
$\beta_{1,2}$	-1.5	-0.225	0.060	0.039	(-0.368, -0.132)	100.0
$\beta_{1,3}$	1.5	0.223	0.060	0.038	(0.135, 0.353)	100.0
$\beta_{2,1}$	-1.0	-0.506	0.132	0.100	(-0.818, -0.292)	100.0
$\beta_{2,2}$	1.0	0.517	0.132	0.101	(0.293, 0.814)	100.0
$\beta_{2,3}$	-2.0	-1.034	0.224	0.144	(-1.519, -0.649)	100.0
$\beta_{3,1}$	1.5	0.127	0.044	0.027	(0.071, 0.226)	99.8
$\beta_{3,2}$	-0.5	-0.041	0.025	0.023	(-0.097, 0.000)	40.4
$\beta_{3,3}$	-1.0	-0.084	0.032	0.025	(-0.163, -0.033)	96.2
$\gamma_{1,0}$	-20.0	20.276	8.274	1.662	(12.872, 44.191)	100.0
$\gamma_{1,1}$	-15.0	-2.224	0.555	0.279	(-3.549, -1.379)	100.0
$\gamma_{2,0}$	10.0	45.427	13.690	3.843	(25.290, 76.279)	100.0
$\gamma_{2,1}$	2.0	1.033	0.329	0.257	(0.484, 1.817)	99.2
$\gamma_{3,0}$	1.5	16.808	7.673	1.234	(10.253, 38.093)	100.0
$\gamma_{3,1}$	-5.0	-0.417	0.130	0.088	(-0.704, -0.229)	100.0
Σ_{11}	55.0	4.562	2.433	1.052	(1.731, 11.141)	100.0
Σ_{12}	10.0	1.421	0.540	0.287	(0.669, 2.746)	100.0
Σ_{13}	24.0	0.609	0.268	0.108	(0.284, 1.334)	100.0
Σ_{22}	60.0	35.444	14.890	8.178	(13.797, 68.303)	100.0
Σ_{23}	-30.0	-2.092	0.732	0.349	(-3.911, -1.057)	100.0
Σ_{33}	100.0	2.956	1.987	0.640	(1.143, 7.973)	100.0
$\Omega_{1,0,0}$	15.0	0.358	0.270	0.188	(0.087, 0.962)	100.0
$\Omega_{1,0,1}$	-7.5	-0.151	0.138	0.109	(-0.466, -0.010)	43.6
$\Omega_{1,1,1}$	7.0	0.162	0.111	0.098	(0.046, 0.425)	100.0
$\Omega_{2,0,0}$	2.0	0.500	0.362	0.331	(0.086, 1.492)	100.0
$\Omega_{2,0,1}$	-1.5	-0.334	0.303	0.285	(-1.117, -0.010)	23.6
$\Omega_{2,1,1}$	1.5	0.372	0.302	0.308	(0.084, 1.176)	100.0
$\Omega_{3,0,0}$	7.0	0.059	0.041	0.033	(0.020, 0.174)	100.0
$\Omega_{3,0,1}$	-5.5	-0.036	0.032	0.029	(-0.122, -0.006)	20.0
$\Omega_{3,1,1}$	5.0	0.053	0.032	0.034	(0.022, 0.139)	100.0
λ_1	NA	0.549	0.062	0.026	(0.439, 0.693)	100.0
λ_2	NA	0.839	0.050	0.027	(0.727, 0.925)	100.0
λ_3	NA	0.449	0.067	0.024	(0.349, 0.617)	100.0

Table 4.Simulation results for the Box-Cox model with an unstructured Σ based on $K = 30$

Parameter	True	Mean	SE	MSD	95% CI	Power
$\beta_{1,1}$	0.5	0.056	0.021	0.016	(0.024, 0.105)	94.4
$\beta_{1,2}$	-1.5	-0.171	0.045	0.022	(-0.276, -0.111)	100.0
$\beta_{1,3}$	1.5	0.170	0.045	0.022	(0.109, 0.266)	100.0
$\beta_{2,1}$	-1.0	-0.497	0.119	0.075	(-0.771, -0.302)	100.0
$\beta_{2,2}$	1.0	0.499	0.121	0.075	(0.315, 0.789)	100.0
$\beta_{2,3}$	-2.0	-1.004	0.215	0.115	(-1.487, -0.653)	100.0
$\beta_{3,1}$	1.5	0.094	0.030	0.014	(0.053, 0.178)	100.0
$\beta_{3,2}$	-0.5	-0.030	0.014	0.012	(-0.062, -0.007)	71.4
$\beta_{3,3}$	-1.0	-0.062	0.022	0.013	(-0.114, -0.032)	99.8
$\gamma_{1,0}$	-20.0	18.729	7.558	1.245	(12.532, 38.828)	100.0
$\gamma_{1,1}$	-15.0	-1.696	0.425	0.171	(-2.676, -1.114)	100.0
$\gamma_{2,0}$	10.0	50.326	12.916	3.938	(31.616, 80.314)	100.0
$\gamma_{2,1}$	2.0	1.007	0.270	0.187	(0.555, 1.604)	100.0
$\gamma_{3,0}$	1.5	15.637	7.406	0.867	(9.906, 37.586)	100.0
$\gamma_{3,1}$	-5.0	-0.310	0.099	0.047	(-0.582, -0.186)	100.0
Σ_{11}	55.0	2.639	1.536	0.498	(1.051, 5.887)	100.0
Σ_{12}	10.0	1.047	0.367	0.170	(0.530, 1.922)	100.0
Σ_{13}	24.0	0.341	0.149	0.047	(0.177, 0.721)	100.0
Σ_{22}	60.0	33.455	14.550	6.637	(13.586, 67.075)	100.0
Σ_{23}	-30.0	-1.509	0.556	0.207	(-2.821, -0.792)	100.0
Σ_{33}	100.0	1.634	1.220	0.273	(0.621, 4.468)	100.0
$\Omega_{1,0,0}$	15.0	0.198	0.133	0.071	(0.064, 0.506)	100.0
$\Omega_{1,0,1}$	-7.5	-0.084	0.064	0.041	(-0.239, -0.016)	82.4
$\Omega_{1,1,1}$	7.0	0.092	0.056	0.038	(0.035, 0.240)	100.0
$\Omega_{2,0,0}$	2.0	0.494	0.310	0.221	(0.132, 1.285)	100.0
$\Omega_{2,0,1}$	-1.5	-0.350	0.245	0.193	(-0.972, -0.049)	67.2
$\Omega_{2,1,1}$	1.5	0.357	0.241	0.206	(0.076, 0.991)	100.0
$\Omega_{3,0,0}$	7.0	0.033	0.023	0.012	(0.012, 0.086)	100.0
$\Omega_{3,0,1}$	-5.5	-0.021	0.018	0.011	(-0.058, -0.005)	80.8
$\Omega_{3,1,1}$	5.0	0.029	0.019	0.013	(0.013, 0.071)	100.0
λ_1	NA	0.502	0.058	0.021	(0.415, 0.645)	100.0
λ_2	NA	0.839	0.045	0.023	(0.746, 0.926)	100.0
λ_3	NA	0.403	0.067	0.018	(0.303, 0.578)	100.0

Table 5.Values of L Measure under Various Models for the Cholesterol Data

φ	Normal	t	$\tau = 20$	$\tau = 25$	Skew t $T = 30$	$\tau = 35$	$T = \infty$
$\Sigma = \text{diag}(\Sigma_{11}, \Sigma_{12}, \Sigma_{13})$							
0.0	2903.15	2799.48	2301.98	2295.19	2296.15	2296.32	2315.58
0.1	3931.54	3824.75	3198.36	3190.21	3191.67	3192.19	3217.58
0.2	5171.96	5065.52	4295.18	4285.55	4287.60	4288.54	4320.84
0.3	6643.73	6541.64	5612.10	5600.88	5603.60	5605.04	5645.03
0.4	8366.21	8272.94	7168.79	7155.85	7159.34	7161.34	7209.78
0.5	10358.71	10279.28	8984.90	8970.14	8974.47	8977.12	9034.77
0.6	12640.57	12580.49	11080.10	11063.40	11068.65	11072.04	11139.65
0.7	15231.14	15196.42	13474.07	13455.29	13461.56	13465.76	13544.06
0.8	18149.74	18146.92	16186.45	16165.47	16172.84	16177.94	16267.68
0.9	21415.70	21451.82	19236.91	19213.61	19222.17	19228.25	19330.15
1.0	25048.37	25130.98	22645.12	22619.37	22629.20	22636.35	22751.13
unstructured Σ							
0.0	2656.13	2490.24	2103.03	2098.22	2096.72	2097.72	2114.88
0.1	3600.37	3414.78	2927.69	2921.81	2920.02	2921.39	2943.53
0.2	4745.28	4541.60	3942.70	3935.63	3933.52	3935.33	3963.05
0.3	6110.19	5890.53	5167.70	5159.32	5156.88	5159.18	5193.10
0.4	7714.43	7481.41	6622.37	6612.57	6609.76	6612.61	6653.34
0.5	9577.31	9334.06	8326.39	8315.03	8311.82	8315.29	8363.44
0.6	11718.16	11468.32	10299.40	10286.37	10282.74	10286.89	10343.07
0.7	14156.31	13904.02	12561.09	12546.26	12542.17	12547.07	12611.88
0.8	16911.07	16660.98	15131.11	15114.35	15109.78	15115.50	15189.54
0.9	20001.78	19759.05	18029.15	18010.32	18005.23	18011.84	18095.72
1.0	23447.75	23218.04	21274.86	21253.83	21248.19	21255.76	21350.08

Note: The reported values are $10^{-4}n^{-3}|L_2(y)|$ where the total sample size $n = 21, 184$.

Table 6.Posterior Estimates under the Skew t Model with an Unstructured Σ and $\tau = 30$ for the Cholesterol Data

Par.	Mean	SD	95% HPD	Par.	Mean	SD	95% HPD
$\gamma_{1,0}$	-20.041	3.222	(-26.402, -13.723)	$\Omega_{1,0,0}$	19.550	10.031	(6.153, 38.535)
$\gamma_{1,1}$	-14.043	0.809	(-15.662, -12.454)	$\Omega_{1,0,1}$	-7.840	5.029	(-17.981, -0.409)
$\gamma_{1,2}$	8.221	4.761	(-1.605, 17.174)	$\Omega_{1,1,1}$	7.676	3.822	(2.495, 14.879)
$\gamma_{1,3}$	-18.852	1.418	(-21.586, -15.974)	$\Omega_{1,2,2}$	61.484	29.571	(20.883, 117.366)
$\gamma_{2,0}$	9.770	1.346	(7.167, 12.415)	$\Omega_{1,2,3}$	-31.530	17.416	(-64.179, -7.053)
$\gamma_{2,1}$	2.084	0.400	(1.285, 2.870)	$\Omega_{1,3,3}$	25.162	12.722	(7.521, 48.872)
$\gamma_{2,2}$	13.428	1.793	(9.935, 16.942)	$\Omega_{2,0,0}$	1.988	1.044	(0.578, 3.973)
$\gamma_{2,3}$	1.230	0.259	(0.726, 1.736)	$\Omega_{2,0,1}$	-1.624	0.906	(-3.328, -0.400)
$\gamma_{3,0}$	0.346	2.531	(-4.517, 5.415)	$\Omega_{2,1,1}$	1.474	0.865	(0.315, 3.104)
$\gamma_{3,1}$	-5.783	0.744	(-7.316, -4.369)	$\Omega_{2,2,2}$	1.055	0.728	(0.098, 2.413)
$\gamma_{3,2}$	1.086	3.350	(-5.490, 7.609)	$\Omega_{2,2,3}$	-0.233	0.354	(-0.968, 0.296)
$\gamma_{3,3}$	-7.577	0.845	(-9.213, -5.868)	$\Omega_{2,3,3}$	0.155	0.190	(0.007, 0.494)
Σ_{11}	59.088	1.556	(56.109, 62.159)	$\Omega_{3,0,0}$	7.510	4.905	(0.955, 16.824)
Σ_{12}	8.479	0.711	(7.122, 9.899)	$\Omega_{3,0,1}$	-5.955	3.896	(-13.445, -0.728)
Σ_{13}	27.489	1.477	(24.718, 30.432)	$\Omega_{3,1,1}$	5.021	3.503	(0.377, 11.597)
Σ_{22}	59.911	1.594	(56.843, 63.059)	$\Omega_{3,2,2}$	10.721	6.381	(2.230, 22.786)
Σ_{23}	-33.192	1.200	(-35.503, -30.811)	$\Omega_{3,2,3}$	-8.059	5.152	(-17.428, -1.037)
Σ_{33}	137.653	3.877	(130.291, 145.380)	$\Omega_{3,3,3}$	6.345	4.535	(0.483, 14.659)
p_{12}	0.143	0.012	(0.120, 0.165)	δ_1	10.735	0.158	(10.429, 11.046)
p_{13}	0.305	0.015	(0.276, 0.335)	δ_2	5.574	0.186	(5.203, 5.924)
p_{23}	-0.366	0.011	(-0.387, -0.343)	δ_3	20.376	0.274	(19.831, 20.915)
ν	5.133	0.184	(4.778, 5.497)				

Table 7.Posterior Estimates under the Skew t model with an Unstructured Σ and $\tau = 30$ for LDL-C

	Par.	Mean	SD	95% HPD
bl_ldlc	$\beta_{1,1}$	-0.032	0.004	(-0.040, -0.025)
bl_hdlc	$\beta_{1,2}$	0.014	0.011	(-0.008, 0.034)
bl_tg	$\beta_{1,3}$	0.007	0.002	(0.003, 0.010)
BMI	$\beta_{1,4}$	0.064	0.020	(0.025, 0.102)
age	$\beta_{1,5}$	-0.158	0.011	(-0.178, -0.137)
duration	$\beta_{1,6}$	0.246	0.299	(-0.354, 0.835)
Female	$\beta_{1,7}$	-1.119	0.233	(-1.585, -0.669)
DM	$\beta_{1,8}$	-2.286	0.320	(-2.903, -1.644)
CHD	$\beta_{1,9}$	0.361	0.306	(-0.263, 0.943)
potency2	$\beta_{1,10}$	-7.264	0.368	(-7.963, -6.518)
potency3	$\beta_{1,11}$	-15.751	0.427	(-16.596, -14.927)
Black	$\beta_{1,12}$	2.384	0.450	(1.509, 3.278)
Hispanic	$\beta_{1,13}$	0.657	0.526	(-0.382, 1.670)
Other	$\beta_{1,14}$	-2.110	0.540	(-3.194, -1.077)
onstatin $\times$ bl_ldlc	$\beta_{1,15}$	-0.074	0.007	(-0.088, -0.061)
onstatin $\times$ bl_hdlc	$\beta_{1,16}$	-0.071	0.017	(-0.105, -0.038)
onstatin $\times$ bl_tg	$\beta_{1,17}$	-0.007	0.003	(-0.013, -0.001)
onstatin $\times$ BMI	$\beta_{1,18}$	-0.063	0.033	(-0.127, 0.003)
onstatin $\times$ age	$\beta_{1,19}$	0.126	0.018	(0.091, 0.161)
onstatin $\times$ duration	$\beta_{1,20}$	0.117	0.720	(-1.282, 1.567)
onstatin $\times$ Female	$\beta_{1,21}$	2.546	0.388	(1.791, 3.304)
onstatin $\times$ DM	$\beta_{1,22}$	-1.158	0.477	(-2.088, -0.225)
onstatin $\times$ CHD	$\beta_{1,23}$	-1.283	0.493	(-2.257, -0.325)
onstatin $\times$ potency2	$\beta_{1,24}$	5.967	0.646	(4.678, 7.226)
onstatin $\times$ potency3	$\beta_{1,25}$	13.716	0.816	(12.071, 15.277)
onstatin $\times$ Black	$\beta_{1,26}$	-3.074	0.814	(-4.619, -1.443)
onstatin $\times$ Hispanic	$\beta_{1,27}$	-2.454	1.000	(-4.425, -0.506)
onstatin $\times$ Other	$\beta_{1,28}$	2.817	0.913	(0.983, 4.579)

Table 8.Posterior Estimates under the Skew t model with an Unstructured Σ and $\tau = 30$ for HDL-C

	Par.	Mean	SD	95% HPD
bl_ldlc	$\beta_{2,1}$	0.003	0.003	(-0.003, 0.010)
bl_hdlc	$\beta_{2,2}$	-0.202	0.010	(-0.220, -0.182)
bl_tg	$\beta_{2,3}$	0.023	0.001	(0.020, 0.026)
BMI	$\beta_{2,4}$	-0.135	0.017	(-0.168, -0.101)
Age	$\beta_{2,5}$	0.081	0.009	(0.062, 0.099)
duration	$\beta_{2,6}$	0.068	0.064	(-0.058, 0.195)
Female	$\beta_{2,7}$	0.567	0.208	(0.162, 0.971)
DM	$\beta_{2,8}$	-2.201	0.270	(-2.734, -1.678)
CHD	$\beta_{2,9}$	-0.916	0.265	(-1.435, -0.398)
potency2	$\beta_{2,10}$	0.772	0.305	(0.157, 1.348)
potency3	$\beta_{2,11}$	0.074	0.354	(-0.618, 0.765)
Black	$\beta_{2,12}$	-2.477	0.369	(-3.214, -1.768)
Hispanic	$\beta_{2,13}$	-1.074	0.450	(-1.934, -0.159)
Other	$\beta_{2,14}$	-0.356	0.490	(-1.312, 0.601)
onstatin $\times$ bl_ldlc	$\beta_{2,15}$	-0.009	0.005	(-0.019, 0.001)
onstatin $\times$ bl_hdlc	$\beta_{2,16}$	0.002	0.015	(-0.027, 0.031)
onstatin $\times$ bl_tg	$\beta_{2,17}$	-0.015	0.002	(-0.020, -0.010)
onstatin $\times$ BMI	$\beta_{2,18}$	0.023	0.027	(-0.031, 0.075)
onstatin $\times$ age	$\beta_{2,19}$	-0.060	0.015	(-0.088, -0.031)
onstatin $\times$ duration	$\beta_{2,20}$	-0.216	0.186	(-0.591, 0.146)
onstatin $\times$ Female	$\beta_{2,21}$	0.764	0.320	(0.137, 1.385)
onstatin $\times$ DM	$\beta_{2,22}$	1.505	0.385	(0.792, 2.293)
onstatin $\times$ CHD	$\beta_{2,23}$	0.404	0.387	(-0.363, 1.157)
onstatin $\times$ potency2	$\beta_{2,24}$	-0.930	0.518	(-1.955, 0.068)
onstatin $\times$ potency3	$\beta_{2,25}$	-0.662	0.641	(-1.919, 0.580)
onstatin $\times$ Black	$\beta_{2,26}$	2.542	0.642	(1.286, 3.808)
onstatin $\times$ Hispanic	$\beta_{2,27}$	-0.498	0.785	(-2.079, 1.003)
onstatin $\times$ Other	$\beta_{2,28}$	-0.697	0.745	(-2.141, 0.762)

Table 9.Posterior Estimates under the Skew t model with an Unstructured Σ and $\tau = 30$ for TG

	Par.	Mean	SD	95% HPD
bl_ldlc	$\beta_{3,1}$	-0.002	0.006	(-0.014, 0.009)
bl_hdlc	$\beta_{3,2}$	0.023	0.018	(-0.011, 0.058)
bl_tg	$\beta_{3,3}$	-0.094	0.003	(-0.099, -0.088)
BMI	$\beta_{3,4}$	0.228	0.032	(0.168, 0.294)
Age	$\beta_{3,5}$	-0.031	0.018	(-0.066, 0.003)
duration	$\beta_{3,6}$	-0.265	0.140	(-0.551, 0.001)
Female	$\beta_{3,7}$	2.654	0.389	(1.877, 3.404)
DM	$\beta_{3,8}$	-0.557	0.508	(-1.553, 0.437)
CHD	$\beta_{3,9}$	0.723	0.499	(-0.279, 1.675)
potency2	$\beta_{3,10}$	-4.181	0.574	(-5.334, -3.095)
potency3	$\beta_{3,11}$	-9.353	0.671	(-10.700, -8.091)
Black	$\beta_{3,12}$	-1.849	0.700	(-3.178, -0.432)
Hispanic	$\beta_{3,13}$	1.383	0.860	(-0.336, 3.039)
Other	$\beta_{3,14}$	-1.257	0.898	(-3.034, 0.525)
onstatin $\times$ bl_ldlc	$\beta_{3,15}$	0.015	0.010	(-0.004, 0.034)
onstatin $\times$ bl_hdlc	$\beta_{3,16}$	-0.087	0.028	(-0.141, -0.034)
onstatin $\times$ bl_tg	$\beta_{3,17}$	-0.028	0.005	(-0.037, -0.019)
onstatin $\times$ BMI	$\beta_{3,18}$	0.047	0.053	(-0.058, 0.150)
onstatin $\times$ age	$\beta_{3,19}$	0.098	0.028	(0.042, 0.151)
onstatin $\times$ duration	$\beta_{3,20}$	0.882	0.333	(0.227, 1.539)
onstatin $\times$ Female	$\beta_{3,21}$	0.279	0.624	(-0.986, 1.450)
onstatin $\times$ DM	$\beta_{3,22}$	2.189	0.745	(0.681, 3.593)
onstatin $\times$ CHD	$\beta_{3,23}$	0.296	0.745	(-1.153, 1.742)
onstatin $\times$ potency2	$\beta_{3,24}$	3.296	1.021	(1.328, 5.299)
onstatin $\times$ potency3	$\beta_{3,25}$	6.352	1.221	(3.928, 8.689)
onstatin $\times$ Black	$\beta_{3,26}$	-0.983	1.264	(-3.490, 1.445)
onstatin $\times$ Hispanic	$\beta_{3,27}$	1.053	1.549	(-1.934, 4.113)
onstatin $\times$ Other	$\beta_{3,28}$	4.217	1.435	(1.354, 6.976)

Table 10.

Posterior Estimates of the Regression Coefficients for Second-Line Therapy under the Skew t Model with an Unstructured Σ and $\tau = 30$

		Par.	Mean	SD	95% HPD
LDL-C	bl_ldlc	$\beta_{1,1} + \beta_{1,15}$	-0.107	0.006	(-0.118, -0.095)
	bl_hdlc	$\beta_{1,2} + \beta_{1,16}$	-0.057	0.013	(-0.083, -0.031)
	bl_tg	$\beta_{1,3} + \beta_{1,17}$	0.000	0.002	(-0.005, 0.004)
	BMI	$\beta_{1,4} + \beta_{1,18}$	0.001	0.027	(-0.052, 0.053)
	Age	$\beta_{1,5} + \beta_{1,19}$	-0.032	0.014	(-0.060, -0.004)
	duration	$\beta_{1,6} + \beta_{1,20}$	0.363	0.660	(-0.924, 1.696)
	Female	$\beta_{1,7} + \beta_{1,21}$	1.428	0.310	(0.836, 2.048)
	DM	$\beta_{1,8} + \beta_{1,22}$	-3.444	0.354	(-4.149, -2.751)
	CHD	$\beta_{1,9} + \beta_{1,23}$	-0.922	0.390	(-1.705, -0.182)
	potency2	$\beta_{1,10} + \beta_{1,24}$	-1.297	0.531	(-2.323, -0.236)
	potency3	$\beta_{1,11} + \beta_{1,25}$	-2.036	0.694	(-3.417, -0.707)
	Black	$\beta_{1,12} + \beta_{1,26}$	-0.689	0.679	(-1.985, 0.666)
	Hispanic	$\beta_{1,13} + \beta_{1,27}$	-1.797	0.844	(-3.438, -0.155)
	Other	$\beta_{1,14} + \beta_{1,28}$	0.707	0.738	(-0.731, 2.135)
HDL-C	bl_ldlc	$\beta_{2,1} + \beta_{2,15}$	-0.005	0.004	(-0.014, 0.002)
	bl_hdlc	$\beta_{2,2} + \beta_{2,16}$	-0.200	0.011	(-0.221, -0.178)
	bl_tg	$\beta_{2,3} + \beta_{2,17}$	0.008	0.002	(0.005, 0.012)
	BMI	$\beta_{2,4} + \beta_{2,18}$	-0.113	0.021	(-0.154, -0.071)
	Age	$\beta_{2,5} + \beta_{2,19}$	0.021	0.011	(0.000, 0.043)
	duration	$\beta_{2,6} + \beta_{2,20}$	-0.147	0.175	(-0.497, 0.198)
	Female	$\beta_{2,7} + \beta_{2,21}$	1.331	0.246	(0.850, 1.815)
	DM	$\beta_{2,8} + \beta_{2,22}$	-0.696	0.275	(-1.229, -0.155)
	CHD	$\beta_{2,9} + \beta_{2,23}$	-0.513	0.286	(-1.071, 0.045)
	potency2	$\beta_{2,10} + \beta_{2,24}$	-0.158	0.419	(-0.980, 0.661)
	potency3	$\beta_{2,11} + \beta_{2,25}$	-0.588	0.532	(-1.650, 0.426)
	Black	$\beta_{2,12} + \beta_{2,26}$	0.065	0.529	(-0.989, 1.084)
	Hispanic	$\beta_{2,13} + \beta_{2,27}$	-1.573	0.642	(-2.848, -0.319)
	Other	$\beta_{2,14} + \beta_{2,28}$	-1.053	0.560	(-2.140, 0.053)
TG	bl_ldlc	$\beta_{3,1} + \beta_{3,15}$	0.013	0.008	(-0.003, 0.028)
	bl_hdlc	$\beta_{3,2} + \beta_{3,16}$	-0.063	0.021	(-0.105, -0.022)
	bl_tg	$\beta_{3,3} + \beta_{3,17}$	-0.122	0.004	(-0.129, -0.114)
	BMI	$\beta_{3,4} + \beta_{3,18}$	0.274	0.042	(0.193, 0.360)
	Age	$\beta_{3,5} + \beta_{3,19}$	0.067	0.022	(0.024, 0.110)
	duration	$\beta_{3,6} + \beta_{3,20}$	0.616	0.302	(0.037, 1.220)
	Female	$\beta_{3,7} + \beta_{3,21}$	2.933	0.490	(1.950, 3.863)

	Par.	Mean	SD	95% HPD
DM	$\beta_{3,8} + \beta_{3,22}$	1.632	0.544	(0.577, 2.716)
CHD	$\beta_{3,9} + \beta_{3,23}$	1.019	0.553	(−0.048, 2.111)
potency2	$\beta_{3,10} + \beta_{3,24}$	−0.885	0.842	(−2.541, 0.767)
potency3	$\beta_{3,11} + \beta_{3,25}$	−3.002	1.022	(−4.993, −0.985)
Black	$\beta_{3,12} + \beta_{3,26}$	−2.831	1.056	(−4.916, −0.810)
Hispanic	$\beta_{3,13} + \beta_{3,27}$	2.435	1.299	(−0.027, 5.035)
Other	$\beta_{3,14} + \beta_{3,28}$	2.960	1.114	(0.725, 5.083)