Research Article

Statistics in Medicine

Received 1 October 2010, Accepted 11 April 2011 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.4291

A Bayesian growth mixture model to examine maternal hypertension and birth outcomes

Brian Neelon,[a*†] Geeta K. Swamy,[b] Lane F. Burgette[c] and Marie Lynn Miranda[a]

Maternal hypertension is a major contributor to adverse pregnancy outcomes, including preterm birth (PTB) and low birth weight (LBW). Although several studies have explored the relationship between maternal hypertension and fetal health, few have examined how the longitudinal trajectory of blood pressure, considered over the course of pregnancy, affects birth outcomes. In this paper, we propose a Bayesian growth mixture model to jointly examine the associations between longitudinal blood pressure measurements, PTB, and LBW. The model partitions women into distinct classes characterized by a mean arterial pressure (MAP) curve and joint probabilities of PTB and LBW. Each class contains a unique mixed effects model for MAP with class-specific regression coefficients and random effect covariances. To account for the strong correlation between PTB and LBW, we introduce a bivariate probit model within each class to capture residual within-class dependence between PTB and LBW. The model permits the association between PTB and LBW to vary by class, so that for some classes, PTB and LBW may be positively correlated, whereas for others, they may be uncorrelated or negatively correlated. We also allow maternal covariates to influence the class probabilities via a multinomial logit model. For posterior computation, we propose an efficient MCMC algorithm that combines full-conditional Gibbs and Metropolis steps. We apply our model to a sample of 1027 women enrolled in the Healthy Pregnancy, Healthy Baby Study, a prospective cohort study of host, social, and environmental contributors to disparities in pregnancy outcomes. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: growth mixture model; latent trajectory model; Bayesian analysis; correlated probit model; maternal hypertension; birth outcomes

1. Introduction

Hypertension in pregnancy is associated with poor birth outcomes, including preterm birth (PTB), low birth weight (LBW), and fetal growth restriction [1]. In healthy pregnant women, blood pressure declines steadily until mid-gestation and then rises until delivery. In contrast, women who develop preeclampsia—a pregnancy-specific syndrome defined by hypertension and elevated urine protein—typically have no early decline in blood pressure, with blood pressure remaining stable during the first half of pregnancy and then rising until delivery [2]. These women also tend to have higher early pregnancy blood pressure levels than those who remain normotensive. Older, non-Hispanic black, and primiparous women are more likely to experience hypertensive disorders during pregnancy than other demographic groups [3, 4]. By monitoring the course of blood pressure during early pregnancy, obstetric providers may be able to identify women at risk for adverse birth outcomes, increase maternal–fetal surveillance, and intervene with appropriate treatments.

By using data from the Healthy Pregnancy, Healthy Baby (HPHB) Study, a prospective cohort study of host, social, and environmental contributors to disparities in pregnancy outcomes, we develop a Bayesian

[a] *Children's Environmental Health Initiative, Nicholas School of the Environment, Duke University, Durham, NC 27708, U.S.A.*
[b] *Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, NC 27710, U.S.A.*
[c] *Department of Statistical Science, Duke University, Durham, NC 27708, U.S.A.*
**Correspondence to: Brian Neelon, Children's Environmental Health Initiative, Nicholas School of the Environment, Duke University, Durham, NC 27708, U.S.A.*
[†] *E-mail: brian.neelon@duke.edu*

growth mixture model [5, 6] to explore the relationship between three outcomes: PTB (gestational age <37 weeks), LBW (birth weight under 2500 g), and maternal mean arterial pressure (MAP), a measure of average blood pressure defined as $(1/3 \times \text{systolic blood pressure}) + (2/3 \times \text{diastolic blood pressure})$. Growth mixture models (GMMs) assume that subjects first fall into one of a finite number of latent classes characterized by a class-specific mean trajectory; then, distributed about these class means, subjects have their own unique longitudinal trajectories defined by a set of random effects with class-specific variance parameters. As such, GMMs can be viewed as finite mixtures of random effects models.

Increasingly, growth mixtures have been used to jointly model discrete and continuous outcomes. Lin *et al.* [7] developed a GMM to relate prostate-specific antigen trajectories to the development of prostate cancer. Muthén *et al.* [8] used a GMM to model adolescent aggression and a binary indicator of juvenile delinquency. Proust-Lima *et al.* [9] proposed a GMM to jointly model multivariate psychological measures and the subsequent occurrence of dementia. Lin *et al.* [10] and Proust-Lima *et al.* [11] developed related models to analyze longitudinal biomarkers and a time to event. More recently, Zhang and Wang [12] used a latent class cure rate model to jointly analyze multiple survival outcomes under a conditional independence assumption. Within the Bayesian framework, Elliott *et al.* [13] used a Bayesian GMM to jointly model daily affect and negative event occurrences during a 35-day study period, and Slaughter *et al.* [14] proposed a related Bayesian latent variable mixture model to examine fetal growth restriction.

We build on this previous work by examining the connections between a longitudinal continuous outcome (MAP) and two strongly correlated, cross-sectional endpoints (PTB and LBW). In this setting, a finite mixture approach is intuitively appealing, because it accounts for the marginal correlation between the repeated measures and cross-sectional outcomes while simultaneously addressing the important clinical question of whether women can be clustered into groups jointly characterized by MAP trajectories and probabilities of PTB and LBW. Within each class, we model MAP by using a mixed effects model with class-specific regression coefficients and random effect covariances. Because PTB and LBW are strongly correlated variables, with correlation coefficients typically 0.50 or higher, we introduce a bivariate probit model within each cluster to capture residual correlation between PTB and LBW. An important feature of this model is that it permits the association between PTB and LBW to vary by class, so that for some classes, PTB and LBW may exhibit strong positive correlations, whereas for others, they may be uncorrelated or negatively correlated. We also allow the mixing weights to depend on maternal covariates through a multinomial logit model. Further, we extend previous work on mixed outcome GMMs [8–10, 12, 15] by developing a Bayesian estimation approach. The advantages of Bayesian inference are well known and include incorporation of prior information, avoidance of asymptotic approximations, and practical estimation of parameter functions. For posterior computation, we propose an efficient MCMC algorithm that alternates between full-conditional Gibbs and Metropolis steps and that can easily be coded using standard statistical software.

We organize the remainder of the paper as follows: Section 2 provides an overview of the HPHB Study; Section 3 describes the proposed model, outlines posterior computation, and proposes strategies for model comparison and assessment of model fit; Section 4 applies the method to data from the HPHB Study; and Section 5 concludes with discussion and future research directions.

2. The Healthy Pregnancy, Healthy Baby Study

The HPHB Study is an ongoing cohort study aimed at determining how host, social, and environmental factors contribute to disparities in pregnancy outcomes. The study is part of the EPA-funded Southern Center on Environmentally Driven Disparities in Birth Outcomes and enrolls pregnant women from the Duke University Obstetrics Clinic and the Durham County Health Department Prenatal Clinic in Durham, North Carolina. One objective of the study was to explore the relationship between hypertensive disorders during pregnancy and birth outcomes.

Beginning in July 2005, women who were English literate, were at least 18 years old, and had a singleton gestation <28 weeks without known congenital anomalies were recruited into the study. Demographic characteristics, medical history, and pregnancy outcomes, including blood pressure measurements at routinely scheduled prenatal visits, were collected via electronic medical record review. Maternal blood samples were collected peripartum to assess environmental exposures such as cotinine, a measure of tobacco exposure. Birth outcomes, including length of gestation and birth weight, were recorded at the time of delivery.

At the time of analysis, over 1200 pregnant women had been enrolled in the study. Because of the small number of Hispanics and other ethnic groups, we restricted our analysis to non-Hispanic black and

non-Hispanic white women. We also excluded women with a history of chronic hypertension. The final sample comprised 1027 women and 10,290 MAP measurements, with a median of nine MAP measurements per participant. The sample proportion of PTB and LBW were 13% and 12%, respectively. These are slightly higher than the national rates, which, in 2008, were 12.3% for PTB and 8.2% for LBW [16]. Table I presents summary statistics of the study participants.

3. Proposed growth mixture model

As noted previously, an ongoing aim of the HPHB Study is to explore factors that affect maternal blood pressure and subsequent birth outcomes. Consequently, we sought to determine whether women could be reasonably grouped into a small number of blood pressure trajectory classes and whether these classes were in turn associated with disparate birth outcomes. Given these aims, it seemed natural to pursue a GMM to jointly model MAP, PTB, and LBW as a function of subject-level predictors.

To this end, let y_{ij} denote the MAP value at the jth clinical visit for woman i, let z_{1i} be a binary indicator that woman i had a preterm delivery, and let z_{2i} be a binary indicator that woman i delivered a LBW infant. To model the three outcomes jointly, we propose the following GMM:

$$\begin{aligned} f(y_{ij}, z_{1i}, z_{2i}|\mathbf{b}_i; \boldsymbol{w}_i, \boldsymbol{x}_{ij}, \boldsymbol{v}_{ij}) &= \sum_{k=1}^{K} \pi_{ik}(\boldsymbol{w}_i)\mathrm{N}(y_{ij}; \eta_{ijk}, \sigma_k^2)\, p(z_{1i}, z_{2i}; \boldsymbol{\psi}_k) \\ &= \sum_{k=1}^{K} \Pr(C_i = k; \boldsymbol{w}_i)\mathrm{N}(y_{ij}; \eta_{ijk}, \sigma_k^2)\, p(z_{1i}, z_{2i}; \boldsymbol{\psi}_k), \\ \eta_{ijk} &= \boldsymbol{x}_{ij}'\boldsymbol{\beta}_k + \boldsymbol{v}_{ij}'\mathbf{b}_i, \quad j = 1, \ldots, n_i;\ i = 1, \ldots, n, \end{aligned} \tag{1}$$

where C_i is a discrete latent variable taking the value k if subject i belongs to class k ($k = 1, \ldots, K$); $\boldsymbol{w}_i$ is an $r \times 1$ vector of subject-level covariates; $\mathrm{N}(\cdot; \mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2; $p(z_{1i}, z_{2i}; \boldsymbol{\psi}_k)$ represents a generic bivariate distribution for $\boldsymbol{z}_i = (z_{1i}, z_{2i})'$, with class-specific parameters $\boldsymbol{\psi}_k$; $\boldsymbol{x}_{ij}$ and $\boldsymbol{v}_{ij}$ are, respectively, $p \times 1$ and $q \times 1$ vectors of fixed and random effect covariates for MAP, including variables for gestational age; β_k is a $p \times 1$ vector of class-k fixed

Table I. Characteristics of the Healthy Pregnancy, Healthy Baby Study participants ($n = 1027$).

Variable	n (%)
Maternal race	
Non-Hispanic white	231 (22)
Non-Hispanic black	796 (78)
Maternal age	
18–20 years	252 (25)
21–34 years	660 (64)
≥35 years	115 (11)
Maternal education	
≤High school	545 (53)
>High school	482 (47)
Parity	
Primiparous	427 (42)
Multiparous	600 (58)
Insurance status	
Private	234 (23)
Other	793 (77)
Preterm birth	129 (13)
Low birth weight	123 (12)
	Mean (SD)
Serum cotinine (ng/mL)	19.44 (52.17)
Mean arterial pressure (mmHg)	88.0 (9.13)

effect regression parameters associated with MAP; and $\mathbf{b}_i \mid C_i = k \sim \mathrm{N}_q(\mathbf{0}, \boldsymbol{\Sigma}_k)$ is a $q \times 1$ vector of random effects for MAP, with class-specific covariance $\boldsymbol{\Sigma}_k$.

Model 1 states that the joint distribution of the three outcomes can be written as a finite mixture that includes a mixed effects model for MAP and a discrete bivariate model for PTB and LBW, each model having its own class-specific parameters. The MAP model is particularly flexible because it permits random variation about the class-specific mean regression curves. Moreover, the random effect covariances are allowed to vary from class to class. Hence, for some classes, there may be extensive variation about the mean trajectory, while for other classes there may be little or no variation, in which case the model essentially reduces to a fixed effects model for that class.

Model 1 also implies that y_{ij}, z_{1i}, and z_{2i} are marginally correlated across the classes; however, we assume conditional on $C_i = k$, y_{ij} to be independent of z_{1i} and z_{2i}. This assumption seems justifiable in our application, because the marginal biserial correlations between MAP and the other two outcomes, although statistically significant, were less than 0.08 in absolute value. This was not true for PTB and LBW, however; in our study, Kendall's tau coefficient for PTB and LBW was 0.49, and the Goodman–Kruskal gamma was 0.90. Although the latent class structure accounts for some of the marginal association between z_{1i} and z_{2i}, so that $\mathrm{corr}(z_{1i}, z_{2i} | C_i = k)$ will in general be closer to the null than the unconditional correlation, $\mathrm{corr}(z_{1i}, z_{2i})$, it is unlikely that it will account for all of the associations. To accommodate this 'residual' within-class dependence between z_{1i} and z_{2i}, we fit a bivariate probit model (cf. [17], Section 4.3) within each class. Specifically, we introduce two underlying normal variables, z^*_{1i} and z^*_{2i}, such that $z_{1i} = 1$ if $z^*_{1i} > 0$ and $z_{2i} = 1$ if $z^*_{2i} > 0$. Further, we assume z^*_{1i} and z^*_{2i} to follow a bivariate normal density with class-specific parameters; that is,

$$\left[\begin{pmatrix} z^*_{1i} \\ z^*_{2i} \end{pmatrix} \middle| C_i = k \right] \overset{iid}{\sim} \mathrm{N}_2(\boldsymbol{\mu}_k, \boldsymbol{R}_k) = \mathrm{N}_2 \left[\begin{pmatrix} \mu_{1k} \\ \mu_{2k} \end{pmatrix}, \begin{pmatrix} 1 & \rho_k \\ \rho_k & 1 \end{pmatrix} \right], \tag{2}$$

where the diagonals of $\boldsymbol{R}_k$ are set to 1 for identifiability. The correlation parameter ρ_k controls the dependence between z_{1i} and z_{2i} and can vary by class. For some classes, ρ_k may be close to 1, indicating strong positive association between z_{1i} and z_{2i}, whereas for other classes, ρ_k may be close to 0, reflecting minimal correlation between the two variables. For the HPHB analysis described below, we assume that z^*_{1i} and z^*_{2i} are independent and identically distributed within each class, and allow covariates to enter our model through the class-membership probabilities, $\pi_{ik}(\boldsymbol{w}_i)$. In general, however, one can allow $\boldsymbol{\mu}_k$ to be function of subject-level predictors.

The joint probabilities for z_{1i} and z_{2i} are obtained by integrating over the bivariate normal density for z^*_{1i} and z^*_{2i}. For example,

$$\Pr(z_{1i} = 1 \,\&\, z_{2i} = 1 \mid C_i = k) = \int_0^\infty \int_0^\infty \mathrm{N}_2(z^*; \boldsymbol{\mu}_k, \boldsymbol{R}_k) dz^*_{1i} dz^*_{2i},$$

where $\boldsymbol{z}^* = (z^*_{1i}, z^*_{2i})'$. The marginal probabilities can be obtained by summing the appropriate joint probabilities. For example, the class-k marginal probability that $z_{1i} = 1$ is $\Pr(z_{1i} = 1 \mid C_i = k) = \Pr(z_{1i} = 1 \,\&\, z_{2i} = 1 \mid C_i = k) + \Pr(z_{1i} = 1 \,\&\, z_{2i} = 0 \mid C_i = k)$.

To complete the model, we assume that the class indicator C_i has a discrete 'categorical' distribution taking the value k with probability $\pi_{ik}(\boldsymbol{w}_i)$, where $\pi_{ik}(\boldsymbol{w}_i)$ is linked to the r-dimensional covariate vector $\boldsymbol{w}_i$ via a multinomial logit model:

$$\begin{aligned} C_i &\sim \mathrm{Cat}\left[\pi_{i1}(\boldsymbol{w}_i), \ldots, \pi_{iK}(\boldsymbol{w}_i)\right], \\ \pi_{ik}(\boldsymbol{w}_i) &= \frac{e^{\boldsymbol{w}_i' \boldsymbol{\gamma}_k}}{\sum_{h=1}^K e^{\boldsymbol{w}_i' \boldsymbol{\gamma}_h}}, \text{ with } \boldsymbol{\gamma}_1 = 0 \text{ for identifiability.} \end{aligned} \tag{3}$$

Finally, for each model fit, we assume that the number of classes K is known. To determine the optimal value of K, we fit several models, each with a unique value of $K \in \{1, \ldots, K_{\max}\}$, and apply a Bayesian model-selection criterion to determine the appropriate choice of K. In Section 6, we suggest alternatives to fixing K for each model.

4. Prior specification, posterior computation, and model selection

4.1. Prior specification

To ensure a well-identified model, we assign weakly informative proper prior distributions to all model parameters. For each β_k, we assume a conjugate $\mathrm{N}_p(\boldsymbol{m}_0, \boldsymbol{S}_0)$ prior; for $\tau_k = \sigma_k^{-2}$, we assume a conjugate $\mathrm{Ga}(\lambda, \delta)$ prior; and for each $\boldsymbol{\Sigma}_k$, we assume a conjugate $\mathrm{IW}(\nu_0, \boldsymbol{D}_0)$ ($\nu_0 \geq q$) distribution. In addition, we assume a $\mathrm{N}_2(\boldsymbol{m}_\mu, \boldsymbol{S}_\mu)$ prior for the bivariate probit means, $\boldsymbol{\mu}_k$, and we allocate $\mathrm{U}(-1, 1)$ priors to the bivariate probit correlations, ρ_k. Finally, for the class-membership parameters, $\boldsymbol{\gamma}_k$, we follow Elliot *et al.* [13] and Garrett and Zeger [18] in recommending a $\mathrm{N}_r[\mathbf{0}, (9/4)\boldsymbol{I}_r]$ prior, where $\boldsymbol{I}_r$ denotes the r-dimensional identity matrix; this in turn induces a prior for $\pi_{ik}(\boldsymbol{w}_i)$ centered at $1/K$ and bounded away from 0 and 1. If there are no class-membership predictors (i.e., $r = 1$), a conjugate $\mathrm{Dirichlet}(e_1, \ldots, e_K)$ prior can be placed directly on the (subject-invariant) class-membership probabilities $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K)'$, which provides convenient closed-form full conditionals. In this context, Frühwirth-Schnatter [19] recommends choosing $e_k > 1$ to bound π_k away from zero.

4.2. Posterior computation

For notational convenience, let $\boldsymbol{\theta}_k = (\boldsymbol{\beta}_k', \sigma_k^2, \boldsymbol{\mu}_k', \rho_k)'$, let $\boldsymbol{C} = (C_1, \ldots, C_n)'$, let $\mathbf{b} = (\mathbf{b}_1', \ldots, \mathbf{b}_n')'$ denote a concatenated vector of random effects for all n subjects, and let $z = (z_1', \ldots, z_n')'$ denote a stacked vector of PTB and LBW observations for the n study participants. Assuming prior independence of the model parameters, the joint posterior is given by

$$\pi(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K; \boldsymbol{C}; \mathbf{b}; \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_K; \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_K | \boldsymbol{y}, z) \propto$$
$$\prod_{k=1}^{K} \left\{ \prod_{i=1}^{n} \pi_{ik}(\boldsymbol{w}_i) \left[\prod_{j=1}^{n_i} \mathrm{N}\left(y_{ij} | \eta_{ijk}, \sigma_k^2\right) \right] p(z_i | \boldsymbol{\mu}_k, \rho_k) \mathrm{N}_q(\mathbf{b}_i | \boldsymbol{\Sigma}_k) \right\}^{\mathrm{I}_{(C_i = k)}}$$
$$\times \pi(\boldsymbol{\beta}_k) \pi(\sigma_k^2) \pi(\boldsymbol{\Sigma}_k) \pi(\boldsymbol{\mu}_k) 1_{\{\rho_k \in (-1,1)\}} \prod_{h=2}^{K} \pi(\boldsymbol{\gamma}_h),$$

where $\mathrm{N}(y_{ij} | \eta_{ijk}, \sigma_k^2)$ denotes the normal likelihood for y_{ij} described in Equation (1); $p(z_i | \boldsymbol{\mu}_k, \rho_k)$ is the discrete joint distribution of z_i conditional on the mean and correlation of the underlying continuous variable z_i^*; $\mathrm{N}_q(\mathbf{b}_i | \boldsymbol{\Sigma}_k)$ denotes the q-dimensional normal prior for $\mathbf{b}_i$ with covariance matrix $\boldsymbol{\Sigma}_k$; $\mathrm{I}_{(C_i = k)}$ is an indicator that subject i belongs to class k; $1_{\{\rho_k \in (-1,1)\}}$ is the prior density for ρ_k; and the remaining $\pi(\cdot)$'s represent the prior distributions for their respective parameters (with $\boldsymbol{\gamma}_1 \equiv 0$ with probability 1 for identifiability).

For posterior computation, we propose an MCMC algorithm that combines draws from full conditionals with Metropolis-based updates. After we assigned initial values to the model parameters, the algorithm iterates between the following steps:

(1) Update $\boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_K$ by using random-walk Metropolis steps;
(2) Sample the class indicators C_i ($i = 1, \ldots, n$) from a discrete categorical distribution with probability vector $\boldsymbol{p}_i = (p_{i1}, \ldots, p_{iK})$ as described in Appendix A;
(3) Given $C_i = k$, update z_{1i}^* and z_{2i}^* as follows:
   (a) Draw $z_{1i}^* | z_{2i}^*, C_i = k$ from its full-conditional normal distribution truncated below (above) by zero for $z_{1i} = 1$ ($z_{1i} = 0$);
   (b) Draw $z_{2i}^* | z_{1i}^*, C_i = k$ from its full-conditional normal distribution truncated below (above) by zero for $z_{2i} = 1$ ($z_{2i} = 0$);
(4) For $k = 1, \ldots, K$, sample the class-specific parameters $\boldsymbol{\beta}_k$, σ_k^2, $\boldsymbol{\Sigma}_k$, and $\boldsymbol{\mu}_k$ from their full conditionals, and update ρ_k by using a random-walk Metropolis step with proposal density restricted to $(-1, 1)$;
(5) Given $C_i = k$, update $\mathbf{b}_i$ by using a full-conditional Gibbs step or, alternatively, a random-walk Metropolis step that is a function of class-k data and parameters.

Appendix A provides details of the algorithm. We monitored convergence by running multiple chains from dispersed initial values and examining standard Bayesian diagnostics, such as trace plots and the Brooks–Gelman–Rubin statistic, $\widehat{R}$, which compares the within-chain variation with the between-chain variation [20]. As a practical rule, a 0.975 quantile for $\widehat{R}$ less than 1.2 is indicative of convergence. In

the application in the succeeding section, we monitored convergence diagnostics by using the R package boa [21].

4.3. Model selection and fit

For model selection, we adopt the deviance information criterion (DIC) developed by Speigelhalter *et al.* [22]. DIC is defined as $\overline{D}(\boldsymbol{\theta}) + p_D$, where $\overline{D}(\boldsymbol{\theta}) = \mathrm{E}[D(\boldsymbol{\theta})|\boldsymbol{y}]$ is the posterior mean of the deviance, $D(\boldsymbol{\theta})$, and $p_D = \overline{D}(\boldsymbol{\theta}) - D(\mathrm{E}[\boldsymbol{\theta}|\boldsymbol{y}])$ is the difference in the posterior mean of the deviance and the deviance evaluated at the posterior mean of the parameters. $\overline{D}(\boldsymbol{\theta})$ is a measure of the model's relative fit, and p_D provides penalty for the model's complexity. Models with smaller DIC are considered preferable. For the application in the succeeding paragraph, we apply a modified DIC recently recommended by Celeux *et al.* [23] for finite mixture models. This modified DIC, termed DIC_3, estimates $\widehat{D}(\boldsymbol{\theta}) = D(\mathrm{E}[\boldsymbol{\theta}|\boldsymbol{y}])$ by using the posterior predictive density of $\boldsymbol{y}$ and is closely related to a measure put forward by Richardson [24] to avoid overfitting the number of mixture components. We refer readers to [23] for details.

To further assess the adequacy of the final model, we use posterior predictive checks [25], in which the observed data are compared with data replicated from the posterior predictive distribution. If the model fits well, the replicated data, $\boldsymbol{y}^{\mathrm{rep}}$, should resemble the observed data $\boldsymbol{y}$. To quantify the similarity, one typically chooses a discrepancy measure, $T = T(\boldsymbol{y}, \boldsymbol{\theta})$, that takes an extreme value if the model conflicts with the observed data. Popular choices for T include sample quantiles and residual-based measures.

The Bayesian predictive p-value denotes the probability that the discrepancy measure based on the predictive sample, $T^{\mathrm{rep}} = T(\boldsymbol{y}^{\mathrm{rep}}, \boldsymbol{\theta})$, is more extreme than the observed measure T. A Monte Carlo estimate of the predictive p-value can be computed by evaluating the proportion of draws in which $T^{\mathrm{rep}} > T$. A p-value close to 0.50 represents adequate model fit, whereas p-values near 0 or 1 indicate lack of fit. The cutoff for determining lack of fit is subjective, although by analogy to the classical p-value, a Bayesian p-value between 0.05 and 0.95 suggests adequate fit. In some cases, a stricter range (such as 0.20, 0.80) may be warranted to guard against overly optimistic conclusions about model fit.

For the proposed GMM, we suggest several test statistics that evaluate the fit of the random effects model for MAP and the bivariate probit model for PTB and LBW. For the bivariate probit model, we propose as test statistics the sample proportions of PTB (T_1) and LBW (T_2), and Kendall's tau correlation between PTB and LBW (T_3). For the MAP model, we adopt the omnibus chi-square measure proposed by Gelman *et al.* [25]

$$T_4 = \sum_k \sum_{i,j} \left[\frac{(y_{ij} - \eta_{ijk})}{\sigma_k^2} \right]^2 \times \mathrm{I}_{(C_i = k)}.$$

We compute the 'observed' T_4 statistic at each MCMC iteration by treating the drawn class indicators, C_i, and random effects, $\mathbf{b}_i$, as observed data. We generate the replicated data by first sampling replicate class indicators C_i^{rep} $(i = 1, \ldots, n)$ by using expression (3); then, conditional on $C_i^{\mathrm{rep}} = k$, we generate $\mathbf{b}_i^{\mathrm{rep}}$ from $\mathrm{N}_q(\mathbf{0}, \boldsymbol{\Sigma}_k)$; finally, we sample y_{ij}^{rep} by using Equation (1).

5. Analysis of the Healthy Pregnancy, Healthy Baby Study

As part of an exploratory analysis, we plotted predicted MAP trajectories from an ordinary mixed model for 50 randomly selected women with at least three measurements (Figure 1). The model included a cubic polynomial for gestational age at the time of visit (in weeks), as well as subject-specific random intercepts and slopes. As Figure 1 indicates, the trajectories are all within the normal MAP range of 70 to 110 mmHg and broadly fall into three groups: the first is with initial MAP values ranging from 95 to 100 mmHg, with relatively flat or monotone increasing curves; the second is with starting values in the upper 80s to lower 90s and U-shape curves with precipitous third-trimester upticks; and the third is with starting values in the lower 80s and shallow U-shaped trajectories. The figure also indicates additional heterogeneity within these clusters, suggesting that a two-class to three-class GMM with cluster-specific random effects might provide optimal fit.

Next we fit the joint GMM for MAP, PTB, and LBW, allowing the number of classes K to range from one to four. Within each class, we fit a random effects model for MAP and a bivariate probit model for PTB and LBW. To fully explore the between-subject heterogeneity, we considered three models for MAP: (i) a fixed effects model with a cubic polynomial for gestational age; (ii) a model with

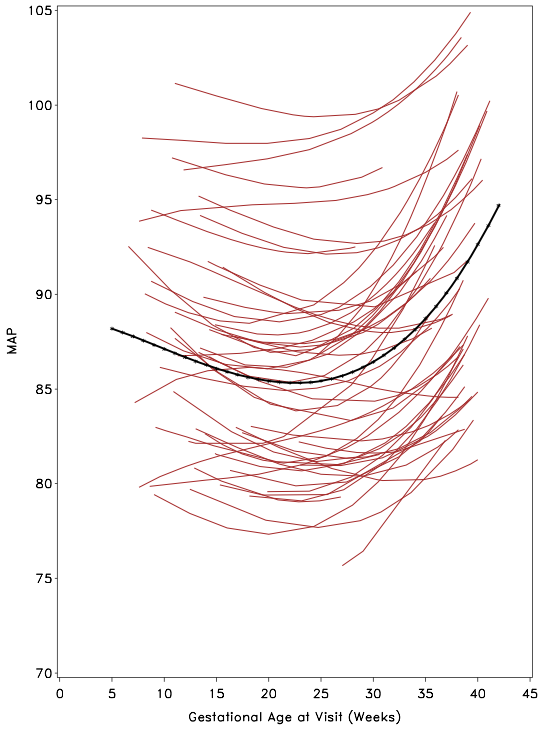

Figure 1. Predicted mean arterial pressure (MAP) curves from cubic polynomial random effects model. The starred black line denotes the average trajectory curve.

an additional subject-specific random intercept; and (iii) a model with random intercepts and slopes in addition to the cubic function. We also fit models with random effects for the quadratic and cubic polynomial terms, but the posterior estimates for the variance components were close to zero and generally failed to converge based on MCMC diagnostics. We therefore restricted our analysis to models with random intercepts and (linear) random slopes. The bivariate probit models included class-specific means and correlation parameters, as described in Equation (2). For models with $K \geqslant 2$ classes, we included several class-membership predictors ($\boldsymbol{w}_i$) in the model: maternal race (non-Hispanic black versus non-Hispanic white), maternal age (<21, 21–34, ⩾35 years), education (>high school (HS) graduate versus ⩽HS graduate), parity (primiparous versus multiparous), insurance status (private versus other), and log(serum cotinine) level.

The models were fit in R version 2.8 [26]. We assigned weakly informative (large-variance) priors for all class-specific regression parameters. For the class-membership parameters, $\boldsymbol{\gamma}_k$ $(k = 1, \ldots, K)$, we assigned $\mathrm{N}_r[\mathbf{0}, (9/4)\boldsymbol{I}_r]$ priors as described in Section 4.1. For each model, we ran two initially dispersed MCMC chains for 50,000 iterations each, discarding the first 25,000 as a burn-in. We retained every 25th draw to reduce autocorrelation. The Brooks–Gelman–Rubin 0.975 quantiles were consistently less than 1.2, indicating adequate convergence of the chains. The trace plots showed excellent mixing with no evidence of 'label switching' common to many finite mixture applications [19], suggesting that the modes of the mixture distribution were distinct enough for the classes to settle into a fixed labeling scheme within each chain. Although formal relabeling methods exist [27], these would likely have little impact on our results, and hence, we did not pursue them further. Occasionally, the class labels required reordering across different chains, but the proper labeling was easily identified in each case so that the chains could be combined for summary purposes.

We present the DIC_3 model-comparison statistics in Table II. For each class, the random slope model had the lowest DIC value; overall, the three-class random slope model provided the best fit. Among the fixed effects models, however, the four-class model had the lowest DIC. This is not surprising, because in the absence of random effects, additional classes are needed to capture the variability in the data. Interestingly, the four-class random slope model had fewer effective parameters than the two-class or three-class models, perhaps reflecting the trade-off between the latent class and the random effects (i.e., fewer random effect parameters are needed to capture heterogeneity under the four-class model).

Table II. Model comparison statistics for Healthy Pregnancy, Healthy Baby Study.

Number of classes	Model description	$\overline{D}$	p_D	DIC_3
	Fixed effects	75725	18	75743
1	Random intercept	66970	326	67296
	Random intercept and slope	65884	589	66473
	Fixed effects	71227	41	71268
2	Random intercept	66421	491	66912
	Random intercept and slope	65162	780	65942
	Fixed effects	69745	64	69809
3	Random intercept	65863	530	66393
	Random intercept and slope	**65004**	**807**	**65811**
	Fixed effects	69102	101	69203
4	Random intercept	66164	551	66715
	Random intercept and slope	65283	764	66047

Note: Bold indicates preferred model.

We present the posterior estimates and 95% CIs for the three-class random slope model in Table III. We report the coefficients for gestational age for a one-trimester change in gestational age. As a check on the functional form of gestational age, we also considered models with quadratic rather than cubic polynomials, but the DIC fits were slightly poorer for the quadratic models (results not reported), and therefore, we present only the results for the three-class cubic model with random slopes.

Figure 2 displays the mean MAP trajectories for the three-class random slope model. Class 1 comprised an estimated 28% of the sample, class 2 contained another 39%, and class 3 comprised the remaining 33%. All three trajectories were within the normal MAP range but exhibited different shapes. Class 1 had an initial MAP value of 93 mmHG, a shallow U-shaped trajectory during the first two trimesters, followed by a sharp increase during the last trimester. Class 2 was characterized by a much deeper trough that reached its nadir at the end of the second trimester. The trajectory for class 3 hovered around 86 mmHG, with a relatively flat curve.

The joint probabilities of PTB and LBW also differed by class (Table IV). Class 1 had the highest probabilities of PTB (0.20) and LBW (0.18), as well as the highest probability of the joint occurrence of both PTB and LBW (0.13). It is likely that some of these women were preeclamptic or delivered infants who were small for gestational age. Class 2 had PTB rate of 0.05, which is well below the national average of 0.12; the LBW rate for class 2 (0.09) was closer to the national average. Class 3 showed a slightly higher than average rates of PTB (0.15) and LBW (0.11).

The largest difference in the probability of PTB occurred between classes 1 and 2 ($\Pr[\text{PTB}_1 - \text{PTB}_2] = 0.15$, 95% CI $= [-0.00, 0.26]$). The difference in Pr(LBW) was also largest between classes 1 and 2, with a value of 0.09 (95% CI $= [0.01, 0.17]$). The correlation between the latent PTB and LBW variables (i.e., $\text{corr}[z^*_{1i}, z^*_{2i}]$) also varied by class: $\rho_1 = 0.84\,(0.70, 0.94)$; $\rho_2 = 0.18\,(-0.69, 0.85)$; $\rho_3 = 0.85\,(0.48, 0.99)$. Hence, even after conditioning on class membership, PTB and LBW remained highly dependent outcomes in classes 1 and 3. However, this association was significantly attenuated in class 2. Thus, given class 2 membership, knowing the probability of PTB provided relatively little additional information about the likelihood of delivering a LBW infant.

Table V presents the class-membership probabilities for two selected covariate profiles: the 'reference group' (non-Hispanic white, age <21 years, multiparous, ⩽HS education, non-private insurance, and median log[cotinine]); and a 'high-risk' group consisting of non-Hispanic black women over age 34 years, delivering their first child, and referenced on other predictors. As expected, the high-risk group was more likely to belong to class 1 and less likely to belong to classes 2 and 3 than women in the reference group. In particular, these women had a 0.26 increased probability (95% CI $= [0.10, 0.43]$) of belonging to class 1 compared with the reference group.

As a final assessment of model fit, we conducted posterior predictive checks by using the four discrepancy measures described in Section 4.3: the sample proportion of PTB (T_1), the sample proportion of LBW (T_2), Kendall's tau correlation between PTB and LBW (T_3), and the omnibus chi-square statistic for MAP (T_4). Figure 3(a)–(c) show the posterior predictive distributions of T_1, T_2 and T_3; the shaded regions correspond to the Bayesian predictive p-values of 0.46, 0.47, and 0.40, respectively, none of which suggests lack of fit. As noted in Section 4.3, the observed and posterior predicted values for

Table III. Parameter estimates and 95% CI for the three-class random slope model.

Class-specific regression parameters				
Class (%)	Outcome variable	Parameter (description)	Posterior mean	95% CI
1 (28%)	MAP	β_{11} (intercept)	93.11	(86.82, 99.61)
		β_{12} (linear GA)†	−0.26	(−11.33, 10.51)
		β_{13} (quadratic GA)	2.88	(−9.23, 3.02)
		β_{14} (cubic GA)	1.13	(0.06, 2.35)
		σ_1 (error SD)	7.05	(6.71, 7.40)
		$\mathbf{\Sigma}_{11_1}$ (Var $[b_{1i}\|C_i = 1]$)	95.05	(64.13, 132.65)
		$\mathbf{\Sigma}_{12_1}$ (Cov$[b_{1i}, b_{2i}\|C_i = 1]$)	−2.42	(−3.62, −1.49)
		$\mathbf{\Sigma}_{22_1}$ (Var$[b_{2i}\|C_i = 1]$)	0.111	(0.08, 0.15)
	PTB and LBW	μ_{11} (E$[z^*_{1i}\|C_i = 1]$)	−0.84	(−1.07, −0.62)
		μ_{21} (E$[z^*_{2i}\|C_i = 1]$)	−0.93	(−1.13, −0.71)
		ρ_1 (corr$[z^*_{1i}, z^*_{2i}\|C_i = 1]$)	0.84	(0.70, 0.94)
2 (39%)	MAP	β_{21} (intercept)	89.69	(85.23, 95.11)
		β_{22} (linear GA)	−2.72	(−11.60, 6.28)
		β_{23} (quadratic GA)	−3.31	(−8.62, 2.01)
		β_{24} (cubic GA)	1.48	(−0.51, 2.49)
		σ_2 (error SD)	4.40	(4.19, 4.60)
		$\mathbf{\Sigma}_{11_2}$ (Var$[b_{1i}\|C_i = 2]$)	47.76	(26.61, 70.88)
		$\mathbf{\Sigma}_{12_2}$ (Cov$[b_{1i}, b_{2i}\|C_i = 2]$)	−0.76	(−1.33, −0.31)
		$\mathbf{\Sigma}_{22_2}$ (Var$[b_{2i}\|C_i = 2]$)	0.04	(0.03, 0.06)
	PTB and LBW	μ_{21} (E$[z^*_{1i}\|C_i = 2]$)	−2.92	(−9.72, −0.95)
		μ_{22} (E$[z^*_{2i}\|C_i = 2]$)	−1.34	(−1.65, −1.06)
		ρ_2 (corr$[z^*_{1i}, z^*_{2i}\|C_i = 2]$)	0.18	(−0.69, 0.85)
3 (33%)	MAP	β_{31} (intercept)	86.59	(81.84, 92.14)
		β_{32} (linear GA)	−3.85	(−13.71, 5.05)
		β_{33} (quadratic GA)	1.83	(−8.62, 2.01)
		β_{34} (cubic GA)	−0.16	(−1.18, 0.70)
		σ_3 (error SD)	4.18	(3.92, 4.52)
		$\mathbf{\Sigma}_{11_3}$ (Var$[b_{1i}\|C_i = 3]$)	71.22	(49.26, 95.21)
		$\mathbf{\Sigma}_{12_3}$ (Cov$[b_{1i}, b_{2i}\|C_i = 3]$)	−1.01	(−1.56, −0.57)
		$\mathbf{\Sigma}_{22_3}$ (Var$[b_{2i}\|C_i = 3]$)	0.04	(0.03, 0.05)
	PTB and LBW	μ_{31} (E$[z^*_{1i}\|C_i = 3]$)	−1.14	(−2.45, −0.53)
		μ_{32} (E$[z^*_{2i}\|C_i = 3]$)	−1.28	(−1.58, −1.04)
		ρ_3 (corr$[z^*_{1i}, z^*_{2i}\|C_i = 3]$)	0.85	(0.48, 0.99)
Class-membership parameters				
	Class 2 versus class 1	γ_{21} (intercept)	1.28	(0.34, 2.12)
		γ_{32} (maternal race = black)	−0.62	(−1.10, −0.11)
		γ_{33} (maternal age = 21-34 years)	−0.38	(−0.82, 0.06)
		γ_{34} (maternal age ⩾ 35 years)	−0.66	(−1.42, −0.09)
		γ_{35} (maternal education > HS)	−0.15	(−0.60, 0.24)
		γ_{36} (primiparous birth)	−0.17	(−0.58, 0.23)
		γ_{27} (private insurance)	−0.04	(−0.58, 0.52)
		γ_{28} (log-serum cotinine [ng/mL])	0.00	(−0.08, 0.09)
	Class 3 versus class 1	γ_{31} (intercept)	1.08	(0.43, 1.80)
		γ_{22} (maternal race = black)	−0.45	(−0.98, −0.10)
		γ_{23} (maternal age = 21–34 years)	−0.41	(−1.02, 0.14)
		γ_{24} (maternal age ⩾ 35 years)	−0.59	(−1.33, 0.22)
		γ_{25} (maternal education > HS)	−0.16	(−0.52, 0.23)
		γ_{26} (primiparous birth)	−0.42	(−0.78, −0.06)
		γ_{37} (private insurance)	0.05	(−0.51, 0.60)
		γ_{38} (log-serum cotinine [ng/mL])	0.00	(−0.10, 0.09)

Note: GA, gestational age; HS, high school; MAP, mean arterial pressure; LBW, low birth weight; PTB, preterm birth; SD, standard deviation.

†Coefficients for GA correspond to a 13-week (i.e., one-trimester) change in GA.

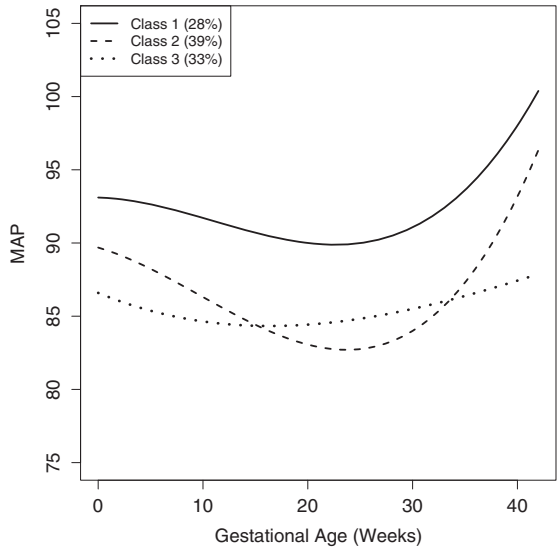

Figure 2. Posterior mean arterial pressure (MAP) trajectories (within-class membership percentages) from the three-class growth mixture model.

Table IV. Posterior probabilities (95% CIs) for preterm birth and low birth weight for the three-class random slope model.

Class	Pr(PTB = no, LBW = no)	Pr(PTB = yes, LBW = no)	Pr(PTB = no, LBW = yes)	Pr(PTB = yes, LBW = yes)
1	0.75 (0.68, 0.81)	0.07 (0.03, 0.12)	0.05 (0.02, 0.08)	0.13 (0.08, 0.18)
2	0.87 (0.77, 0.94)	0.04 (0.00, 0.12)	0.08 (0.03, 0.12)	0.01 (0.00, 0.08)
3	0.82 (0.70, 0.93)	0.07 (0.00, 0.17)	0.03 (0.00, 0.08)	0.08 (0.01, 0.14)

Note: LBW, low birth weight; PTB, preterm birth.

Table V. Predicted class-membership probabilities (95% CIs) by covariate profile.

	Class-membership probabilities (95% CI)		
Covariate profile	Class 1	Class 2	Class 3
Reference group	0.13 (0.07, 0.23)	0.48 (0.29, 0.64)	0.39 (0.24, 0.54)
High-risk group	0.39 (0.25, 0.56)	0.34 (0.20, 0.49)	0.27 (0.17, 0.40)

T_4 vary across MCMC samples (both being functions of the posterior parameter draws). Therefore, Figure 3(d) provides a two-dimensional scatterplot of the posterior predicted values based on $\boldsymbol{y}^{\text{rep}}$ versus the observed values based on $\boldsymbol{y}$. The Bayesian predictive p-value of 0.49 corresponds to the proportion of samples above the diagonal and again suggests adequate fit of the three-class random slope model.

6. Discussion

This study is among the first to examine how gestational blood pressure, considered functionally across the entire course of pregnancy, affects a range of birth outcomes. For our analysis, we developed a flexible GMM to jointly model MAP trajectories, PTB, and LBW. Within each latent class, we fit a mixed regression model for the longitudinal MAP values and a bivariate probit model for the highly dependent PTB and LBW outcomes. The model permits flexible modeling of gestational trends, class-membership probabilities, and within-class correlation structures. We adopted a Bayesian estimation approach and proposed an efficient MCMC algorithm that relies on closed-form Gibbs steps and Metropolis steps.

Our analysis revealed three latent classes: (i) a class characterized by a high initial MAP, a shallow U-shaped trajectory during early pregnancy with a sharp increase in trimester 3, and elevated risks of PTB and LBW; (ii) a class characterized by a deep U-shaped trajectory and PTB and LBW rates at or below the national averages; and (iii) a class marked by low MAP values, a generally flat trajectory, and

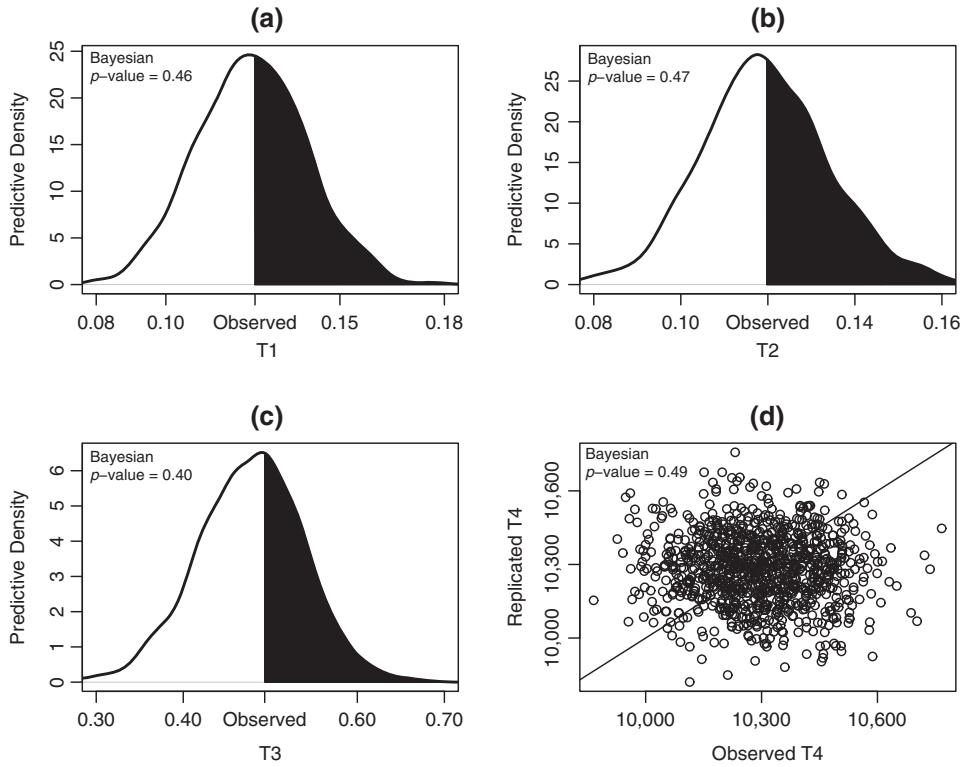

Figure 3. Results from posterior predictive checks for the three-class random slope model. (a)–(c) show the posterior distribution of the proportion PTB, LBW, and corr(PTB, LBW). (d) provides a scatterplot of predicted versus observed chi-square discrepancy measure, T_4 across MCMC samples. In (a)–(c), the Bayesian predictive p-values (0.46, 0.47, and 0.40, respectively) correspond the area in the shaded region. In (d), the Bayesian predictive p-value (0.49) represents the proportion of samples above the diagonal.

rates of PTB and LBW slightly above the national averages. Non-Hispanic black, primiparous women over age 34 years were more likely than other demographic groups to be in class 1. Because the MAP values in our sample typically fell within the normal range of 70–110 mmHg, cross-sectional snapshots or summary measures of gestational blood pressure may not have provided sufficient information to identify women at risk for adverse birth outcomes. The shape of the trajectory curve is also informative. For example, at the onset of the second trimester (about 14 weeks), classes 2 and 3 had similar average MAP values, each around 85 mmHg. However, the longitudinal trajectory for the two classes differed substantially. The deep, U-shaped trajectory typically associated with healthy pregnancy outcomes characterized class 2[2]. It is unclear why the rate of LBW is higher than the rate of PTB in this class, although one possibility is that maternal MAP alters placental perfusion, which in turn directly impacts fetal growth and hence LBW rather than PTB. Class 3, in contrast, lacked the second-trimester nadir in MAP, which could account for the higher rates of PTB and LBW compared with class 2. Class 1 likely represents women with high initial MAP values who have minimal to no mid-trimester drop in MAP and subsequently go on to have high rates of PTB and LBW. By monitoring changes in blood pressure during early pregnancy, obstetric providers may be able to identify women at risk for adverse pregnancy outcomes, as well as those who are candidates for therapeutic interventions, such as low-dose aspirin therapy, which has been shown to significantly reduce the risk of preeclampsia, stillbirth, and PTB [28]. They can also recommend more frequent prenatal visits, fetal ultrasound, and increased fetal cardiac monitoring as ways of reducing the risk of adverse outcomes such as perinatal death.

As part of an ongoing work, the HPHB investigators are exploring additional genetic and environmental factors, such as air pollution, that may affect the class-membership probabilities. As a straightforward extension, one could introduce such predictors directly into the within-class correlated probit models for PTB and LBW. This would allow for an interaction between these predictors and the latent class indicators, thus permitting the predictor effects on PTB and LBW to vary by class. For example, in some classes, air pollution might increase the probability of both PTB and LBW; for other classes, it may

increase the probability of only one of these outcomes; and for some classes, air pollution may have no impact at all. The model can also be extended to accommodate multiple longitudinal outcomes by applying the factor-analytic approaches recently described in [9, 29].

To estimate the number of latent classes, we adopted a model-comparison approach based on DIC, which has been used successfully in a number of recent studies [13, 29–31]. A practical limitation of this approach, however, is that the analysis must be repeated for different choices of K, the number of classes under consideration. In our analysis, we conducted analyses for $K = 1, \ldots, 4$. Although this approach is straightforward to implement, it fails to fully incorporate the uncertainty associated with the unknown number of classes K. One way around this is to treat K as a random variable and employ reversible jump MCMC [32], which shifts between models of different dimensions $K \in 1, \ldots, K_{max}$. The drawback here is potentially the slow mixing of the MCMC chain. An alternative strategy is to specify an infinite mixture via a Dirichlet process (DP) prior on the number of classes [33]. The DP prior avoids underfitting the number of classes (because in principle, an infinite number of classes are allowed); conversely, the DP prior shrinks back to a lower dimension if the data suggest that few classes are required, thereby avoiding overfitting. However, from a clinical standpoint, it may be difficult to assign an interpretation to more than five or six classes, even if additional classes provide slightly better fit. In such cases, it may be preferable to place a small upper bound on the number of classes, as we did in our analysis.

The GMM developed here not only allows subjects to be clustered into clinically relevant classes, it also provides flexible modeling of the longitudinal trajectories via within-class random effects. In our analysis, we imposed standard normality assumptions on these random effects. One can relax this assumption, however, by using Bayesian nonparametric methods such as DP priors to model the random effects. Moreover, if categorizing subjects into distinct trajectory classes is not of primary concern, one could pursue a more fully nonparametric functional data approach as described in [34, 35].

Nevertheless, the approach presented here should prove useful to practitioners who wish to jointly model longitudinal continuous and cross-sectional correlated binary outcomes, classify subjects into clinically interpretable clusters, and accommodate predictor effects, all within a computationally tractable Bayesian framework.

Appendix A. MCMC algorithm for growth mixture model

Here, we describe the MCMC steps for the GMM described in Section 3. Starting with initial values for model parameters, the algorithm iterates through the following steps until convergence (as determined by MCMC diagnostics):

(1) Update $\boldsymbol{\gamma}_k$: The full conditional for r-dimensional vector $\boldsymbol{\gamma}_k$ $(k = 2, \ldots, K)$ is given by

$$\begin{aligned}\pi(\boldsymbol{\gamma}_k|\cdot) &\propto \prod_{i=1}^{n} [\Pr(C_i = k|\boldsymbol{\gamma}_k)]^{\mathrm{I}(C_i=k)}\, \pi(\boldsymbol{\gamma}_k) \\ &= \prod_{i:\, C_i=k} \left(\frac{e^{\boldsymbol{w}_i'\boldsymbol{\gamma}_k}}{\sum_{h=1}^{K} e^{\boldsymbol{w}_i'\boldsymbol{\gamma}_h}} \right) \mathrm{N}_r[\boldsymbol{\gamma}_k; \mathbf{0}, (9/4)\boldsymbol{I}_r],\end{aligned}$$

where $\mathrm{N}_r(\boldsymbol{\gamma}_k; \cdot)$ is an r-dimensional normal distribution evaluated at $\boldsymbol{\gamma}_k$. Because this full conditional does not have a closed analytic form, we update $\boldsymbol{\gamma}_k$ by using a random-walk Metropolis algorithm based on a multivariate-$t_3(s_g \boldsymbol{T}_k)$ proposal density centered at the previous value, $\boldsymbol{\gamma}_k^{\text{old}}$, where the parameter s_g scales the covariance to achieve an optimal acceptance rates and $\boldsymbol{T}_k$ is a class-specific scale matrix. To improve mixing, we apply the adaptive proposal developed by Harrio *et al.* [36], which uses the empirical covariance from an extended burn-in period to tune $\boldsymbol{T}_k$ so that it emulates the true posterior covariance. One updating option is to generate separate proposals for each $\boldsymbol{\gamma}_k$ and then accept/reject all $K - 1$ vectors simultaneously in a single acceptance step. An alternative is to accept/reject each proposal in its own acceptance step or perhaps to update all $r \times (K-1)$ parameters simultaneously by using a single proposal density and acceptance step. This latter approach may require an accurate estimate of the $r(K-1) \times r(K-1)$ posterior covariance. Our experience suggests that option 1 works well and is easy to implement.

(2) Update C_i: For $i = 1, \ldots, n$, draw C_i from its full conditional

$$\pi(C_i|\cdot) = \Pr(C_i = k|\cdot) = \text{Cat}(p_{i1}, \ldots, p_{iK}), \text{ where}$$
$$p_{ik} = \frac{\pi_{ik}(\boldsymbol{w}_i)\left[\prod_{j=1}^{n_i} f(y_{ij}|\eta_{ijk}, \sigma_k^2)\right] \text{N}_2(\boldsymbol{z}_i; \boldsymbol{\mu}_k, \rho_k)\text{N}_q(\mathbf{b}_i; \mathbf{0}, \boldsymbol{\Sigma}_k)}{\sum_{h=1}^{K} \pi_{ih}(\boldsymbol{w}_i)\left[\prod_{j=1}^{n_i} f(y_{ij}|\eta_{ijh}, \sigma_h^2)\right] \text{N}_2(\boldsymbol{z}_i; \boldsymbol{\mu}_h, \rho_h)\text{N}_q(\mathbf{b}_i; \mathbf{0}, \boldsymbol{\Sigma}_h)},$$

where $\pi_{ik}(\boldsymbol{w}_i) = e^{\boldsymbol{w}_i'\boldsymbol{\gamma}_k} / \sum_{h=1}^{K} e^{\boldsymbol{w}_i'\boldsymbol{\gamma}_h}$. If there are no class-membership covariates [i.e., $r = 1$ in Step (1)], then update π_k directly from a Dirichlet$(n_1 + e_1, \ldots, n_K + e_k)$ distribution, where $e_1, \ldots, e_K$ are prior hyperparameters and $n_k = \sum_{i=1}^{n} \text{I}_{(C_i=k)}$ is the number of subject in class k.

(3) Update z_{1i}^* and z_{2i}^*: Given $C_i = k$ and current values of z_{2i}^*, μ_{2k}, and ρ_k, update z_{1i}^* from $\text{N}(E_{1k}, 1 - \rho_k^2)$ truncated below by 0 if $z_{1i} = 1$ and above by 0 if $z_{1i} = 0$, where $E_{1k} = \mu_{1k} + \rho_k(z_{2i}^* - \mu_{2k})$. Update z_{2i}^* in an analogous manner conditional on z_{1i}^*, μ_{1k}, and ρ_k.

(4) For $k = 1, \ldots, K$, update the following class-specific parameters:

(a) $\boldsymbol{\beta}_k$: Assuming a $\text{N}_p(\boldsymbol{m}_0, \boldsymbol{S}_0)$, update $\boldsymbol{\beta}_k$ from its $\text{N}_p(\boldsymbol{\mu}_{\beta_k}, \boldsymbol{V}_{\beta_k})$ full conditional, where

$$\boldsymbol{V}_{\beta_k} = \left[\boldsymbol{S}_0 + \sigma_k^{-2}(\boldsymbol{X}_k'\boldsymbol{X}_k)\right]^{-1} \text{ and}$$
$$\boldsymbol{\mu}_{\beta_k} = \boldsymbol{V}_{\beta_k}\left\{\boldsymbol{S}_0^{-1}\boldsymbol{m}_0 + \sigma_k^{-2}\boldsymbol{X}_k'\left[\boldsymbol{y}_k - \boldsymbol{V}_k\mathbf{b}_k\right]\right\}.$$

Here, $\boldsymbol{X}_k$ denotes an $N_k \times p$ fixed effects design matrix for the N_k observations in class k, $\boldsymbol{y}_k$ is an $N_k \times 1$ response vector for class k, $\boldsymbol{V}_k$ is an $N_k \times qn_k$ random effects design matrix for subjects in class k, and $\mathbf{b}_k$ is a $qn_k \times 1$ stacked vector of random effects for the n_k subjects in class k;

(b) σ_k^2: Assuming a $\text{Ga}(\lambda, \delta)$ prior for $\tau_k = \sigma_k^{-2}$, we first draw τ_k from its $\text{Ga}(l, d)$ full conditional and then set $\sigma_k^2 = 1/\tau_k$, where $l = \lambda + N_k/2$, $d = \delta + \frac{1}{2}(\boldsymbol{y}_k - \boldsymbol{X}_k\beta_k - \boldsymbol{V}_k\mathbf{b}_k)'(\boldsymbol{y}_k - \boldsymbol{X}_k\beta_k - \boldsymbol{V}_k\mathbf{b}_k)$, and $\boldsymbol{y}_k$, $\boldsymbol{X}_k$, $\boldsymbol{V}_k$ and $\mathbf{b}_k$ are defined in the previous step;

(c) $\boldsymbol{\Sigma}_k$: Assuming an $\text{IW}(\nu_0, \boldsymbol{D}_0)$ prior, update $\boldsymbol{\Sigma}_k$ from its full conditional

$$\pi(\boldsymbol{\Sigma}_k|\cdot) = \pi(\boldsymbol{\Sigma}_k|\mathbf{b}_k) = \text{IW}(\nu_0 + n_k, \boldsymbol{D}_0 + \boldsymbol{B}_k'\boldsymbol{B}_k)$$

where n_k denotes the number of subjects in class k, and $\boldsymbol{B}_k$ denotes the corresponding $n_k \times q$ matrix of random effects;

(d) $\boldsymbol{\mu}_k$: Assuming a $\text{N}_2(\boldsymbol{m}_\mu, \boldsymbol{S}_\mu)$ prior, update $\boldsymbol{\mu}_k$ from its $\text{N}_2(\boldsymbol{E}_{\mu_k}, \boldsymbol{V}_{\mu_k})$ full conditional, where

$$\boldsymbol{V}_{\mu_k} = \left[\boldsymbol{S}_\mu + (\boldsymbol{U}_k'\boldsymbol{\Omega}_k\boldsymbol{U}_k)\right]^{-1} \text{ and}$$
$$\boldsymbol{E}_{\mu_k} = \boldsymbol{V}_{\mu_k}\left(\boldsymbol{S}_\mu^{-1}\boldsymbol{m}_\mu + \boldsymbol{U}_k'\boldsymbol{\Omega}_k z^*\right).$$

Here, $\boldsymbol{U}_k$ is a $2n_k \times 2$ matrix with $(1, 0)$ and $(0, 1)$ alternating by row, $\boldsymbol{\Omega}_k$ is a $2n_k \times 2n_k$ block diagonal matrix with each 2×2 block equal to $\begin{pmatrix} 1 & \rho_k \\ \rho_k & 1 \end{pmatrix}^{-1}$, and z^* is a $2n_k \times 1$ stacked vector of alternating z_1^* and z_2^* values for the n_k subjects in class k;

(e) ρ_k: Assuming a uniform prior on $(-1, 1)$, update ρ_k by using a random-walk Metropolis step centered at the previous value of ρ_k and restricted to the interval $(-1, 1)$. For our analysis, we used a truncated normal proposal restricted to $(-1, 1)$.

(5) Update $\mathbf{b}_i$: Because of conjugacy, the $q \times 1$ random effect vector $\mathbf{b}_i$ has a closed-form full conditional so that a Gibbs step is one updating option. In general, we have found a Metropolis step to be computationally faster, although its higher autocorrelation requires somewhat longer run times. For our application, we updated $\mathbf{b}_i$ by using a random-walk Metropolis algorithm with a bivariate $t_3(s_b\boldsymbol{\Sigma}_k)$ proposal centered at the previous value, $\mathbf{b}_i^{\text{old}}$, and s_b denotes a scaling factor used to achieve optimal acceptance rates.

Acknowledgements

This work was supported by grant number RD-83329301-4 from the US Environmental Protection Agency, Research Triangle Park, North Carolina, and was conducted in accordance with a human subjects research protocol approved by the Duke University's Institutional Review Board.

References

1. Allen VM, Joseph KS, Murphy KE, Magee LA, Ohlsson A. The effect of hypertensive disorders in pregnancy on small for gestational age and stillbirth: a population based study. *BMC Pregnancy Childbirth* 2004; **4**(17):1–8.
2. Hermida RC, Ayala DE, Iglesias M. Predictable blood pressure variability in healthy and complicated pregnancies. *Hypertension* 2001; **38**:736–741.
3. Miranda ML, Swamy GK, Edwards S, Maxson P, Gelfand A, James S. Disparities in maternal hypertension and pregnancy outcomes: evidence from North Carolina, 1994–2003. *Public Health Reports* 2010; **125**:579–587.
4. American Congress of Obstetricians and Gynecologists. Diagnosis and management of preeclampsia and eclampsia. *Obstetrics and Gynecology* 2002; **99**:159–167.
5. Verbeke G, Lesaffre E. A linear mixed-effects model with heterogeneity in the random-effects population. *Journal of the American Statistical Association* 1996; **91**:217–221.
6. Muthén B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* 1999; **55**:463–469.
7. Lin H, McCulloch CE, Turnbull BW, Slate EH, Clark LC. A latent class mixed model for analyzing biomarker trajectories with irregularly scheduled observations. *Statistics in Medicine* 2000; **19**:1303–1318.
8. Muthén B, Brown CH, Booil Jo KM, Khoo S-T, Yang C-C, Wang C-P, Kellam SG, Carlin JB, Liao J. General growth mixture modeling for randomized preventive interventions. *Biostatistics* 2002; **3**:459–475.
9. Proust-Lima C, Letenneur L, Jacqmin-Gaddam H. A nonlinear latent class model for joint analysis of multivariate longitudinal data and a binary outcome. *Statistics in Medicine* 2007; **26**:2229–2245. DOI: 10.1002/sim.2659.
10. Lin H, McCulloch C E, Turnbull B W, Slate E H. Latent class models for joint analysis of longitudinal biomarker and event process data: application to longitudinal prostate-specific antigen readings and prostate cancer. *Journal of the American Statistical Association* 2002; **97**:53–65.
11. Proust-Lima C, Joly P, Dartigues J-F, Jacqmin-Gaddam H. Joint modelling of multivariate longitudinal outcomes and a time-to-event: a nonlinear latent class approach. *Computational Statistics & Data Analysis* 2009; **53**:1142–1154.
12. Zhang J J, Wang M. Latent class joint model of ovarian function suppression and DFS for premenopausal breast cancer patients. *Statistics in Medicine* 2010; **29**:2310–2324. DOI: 10.1002/sim.3977.
13. Elliott MR, Gallo JJ, Ten Have TR, Bogner HR, Katz IR. Using a Bayesian latent growth curve model to identify trajectories of positive affect and negative events following myocardial infarction. *Biostatistics* 2005; **6**:119–143.
14. Slaughter JC, Herring AH, Thorp JM. A Bayesian latent variable mixture model for longitudinal fetal growth. *Biometrics* 2009; **65**:1233–1242.
15. Gueorguieva RV, Agresti A. A correlated probit model for joint modeling of clustered binary and continuous responses. *Journal of the American Statistical Association* 2001; **96**:1102–1112.
16. Hamilton BE, Martin JA, Ventura SJ. Births: preliminary data for 2008. *National Vital Statistic Reports* 2010; **58**:1–17.
17. Rossi PE, Allenby GE, McCullough R. *Bayesian Statistics in Marketing*. John Wiley & Sons: Chichester, 2005.
18. Garrett ES, Zeger SL. Latent class model diagnosis. *Biometrics* 2000; **56**:1055–1067.
19. Frühwirth-Schnatter S. *Finite Mixture and Markov Switching Models*. Springer: New York, 2006.
20. Gelman A, Carlin J B, Stern H S, Rubin D B. *Bayesian Data Analysis* (2nd edn). Chapman & Hall: Boca Raton, 2004.
21. Smith BJ. boa: An R package for MCMC output convergence assessment and posterior inference. *Journal of Statistical Software* 2007; **21**:1–37.
22. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B* 2002; **64**:583–639.
23. Celeux G, Forbes F, Robert CP, Titterington DM. Deviance information criteria for missing data models. *Bayesian Analysis* 2006; **1**:651–674.
24. Richardson S. Discussion of Spiegelhalter *et al*. *Journal of the Royal Statistical Society, Series B* 2002; **64**:626–627.
25. Gelman A, Meng XL, Stern HS. Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica* 1996; **6**:733–807.
26. R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, 2008. ISBN 3-900051-07-0. http://www.R-project.org.
27. Stephens M. Dealing with label switching in mixture models. *Journal of the Royal Statistical Society, Series B* 2000; **62**:795–809.
28. Duley L, Henderson-Smart D, Knight M, King J. Antiplatelet drugs for prevention of pre-eclampsia and its consequences: systematic review. *British Medical Journal* 2001; **322**:329–333.
29. Leiby BE, Sammel MD, Ten Have TR, Lynch KG. Identification of multivariate responders and non-responders by using Bayesian growth curve latent class models. *Journal of the Royal Statistical Society, Series C* 2009; **58**:505–524.
30. White JW, Standish JD, Thorrold SR, Warner RR. Markov chain Monte Carlo methods for assigning larvae natal sites using natural geochemical tags. *Ecological Applications* 2008; **18**:1901–1913.
31. Neelon B, O'Malley AJ, Normand S-LT. A Bayesian two-part latent class model for longitudinal medical expenditure data: assessing the impact of mental health and substance abuse parity. *Biometrics* 2011; **67**:280–289. DOI: 10.1111/j.1541-0420.2010.01439.x.

32. Green PJ. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 1995; **82**:711–732.
33. Ferguson TS. A Bayesian analysis of some nonparametric problems. *Annals of Statistics* 1973; **1**:209–30.
34. Crainiceanu CM, Goldsmith AJ. Bayesian functional data analysis using WinBUGS. *Journal of Statistical Software* 2010; **32**:1–33.
35. MacLehose RF, Dunson DB. Nonparametric Bayes kernel-based priors for functional data analysis. *Statistica Sinica* 2009; **19**:611–629.
36. Harrio H, Saksman E, Tamminen J. Componentwise adaptation for high dimensional MCMC. *Computational Statistics* 2005; **20**:165–273.