

Bayesian flexible hierarchical skew heavy-tailed multivariate meta regression models for individual patient data with applications

SUNGDUK KIM, MING-HUI CHEN, JOSEPH IBRAHIM*,
ARVIND SHAH, AND JIANXIN LIN

A flexible class of multivariate meta-regression models are proposed for Individual Patient Data (IPD). The methodology is well motivated from 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. The research 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. The proposed multivariate meta-regression models allow for different skewness parameters and different degrees of freedom for the multivariate outcomes from different trials under a general class of skew t -distributions. The theoretical properties of the proposed models are examined and an efficient Markov chain Monte Carlo (MCMC) sampling algorithm is developed for carrying out Bayesian inference under the proposed multivariate meta-regression model. In addition, the Conditional Predictive Ordinates (CPOs) are computed via an efficient Monte Carlo method. Consequently, the logarithm of the pseudo marginal likelihood and Bayesian residuals are obtained for model comparison and assessment, respectively. A detailed analysis of the IPD meta data from the 26 Merck clinical trials is carried out to demonstrate the usefulness of the proposed methodology.

KEYWORDS AND PHRASES: Collapsed Gibbs, CPO Identity II, Heterogeneity, Multi-dimensional random effects, Multiple trials, Outlying trials.

1. INTRODUCTION

According to the National Vital Statistics Reports (Heron 2019) cardiovascular disease (CVD) continues to be the leading cause of death for both men and women. This is the case in the U.S. and worldwide. More than half of all people who die due to heart disease are men. It has been confirmed that increased low density lipoprotein cholesterol (LDL-C) is an

independent risk factor for CVD. LDL-C lowering has been consistently shown to reduce the risk of CVD. One large meta-analysis (Baigent et al. 2010) of statin clinical trials shows a progressive reduction in risk of major CVD events with lower on-treatment LDL-C levels. Although LDL-C is a primary cause of CVD, other risk factors contribute, as well, for example, high-density lipoprotein cholesterol (HDL-C) and triglycerides (TG). Large cohort studies show a strong and inverse relationship of HDL-C levels with the risk of incident CVD independent of other lipids. HDL-C is positively associated with a decreased risk of coronary heart disease (CHD). As defined by the US National Cholesterol Education Program Adult Treatment Panel III guidelines (ATP III 2001), an HDL-C level of 60 mg/dL or greater is a negative risk factor. A long-standing association exists between elevated triglyceride levels and CVD (Austin et al. 1998; Sarwar et al. 2007). In a meta-analysis of 17 studies, increased triglyceride levels were associated with increased coronary disease risk in both men and women, after adjustment for HDL-C and other risk factors (Hokanson and Austin 1996). A randomized, controlled clinical trial REDUCE-IT (Bhatt et al. 2019) has demonstrated that intervention to low triglyceride level is associated with reduced CVD events. Among lipid-lowering drugs, statins are the cornerstone of therapy. Ezetimibe is the most commonly used nonstatin agent. It lowers LDL-C levels by 13% to 20% and has a low incidence of side effects (Cannon et al. 2015; Kashani et al. 2008).

Meta-regression (MR) of individual patient data (IPD) is an effective modeling tool for explaining heterogeneity between trials, synthesizing evidence across studies, investigating individual-level interactions, or identifying subgroups (Simmonds and Higgins 2007; Ritz et al. 2008; Kim et al. 2013; Riley et al. 2015; Burke et al. 2017; Belias et al. 2019; Ibrahim et al. 2019). In dealing with IPD multivariate meta-data, it is often the case that the data may be highly skewed and or have heavy-tailed and non-normal distributions to properly model certain response variables which may have skew and/or heavy-tailed distributions.

The modeling framework proposed here is motivated from multivariate IPD data from 26 Merck clinical trials for cholesterol lowering drugs. In our application, we consider a

*Corresponding author.

three-dimensional continuous response, in which some components of the response variable are heavy-tailed and/or skew distributions, and some components may have symmetric and/or light tailed distributions. We do not know which are which in advance and our hope is to develop a model that accommodates this flexibility. Thus, in these settings, one needs more complex models than the traditional linear mixed model. There is abundant literature on using skew and/or heavy tailed distributions for modeling univariate and or multivariate data (Chen et al. 1999; Branco and Dey 2001; Azzalini and Capitanio 2003; Sahu et al. 2003; Genton 2004; Adcock 2004; Kim et al. 2008; Arellano-Valle and Genton 2010; Chang and Zimmerman 2016), but sparse in the multivariate MR setting (Kim et al. 2013; Ibrahim et al. 2019). One of the challenges of modeling skew and heavy tailed distributions in the MR setting is that one needs to develop a flexible class of models that allow certain components of the multivariate response to have skew and/or heavy-tailed while allowing for other components to have symmetric and/or light-tailed distributions, while at the same time, capturing heterogeneity between trials via appropriate random effects. Since the skewness and heaviness of the tails will not be known in advance, one needs to also model the skew parameters and scale parameters appropriately to correctly capture the data structure in this complex multivariate MR setting.

Instead of using a Box-Cox transformation on the multivariate response variables as in Kim et al. (2013), we extend the multivariate skew MR models of Ibrahim et al. (2019) to develop a flexible class of multivariate MR models that accommodate skewness and heavy tailed distributions for multivariate meta-data. Under the proposed models, we first assume that the skewness parameters, the covariance matrices of the multivariate responses, and the degrees of freedom in the multivariate t distributions for the error terms are different across trials at the first stage, and then assume hierarchical priors for the multivariate random effects, the skewness parameters, the covariance matrices, and the degrees of freedom at the second stage. The proposed models are very flexible and general. As empirically shown in Section 5, the proposed model leads to a substantial gain in the goodness-of-fit of the multivariate IPD data from the 26 Merck clinical trials. Due to the complexity and computational challenge of the proposed models, an efficient Markov chain Monte Carlo (MCMC) sampling algorithm is developed for sampling from the posterior distribution of the model parameters. Moreover, one needs to develop and use model assessment tools to examine and find the best fitting models. In this paper, we consider the Conditional Predictive Ordinate (CPO) and develop its computational implementation for the proposed model. It is shown that CPO along with our flexible modeling framework identifies a more suitable model than a traditional MR model.

The rest of the paper is organized as follows. In Section 2, we discuss the Merck cholesterol data in detail. Sections

3.1–3.2 lay out the modeling details of our proposed multivariate skew heavy-tailed random effects meta-regression model, the prior distributions, properties of the proposed model as well as a flow diagram explaining all of the components of the model. Section 3.3 develops the likelihood function and joint posterior distribution of all the parameters. Section 4.1 develops new analytic and computational results for CPO, and Section 4.2 gives details of the MCMC computational development. Section 5 presents a detailed analysis of the cholesterol data showing that our proposed model provides a better interpretation and fit over the existing MR model. We close the paper with a discussion in Section 6.

2. THE CHOLESTEROL DATA

We consider the individual patient data (IPD) from 26 Merck-sponsored double-blind, randomized, active or placebo-controlled clinical trials on adult patients with primary hypercholesterolemia, which were analyzed by Kim et al. (2013) and Ibrahim et al. (2019). The IPD data considered in our analyses are a subset of the meta-data published in Leiter et al. (2011). The citations of the primary papers published in clinical journals for these 26 trials can be found in Leiter et al. (2011). A detailed summary of the covariates was provided in Tables 1 and 2 of Kim et al. (2013).

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. Statin potency describes the chemical/medicinal strength of the statin, and it was categorized into three potency classes (low, medium, and high). The covariates include treatment (trt: 0 = Statin and 1 = Statin + EZE), baseline LDL-C (bl-ldlc), baseline HDL-C (bl-hdlc), baseline TG (bl-tg), age, race (White (reference), Black, Hispanic, and Other), gender (Female: 0 = male (reference), 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). Tables 1 and 2 of Kim et al. (2013) show a considerable amount of heterogeneity in the covariates across the trials. Therefore, 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. As empirically

shown in Kim et al. (2013) and Ibrahim et al. (2019), skew and heavy-tailed distributions are needed for modeling these three primary outcome variables.

3. MULTIVARIATE META REGRESSION MODELS

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 $\mathbf{y}_{ik} = (y_{i1k}, \dots, y_{iJk})'$ 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 $\text{onstatain}_k = 1$ if patients were on statin and 0 if not on statin prior to the trial. Furthermore, let \mathbf{x}_{ijk} denote a p_j -dimensional vector of covariates for the j^{th} response corresponding to the i^{th} patient, and $\boldsymbol{\beta}_j = (\beta_{j1}, \dots, \beta_{jp_j})'$ is the vector of fixed effects regression coefficients corresponding to the p_j covariates. Let \mathbf{w}_{ijk} denote a q_j -dimensional vector for the random effects.

3.1 Preliminary

The multivariate skew meta-regression model proposed by Ibrahim et al. (2019) is given by

$$\begin{aligned} y_{ijk} &= [\gamma_{jk0} + \gamma_{jk1}\text{trt}_{ik}](1 - \text{onstatain}_k) \\ &\quad + [\gamma_{jk2} + \gamma_{jk3}\text{trt}_{ik}]\text{onstatain}_k \\ &\quad + \delta_j(z_{ijk} - E[z_{ijk}]) + \mathbf{x}'_{ijk}\boldsymbol{\beta}_j + \epsilon_{ijk}, \\ &\quad (\epsilon_{i1k}, \dots, \epsilon_{iJk})' | \lambda_{ik} \sim N(\mathbf{0}, \lambda_{ik}^{-1}\boldsymbol{\Sigma}), \\ (3.1) \quad \lambda_{ik} &\sim \text{Gamma}(\nu/2, \nu/2), \end{aligned}$$

where $z_{ijk} | \psi_{ik}$ follows an exponential distribution with mean $\frac{1}{\psi_{ik}}$, $\psi_{ik} \sim \text{Gamma}(\tau + 1, \tau)$, where $\text{Gamma}(a, b)$ denotes the gamma distribution with mean a/b and variance a/b^2 , and $\boldsymbol{\gamma}_j = (\gamma_{jk0}, \gamma_{jk1}, \gamma_{jk2}, \gamma_{jk3})' \sim N_4(\boldsymbol{\gamma}_j, \boldsymbol{\Omega}_j)$ with $\boldsymbol{\gamma}_j = (\gamma_{j0}, \gamma_{j1}, \gamma_{j2}, \gamma_{j3})'$. The model in (3.1) assumed that z_{i1k}, \dots, z_{iJk} are dependent as well as the same covariance matrices, skewness parameters, and degrees of freedom across K trials. Let $\lambda_k = \sum_{i=1}^{n_k} \lambda_{ik}$. Following Ibrahim et al. (2019), we define

$$(3.2) \quad \lambda_k^* = (\lambda_k - n_k)/(n_k/(\nu/2))^{1/2}.$$

Then, we have $E[\lambda_k^*] = 0$ and $\text{Var}(\lambda_k^*) = 1$ a priori, where the expectation and variance are taken with respect to $\text{Gamma}(n_k\nu/2, \nu/2)$, and consequently, λ_k^* approximately follows the standard $N(0, 1)$ distribution. Figure 4 of Ibrahim et al. (2019) shows the boxplots of the λ_k^* 's from the posterior distribution under the skew t model in (3.1) with $\tau = 30$ and an unstructured $\boldsymbol{\Sigma}$ for the cholesterol data. From these boxplots, it was found that the posterior dis-

tributions of the λ_k^* 's substantially depart from their prior distributions for three trials. The reasons for such a discrepancy between the prior and posterior distributions of λ_k^* could be two-fold: (i) outlying trials or (ii) lack of fit due to the use of the same degrees of freedom ν across all trials.

3.2 Hierarchical skew heavy-tailed multivariate meta regression models

We propose the following flexible hierarchical skew heavy-tailed multivariate meta regression models as follows.

Stage 1: Model for Multivariate Responses

$$(3.3) \quad y_{ijk} = \mathbf{x}'_{ijk}\boldsymbol{\beta}_j + \mathbf{w}'_{ik}\boldsymbol{\gamma}_{jk} + \delta_{jk}(z_{ijk} - E[z_{ijk}]) + \epsilon_{ijk},$$

where $\boldsymbol{\gamma}_{jk} = (\gamma_{j1k}, \dots, \gamma_{jq_jk})'$ represents the vector of q_j -dimensional random effects for the j^{th} response, δ_{jk} is a skewness parameter for the j^{th} response in the k^{th} trial, z_{ijk} is the skewness latent variable with the expected value $E[z_{ijk}]$, and $\boldsymbol{\epsilon}_{ik} = (\epsilon_{i1k}, \dots, \epsilon_{iJk})'$ is the vector of error terms. We assume

$$(3.4) \quad \begin{aligned} \boldsymbol{\epsilon}_{ik} | \lambda_{ik} &\sim N(\mathbf{0}, \lambda_{ik}^{-1}\boldsymbol{\Sigma}_k) \quad \text{and} \\ \lambda_{ik} &\sim \text{Gamma}(\nu_k/2, \nu_k/2), \end{aligned}$$

where $\boldsymbol{\Sigma}_k$ is a $J \times J$ positive definite covariance matrix and $\nu_k > 0$ is an unknown parameter. In (3.4), ν_k corresponds the degrees of freedom and the variance of $\boldsymbol{\epsilon}_{ik}$ is finite for all $\nu_k > 2$. Also, in (3.3), we assume that

$$(3.5) \quad z_{ijk} | \psi_{ik} \sim \mathcal{E}(\psi_{ik}) \quad \text{and} \quad \psi_{ik} \sim \text{Gamma}(\tau + 1, \tau),$$

where $\mathcal{E}(\psi_{ik})$ denotes an exponential distribution with mean $\frac{1}{\psi_{ik}}$ and $\text{Gamma}(\tau + 1, \tau)$ denotes a Gamma distribution with mean $\frac{\tau+1}{\tau}$ and variance $\frac{\tau+1}{\tau^2}$. Under this assumption, $E[z_{ijk}] = 1$ and $\text{Var}(z_{ijk}) = \frac{\tau+1}{\tau-1}$ for $\tau > 1$. Let $\mathbf{z}_{ik}^\delta = (z_{i1k}^\delta, \dots, z_{iJk}^\delta)' = (\delta_{1k}(z_{i1k} - E[z_{i1k}]), \dots, \delta_{Jk}(z_{iJk} - E[z_{iJk}]))'$. The covariance matrix of \mathbf{z}_{ik}^δ is given by

$$(3.6) \quad \text{Var}(\mathbf{z}_{ik}^\delta) = \frac{\tau}{\tau-1} \text{diag}(\delta_{1k}^2, \dots, \delta_{Jk}^2) + \frac{1}{\tau-1} \boldsymbol{\delta}_k \boldsymbol{\delta}_k',$$

where $\boldsymbol{\delta}_k = (\delta_{1k}, \dots, \delta_{Jk})'$ for $k = 1, \dots, K$. From (3.6), we note that the correlations of the responses, y_{ijk} , depend on the k^{th} skewness parameter δ_{jk} , and the z_{ijk} 's are independent when $\tau \rightarrow \infty$. We further assume that $\boldsymbol{\gamma}_{jk}$, z_{ijk} , and $\boldsymbol{\epsilon}_{ik}$ are independent.

At Stage 2, models for the random effects, covariance matrices, skewness parameters, and degrees of freedom are specified as follows.

Stage 2a: Model for Random Effects

$$(3.7) \quad \boldsymbol{\gamma}_{jk} \sim N(\boldsymbol{\gamma}_j, \boldsymbol{\Omega}_j),$$

where $\boldsymbol{\gamma}_j = (\gamma_{j1}, \dots, \gamma_{jq_j})'$ is the overall mean vector of γ_{jk} and Ω_j is a $q_j \times q_j$ covariance matrix of the random effects γ_{jk} .

Stage 2b: Model for Covariance Matrices

$$(3.8) \quad \Sigma_k^{-1} \sim \text{Wishart}_J(v, (v - J - 1)\Sigma),$$

where Σ is a $J \times J$ overall covariance matrix and $v > J + 1$. Also Σ_k has prior expectation $E[\Sigma_k | \Sigma] = \Sigma$ when $v > J + 1$. The model (3.8) is attractive as it allows for ‘‘borrowing of strength’’ across trials through the common second-level covariance matrix Σ and it also accounts for the heterogeneity of the within-study covariance matrices among different trials at the same time. The parameter v in (3.8) controls the amount of borrowing across trials. The larger the value of v , the more the within-trial covariance matrices borrow strength from different trials. Note that $\text{Wishart}_J(d_0, S_0)$ denotes the Wishart distribution with mean $d_0 S_0$. That is, $\pi(\Sigma_k^{-1} | d_0, S_0) \propto |\Sigma_k^{-1}|^{(d_0 - J - 1)/2} \exp(-\frac{1}{2} \text{tr}(S_0^{-1} \Sigma_k^{-1}))$.

Stage 2c: Model for Skewness Parameters

$$(3.9) \quad \delta_{jk} \sim N\left(\delta_j, \sigma_{\delta_j}^2\right),$$

where $-\infty < \delta_j < \infty$ is the overall skewness parameter for the j^{th} response and $\sigma_{\delta_j}^2 > 0$ is the variance parameter, controlling the amount of ‘‘borrowing’’ across trials for within-trial skewness parameters for the j^{th} response.

Stage 2d: Model for Degrees of Freedom

$$(3.10) \quad \nu_k \sim \text{Gamma}(\nu_a, \nu_a / \nu_b),$$

where $\nu_a > 0$ controls the amount of borrowing across trials and $\nu_b > 0$ is the overall degrees of freedom. Under (3.10), the prior mean of ν_k is $E[\nu_k] = \nu_b$.

At Stage 3, the prior distributions of the hyperparameters for the random effects, covariance matrices, skewness parameters, and degrees of freedom, which are proposed at Stage 2, as well as the regression coefficients, are specified as follows. Let $\boldsymbol{\delta} = (\delta_1, \dots, \delta_J)'$ and $\boldsymbol{\sigma}_{\delta}^2 = (\sigma_{\delta_1}^2, \dots, \sigma_{\delta_J}^2)'$. We assume that $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\delta}^*$, Σ^* , $\boldsymbol{\nu}^*$, τ , and Ω are independent *a priori*.

Stage 3a: Prior distributions of the Hyperparameters for Random Effects

$$(3.11) \quad \boldsymbol{\gamma} \sim N_{qJ}(0, c_1 I_{qJ}) \text{ and } \Omega_j^{-1} \sim \text{Wishart}_{q_j}(d_1, S_1),$$

where $q = \sum_{j=1}^J q_j$.

Stage 3b: Prior distributions of the Hyperparameters for Covariance Matrices

$$(3.12) \quad v \sim \text{Gamma}(a_1, b_1) \text{ and } \Sigma \sim \text{Wishart}_J(d_2, S_2).$$

Stage 3c: Prior distributions of the Hyperparameters for Skewness Parameters and Latent Variables

$$(3.13) \quad \boldsymbol{\delta} \sim N_J(0, c_2 I_J), \sigma_{\delta_j}^2 \sim \text{IGamma}(a_2, b_2), \text{ and } \tau \sim \text{Gamma}(a_3, b_3) 1(\tau > 1),$$

where $\text{IGamma}(a, b)$ denotes the inverse gamma distribution with mean $b/(a - 1)$ when $a > 1$ and variance $b^2/[(a - 1)^2(a - 2)]$ when $a > 2$.

Stage 3d: Prior distributions of the Hyperparameters for Degrees of Freedom

$$(3.14) \quad \nu_a \sim \text{Gamma}(a_4, b_4) \text{ and } \nu_b \sim \text{IGamma}(a_5, b_5).$$

Stage 3e: Prior distribution for Fixed Effects Regression Coefficients

$$(3.15) \quad \boldsymbol{\beta} \sim N_p(0, c_3 I_p),$$

where $p = \sum_{j=1}^J p_j$.

The multivariate meta-regression model defined in (3.3), (3.4), (3.6), (3.7), (3.9), and (3.10) is very general and flexible, and it includes as special cases the multivariate normal meta-regression model, the multivariate t meta-regression model, and the multivariate skew t meta-regression model. Furthermore, this proposed model also incorporates the different covariance matrices, skewness parameters, and degrees of freedom across the K trials. In the analysis, the hyperparameters of the prior distribution at Stage 3 were specified as $c_1 = 100$, $c_2 = 100$, $c_3 = 100$, $d_1 = q_j + 0.1$, $S_1 = 0.1$, $d_2 = J + 0.1$, $S_2 = 0.1$, $a_1 = 1$, $b_1 = 0.1$, $a_2 = 0.1$, $b_2 = 0.1$, $a_3 = 1$, $b_3 = 0.1$, $a_4 = 1$, $b_4 = 0.1$, $a_5 = 0.1$, and $b_5 = 0.1$. These choices of the hyperparameters lead to noninformative priors. The flow diagram of the proposed model is given in Figure 1. Simultaneous estimation of all parameters is not easy and requires a sophisticated and computationally intensive MCMC sampling algorithm.

3.3 The likelihood function and posterior distribution

Let $X_{ik} = \text{diag}(\mathbf{x}'_{i1k}, \dots, \mathbf{x}'_{iJk})$, $W_{ik} = \text{diag}(\mathbf{w}'_{ik}, \dots, \mathbf{w}'_{ik})$, $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \dots, \boldsymbol{\beta}'_J)'$, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}'_1, \dots, \boldsymbol{\gamma}'_J)'$, $\boldsymbol{\delta}^* = (\boldsymbol{\delta}'_1, \dots, \boldsymbol{\delta}'_K)$, $\Sigma^* = (\Sigma_1, \dots, \Sigma_K)$, $\boldsymbol{\nu}^* = (\nu_1, \dots, \nu_K)'$, $\boldsymbol{\gamma}_k^R = (\boldsymbol{\gamma}'_{1k}, \dots, \boldsymbol{\gamma}'_{Jk})'$, and $\boldsymbol{\psi} = (\psi_{11}, \dots, \psi_{n_K K})'$. Also let $\mathbf{y}_{ik} = (y_{i1k}, \dots, y_{iJk})'$, $\mathbf{y} = (\mathbf{y}'_{11}, \dots, \mathbf{y}'_{n_K K})'$, $W = (W_{11}, \dots, W_{n_K K})$, and $X = (X_{11}, \dots, X_{n_K K})$. Furthermore, we let $D_{\text{obs}} = (\mathbf{y}, X, W)$ denote the observed data. Then the complete-data likelihood function is given by

$$\begin{aligned} L(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}^*, \Sigma^*, \boldsymbol{\nu}^*, \tau, \Omega, \boldsymbol{\gamma}^R, \mathbf{z}, \boldsymbol{\psi}, \boldsymbol{\lambda} | D_{\text{obs}}) \\ = \prod_{k=1}^K \prod_{i=1}^{n_k} f(\mathbf{y}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \mathbf{z}_{ik}, \lambda_{ik}, X_{ik}, W_{ik}) \end{aligned}$$

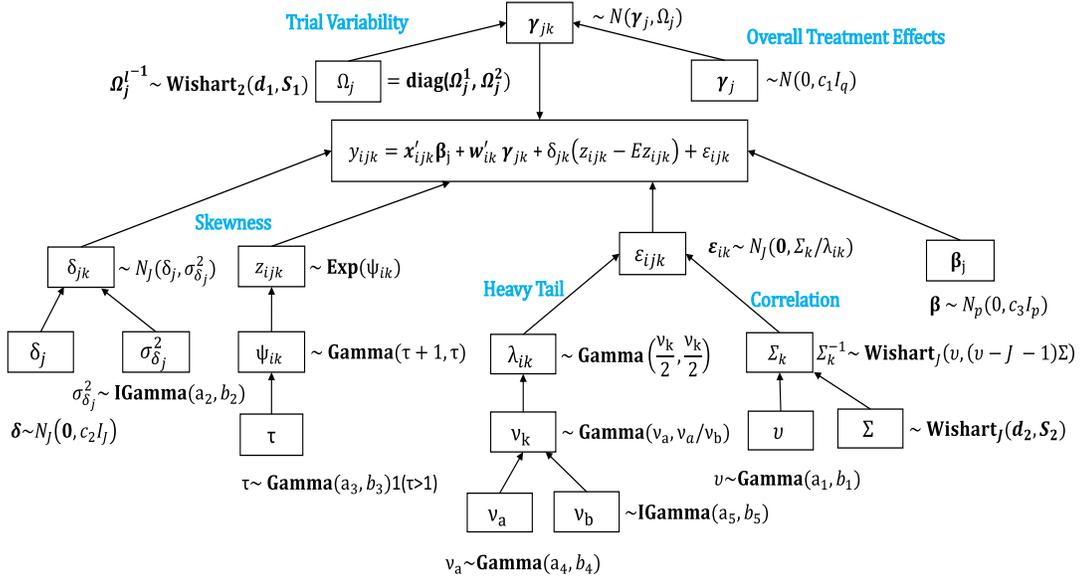


Figure 1. Flow Diagram of the Proposed Model.

$$\begin{aligned}
& \times \left(\frac{\nu_k}{2} \right)^{\frac{\nu_k}{2}} \lambda_{ik}^{\frac{\nu_k}{2}-1} \exp \left\{ -\frac{\nu_k}{2} \lambda_{ik} \right\} \\
& \times \prod_{k=1}^K \prod_{i=1}^{n_k} \left[\prod_{j=1}^J \psi_{ijk} \exp(-\psi_{ijk} z_{ijk}) \right] \\
& \times \frac{\tau^{\tau+1}}{\Gamma(\tau+1)} \psi_{ik}^{\tau} \exp \left\{ -\tau \psi_{ik} \right\} \\
& \times \prod_{k=1}^K \prod_{j=1}^J (2\pi)^{-\frac{4}{2}} |\Omega_j|^{-\frac{1}{2}} \\
(3.16) \quad & \times \exp \left\{ -\frac{1}{2} (\gamma_{jk} - \gamma_j)' \Omega_j^{-1} (\gamma_{jk} - \gamma_j) \right\},
\end{aligned}$$

where $f(\mathbf{y}_{ik} | \boldsymbol{\beta}, \Delta_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \mathbf{z}_{ik}, \lambda_{ik}, X_{ik}, W_{ik}) = \frac{|\Sigma_k|^{-1/2}}{(2\pi)^{J/2}} \lambda_{ik}^{J/2} \exp \left\{ -\frac{\lambda_{ik}}{2} (\mathbf{y}_{ik} - X_{ik} \boldsymbol{\beta} - W_{ik} \boldsymbol{\gamma}_k^R - \Delta_k \mathbf{z}_{ik}^*)' \Sigma_k^{-1} (\mathbf{y}_{ik} - X_{ik} \boldsymbol{\beta} - W_{ik} \boldsymbol{\gamma}_k^R - \Delta_k \mathbf{z}_{ik}^*) \right\}$ with $\mathbf{z}_{ik}^* = \mathbf{z}_{ik} - E[\mathbf{z}_{ik}]$, $\Delta_k = \text{diag}(\delta_{1k}, \dots, \delta_{Jk})$, $\Omega = \text{diag}(\Omega_1, \dots, \Omega_J)$, $\mathbf{z} = (\mathbf{z}'_{11}, \dots, \mathbf{z}'_{n_K K})'$, $\boldsymbol{\gamma}^R = ((\boldsymbol{\gamma}_1^R)', \dots, (\boldsymbol{\gamma}_K^R)')$, and $\boldsymbol{\lambda} = (\lambda_{11}, \dots, \lambda_{n_K K})$. Then using the complete-data likelihood function in (3.16) and prior distributions in Section 3.2, the joint posterior distribution of all the parameters is given by

$$\begin{aligned}
& \pi(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}^*, \boldsymbol{\delta}, \boldsymbol{\sigma}_{\delta}^2, \Sigma^*, v, \Sigma, \boldsymbol{\nu}^*, \nu_a, \nu_b, \tau, \Omega, \boldsymbol{\gamma}^R, \mathbf{z}, \boldsymbol{\psi}, \boldsymbol{\lambda} | D_{\text{obs}}) \\
& \propto L(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}^*, \Sigma^*, \boldsymbol{\nu}^*, \tau, \Omega, \boldsymbol{\gamma}^R, \mathbf{z}, \boldsymbol{\psi}, \boldsymbol{\lambda} | D_{\text{obs}}) \\
& \quad \times \pi(\boldsymbol{\beta}) \pi(\boldsymbol{\gamma}), \pi(\boldsymbol{\delta}^* | \boldsymbol{\delta}, \boldsymbol{\sigma}_{\delta}^2) \pi(\boldsymbol{\delta}) \pi(\boldsymbol{\sigma}_{\delta}^2) \\
& \quad \times \pi(\Sigma^* | v, \Sigma) \pi(v) \pi(\Sigma) \pi(\Omega) \\
& \quad \times \pi(\boldsymbol{\nu}^* | \nu_a, \nu_b) \pi(\nu_a) \pi(\nu_b) \pi(\tau).
\end{aligned}
(3.17)$$

4. BAYESIAN INFERENCE

4.1 Bayesian model comparison via CPO's

Given the rich specification of our proposed model, it is of interest to compare the performance of various special cases of the general multivariate skew meta-regression model proposed in Section 3.2. That is, we need methods for checking whether a skew and/or heavy-tailed distribution is needed for modeling the \mathbf{y}_{ik} 's. Also, we need to investigate whether the skewness, variance, and degrees of freedom are varying across trials. To this end, we carry out the model comparison using the logarithm of pseudo-marginal likelihood (LPML) proposed by Ibrahim et al. (2001). The LPML is a well-established Bayesian model comparison criterion based on the conditional predictive ordinate (CPO) statistic. The CPO statistic for the i th subject in the k th trial is the marginal posterior predictive density of \mathbf{y}_{ik} . As suggested in Ibrahim et al. (2001), a natural summary statistic of the CPO $_{ik}$'s is the LPML defined as

$$(4.1) \quad \text{LPML} = \sum_k^K \sum_i^I \log(\text{CPO}_{ik}).$$

We use LPML as a criterion-based measure for model selection. The larger the LPML, the better the fit of a given model. From (3.16), the marginal likelihood function is given by

$$\begin{aligned}
& f(\mathbf{y}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik}) \\
& = \int_{\lambda_{ik} \in \mathbb{R}_+} \int_{\psi_{ik} \in \mathbb{R}_+} \int_{\mathbf{z}_{ik} \in \mathbb{R}_+^J} f(\mathbf{y}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \mathbf{z}_{ik}, \\
& \quad \lambda_{ik}, X_{ik}, W_{ik})
\end{aligned}$$

$$\begin{aligned}
& \times \prod_{j=1}^J [\psi_{ik} \exp\{-\psi_{ik} z_{ijk}\}] d\mathbf{z}_{ik} \\
& \quad \times \frac{\tau^{\tau+1}}{\Gamma(\tau+1)} \psi_{ik}^{\tau} \exp\{-\tau \psi_{ik}\} d\psi_{ik} \\
& \times \frac{(\frac{\nu_k}{2})^{\frac{\nu_k}{2}}}{\Gamma(\frac{\nu_k}{2})} \lambda_{ik}^{\frac{\nu_k}{2}-1} \exp\{-\frac{\nu_k}{2} \lambda_{ik}\} d\lambda_{ik} \\
& = \frac{|\Sigma_k|^{-\frac{1}{2}} \Gamma(\frac{\nu_k+J}{2})}{\pi^{\frac{J}{2}} \Gamma(\frac{\nu_k}{2})} \nu_k^{\frac{\nu_k}{2}} \frac{\Gamma(\tau+J+1)}{\Gamma(\tau+1)} \tau^{\tau+1} \\
& \times \int_{\mathbf{z}_{ik} \in R_+^J} (A_{ik} + \nu_k)^{-\frac{\nu_k+J}{2}} \left(\sum_{j=1}^J z_{ijk} + \tau \right)^{-(J+\tau+1)} d\mathbf{z}_{ik},
\end{aligned} \tag{4.2}$$

where $A_{ik} = (\mathbf{y}_{ik} - X_{ik} \boldsymbol{\beta} - W_{ik} \boldsymbol{\gamma}_k^R - \Delta_k \mathbf{z}_{ik}^*)' \Sigma_k^{-1} (\mathbf{y}_{ik} - X_{ik} \boldsymbol{\beta} - W_{ik} \boldsymbol{\gamma}_k^R - \Delta_k \mathbf{z}_{ik}^*)$. Let $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}^*, \boldsymbol{\delta}, \boldsymbol{\sigma}_\delta^2, \Sigma^*, \nu, \Sigma, \boldsymbol{\nu}^*, \nu_a, \nu_b, \tau, \Omega, \boldsymbol{\gamma}^R)$ denote the collection of parameters and the random effects. Let $D_{\text{obs}}^{(-ik)}$ denote the observed data with the i th patient in the k th trial deleted. Let Θ and \mathcal{Z} be the parameter spaces corresponding to $\boldsymbol{\theta}$ and \mathbf{z} , respectively. Also let $\pi(\boldsymbol{\theta} | D_{\text{obs}}^{(-ik)})$ denote the posterior of $\boldsymbol{\theta}$ given $D_{\text{obs}}^{(-ik)}$. The following proposition gives a computational form of the CPO statistic.

Proposition 4.1. *Let $h_{ik}(\mathbf{z}_{ik})$ be a normalized weight function satisfying $\int_{R_+^J} h_{ik}(\mathbf{z}_{ik}) d\mathbf{z}_{ik} = 1$. For the i th patient in the k th trial, CPO_{ik} can be written as*

$$\begin{aligned}
\text{CPO}_{ik} &= \int_{\Theta} f(\mathbf{y}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik}) \\
& \quad \times \pi(\boldsymbol{\theta} | D_{\text{obs}}^{(-ik)}) d\boldsymbol{\theta} \\
& = \left[\int_{\Theta} \int_{\mathcal{Z}} f_{\mathbf{y}_{ik}, \mathbf{z}_{ik}} \times \pi(\boldsymbol{\theta}, \mathbf{z} | D_{\text{obs}}) d\mathbf{z} d\boldsymbol{\theta} \right]^{-1},
\end{aligned} \tag{4.3}$$

where $f_{\mathbf{y}_{ik}, \mathbf{z}_{ik}} = \frac{h_{ik}(\mathbf{z}_{ik})}{f(\mathbf{y}_{ik}, \mathbf{z}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik})}$ and $f(\mathbf{y}_{ik}, \mathbf{z}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik}) = (A_{ik} + \nu_k)^{-\frac{\nu_k+J}{2}} (\sum_{j=1}^J z_{ijk} + \tau)^{-(J+\tau+1)} \frac{|\Sigma_k|^{-\frac{1}{2}} \Gamma(\frac{\nu_k+J}{2})}{\pi^{\frac{J}{2}} \Gamma(\frac{\nu_k}{2})} \nu_k^{\frac{\nu_k}{2}} \frac{\Gamma(\tau+J+1)}{\Gamma(\tau+1)} \tau^{\tau+1}$.

The proof of this proposition is given in Appendix A.

REMARK 4.1. Using the CPO Identity I in Zhang et al. (2017), CPO_{ik} can be written as

$$\text{CPO}_{ik} = \frac{1}{\int_{\Theta} \frac{1}{f(\mathbf{y}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik})} \pi(\boldsymbol{\theta} | D_{\text{obs}}) d\boldsymbol{\theta}},$$

where $f(\mathbf{y}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik})$ is given in (4.2). This identity requires a J -dimensional integration over R_+^J . Compared to the CPO identity I, (4.3) is more computationally attractive.

REMARK 4.2. The CPO_{ik} given in (4.3) uses the CPO Identity II in Zhang et al. (2017). Now, let $\{(\boldsymbol{\theta}^{(b)}, \mathbf{z}^{(b)}), b = 1, \dots, B\}$ denote a Gibbs sample of $(\boldsymbol{\theta}, \mathbf{z} | D_{\text{obs}})$. Using Proposition 4.1, a Monte Carlo estimator of CPO_{ik} in (4.3) is given by

$$\widehat{\text{CPO}}_{ik} = B \left[\sum_{b=1}^B f_{\mathbf{y}_{ik}, \mathbf{z}_{ik}}^{(b)} \right]^{-1}, \tag{4.4}$$

where $f_{\mathbf{y}_{ik}, \mathbf{z}_{ik}}^{(b)} = \frac{h_{ik}(\mathbf{z}_{ik}^{(b)})}{f(\mathbf{y}_{ik}, \mathbf{z}_{ik}^{(b)} | \boldsymbol{\beta}^{(b)}, \boldsymbol{\delta}_k^{(b)}, \Sigma_k^{(b)}, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik})}$.

The Monte Carlo error of $\widehat{\text{CPO}}_{ik}$ given in (4.4) depends on the choice of the weight function $h_{ik}(\mathbf{z}_{ik})$. Following Theorem 1 in Zhang et al. (2017), the optimal weight function is defined as the weight function minimizing the variance of the Monte Carlo estimator. Here, we have

$$h_{ik, \text{opt}}(\mathbf{z}_{ik}) = \frac{f(\mathbf{y}_{ik}, \mathbf{z}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik})}{f(\mathbf{y}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik})}. \tag{4.5}$$

However, the optimal weight function is not analytically available and difficult to compute since the denominator involves a J -dimensional integration. We can, instead, use a multivariate normal density as a possible choice of $h_{ik}(\mathbf{z}_{ik})$, which is constructed via the Laplace approximation to the joint density $f(\mathbf{y}_{ik}, \mathbf{z}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik})$. Let $\mathbf{u}_{ik} = \log(\mathbf{z}_{ik})$ for $i = 1, \dots, n_k, k = 1, \dots, K$, then $\mathbf{z}_{ik} = \exp(\mathbf{u}_{ik})$. Here, log and exp functions applied to a vector means that the operations are applied to every element of the vector. Denote $g(\mathbf{u}_{ik}) = \log \{f(\mathbf{y}_{ik}, \exp(\mathbf{u}_{ik}) | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik}) \prod_{j=1}^J \exp(u_{ijk})\}$. Let \mathbf{u}_{ik}^0 be the stationary point, where $\nabla g(\mathbf{u}_{ik}) = 0$, and $B = -\nabla^2 g(\mathbf{u}_{ik})|_{\mathbf{u}=\mathbf{u}_{ik}^0}$. Then, the optimal weight function $h_{ik, \text{opt}}(\mathbf{z}_{ik})$ is approximately given by $\frac{f(\log(\mathbf{z}_{ik}) | \mathbf{u}_{ik}^0, B^{-1})}{\prod_{j=1}^J z_{ijk}^0}$, where $f(\cdot | \mathbf{u}_{ik}^0, B^{-1})$ denotes the density of a $N(\mathbf{u}_{ik}^0, B^{-1})$ distribution. A detailed derivation of this approximated optimal weight function is given in Appendix B.

4.2 Computational development

We consider the following one-to-one transformations: $\boldsymbol{\gamma}_k^{*R} = (\boldsymbol{\gamma}_{1k}^{*'}, \boldsymbol{\gamma}_{2k}^{*'}, \dots, \boldsymbol{\gamma}_{Jk}^{*'})' = \boldsymbol{\gamma}_k^R - \boldsymbol{\gamma}$ and $\boldsymbol{\delta}_k^* = (\boldsymbol{\delta}_{1k}^{*'}, \boldsymbol{\delta}_{2k}^{*'}, \dots, \boldsymbol{\delta}_{Jk}^{*'})' = \boldsymbol{\delta}_k - \boldsymbol{\delta}$ for $k = 1, \dots, K$. Thus, $\boldsymbol{\gamma}_{jk}^* = \boldsymbol{\gamma}_{jk} - \boldsymbol{\gamma}_j$ and $\boldsymbol{\delta}_{jk}^* = \boldsymbol{\delta}_{jk} - \boldsymbol{\delta}_j$ for $j = 1, \dots, J$ and $k = 1, \dots, K$. Write $\boldsymbol{\gamma}^{*R} = ((\boldsymbol{\gamma}_1^{*R})', \dots, (\boldsymbol{\gamma}_K^{*R})')'$ and $\boldsymbol{\delta}^{*} = (\boldsymbol{\delta}_1^{*'}, \dots, \boldsymbol{\delta}_K^{*'})'$. Let $\Delta_k^* = \text{diag}(\boldsymbol{\delta}_{1k}^*, \dots, \boldsymbol{\delta}_{Jk}^*)$ and $\Delta = \text{diag}(\delta_1, \dots, \delta_J)$. Also, let $X_{ik}^* = (X_{ik}, W_{ik})$ for $i = 1, \dots, n$ and $k = 1, \dots, K$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')'$. We present a detailed development of the MCMC sampling algorithm. Although the analytic evaluation of the joint posterior distribution of $(\boldsymbol{\theta}, \boldsymbol{\delta}^{*}, \boldsymbol{\delta}, \boldsymbol{\sigma}_\delta^2, \Sigma^*, \nu, \Sigma, \boldsymbol{\nu}^*, \nu_a, \nu_b, \tau, \Omega, \boldsymbol{\gamma}^{*R}, \mathbf{z}, \boldsymbol{\psi}, \boldsymbol{\lambda})$ based on the observed data D_{obs} given in Equation (3.17) is not possible, the proposed model allows us to develop an efficient MCMC sampling algorithm to sample from (3.17).

The MCMC sampling algorithm requires sampling from the following full conditional distributions in turn: (i) $[\gamma^{*R}, \theta, \delta^{**}, \delta, \sigma_\delta^2 \mid \Sigma^*, v, \Sigma, \nu^*, \nu_a, \nu_b, \tau, \Omega, z, \psi, \lambda, D_{\text{obs}}]$; (ii) $[\lambda, \nu^*, \nu_a, \nu_b \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \tau, \Omega, \gamma^{*R}, z, \psi, D_{\text{obs}}]$; (iii) $[\Sigma^*, v, \Sigma \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \nu^*, \nu_a, \nu_b, \tau, \Omega, \gamma^{*R}, z, \psi, \lambda, D_{\text{obs}}]$; (iv) $[z \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \nu^*, \nu_a, \nu_b, \tau, \Omega, \gamma^{*R}, \psi, \lambda, D_{\text{obs}}]$; (v) $[\psi, \tau \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \nu^*, \nu_a, \nu_b, \Omega, \gamma^{*R}, z, \lambda, D_{\text{obs}}]$; and (vi) $[\Omega \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \nu^*, \nu_a, \nu_b, \tau, \gamma^{*R}, z, \psi, \lambda, D_{\text{obs}}]$. For (i), we apply the collapsed Gibbs technique of Liu (1994) and Chen et al. (2000) through the identity

$$\begin{aligned} & [\gamma^{*R}, \theta, \delta^{**}, \delta, \sigma_\delta^2 \mid \Sigma^*, v, \Sigma, \nu^*, \nu_a, \nu_b, \tau, \Omega, z, \psi, \lambda, D_{\text{obs}}] \\ &= [\gamma^{*R} \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \nu^*, \nu_a, \nu_b, \tau, \Omega, z, \psi, \lambda, D_{\text{obs}}] \\ &\times [\delta^{**} \mid \theta, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \nu^*, \nu_a, \nu_b, \tau, \Omega, z, \psi, \lambda, D_{\text{obs}}] \\ (4.6) \\ &\times [\theta, \delta, \sigma_\delta^2 \mid \Sigma^*, v, \Sigma, \nu^*, \nu_a, \nu_b, \tau, \Omega, z, \psi, \lambda, D_{\text{obs}}]. \end{aligned}$$

That is, we sample δ^{**} after collapsing out γ^{*R} , and also sample θ , δ , and σ_δ^2 after collapsing out γ^{*R} and δ^{**} . For (ii), we apply the collapsed Gibbs technique of Liu (1994) and Chen et al. (2000) through the identity

$$\begin{aligned} & [\lambda, \nu^*, \nu_a, \nu_b \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \tau, \Omega, \gamma^{*R}, z, \psi, D_{\text{obs}}] \\ &= [\lambda \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \nu^*, \nu_a, \nu_b, \tau, \Omega, \gamma^{*R}, z, \psi, D_{\text{obs}}] \\ &\times [\nu_b \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \nu^*, \nu_a, \tau, \Omega, \gamma^{*R}, z, \psi, D_{\text{obs}}] \\ (4.7) \\ &\times [\nu^*, \nu_a \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \tau, \Omega, \gamma^{*R}, z, \psi, D_{\text{obs}}]. \end{aligned}$$

That is, we sample ν_b after collapsing out λ , and also sample ν^* and ν_a after collapsing out λ and ν_b . For (iii), we apply the collapsed Gibbs technique of Liu (1994) and Chen et al. (2000) through the identity

$$\begin{aligned} & [\Sigma^*, v, \Sigma \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \nu^*, \nu_a, \nu_b, \tau, \Omega, \gamma^{*R}, z, \psi, \lambda, D_{\text{obs}}] \\ &= [\Sigma^* \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, v, \Sigma, \nu^*, \nu_a, \nu_b, \tau, \Omega, \gamma^{*R}, z, \psi, \lambda, D_{\text{obs}}] \\ (4.8) \\ &\times [v, \Sigma \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \nu^*, \nu_a, \nu_b, \tau, \Omega, \gamma^{*R}, z, \psi, \lambda, D_{\text{obs}}]. \end{aligned}$$

That is, we sample v and Σ after collapsing out Σ^* . For (v), we apply the collapsed Gibbs technique of Liu (1994) and Chen et al. (2000) through the identity

$$[\psi, \tau \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \nu^*, \nu_a, \nu_b, \Omega, \gamma^{*R}, z, \lambda, D_{\text{obs}}]$$

$$\begin{aligned} &= [\psi \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \nu^*, \nu_a, \nu_b, \tau, \Omega, \gamma^{*R}, z, \lambda, D_{\text{obs}}] \\ (4.9) \\ &\times [\tau \mid \theta, \delta^{**}, \delta, \sigma_\delta^2, \Sigma^*, v, \Sigma, \nu^*, \nu_a, \nu_b, \Omega, \gamma^{*R}, \psi, z, \lambda, D_{\text{obs}}]. \end{aligned}$$

All the full conditional distributions discussed above are presented in Section S1 of the Supplementary Materials (http://intjpress.com/site/pub/files/_supp/sii/2020/0013/0004/SII-2020-0013-0004-s004.pdf).

5. ANALYSIS OF THE CHOLESTEROL DATA

We re-analyze the cholesterol data discussed in Section 2. In (3.3), \mathbf{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 as in Ibrahim et al. (2019). We model these three outcome variables LDL-C, HDL-C, and TG jointly via (3.3), (3.4), (3.5), and (3.6) with $J = 3$ and $K = 26$ in conjunction with the models specified at Stage 2 and the priors specified at Stage 3. We standardized all the fourteen covariates for numerical stability in the posterior computations.

As shown in Ibrahim et al. (2019), the multivariate skew meta-regression model with an unstructured covariance matrix for the multivariate outcome variables outperformed the symmetric normal and t models as well as the skew models with a diagonal covariance matrix. Thus, for the cholesterol data, we only fit the flexible multivariate meta-regression models defined by (3.3), (3.4), (3.5), and (3.6) with different fixed values of τ in (3.6) and random τ with prior specified in (3.13). We also consider different fixed values of ν_a in (3.10) and random ν_a with a prior specified in (3.14). In total, we consider 14 different models, including the one, which was the best model considered in Ibrahim et al. (2019), and the values of LPML are reported in Table 1. We see, from Table 1, that the LPML values under the proposed flexible skew t models are greater than the one (LPML = -262517.61) under the skew t model with $\tau = 30$ (Ibrahim et al. 2019). The best flexible skew t model is the one with random ν_a and τ , which has LPML = -261321.15 while the second best model is the one with $\nu_a = 25$ and $\tau = 5$. The values of LPML for these models are -261321.15 and -261322.38, which are very close.

Table 1. Values of LPML Measure under Various Models for the Cholesterol Data

	Flexible skew t model				
	$\tau = 5$	$\tau = 10$	$\tau = 20$	$\tau = 30$	random ν_a & τ
$\nu_a = 20$	-261324.67	-261391.54	-261423.48	-261455.18	-261321.15
$\nu_a = 25$	-261322.38	-261346.60	-261413.49	-261448.32	
$\nu_a = 30$	-261323.05	-261358.94	-261415.45	-261450.02	
skew t model with $\tau = 30$					
-262517.61					

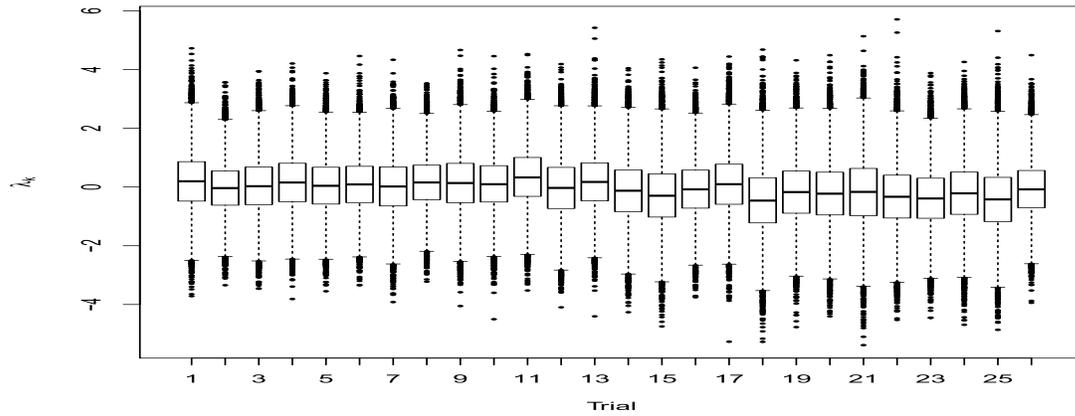


Figure 2. Boxplots of the λ_k^* 's from the posterior distribution under flexible skew model with random ϕ and τ for the cholesterol data.

Table 2. Posterior Estimates under the flexible skew t model with random ν_a and τ for LDL-C

	Par.	Mean	SD	95% HPD
bl_ldlc	$\beta_{1,1}$	-0.031	0.003	(-0.037, -0.025)
bl_hdlc	$\beta_{1,2}$	0.013	0.009	(-0.005, 0.032)
bl_tg	$\beta_{1,3}$	0.007	0.001	(0.004, 0.010)
BMI	$\beta_{1,4}$	0.069	0.017	(0.035, 0.102)
age	$\beta_{1,5}$	-0.144	0.009	(-0.162, -0.126)
duration	$\beta_{1,6}$	0.338	0.282	(-0.217, 0.895)
Female	$\beta_{1,7}$	-1.239	0.206	(-1.621, -0.812)
DM	$\beta_{1,8}$	-2.315	0.278	(-2.853, -1.769)
CHD	$\beta_{1,9}$	0.258	0.270	(-0.301, 0.763)
potency2	$\beta_{1,10}$	-7.319	0.327	(-7.954, -6.669)
potency3	$\beta_{1,11}$	-15.738	0.375	(-16.438, -14.968)
Black	$\beta_{1,12}$	2.220	0.394	(1.433, 2.975)
Hispanic	$\beta_{1,13}$	0.397	0.456	(-0.519, 1.269)
Other	$\beta_{1,14}$	-2.077	0.490	(-3.046, -1.146)
onstatin \times bl_ldlc	$\beta_{1,15}$	-0.058	0.007	(-0.072, -0.045)
onstatin \times bl_hdlc	$\beta_{1,16}$	-0.074	0.017	(-0.109, -0.041)
onstatin \times bl_tg	$\beta_{1,17}$	-0.005	0.003	(-0.011, 0.001)
onstatin \times BMI	$\beta_{1,18}$	-0.068	0.034	(-0.135, -0.004)
onstatin \times age	$\beta_{1,19}$	0.108	0.018	(0.074, 0.145)
onstatin \times duration	$\beta_{1,20}$	0.155	0.633	(-1.087, 1.398)
onstatin \times Female	$\beta_{1,21}$	2.722	0.396	(1.956, 3.497)
onstatin \times DM	$\beta_{1,22}$	-0.654	0.478	(-1.598, 0.281)
onstatin \times CHD	$\beta_{1,23}$	-0.862	0.504	(-1.855, 0.106)
onstatin \times potency2	$\beta_{1,24}$	6.144	0.655	(4.873, 7.431)
onstatin \times potency3	$\beta_{1,25}$	13.924	0.835	(12.336, 15.597)
onstatin \times Black	$\beta_{1,26}$	-2.630	0.851	(-4.272, -0.953)
onstatin \times Hispanic	$\beta_{1,27}$	-2.094	1.020	(-4.151, -0.139)
onstatin \times Other	$\beta_{1,28}$	3.108	0.962	(1.213, 4.981)

We extend λ_k^* in (3.2) to

$$(5.1) \quad \lambda_k^* = (\lambda_k - n_k)/(n_k/(\nu_k/2))^{1/2}$$

to account for different degrees of freedom for the k th trial under the proposed flexible skew meta-regression model. Figure 2 shows the boxplots of the λ_k^* 's defined in (5.1) under the flexible skew t model random ν_a and τ . The boxplots

of the λ_k^* 's under the flexible skew t model with $\nu_a = 25$ and $\tau = 5$ are shown in Figure S.1 of the Supplementary Materials. We see from Figure 2 and Figure S.1 that all of these boxplots had a median close to zero and no obvious outlying trials were found from these two figures. These two figures are quite different than Figure 2 of Ibrahim et al. (2019), in which boxplots corresponding to trials 8 and 25 were quite different than the rest of the 24 boxplots and these boxplots

Table 3. Posterior Estimates under the flexible skew t model with random ν_a and τ for HDL-C

	Par.	Mean	SD	95% HPD
bl_ldlc	$\beta_{2,1}$	0.002	0.003	(-0.004, 0.008)
bl_hdlc	$\beta_{2,2}$	-0.197	0.010	(-0.217, -0.178)
bl_tg	$\beta_{2,3}$	0.024	0.001	(0.021, 0.027)
BMI	$\beta_{2,4}$	-0.145	0.017	(-0.179, -0.112)
age	$\beta_{2,5}$	0.075	0.009	(0.057, 0.094)
duration	$\beta_{2,6}$	0.065	0.063	(-0.057, 0.190)
Female	$\beta_{2,7}$	0.523	0.212	(0.114, 0.941)
DM	$\beta_{2,8}$	-2.270	0.278	(-2.819, -1.728)
CHD	$\beta_{2,9}$	-0.985	0.273	(-1.536, -0.465)
potency2	$\beta_{2,10}$	0.790	0.317	(0.176, 1.424)
potency3	$\beta_{2,11}$	0.256	0.367	(-0.445, 0.989)
Black	$\beta_{2,12}$	-2.448	0.371	(-3.187, -1.730)
Hispanic	$\beta_{2,13}$	-1.097	0.449	(-1.972, -0.211)
Other	$\beta_{2,14}$	-0.050	0.504	(-1.051, 0.920)
onstatin \times bl_ldlc	$\beta_{2,15}$	-0.007	0.005	(-0.017, 0.003)
onstatin \times bl_hdlc	$\beta_{2,16}$	-0.006	0.014	(-0.035, 0.021)
onstatin \times bl_tg	$\beta_{2,17}$	-0.016	0.002	(-0.020, -0.011)
onstatin \times BMI	$\beta_{2,18}$	0.027	0.027	(-0.026, 0.081)
onstatin \times age	$\beta_{2,19}$	-0.059	0.014	(-0.088, -0.031)
onstatin \times duration	$\beta_{2,20}$	-0.245	0.194	(-0.638, 0.128)
onstatin \times Female	$\beta_{2,21}$	0.786	0.324	(0.161, 1.427)
onstatin \times DM	$\beta_{2,22}$	1.719	0.390	(0.914, 2.462)
onstatin \times CHD	$\beta_{2,23}$	0.464	0.395	(-0.299, 1.248)
onstatin \times potency2	$\beta_{2,24}$	-1.052	0.526	(-2.094, -0.029)
onstatin \times potency3	$\beta_{2,25}$	-0.946	0.643	(-2.176, 0.338)
onstatin \times Black	$\beta_{2,26}$	2.571	0.634	(1.328, 3.819)
onstatin \times Hispanic	$\beta_{2,27}$	-0.436	0.769	(-1.967, 1.038)
onstatin \times Other	$\beta_{2,28}$	-0.895	0.742	(-2.300, 0.579)

were much more heterogenous than those in Figures 2 and Figure S.1 under the proposed models. Thus, the outlying trials identified in Ibrahim et al. (2019) could be due to a lack of fit.

The posterior estimates, including posterior means, posterior standard deviations (SDs), and 95% highest posterior density (HPD) intervals of the parameters under the flexible skew t model with random ν_a and τ are reported in Tables 2–6 and Tables S.1–S.3. Those posterior estimates under the flexible skew t model with $\nu_a = 25$ and $\tau = 5$ are reported in Tables S.4–S.11 of the Supplementary Materials. The posterior means and the 95% HPD intervals of the 28 regression coefficients in Tables 2–4 (or Tables S.4–S.6) under the proposed flexible skew models and the skew t model with $\tau = 30$ (Ibrahim et al., 2019) are also plotted in Figure 3 and Figure S.2 of the Supplementary Materials. We call a posterior estimate “statistically significant at a significance level of 0.05” if the corresponding 95% HPD interval does not contain 0. Under this notion, we see from Figures 2 and S.2 that significant posterior estimates are consistent between the proposed model and the model of Ibrahim et al. (2019) for most of the regression coefficients except for a few coefficients. For example, for LDL-C, onstatin \times DM ($\beta_{1,22}$) was not significant with 95% HPD intervals (-1.598, 0.281) and (-1.553, 0.322) under the proposed models with random

ν_a and τ (Table 2) and with $\nu_a = 25$ and $\tau = 5$ (Table S.4), respectively, while it was significant with 95% HPD interval (-2.088, -0.225) under the skew t model with $\tau = 30$ (Table 7 of Ibrahim et al. (2019)). For HDL-C, onstatin \times potency2 ($\beta_{2,24}$) was nearly significant with HPD intervals (-2.094, -0.029) and (-2.053, 0.000) under the proposed models with random ν_a and τ (Table 3) and with $\nu_a = 25$ and $\tau = 5$ (Table S.5), respectively, while it was not significant with 95% HPD interval (-1.955, 0.068) under the skew t model with $\tau = 30$ (Table 8 of Ibrahim et al. (2019)). For TG, onstatin \times duration ($\beta_{3,20}$) was not significant with HPD intervals (-0.242, 1.380) and (-0.215, 1.390) under the proposed models with random ν_a and τ (Table 4) and with $\nu_a = 25$ and $\tau = 5$ (Table S.6), respectively, while it was significant with 95% HPD interval (0.227, 1.539) under the skew t model with $\tau = 30$ (Table 9 of Ibrahim et al. (2019)). Therefore, different models may identify different sets of significant covariates. Since the proposed flexible skew model fits the cholesterol data much better than the skew t model with $\tau = 30$, the latter model may potentially incorrectly identify the association between the outcome variables and covariates, yielding a misleading conclusion in terms of the clinical importance of covariates.

The results shown in Table 5 under the flexible skew model with random ν_a and τ , Table S.7 under the flexi-

Table 4. Posterior Estimates under the flexible skew t model with random ν_a and τ for TG

	Par.	Mean	SD	95% HPD
bl_ldlc	$\beta_{3,1}$	-0.002	0.006	(-0.013, 0.009)
bl_hdlc	$\beta_{3,2}$	0.030	0.017	(-0.002, 0.063)
bl_tg	$\beta_{3,3}$	-0.093	0.003	(-0.098, -0.088)
BMI	$\beta_{3,4}$	0.237	0.029	(0.181, 0.295)
age	$\beta_{3,5}$	-0.032	0.016	(-0.064, 0.000)
duration	$\beta_{3,6}$	-0.186	0.182	(-0.566, 0.140)
Female	$\beta_{3,7}$	2.570	0.361	(1.879, 3.279)
DM	$\beta_{3,8}$	-0.549	0.480	(-1.465, 0.417)
CHD	$\beta_{3,9}$	0.680	0.471	(-0.243, 1.591)
potency2	$\beta_{3,10}$	-4.303	0.572	(-5.422, -3.179)
potency3	$\beta_{3,11}$	-9.458	0.666	(-10.775, -8.178)
Black	$\beta_{3,12}$	-1.997	0.665	(-3.334, -0.744)
Hispanic	$\beta_{3,13}$	1.045	0.805	(-0.512, 2.615)
Other	$\beta_{3,14}$	-1.012	0.838	(-2.697, 0.573)
onstatin \times bl_ldlc	$\beta_{3,15}$	0.015	0.010	(-0.004, 0.036)
onstatin \times bl_hdlc	$\beta_{3,16}$	-0.093	0.028	(-0.150, -0.039)
onstatin \times bl_tg	$\beta_{3,17}$	-0.026	0.005	(-0.036, -0.017)
onstatin \times BMI	$\beta_{3,18}$	0.038	0.054	(-0.064, 0.146)
onstatin \times age	$\beta_{3,19}$	0.096	0.029	(0.040, 0.153)
onstatin \times duration	$\beta_{3,20}$	0.605	0.413	(-0.242, 1.380)
onstatin \times Female	$\beta_{3,21}$	0.262	0.639	(-0.975, 1.513)
onstatin \times DM	$\beta_{3,22}$	2.221	0.766	(0.715, 3.684)
onstatin \times CHD	$\beta_{3,23}$	0.922	0.804	(-0.690, 2.482)
onstatin \times potency2	$\beta_{3,24}$	3.468	1.063	(1.384, 5.549)
onstatin \times potency3	$\beta_{3,25}$	6.237	1.302	(3.623, 8.732)
onstatin \times Black	$\beta_{3,26}$	-1.101	1.316	(-3.662, 1.482)
onstatin \times Hispanic	$\beta_{3,27}$	0.896	1.603	(-2.240, 4.007)
onstatin \times Other	$\beta_{3,28}$	3.854	1.458	(1.039, 6.738)

Table 5. Posterior Estimates under the flexible skew t model with random ν_a and τ

Par.	Mean	SD	95% HPD	Par.	Mean	SD	95% HPD
$\gamma_{1,0}$	-22.112	3.034	(-28.094, -16.127)	$\Omega_{1,0,0}$	20.133	10.345	(6.526, 39.943)
$\gamma_{1,1}$	-14.129	0.827	(-15.748, -12.455)	$\Omega_{1,0,1}$	-9.158	5.367	(-19.859, -1.405)
$\gamma_{1,2}$	6.229	4.290	(-2.232, 14.651)	$\Omega_{1,1,1}$	8.197	3.900	(2.791, 15.742)
$\gamma_{1,3}$	-19.107	1.410	(-21.849, -16.293)	$\Omega_{1,2,2}$	43.791	21.636	(13.970, 84.722)
$\gamma_{2,0}$	9.918	1.354	(7.291, 12.594)	$\Omega_{1,2,3}$	-26.458	14.770	(-54.154, -5.612)
$\gamma_{2,1}$	2.111	0.430	(1.265, 2.956)	$\Omega_{1,3,3}$	24.157	12.438	(7.482, 47.977)
$\gamma_{2,2}$	13.850	1.818	(10.345, 17.446)	$\Omega_{2,0,0}$	1.957	1.085	(0.530, 4.033)
$\gamma_{2,3}$	1.256	0.260	(0.763, 1.773)	$\Omega_{2,0,1}$	-1.734	0.990	(-3.616, -0.395)
$\gamma_{3,0}$	-1.626	2.679	(-6.795, 3.708)	$\Omega_{2,1,1}$	1.680	0.992	(0.344, 3.587)
$\gamma_{3,1}$	-5.918	0.704	(-7.304, -4.523)	$\Omega_{2,2,2}$	1.182	0.817	(0.121, 2.657)
$\gamma_{3,2}$	4.180	3.665	(-2.717, 11.540)	$\Omega_{2,2,3}$	-0.188	0.355	(-0.930, 0.366)
$\gamma_{3,3}$	-7.923	0.738	(-9.373, -6.448)	$\Omega_{2,3,3}$	0.137	0.161	(0.007, 0.419)
Σ_{11}	72.921	4.189	(64.870, 81.231)	$\Omega_{3,0,0}$	9.561	6.304	(1.245, 21.757)
Σ_{12}	13.629	3.067	(7.729, 19.587)	$\Omega_{3,0,1}$	-6.229	4.040	(-14.125, -0.545)
Σ_{13}	22.812	4.459	(14.018, 31.508)	$\Omega_{3,1,1}$	4.463	3.107	(0.218, 10.459)
Σ_{22}	96.721	5.193	(86.786, 107.018)	$\Omega_{3,2,2}$	9.301	5.919	(1.423, 20.324)
Σ_{23}	-37.165	5.046	(-47.266, -27.514)	$\Omega_{3,2,3}$	-5.427	4.046	(-13.380, -0.181)
Σ_{33}	182.385	10.700	(161.945, 203.932)	$\Omega_{3,3,3}$	3.574	3.221	(0.017, 9.468)
δ_1	9.651	0.612	(8.438, 10.865)	$\sigma_{\delta_1}^2$	8.242	3.181	(3.345, 14.545)
δ_2	-1.522	0.189	(-1.907, -1.166)	$\sigma_{\delta_2}^2$	0.154	0.133	(0.016, 0.403)
δ_3	19.112	0.820	(17.518, 20.752)	$\sigma_{\delta_3}^2$	13.696	5.151	(5.273, 23.948)
ϕ	19.672	8.486	(6.197, 36.399)	ν_0	33.921	4.472	(25.751, 43.207)
v	7.863	0.523	(6.886, 8.926)	τ	4.350	0.290	(3.787, 4.914)

Table 6. Posterior Estimates under the flexible skew t model with random ν_a and τ

Par.	Mean	SD	95% HPD	Par.	Mean	SD	95% HPD
v_1	7.789	1.189	(5.593, 10.179)	v_{14}	7.208	0.961	(5.508, 9.191)
v_2	8.694	1.554	(5.998, 11.880)	v_{15}	5.408	0.619	(4.258, 6.645)
v_3	7.511	1.514	(4.649, 10.531)	v_{16}	8.761	1.418	(6.285, 11.631)
v_4	8.520	1.328	(6.243, 11.313)	v_{17}	8.418	1.772	(5.182, 11.890)
v_5	9.176	1.536	(6.449, 12.285)	v_{18}	9.272	1.730	(6.292, 12.851)
v_6	8.560	1.419	(5.989, 11.384)	v_{19}	10.479	2.113	(6.847, 14.652)
v_7	9.728	1.521	(6.958, 12.749)	v_{20}	8.379	1.506	(5.662, 11.390)
v_8	7.000	0.581	(5.900, 8.141)	v_{21}	6.626	1.067	(4.680, 8.731)
v_9	8.150	0.906	(6.412, 9.894)	v_{22}	6.671	1.198	(4.527, 9.110)
v_{10}	10.103	1.113	(8.024, 12.286)	v_{23}	6.417	1.093	(4.403, 8.575)
v_{11}	7.910	0.978	(6.087, 9.844)	v_{24}	7.593	1.362	(5.128, 10.317)
v_{12}	5.809	1.173	(3.609, 8.091)	v_{25}	6.825	1.003	(4.937, 8.806)
v_{13}	6.838	1.114	(4.860, 9.160)	v_{26}	6.235	0.923	(4.537, 8.080)

ble skew model with $\nu_a = 25$ and $\tau = 5$, and Table 6 of Ibrahim et al. (2019) 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. For first-line therapy, the 95% HPD intervals were (-15.771, -12.492), (-15.748, -12.455), and (-15.662, -12.454) for the percent change from baseline in LDL-C ($\gamma_{1,1}$); (1.265, 2.956), (1.302, 2.952) and (1.285, 2.870) for the percent change from baseline in HDL-C ($\gamma_{2,1}$); and (-7.304, -4.523), (-7.357, -4.548), and (-7.316, -4.369) for the percent change from baseline in TG ($\gamma_{3,1}$), respectively, under the flexible skew model with random ν_a and τ , the flexible skew model with $\nu_a = 25$ and $\tau = 5$, and the skew model with $\tau = 30$. For second-line therapy, these 95% HPD intervals were (-21.849, -16.293), (-21.840, -16.260), and (-21.586, -15.974) for the percent change from baseline in LDL-C ($\gamma_{1,3}$); (0.763, 1.773), (0.740, 1.751), and (0.726, 1.736) for the percent change from baseline in HDL-C ($\gamma_{2,3}$); and (-9.373, -6.448), (-9.306, -6.452), and (-9.213, -5.868) for the percent change from baseline in TG ($\gamma_{3,3}$), respectively, under the above three models.

Tables 6 and S.8 as well as Figures 4 and S.3 show that the posterior estimates of the degrees of freedom, v_k 's, vary across trials, with the posterior estimates from 5.408 to 10.479 (Table 6) and 5.452 to 9.778 (Table S.8). Tables S.1 and S.9 and Figures 5 and S.4 indicate that the skewness parameters, $\delta_{k,j}$'s, are very heterogeneous for outcome variables LDL-C and TG and relatively homogenous for HDL-C among the 26 trials. Finally, we see from Tables S.2, S.3, S.10, and S.11 that the magnitudes of the covariances and the variances are different across these 26 trials, however, interestingly, the signs of the correlations among the three outcome variables (LDL-C, HDL-C, and TG) are consistent across trials. These posterior estimates suggest that the skewness parameters, the covariance matrices of the multivariate responses, and the degrees of freedom in the multivariate t distributions for the error terms should be different across trials, which further empirically confirms the

finding according to the LPML criterion that the flexible skew model did fit the cholesterol data better than the skew t model with $\tau = 30$.

To compute posterior estimates, including posterior means, posterior SDs, 95% HPD intervals, LPMLs, and boxplots of λ_k^* 's, we used 20,000 MCMC samples, which were taken from every fifth iteration, after a burn-in of 20,000 iterations. The convergence of the MCMC sampling algorithm was checked using several diagnostic procedures discussed in Chen et al. (2000). The HPD intervals were computed via the Monte Carlo method developed by Chen and Shao (1999). Computer code was written for the FORTRAN 95 compiler using IMSL subroutines with double precision accuracy. The FORTRAN code is available from the authors upon request.

6. DISCUSSION

In this paper, we have proposed a general and flexible multivariate skew and heavy-tailed meta-regression model for modeling individual level patient meta-data, which is a novel extension of the multivariate skew meta-regression model of Ibrahim et al. (2019). Due to the complexity of the proposed model, we have also developed an efficient MCMC sampling algorithm using the collapsed Gibbs technique of Liu (1994) and Chen et al. (2000) 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 logarithm of the pseudo marginal likelihood for model comparison. As was seen from the analysis of the (LDL-C, HDL-C, TG) data, the proposed model has substantially improved goodness-of-fit over the one developed in Ibrahim et al. (2019).

An extension to network meta-regression (NMR) of the proposed model for individual level patient network meta-data can be developed. Under the network regression setting, multiple treatments are compared using both direct comparisons of interventions within randomized controlled trials and indirect comparisons across trials based on a com-

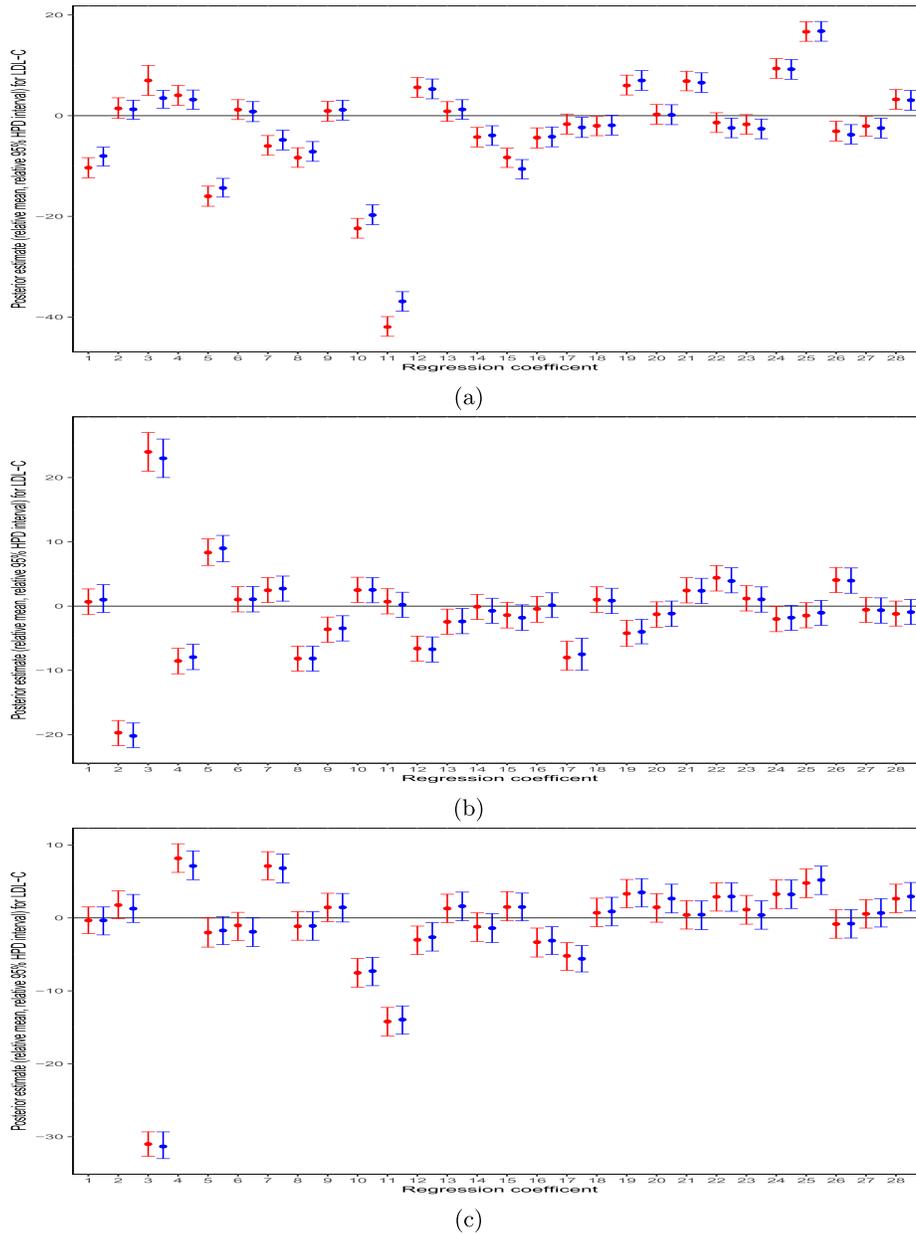


Figure 3. Plots of the relative posterior estimates (mean/SD and 95% HPD interval/SD) for β under the skew t model with $\tau = 30$ (blue) and flexible skew model with random ν_α and τ (red) for the cholesterol data.

mon comparator, accounting for covariates. Such an extension is potentially useful and significant in comparing and assessing the effects of cholesterol lowering drugs. A detailed development of this extension is beyond the scope of the current paper.

Supplementary materials available on the journal website consist of the full conditional distributions, additional tables (Tables S.1–S.3) of the posterior estimates under the flexible skewed model with random ν_α and τ for the cholesterol data, and the posterior estimates (Table S.4–S.11; Figure S.1–S.4) under the flexible skewed model with $\nu_\alpha = 25$ and $\tau = 5$.

APPENDIX

Appendix A. Proof of Proposition 4.1

The CPO_{ik} statistic for the i th patient in the k th trial is defined as

$$\begin{aligned} CPO_{ik} &= \int_{\Theta} f(\mathbf{y}_{ik} | X_{ik}, W_{ik}, \boldsymbol{\theta}) \pi(\boldsymbol{\theta} | D_{\text{obs}}^{(-ik)}) d\boldsymbol{\theta} \\ &= \int_{\Theta} f(\mathbf{y}_{ik} | X_{ik}, W_{ik}, \boldsymbol{\theta}) \end{aligned}$$

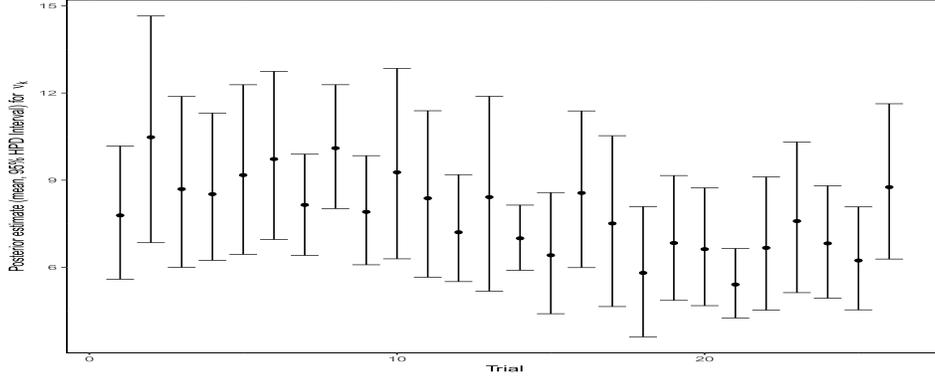


Figure 4. Plots of the posterior estimates (mean and 95% HPD interval) for ν_k under the flexible skew model with random ν_α and τ for the cholesterol data.

$$\begin{aligned}
 & \times \frac{f(\mathbf{y}_{-ik} | X_{-ik}, W_{-ik}, \boldsymbol{\theta}) \pi(\boldsymbol{\theta})}{\int_{\Theta} f(\mathbf{y}_{-ik} | X_{-ik}, W_{-ik}, \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} d\boldsymbol{\theta} \\
 \text{(A.1)} \quad & = \int_{\Theta} \int_{\mathcal{Z}} \frac{\left[\prod_{k=1}^K \prod_{i=1}^{n_k} f_{\mathbf{y}_{ik}, \mathbf{z}_{ik}}^A \right] \pi(\boldsymbol{\theta})}{\int_{\Theta} f(\mathbf{y}_{-ik} | X_{-ik}, W_{-ik}, \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} dz d\boldsymbol{\theta}, \\
 & = \frac{f(\mathbf{y}_{ik}, \mathbf{z}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik})}{f(\mathbf{y}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik})} \\
 & = \frac{f(\mathbf{y}_{ik}, \mathbf{z}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik})}{f^M(\mathbf{y}_{ik}, \mathbf{z}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik})},
 \end{aligned}$$

where $f_{\mathbf{y}_{ik}, \mathbf{z}_{ik}}^A = f(\mathbf{y}_{ik} | X_{ik}, W_{ik}, \boldsymbol{\theta}, \mathbf{z}_{ik}) \pi(\mathbf{z}_{ik} | \tau)$.

Let \mathcal{Z}_{-ik} denote the integration region of \mathbf{z}_{-ik} . Note that

$$\begin{aligned}
 & \int_{\Theta} f(\mathbf{y}_{-ik} | X_{-ik}, W_{-ik}, \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
 & = \int_{\Theta} \int_{\mathcal{Z}_{-ik}} f(\mathbf{y}_{-ik} | X_{-ik}, W_{-ik}, \boldsymbol{\theta}, \mathbf{z}_{-ik}) \\
 & \quad \times \pi(\mathbf{z}_{-ik} | \tau) \pi(\boldsymbol{\theta}) dz_{-ik} d\boldsymbol{\theta} \\
 & = \int_{\Theta} \int_{\mathcal{Z}} h_{ik}(\mathbf{z}_{ik}) f(\mathbf{y}_{-ik} | X_{-ik}, W_{-ik}, \boldsymbol{\theta}, \mathbf{z}_{-ik}) \\
 \text{(A.2)} \quad & \times \pi(\mathbf{z}_{-ik} | \tau) \pi(\boldsymbol{\theta}) dz d\boldsymbol{\theta},
 \end{aligned}$$

where $h_{ik}(\mathbf{z}_{ik})$ is any weight density function such that $\int_{\mathcal{Z}_{ik}} h_{ik}(\mathbf{z}_{ik}) dz_{ik} = 1$, and

$$\begin{aligned}
 & \frac{f(\mathbf{y}_{-ik} | X_{-ik}, W_{-ik}, \boldsymbol{\theta}, \mathbf{z}_{-ik}) \pi(\mathbf{z}_{-ik} | \tau)}{1} \\
 & = \frac{f(\mathbf{y}_{ik} | X_{ik}, W_{ik}, \boldsymbol{\theta}, \mathbf{z}_{ik}) \pi(\mathbf{z}_{ik} | \tau)}{f(\mathbf{y}_{ik} | X_{ik}, W_{ik}, \boldsymbol{\theta}, \mathbf{z}_{ik}) \pi(\mathbf{z}_{ik} | \tau)} \\
 \text{(A.3)} \quad & \times \left[\prod_{k=1}^K \prod_{i=1}^{n_k} f(\mathbf{y}_{ik} | X_{ik}, W_{ik}, \boldsymbol{\theta}, \mathbf{z}_{ik}) \pi(\mathbf{z}_{ik} | \tau) \right].
 \end{aligned}$$

Thus, combining (A.1), (A.2), and (A.3) yields (4.3).

Appendix B. Derivation of an approximation of the optimal weight function

The optimal weight function is given by

$$\begin{aligned}
 & h_{ik, \text{opt}}(\mathbf{z}_{ik}) \\
 & = f(\mathbf{z}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, \mathbf{y}_{ik}, X_{ik}, W_{ik})
 \end{aligned}$$

where $f^M(\mathbf{y}_{ik}, \mathbf{z}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik}) = \int_{\mathcal{Z}_{ik} \in R_+^J} f(\mathbf{y}_{ik}, \mathbf{z}_{ik} | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik}) dz_{ik}$. Let $\mathbf{u}_{ik} = \log(\mathbf{z}_{ik})$ for $i = 1, \dots, n_k$, $k = 1, \dots, K$, then $\mathbf{z}_{ik} = \exp(\mathbf{u}_{ik})$. We note that the log and exp functions applied to a vector means that the operations are applied to every element of the vector. Then, we have

$$h_{ik, \text{opt}}(\mathbf{z}_{ik}) = \frac{f_{\mathbf{y}_{ik}, \mathbf{u}_{ik}}^A}{\int_{\mathbf{u}_{ik} \in R^J} f_{\mathbf{y}_{ik}, \mathbf{u}_{ik}}^A \times \prod_{j=1}^J \exp(u_{ijk}) d\mathbf{u}_{ik}},$$

where $f_{\mathbf{y}_{ik}, \mathbf{u}_{ik}}^A = f(\mathbf{y}_{ik}, \exp(\mathbf{u}_{ik}) | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik})$, and

$$\begin{aligned}
 & h_{ik, \text{opt}}(\mathbf{z}_{ik}) \prod_{j=1}^J \exp(u_{ijk}) \\
 \text{(A.4)} \quad & = \frac{f_{\mathbf{y}_{ik}, \mathbf{u}_{ik}}^A \times \prod_{j=1}^J \exp(u_{ijk})}{\int_{\mathbf{u}_{ik} \in R^J} f_{\mathbf{y}_{ik}, \mathbf{u}_{ik}}^A \times \prod_{j=1}^J \exp(u_{ijk}) d\mathbf{u}_{ik}}.
 \end{aligned}$$

The right hand side of (A.4) can be approximated by a multivariate normal density which is constructed via the Laplace approximation to the joint density $f(\mathbf{y}_{ik}, \exp(\mathbf{u}_{ik}) | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik}) \prod_{j=1}^J \exp(u_{ijk})$. Denote $g(\mathbf{u}_{ik}) = \log \{ f(\mathbf{y}_{ik}, \exp(\mathbf{u}_{ik}) | \boldsymbol{\beta}, \boldsymbol{\delta}_k, \Sigma_k, \boldsymbol{\gamma}_k^R, \tau, \nu_k, X_{ik}, W_{ik}) \prod_{j=1}^J \exp(u_{ijk}) \}$. Letting \mathbf{u}_{ik}^0 be the stationary point where $\nabla g(\mathbf{u}_{ik}) = 0$, and $B = -\nabla^2 g(\mathbf{u}_{ik})|_{\mathbf{u}=\mathbf{u}_{ik}^0}$, the function $h_{ik, \text{opt}}(\mathbf{z}_{ik}) \prod_{j=1}^J \exp(u_{ijk})$ can be approximated by $\phi(\log(\mathbf{z}_{ik}) | \mathbf{u}_{ik}^0, B^{-1})$, where $\phi(\log(\mathbf{z}_{ik}) | \mathbf{u}_{ik}^0, B^{-1})$ denotes the density function of $N(\mathbf{u}_{ik} | \mathbf{u}_{ik}^0, B^{-1})$. Thus, $h_{ik, \text{opt}}(\mathbf{z}_{ik})$ is approximated by the function

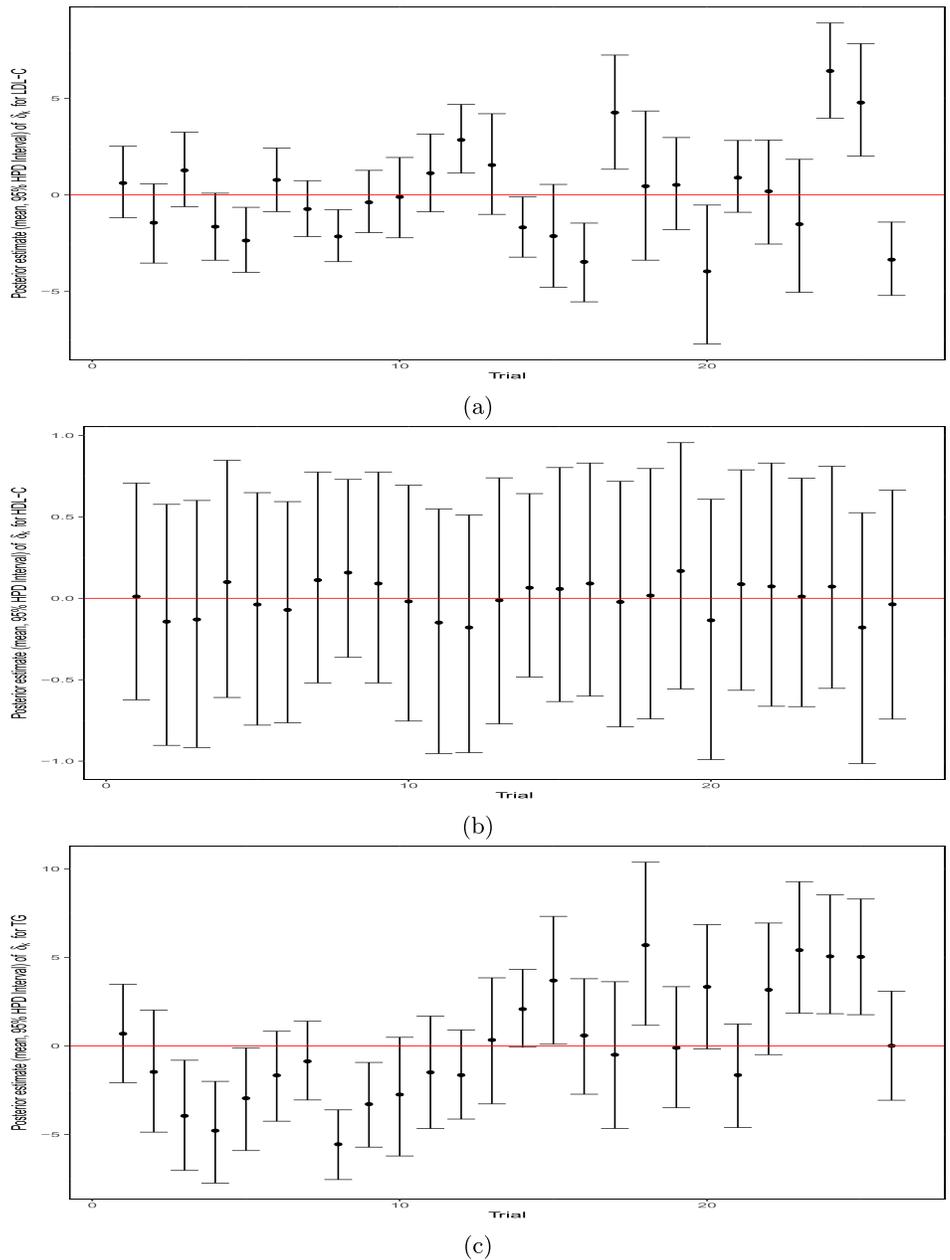


Figure 5. Plots of the posterior estimates (mean and 95% HPD interval) for δ_k under the flexible skew model with random ν_α and τ for the cholesterol data.

$\frac{\phi(\log(\mathbf{z}_{ik})|\mathbf{u}_{ik}, B^{-1})}{\prod_{j=1}^J z_{ijk}}$. From equation (4.2), we have

$$\begin{aligned}
 g(\mathbf{u}_{ik}) &= \log \left\{ f_{\mathbf{y}_{ik}, \mathbf{u}_{ik}}^A \times \prod_{j=1}^J \exp(u_{ijk}) \right\} \\
 &\propto -\frac{\nu_k + J}{2} \log(A_{ik}(\mathbf{u}_{ik}) + \nu_k) \\
 &\quad - (J + \tau + 1) \log \left(\sum_{j=1}^J \exp(u_{ijk}) + \tau \right)
 \end{aligned}$$

$$(A.5) \quad + \sum_{j=1}^J u_{ijk},$$

where $A_{ik}(\mathbf{u}_{ik}) = (\mathbf{y}_{ik} - X_{ik}\boldsymbol{\beta} - W_{ik}\boldsymbol{\gamma}_k^R - \Delta_k(\exp(\mathbf{u}_{ik}) - E\mathbf{z}_{ik}))' \Sigma_k^{-1} (\mathbf{y}_{ik} - X_{ik}\boldsymbol{\beta} - W_{ik}\boldsymbol{\gamma}_k^R - \Delta_k(\exp(\mathbf{u}_{ik}) - E\mathbf{z}_{ik}))$. Using $g(\mathbf{u}_{ik})$ in (A.5), we obtain $\nabla g(\mathbf{u}_{ik})$ and $\nabla^2 g(\mathbf{u}_{ik})$ after some algebra.

ACKNOWLEDGEMENTS

Drs. Chen and Ibrahim'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.

Received 3 January 2020

REFERENCES

- ADCOCK, C. (2004). *Capital Asset Pricing for UK Stocks under the Multivariate Skew-Normal Distribution*. Boca Raton, FL: Chapman & Hall/CRC. [MR2155330](#)
- ARELLANO-VALLE, R. and GENTON, M. (2010). Multivariate unified skew-elliptical distributions. *Chilean Journal of Statistics*, 1:17–33. [MR2756082](#)
- ATP III (2001). Executive summary of the third report of the national cholesterol education program (ncep) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel iii). *Journal of the American Medical Association*, 285:2486–2497.
- AUSTIN, M., HOKANSON, J., and EDWARDS, K. (1998). Hypertriglyceridemia as a cardiovascular risk factor. *The American Journal of Cardiology*, 81:7B–12B.
- AZZALINI, A. and CAPITANIO, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of The Royal Statistical Society, Series B*, 65:367–389. [MR1983753](#)
- BAIGENT, C., BLACKWELL, L., ET AL., and for CHOLESTEROL TREATMENT TRIALISTS' (CTT) COLLABORATION (2010). Efficacy and safety of more intensive lowering of ldl cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials for cholesterol treatment trialists' (ctt) collaboration. *Lancet*, 376:1670–1681.
- BELIAS, M., ROVERS, M., REITSMA, J., DEBRAY, T., and INTHOUT, J. (2019). Statistical approaches to identify subgroups in meta-analysis of individual participant data: a simulation study. *BMC Medical Research Methodology*, 19:183.
- BHATT, D., STEG, P., MILLER, M., BRINTON, E., JACOBSON, T., KETCHUM, S., JR., R. D., JULIANO, R., JIAO, L., GRANOWITZ, C., TARDIF, J.-C., BALLANTYNE, C., and for the REDUCE-IT INVESTIGATORS (2019). Cardiovascular risk reduction with icosapent ethyl for hypertriglyceridemia. *The New England Journal of Medicine*, 380:11–22.
- BRANCO, M. and DEY, D. (2001). A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis*, 79:99–113. [MR1867257](#)
- BURKE, D., ENSOR, J., and RILEY, R. (2017). Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Statistics in Medicine*, 36:855–875. [MR3597661](#)
- CANNON, C., BLAZING, M., GIUGLIANO, R., MCCAGG, A., WHITE, J. A., THEROUX, P., DARIUS, H., LEWIS, B., OPHUIS, T. O., JUKEMA, J., FERRARI, G. D., RUZYLO, W., LUCCA, P. D., IM, K., BOHULA, E., REIST, C., WIVIOTT, S., TERSHAKOVIC, A., MUSLINER, T., BRAUNWALD, E., CALIFF, R., and for the IMPROVE-IT INVESTIGATORS (2015). Ezetimibe added to statin therapy after acute coronary syndromes. *The New England Journal of Medicine*, 372:2387–2397.
- CHANG, S. C. and ZIMMERMAN, D. L. (2016). Skew-normal antedependence models for skewed longitudinal data. *Biometrika*, 103:363–376. [MR3509892](#)
- CHEN, M.-H., DEY, D., and SHAO, Q. (1999). A new skewed link model for dichotomous quantal response data. *Journal of the American Statistical Association*, 94:1172–1186. [MR1731481](#)
- CHEN, M.-H. and SHAO, Q. (1999). Monte carlo estimation of bayesian credible and hpd intervals. *Journal of Computational and Graphical Statistics*, 8:69–92. [MR1705909](#)
- CHEN, M.-H., SHAO, Q., and IBRAHIM, J. (2000). *Monte Carlo Methods in Bayesian Computation*. New York: Springer. [MR1742311](#)
- GENTON, M. (2004). *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. Boca Raton, FL: Chapman & Hall/CRC. [MR2155324](#)
- HERON, M. (2019). Deaths: Leading causes for 2017. *National Vital Statistics Reports*, 68:June 24.
- HOKANSON, J. and AUSTIN, M. (1996). Plasma triglyceride level is a risk factor for cardiovascular disease independent of high-density lipoprotein cholesterol level: a meta-analysis of population-based prospective studies. *Journal of Cardiovascular Risk*, 3:213–219.
- IBRAHIM, J., CHEN, M.-H., and SINHA, D. (2001). *Bayesian Survival Analysis*. New York: Springer. [MR1876598](#)
- IBRAHIM, J., KIM, S., CHEN, M.-H., SHAH, A., and LIN, J. (2019). Bayesian multivariate skew meta-regression models for individual patient data. *Statistical Methods in Medical Research*, 28:3415–3436. [MR4002707](#)
- KASHANI, A., SALLAM, T., BHEEMREDDY, S., MANN, D., WANG, Y., and FOODY, J. (2008). Review of side-effect profile of combination ezetimibe and statin therapy in randomized clinical trials. *The American Journal of Cardiology*, 101:1606–1613.
- KIM, S., CHEN, M.-H., and DEY, D. (2008). Flexible generalized t-link models for binary response data. *Biometrika*, 95:93–106. [MR2409717](#)
- KIM, S., CHEN, M.-H., IBRAHIM, J., SHAH, A., and LIN, J. (2013). 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*, 32:3972–3990. [MR3102429](#)
- LEITER, L., BETTERIDGE, D., FARNIER, M., GUYTON, J., LIN, J., SHAH, A., JOHNSON-LEVONAS, A., and BRUDI, P. (2011). 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*, 13:615–628.
- LIU, J. (1994). The collapsed gibbs sampler in bayesian computations with applications to a gene regulation problem. *Journal of the American Statistical Association*, 89:958–966. [MR1294740](#)
- RILEY, R., PRICE, M., JACKSON, D., WARDLE, M., GUEYFFIER, F., WANG, J., STAESSEN, J., and WHITE, I. (2015). Multivariate meta-analysis using individual participant data. *Research Synthesis Methods*, 6:157–174.
- RITZ, J., DEMIDENKO, E., and SPIEGELMAN, D. (2008). Multivariate meta-analysis for data consortia, individual patient meta-analysis, and pooling projects. *Journal of Statistical Planning and Inference*, 138:1919–1933. [MR2406414](#)
- SAHU, S., DEY, D., and BRANCO, M. (2003). A new class of multivariate skew distributions with applications to bayesian regression models. *Canadian Journal of Statistics*, 31:129–150. [MR2016224](#)
- SARWAR, N., DANESH, J., EIRIKSDOTTIR, G., SIGURDSSON, G., WAREHAM, N., BINGHAM, S., BOEKHOLDT, S., KHAW, K., and GUDNASON, V. (2007). Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 western prospective studies. *Circulation*, 115:450–458.
- SIMMONDS, M. and HIGGINS, J. (2007). Covariate heterogeneity in meta-analysis: Criteria for deciding between meta-regression and individual patient data. *Statistics in Medicine*, 26:2982–2999. [MR2370988](#)
- ZHANG, D., CHEN, M.-H., IBRAHIM, J., BOYE, M., and SHEN, W. (2017). Bayesian model assessment in joint modeling of longitudinal and survival data with applications to cancer clinical trials. *Journal of Computational and Graphical Statistics*, 26:121–133. [MR3610413](#)

Sungduk Kim
Biostatistics Branch
Division of Cancer Epidemiology and Genetics
National Cancer Institute
Rockville, MD
USA
E-mail address: kims2@mail.nih.gov

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

Joseph Ibrahim
Department of Biostatistics
University of North Carolina at Chapel Hill
Chapel Hill, NC
USA
E-mail address: ibrahim@bios.unc.edu

Arvind Shah
Clinical Biostatistics
Merck & Co., Inc.
Rahway, NJ
USA
E-mail address: arvind.shah@merck.com

Jianxin Lin
Clinical Biostatistics
Merck & Co., Inc.
Rahway, NJ
USA
E-mail address: jianxin.lin@merck.com