Published in final edited form as:

Stat Methods Med Res. 2016 April ; 25(2): 520–537. doi:10.1177/0962280212460444.

Modeling Clinical Outcome Using Multiple Correlated Functional Biomarkers: A Bayesian Approach

Qi Long*, Xiaoxi Zhang[†], Yize Zhao*, Brent A. Johnson*, and Roberd M. Bostick[‡]

*Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA 30322.

[†]Pfizer Inc., New York, NY 10017.

[‡]Department of Epidemiology, Emory University, Atlanta, GA 30322

Abstract

In some biomedical studies, biomarkers are measured repeatedly along some spatial structure or over time and are subject to measurement error. In these studies, it is often of interest to evaluate associations between a clinical endpoint and these biomarkers (also known as functional biomarkers). There are potentially two levels of correlation in such data, namely, between repeated measurements of a biomarker from the same subject and between multiple biomarkers from the same subject; none of the existing methods accounts for correlation between multiple functional biomarkers. We propose a Bayesian approach to model a clinical outcome of interest (e.g., risk for colorectal cancer) in the presence of multiple functional biomarkers while accounting for potential correlation. Our simulations show that the proposed approach achieves good performance in finite samples under various settings. In the presence of substantial or moderate correlation, the proposed approach outperforms an existing approach that does not account for correlation. The proposed approach is applied to a study of biomarkers of risk for colorectal neoplasms and our results show that the risk for colorectal cancer is associated with two functional biomarkers, APC and TGF- α , in particular, with their values in the region between the proliferating zone and the differentiating zone of colorectal crypts.

Keywords

Bayesian models; Correlated data; Functional biomarker; Measurement error

1 Introduction

In some biomedical studies, researchers are interested in evaluating associations between a real-valued scalar clinical outcome, e.g., risk for colorectal cancer, and biomarkers that are repeatedly measured and are subject to measurement error. In contrast to standard scalar biomarkers, we refer to these repeatedly measured biomarkers as functional biomarkers, or functional predictors, which are used interchangeably in this paper.

In practice, functional biomarkers are often measured along certain spatial structures. One such example is a recent study of biomarkers of risk for colorectal cancer, which motivates our current work. The Markers of Adenomatous Polyps II study (MAP II)^{1,2} was a pilot case-control study to validate a panel of plausible protein biomarkers of risk for colorectal

cancer^{1,2}, which describe molecular phenotypes of the normal-appearing colorectal epithelium. These biomarkers represent highlights of features that are related to the known molecular basis of the earliest stages of colorectal carcinogenesis. This study recruited adult subjects who had no history of previous colorectal adenoma or malignancy of any type and were scheduled for elective outpatient colonoscopy. Participants who were found to have one or more adenomas – immediate precursors to colorectal cancer – were considered “cases” (at higher risk for colorectal cancer), and participants with no adenoma were considered “controls” (at lower risk for colorectal cancer). In the MAP II study, it is of particular interest to examine the protein biomarker level along the length of a U-shaped microscopic structure in the human colon, the colorectal crypt; cf. Figure 1 in Daneial et al.¹ It is well known that the colorectal crypt is an elegant model of the regulation of cell proliferation, differentiation, and apoptosis in a continuously renewing epithelium. In this study, biopsy samples from patients were processed immunohistochemically (i.e., stained), and each sample often contains multiple crypts. Using the stained samples, the protein biomarker distribution along each crypt was constructed from the stain density data. Each crypt was divided into two symmetric hemi-crypts, and the entire length of each hemi-crypt was then standardized into 50 segments which are numbered in an ascending order from the base to the top of a crypt, and then the mean staining density was plotted across the crypt against the segment location; cf. Figures 2 and 3 in Daniel et al.¹ One primary goal of the MAP II study is to assess the association between risk for colorectal cancer and the functional markers, and eventually to build risk profiles using these functional biomarkers. Evidently, the biomarker distributions along the length of the colon crypts form a natural one-dimensional spatial data structure, and hence are examples of functional biomarkers measured over space. In this motivating example, functional biomarkers were measured at discrete design points and were subject to measurement error³, which is usually the case in practice. In addition, each functional biomarker is repeatedly measured long the length of crypts for each subject, therefore these measurements are expected to be correlated. In the presence of multiple functional biomarkers, it is further expected that the measurements for different functional biomarkers are also correlated when measured from the same subject.

In other settings, functional biomarkers can also be measured repeatedly over time. One such example is a randomized trial of D-penicillamine in patients with primary biliary cirrhosis (PBC) which was conducted by the Mayo Clinic⁴, where it was of interest to estimate the association between the survival outcome and biomarker levels that were measured over time. The methods proposed in the current paper are also applicable in these settings.

Simple approaches can be adopted to model a clinical outcome of interest using functional biomarkers. In the MAP II study^{1,2}, some ad-hoc summary measures of the functional biomarkers were first created including the mean over all segments, the means over portions of crypts and certain ratios of these means, and then used in regression models to evaluate their associations with risk for colorectal cancer. Admittedly, these summary measures are often created based on scientific evidence and this approach is appealing due to its ease of implementation; however, it has some limitations. In particular, the summary measures are not generated through data-driven approaches and may not capture the features of functional biomarkers that are truly associated with the outcome. As another simple approach, one can

simply treat the observed value of a functional biomarker at each segment t as an independent predictor for the MAP II study. Then one can proceed to use a logistic regression model with 50 predictors that correspond to the 50 segments. Apparently, this approach ignores the inherent functional data structure and is not directly applicable in the case of an unbalanced design, i.e., the biomarker measurements are taken at different locations (or time points) across subjects. In this article, we do not entertain these approaches as viable alternatives.

In recent years, more sophisticated approaches have been proposed, in particular, generalized functional linear models (GFLM, for short) have drawn considerable interest^{5–11}. In GFLM, functional predictors are modeled using nonparametric regression techniques such as splines, and then they are related to the clinical outcome of interest through a model that is similar to generalized linear models¹². Existing works on GFLM primarily focus on developing estimators that are obtained ignoring correlation among repeated measures data and on the optimal convergence rates of estimators rather than making statistical inference. Just as important, none of the existing methods accounts for correlation between multiple functional biomarkers that are measured from the same subject. Besides, most of the existing works are restricted to the case of no measurement error. While these estimators remain consistent in the presence of correlation, they are potentially less efficient compared to an estimator that accounts for such correlation.

For the problem of our interest, we have recently proposed a semi-parametric approach to fit GFLM¹¹, which avoids some strong parametric assumptions and is shown to outperform the maximum likelihood approach proposed by James⁹ when these assumptions are violated in the data. However, our subsequent research reveals that this semi-parametric approach encounters the issue of multiple roots in the presence of multiple biomarkers, which presents a challenge of choosing the consistent root, a well-known problem in the conditional score literature; furthermore, it is difficult to extend the semi-parametric method to account for correlation between multiple functional biomarkers that are measured from the same subject. In addition, it is noteworthy that Long et al.¹¹ show that the “so-called” two-step method, which obtains an estimated curve for each functional biomarker first and then fits GFLM using the estimated functional biomarker values, is outperformed by the maximum likelihood approach under various settings; hence we do not entertain this two-step method in the current paper.

Of note, colon crypt data similar to those in the MAP II study have been analyzed by Morris et al.¹³, Apanasovich et al.¹⁴, Li et al.¹⁵, and Baladandayuthapani et al.¹⁶ among others. These studies, however, are fundamentally different from ours in that their primary goal was to model and estimate the biomarker functions along the length of crypts and/or test correlations among these biomarker functions, and crypt-level biomarker values are used as the outcome of interest. In our model of interest, the outcome is a clinical endpoint and the functional biomarkers are treated as predictors. Consequently, we adopt the framework of GFLM and use subject-level biomarker data.

In this article, we investigate a Bayesian joint modeling approach to overcome some limitations of existing methods. The proposed approach models the clinical outcome of

interest and the functional biomarkers jointly and accounts for correlations among repeated measurements of a same biomarker and correlations among multiple biomarkers measured from the same subject. The remainder of this article is organized as follows. In Section 2, we first introduce the model for a single functional biomarker and then extend it to the case of multiple functional predictors, followed by the details of Bayesian inference. In Section 3, we conduct simulation studies to evaluate the finite sample performance of the proposed Bayesian method under several settings. In Section 4, we apply the proposed Bayesian method to the MAP II colorectal cancer study. We conclude this article with some discussion remarks in Section 5.

2 Methodology

2.1 The Case of a Single Functional Biomarker

We first consider the case of a single functional biomarker. Suppose that a functional biomarker, $Z(\cdot)$, is measured over time or space. Given $Z(\cdot)$, we assume that the scalar outcome of interest, namely, y , follows a distribution from an exponential family with density:

$$\exp \{ [y\theta(x, Z(\cdot)) - b(\theta(x, Z(\cdot)))] / a(\phi) + c(y, \phi) \}, \quad (1)$$

where θ is the natural parameter, ϕ is a dispersion parameter, and $a(\cdot)$, $b(\cdot)$ and $c(\cdot)$ are real-valued smooth functions. We further assume that $\mu = E(Y)$ is related to the functional predictor, $Z(\cdot)$, and other scalar predictors, $\mathbf{x}$, through:

$$g(\mu) = \alpha_0 + \alpha_1^T \mathbf{x} + \int Z(t) \beta(t) dt, \quad (2)$$

where $g(\cdot)$ is a monotonic real-valued link function, $\mathbf{x}$ is a set of scalar predictors, and α_1 is a vector of parameters associated with $\mathbf{x}$. $\beta(\cdot)$ can be considered as a weight function and is analogous to regression coefficient, β , in a generalized linear model¹². The objective is to estimate $\beta(\cdot)$ in Equation (2).

For each subject i ($i = 1, \dots, n$), $Z_i(t)$ is modeled as a smooth curve from a given functional family, for example, cubic splines. The resulting parameterization for Z_i is $Z_i(t) = \mathbf{s}(t)^T \boldsymbol{\gamma}_i$, where $\mathbf{s}(t) = (s_1(t), \dots, s_q(t))^T$ represent the q -dimensional basis functions evaluated at t and $\boldsymbol{\gamma}_i = (\gamma_{i1}, \dots, \gamma_{iq})^T$ are the q -dimensional unknown coefficients associated with $Z_i(\cdot)$.

Combining this with Equation (2), it follows that $g(\mu_i) = \alpha_0 + \alpha_1^T \mathbf{x}_i + \boldsymbol{\eta}^T \boldsymbol{\gamma}_i$, where $\boldsymbol{\eta} = \int \beta(t) \mathbf{s}(t) dt$. To ensure identifiability of $\beta(t)$, $\beta(t)$ is assumed to belong to the same function space as $Z_i(t)$, i.e., the space spanned by $\mathbf{s}(t)$; in other words, $\beta(t) = \mathbf{s}(t)^T \boldsymbol{\nu}$. It follows that $\boldsymbol{\eta} = (\int \mathbf{s}(t) \mathbf{s}(t)^T dt) \boldsymbol{\nu}$, and $\boldsymbol{\nu}$ and hence $\beta(\cdot)$ can be computed given $\boldsymbol{\eta}$, when $\mathbf{s}(t)$ is a set of orthogonal basis functions, $\boldsymbol{\eta} = \boldsymbol{\nu}$. In other words, estimating $\beta(\cdot)$ is equivalent to estimating $\boldsymbol{\eta}$ and this fact will be used throughout the remainder of the article.

The functional biomarkers are measured only at discrete time points or locations and are subject to measurement error, that is, the observed biomarker values are

$$\begin{aligned} z_i(t_{ik}) &= Z_i(t_{ik}) + \epsilon_{ik} \\ &= \mathbf{s}(t_{ik})^T \boldsymbol{\gamma}_i + \epsilon_{ik} \end{aligned}$$

at m_i design points $(t_{i1}, \dots, t_{im_i})$, where $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \dots, \epsilon_{im_i})^T \sim N(0, \Omega_i)$. For subject i , we let y_i denote the response of interest, $\mathbf{z}_i$ denote the m_i -vector of biomarker measurements, and $\boldsymbol{\epsilon}_i$ denote the measurement error. Furthermore, let $\mathbf{S}_i = (\mathbf{s}(t_{i1}), \dots, \mathbf{s}(t_{im_i}))$ ($q \times m_i$) be the corresponding spline basis matrix for subject i evaluated at m_i design points. Then the model of interest can be summarized as

$$\begin{aligned} f(y_i; \theta_i, \phi) &= \exp \left\{ \frac{y_i \theta_i - b(\theta_i)}{a(\phi)} + c(y, \phi) \right\}, \\ g(\mu_i) &= \alpha_0 + \boldsymbol{\alpha}_1^T \mathbf{x}_i - \boldsymbol{\eta}^T \boldsymbol{\gamma}_i, \\ \mathbf{z}_i &= \mathbf{S}_i^T \boldsymbol{\gamma}_i + \boldsymbol{\epsilon}_i, \end{aligned}$$

where $\boldsymbol{\eta} = \int \beta(t) \mathbf{s}(t) dt$. It is also assumed that $\mathbf{z}_i$ and y_i are independent given $\boldsymbol{\gamma}_i$ and $\boldsymbol{\epsilon}_i$ is independent of y_i . Note that Ω_i is not assumed to be diagonal; in other words, measurements at two different design points from the same subject, say z_{ik1} and z_{ik2} , can be correlated. Subsequently, the joint likelihood of the observed data can be written as follows:

$$l(\phi, \boldsymbol{\eta}, \boldsymbol{\gamma}_i; \mathbf{y}_i, \mathbf{z}_i) = \exp \left\{ \frac{y_i \theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right\} \times \frac{1}{(2\pi)^{m_i/2} \sqrt{|\Omega_i|}} \exp \left\{ -\frac{1}{2} (\mathbf{z}_i - \mathbf{S}_i^T \boldsymbol{\gamma}_i)^T \Omega_i^{-1} (\mathbf{z}_i - \mathbf{S}_i^T \boldsymbol{\gamma}_i) \right\},$$

For simplicity, we use the canonical link function, i.e., $\theta_i = g(\mu_i) = \alpha_0 + \boldsymbol{\alpha}_1^T \mathbf{x}_i + \boldsymbol{\eta}^T \boldsymbol{\gamma}_i$ throughout this article. The extensions to other link functions are straightforward.

2.2 Extension to Multiple Functional Biomarkers

We now extend the model described in Section 2.1 to the case of multiple biomarkers. Suppose that for subject i , p functional biomarkers indexed by j , namely, Z_{ij} ($j = 1, 2, \dots, p$) are collected at the same set of design points $(\{t_{ik}, k = 1, \dots, m_i\})$, and these biomarkers may be measured at different design points for different subjects. Z_{ij} 's are subject to measurement error, that is, the observed functional biomarker values are

$$z_{ijk} = Z_{ij}(t_{ik}) = Z_{ij}(t_{ik}) + \tilde{\epsilon}_{ijk},$$

where $\tilde{\epsilon}_{ijk}$'s are potentially correlated. $\tilde{\epsilon}_{ijk}$ captures two levels of correlation and can be partitioned into two components, namely, $\tilde{\epsilon}_{ijk} = b_i + \epsilon_{ijk}$, where $\boldsymbol{\epsilon}_{ij} = (\epsilon_{ij1}, \dots, \epsilon_{ijm_i})^T \sim N(0, \Omega_{ij})$ and $b_i \sim N(0, \sigma_b^2)$. $\boldsymbol{\epsilon}_{ij}$ is not assumed to be diagonal and characterizes correlation between repeated measurements of biomarker j from subject i , and b_i is a subject level random effect and characterizes additional correlation between measurements of different biomarkers from subject i . $\boldsymbol{\epsilon}_{ij}$'s and b_i 's are assumed to be independent of each other. Along similar lines in Section 2.1, $Z_{ij}(t)$ is modeled as a smooth curve from a given functional family, that is, $Z_{ij}(t) = \mathbf{s}(t)^T \boldsymbol{\gamma}_{ij}$. Then, the observed biomarker measurements can be rewritten as

$$z_{ijk} = s(t_{ik})^T \gamma_{ij} + b_i + \epsilon_{ijk}. \quad (3)$$

Similar to as in Section 2.1, to ensure identifiability of $\beta_j(t)$, $\beta_j(t)$ is assumed to belong to the same function space as $Z_{ij}(t)$, i.e., the space spanned by $s(t)$; in other words, $\beta_j(t) = s(t)^T \mathbf{v}_j$.

It follows that Equation (2), which relates $\mu_i = E(Y_i)$ to the functional predictor $Z_{ij}(\cdot)$ and other scalar predictors $\mathbf{x}_i$, can be rewritten as

$$\begin{aligned} g(\mu_i) &= \alpha_0 + \boldsymbol{\alpha}_1^T \mathbf{x}_i + \sum_{j=1}^p \int Z_{ij}(t) \beta_j(t) dt \\ &= \alpha_0 + \boldsymbol{\alpha}_1^T \mathbf{x}_i + \sum_{j=1}^p \boldsymbol{\eta}_j^T \boldsymbol{\gamma}_{ij}, \end{aligned} \quad (4)$$

where $\boldsymbol{\eta}_j = \int \beta_j(t) s(t) dt = \int s(t) s(t)^T dt \mathbf{v}_j$ and $\boldsymbol{\gamma}_{ij} = (\gamma_{ij1}, \gamma_{ij2}, \dots, \gamma_{ijq})^T$ are the set of parameters associated with the q basis functions for the j th biomarker of the i th individual. It follows that the complete model for multiple functional biomarkers can be summarized as follows

$$f(y_i; \theta_i, \phi) = \exp \left\{ \frac{y_i \theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right\}, \quad (5)$$

$$g(\mu_i) = \alpha_0 + \boldsymbol{\alpha}_1^T \mathbf{x}_i + \sum_{j=1}^p \boldsymbol{\eta}_j^T \boldsymbol{\gamma}_{ij}, \quad (6)$$

$$\mathbf{z}_i = \mathbf{S}_i^* \boldsymbol{\gamma}_i + \mathbf{W}_i b_i + \epsilon_i, \quad (7)$$

where $\mathbf{z}_i = (\mathbf{z}_{i1}^T, \dots, \mathbf{z}_{ip}^T)^T$ ($pm_i \times 1$) is the collection of all biomarker measurements for subject i with $\mathbf{z}_{ij} = (z_{ij1}, \dots, z_{ijm_j})^T$ being the collection of m_j repeated measurements of biomarker j for i th subject, $\boldsymbol{\gamma}_i = (\gamma_{i1}, \dots, \gamma_{ip})^T$ ($pq \times 1$),

$$\mathbf{S}_i^* = \text{diag} \left(\overbrace{\mathbf{S}_i^T, \mathbf{S}_i^T, \dots, \mathbf{S}_i^T}^p \right) \quad (pm_i \times pq) \quad \text{with } \mathbf{S}_i = [s(t_{i1}), \dots, s(t_{im_j})]^T (m_j \times q), \mathbf{W}_i =$$

$\mathbf{1}_{pm_j}$ and $\epsilon_i = (\epsilon_{i1}, \dots, \epsilon_{ip})^T$ ($pm_i \times 1$) with a variance

$\Omega_i^* = \text{diag}(\Omega_{i1}, \dots, \Omega_{ip})$ ($pm_i \times pm_i$). Generically, let $\text{diag}(B \times p)$ denote a block diagonal matrix with B in the diagonal p times and $\mathbf{1}_p$ denote a p -vector of 1.

Since $\beta_j(t)$ is determined by $\mathbf{v}_j$ which in turn is determined by $\boldsymbol{\eta}_j$, estimating $\beta_j(t)$ is

equivalent to estimating $\boldsymbol{\eta}_j$. Consequently, let $\boldsymbol{\xi} = (\alpha_0, \boldsymbol{\alpha}_1^T, \boldsymbol{\eta}_1^T, \boldsymbol{\eta}_2^T, \dots, \boldsymbol{\eta}_p^T)^T$, denote the set of parameters of primary interest. The joint likelihood of the data can be written as

$$\begin{aligned} \ell(\xi, \phi, \gamma_i, \Omega_{ij}, \mathbf{b}_i, \sigma_b^2; \mathbf{y}, \mathbf{z}) &\propto \prod_{i=1}^n \left[\exp \left\{ \frac{y_i \theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right\} \times \prod_{j=1}^p |\Omega_{ij}|^{-\frac{1}{2}} \exp \right. \\ &\quad \left. \left\{ -\frac{1}{2} (\mathbf{z}_{ij} - \mathbf{S}_i^T \gamma_{ij} - \mathbf{1}_{m_i} b_i)^T \Omega_{ij}^{-1} (\mathbf{z}_{ij} - \mathbf{S}_i^T \gamma_{ij} - \mathbf{1}_{m_i} b_i) \right\} \times (\sigma_b^2)^{-\frac{1}{2}} \exp \left\{ -\frac{b_i^2}{2\sigma_b^2} \right\} \right], \end{aligned} \quad (8)$$

where θ_i 's are functions of ξ .

2.3 Bayesian Inference

We now specify the prior distributions for the parameters involved in the joint likelihood (8). In all cases, we choose conjugate prior distributions when possible; when conjugate priors are not available, we choose prior distributions as general as possible while still being proper. The prior distributions for the parameters of interest are specified as follows

$$\begin{aligned} \gamma_{ij} &\sim N(\mu_j, \Gamma_j) \quad (j=1, \dots, p), \\ \xi &\sim N(\mu_\xi, \Sigma_\xi), \\ \sigma_b^2 &\sim \text{Inv} - \chi^2(\nu_b, s_b^2). \end{aligned}$$

We further specify priors on the hyperparameters (μ_j, Γ_j) as follows

$$\begin{aligned} \Gamma_j^{-1} &\sim \text{Wishart}(\mathbf{S}_\Gamma, \nu_\Gamma), \\ \mu_j | \Gamma_j &\sim N(\mu_0, \Sigma_0). \end{aligned}$$

To complete the specification of the Bayesian model, we also need to specify the prior distributions of ϕ and Ω_{ij} in the joint likelihood (8). When ϕ and ξ are distinct, such as in the case of y being Gaussian, we propose to use either a flat prior distribution or a conjugate prior distribution for ϕ . For Ω_{ij} , we choose the prior distribution based on the data structure. For a balanced design, i.e., $m_i = m$, one may assume that $\Omega_{ij} = \Omega_j$, i.e., the covariance matrix of ε_{ij} is the same for different subjects. Subsequently, we can specify a conjugate prior distribution for Ω_j , i.e., $\Omega_j^{-1} \sim \text{Wishart}(\mathbf{S}_\Omega, \nu_\Omega)$ ($j = 1, \dots, p$). For an unbalanced design, i.e., m_i is different for different subjects, we assume that Ω_{ij} follows an AR(1) structure, i.e., the (k_1, k_2) 'th element of Ω_{ij} is $\phi_j \rho_j^{|t_{ik_1} - t_{ik_2}|}$, where $\phi_j (\phi_j > 0)$ is the biomarker specific variance and $\rho_j (|\rho_j| < 1)$ is the biomarker specific correlation coefficient. Subsequently, we can either assume flat prior distributions for ϕ_j 's and ρ_j 's, or alternatively impose a hierarchical structure on ϕ_j 's and ρ_j 's, both of which are straightforward to implement. In this article, we adopt the first approach for the prior distributions of ϕ_j 's and ρ_j 's. Nevertheless, other specifications of Ω_{ij} that avoid over-parameterization can also be reasonable choices.

We adopt Markov Chain Monte Carlo algorithms to approximate the posterior distributions of the parameters of interest, in particular, ξ . Specifically, we use the Gibbs sampler to draw samples from the posterior distribution. When the outcome is not Gaussian or the data are

not balanced, some parameters in the likelihood (8) do not have conjugate priors. For example, when the data are not balanced and an AR(1) structure is assumed for the correlation matrix, the parameters, ρ_j 's, do not have conjugate priors. Consequently, an adaptive Metropolis rejection sampling (ARMS)¹⁷ algorithm is used to obtain the corresponding full conditionals. The full conditionals of the proposed model are provided in the Appendix for several special cases, which include Gaussian and binary outcomes, and balanced and unbalanced data. The numerical algorithms for generating the posterior distributions are implemented in R.

After obtaining the posterior distribution of ξ , in particular, η_j ($j = 1, \dots, p$), one can construct the posterior predictive distribution of $\beta_f(\cdot)$ ($j = 1, \dots, p$) using the relationship between β_j and η_j noted before, e.g, if $s(\cdot)$ is a set of ortho-normal basis functions,

$\beta_j(\cdot) = \eta_j^T s(\cdot)$. Then for a set of I pre-specified design points, say $\mathbf{t} = (t_1, \dots, t_I)$, the posterior predictive distribution of $\beta_f(\mathbf{t})$ can be approximated by plugging in the samples of the posterior distribution for η_j .

3 Simulation Studies

We conduct simulation studies to evaluate the performance of the Bayesian models, and we examine the impact of the sample size, namely, the number of subjects (n) and the number of repeated biomarker measurements (m_i) as well as the magnitude of correlation. We consider two types of outcome variables (y), Gaussian and binary. We also investigate two types of data structure, balanced and unbalanced. In all setups, we consider models that involve two ($p = 2$) functional biomarkers. Our simulation results are summarized over 1000 Monte Carlo data sets using the following measures: the bias of the estimated ξ based on the posterior mean (bias), the square root of the mean squared error (SMSE), the mean of the posterior standard deviation for the Bayesian method (pSD), the Monte Carlo standard deviation of the posterior means (SD), and the coverage rates of the 95% posterior credible intervals (CR).

We first consider the case of Gaussian outcomes. The true weight function ($\beta_f(\cdot)$, $j = 1, 2$) and functional biomarkers ($Z_{ij}(\cdot)$, $j = 1, 2$) are assumed to belong to the functional space spanned by the set of first q Legendre orthogonal polynomial basis functions ($q = 6$); for $Z_{ij}(\cdot)$ ($j = 1, 2$), γ_{j1} and γ_{j2} are sampled from a multivariate normal distribution with a mean of μ_1 and μ_2 , respectively, and a variance-covariance matrix of $\mathbf{I}_q$. In the case that y follows Gaussian distributions, we set $\mu_1 = (-0.5, 0.5, 0.5, -0.7, -0.5, 0.6)^T$ and $\mu_2 = (0.5, -0.6, 0.5, -0.4, 0.6, -0.7)^T$. For each subject i , z_{ij} is generated using Equation (3) for given values of σ_b^2 and Ω_{ij} . As noted in Section 2.2, estimating $\beta_f(\cdot)$ is equivalent to estimating η_j ; thus, in all simulations, we use η_j to specify the underlying true model and use the estimated η_j to evaluate model performance. To specify ξ , we let $\alpha_0 = -0.5$, $\alpha_1 = 0.5$, and $\eta_1 = (0.3, -0.5, 0.6, -0.3, -0.4, 0.5)^T$ and $\eta_2 = (-0.6, 0.5, -0.5, -0.5, 0.4, 0.5)^T$ ($q = 6$) for all cases. Then a linear regression model is assumed for Equation (6), and y is generated from a Gaussian distribution with a mean computed using Equation (6) and a variance of $\sigma_y^2 = 0.2$.

For the case of Gaussian outcomes, we consider two types of data structures. For a balanced design, t_{ik} 's ($k = 1, \dots, m$) are generated from equally spaced points on $[-1, 1]$, and Ω_j is a randomly generated positive definite matrix with its maximum element equal to a constant denoted by $\max(\Omega_j)$. For an unbalanced design, t_{ik} ($k = 1, \dots, m_i$) are generated from a uniform distribution on $[-1, 1]$, and Ω_{ij} is generated from an AR(1) structure for each biomarker j such that $\text{cov}(z_{ijk}, x_{ijl}) = \phi_j \rho_j^{10(t_{ik} - t_{il})}$, where ϕ_j defines the size of the measurement error and ρ_j defines the strength of the correlation.

Table 1 presents the simulation results in the case of Gaussian outcomes and a balanced data structure, where $n = 200$ and $m_j = 20$. In this setting, we also compare the Bayesian estimator with the estimator proposed in James⁹ that assumes independence between repeatedly measured biomarker values and is denoted as the independence model in Table 1. Different sizes of measurement error and correlation are considered. Compared to the independence model, the proposed Bayesian approach exhibits mostly less bias, smaller sampling variance (SD), smaller SMSE, and considerably better coverage properties. The difference in SMSE ranges from moderate, when the size of measurement error and of correlation are small, to substantial, when they are large. While the estimator from the independence model shows mostly negligible bias when the size of measurement error is moderate (i.e., $\max(\Omega_j) = 0.2$), its standard error vastly underestimates the true variation represented by SD and the coverage rates of the 95% Wald confidence intervals are well below the nominal level. As the size of measurement error and of correlation increase (i.e., $\max(\Omega_j) = 0.5$), both estimators show degradation of finite sample performance: while the Bayesian estimator starts to exhibit appreciable bias for α_0 , η_{15} and η_{16} resulting in coverage rates below the nominal level, the performance for the remaining parameters are still acceptable; on the other hand, the deterioration of performance is more pronounced for the independence model. When the design is unbalanced, the comparisons between the two methods remain similar.

In the case of binary outcomes, we set the means of γ_{11} and γ_{12} as $\boldsymbol{\mu}_1 = (0.3, 0.1, 0.3, 0.2, 0.1, 0.2)^T$ and $\boldsymbol{\mu}_2 = (-0.2, -0.1, -0.2, -0.1, -0.2, -0.1)^T$, and we consider the balanced data structure. A logistic model is assumed for Equation (6) and y is drawn from a Bernoulli distribution with the success probability computed using Equation (6). All the other setups are kept the same as in the case of Gaussian outcomes and the estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\eta}$ are again used to evaluate model performance. Table 2 shows that the performance of the proposed Bayesian model remains satisfactory for binary outcomes when the sample size is moderate.

4 Analysis of MAP II Data

We apply the proposed Bayesian model to the MAP II study. Following our previous notation, we let y denote the binary outcome of interest, namely, the presence ($y = 1$) or absence ($y = 0$) of colorectal adenomas. In this study, each subject had biomarkers measured in multiple hemi-crypts ("scored"), and we use the average value at each of the 50 segments the hemi-crypt is divided into over multiple hemi-crypts scored from subject i as the observed $z_{ij}(t_k)$. Furthermore, using similar arguments as in Prentice and Pyke¹⁸, it can be readily shown that case-control studies such as the MAP II study can be analyzed using our

model as if they are prospective studies. While the interpretation of the intercept term, namely α_0 , is no longer valid, the interpretation for η and hence the weight functions, $\beta(\cdot)$, still holds.

Two functional biomarkers are of particular interest, namely, the expression level of APC, a known tumor suppressor gene, and the expression level of TGF- α , a known mediator of oncogenesis and malignant progression in colorectal carcinogenesis. Previously, data collected in this study were analyzed in Dash et al.² and Daniel et al.¹ for APC and TGF- α , respectively, and it was shown that the controls on average have higher APC levels and lower TGF- α along the length of colon crypts, and larger differences are observed between segments 15 and 35 for both APC and TGF- α . In these papers, the primary analysis was to associate ad-hoc summary measures of the functional biomarkers with the case/control status. The ad-hoc summary measures included the mean over all segments, the means over portions of crypts, and certain ratios of these means. It is shown that some of these summary measures of APC and TGF- α are associated with the case/control status. We note that association between risk for colorectal cancer and each biomarker is examined separately.

We reanalyze the colorectal adenoma data including both biomarkers using the proposed Bayesian model and only the data for subjects on whom both biomarkers were measured are used in the current analysis. We include patient age as a scalar covariate, i.e., x . For the clinical outcome, the risk of colorectal adenomas, we use a logistic model, i.e.,

$$\text{logit} [P(y=1|x, Z_1(\cdot), Z_2(\cdot))] = \alpha_0 + \alpha_1 x + \int \beta_1(t) Z_1(t) dt + \int \beta_2(t) Z_2(t) dt,$$

where x is the age of a patient, $Z_1(t)$ and $Z_2(t)$ are the underlying true distributions of the expression level along the length of hemi-crypts for APC and TGF- α , respectively. Legendre polynomial functions are used for the basis functions $s(\cdot)$ for $Z_1(\cdot)$ and $Z_2(\cdot)$, which is deemed appropriate based on the observed biomarker values along the length of crypts. Since Legendre polynomial functions are orthogonal, it is straightforward to obtain the estimate of $\beta_j(\cdot)$ using the estimate of η_j ($j = 1, \dots, p$), i.e., $\beta_j(\cdot) = s(\cdot)^T \eta_j$. To perform statistical inference, we need to choose an optimal q ; to achieve this goal, the Deviance Information Criterion (DIC)¹⁹ is computed for a range of values for q ($q = 2, 3, 4, 5$) and the q value with the minimum DIC ($q = 3$) is chosen for our final model, since a smaller DIC value indicates a better fit when comparing models.

Table 3 presents the parameter estimates (the posterior means) for ξ and their posterior SDs using the Bayesian approach, and it shows that the effect of age (α_1) on the risk of colon adenomas is significant and the risk increases as age increases. Table 3 also shows that the coefficients that correspond to the quadratic basis function for $\beta(\cdot)$, namely, η_{13} and η_{23} , are different from 0 with their 95% credible intervals excluding 0, confirming the need to include the quadratic term for both weight functions. Figures 1 and 2 display the estimated weight functions $\hat{\beta}_1(\cdot)$ and $\hat{\beta}_2(\cdot)$ for APC and TGF- α , respectively, and their 95% credible intervals. Except for the regions near the base and apex segments of crypts, the estimated weight function is mostly negative for APC and mostly positive for TGF- α , which indicates that a lower expression level of APC is associated with a higher risk of colorectal adenoma

($y = 1$) and a higher expression level of TGF- α is associated with a higher risk of colorectal adenoma. In these regions, the APC expression near Segment 25 and the TGF- α expression near segment 17 had the largest absolute weights and, therefore, their associations with risk of colorectal adenoma are strongest. In addition, the 95% credible interval of the weight function for APC and TGF- α excludes 0 in Segments 15-36 and 9-29, respectively; in other words, the weight functions in these locations are significantly different from 0. The estimated weight function $\hat{\beta}_1(\cdot)$ is significantly greater than 0 in segments close to the base and apex of the crypts, and the estimated weight function $\hat{\beta}_2(\cdot)$ is significantly less than 0 in segments close to the apex of the crypts. It has been noted that the measurements close to the base and apex of the crypts are problematic due to the limitation of the procedure and are subject to unusually large variation compared to the other parts of the crypts; hence, caution is advised when interpreting these results. Consequently, we primarily focus on the results in the middle segments. It is well known in the literature²⁰ that the bottom 60% of the crypt is the proliferating zone, the upper 40% is the differentiating zone, and the biomarker distributions are likely to change around this transitional region. Our results provide further evidence that it is important to incorporate the biomarker expression level near this transition region for both biomarkers when constructing a predictive profile for risk for colorectal neoplasms. Compared to the results obtained in our previous study for modeling risk for colorectal cancer using the APC functional biomarker (Figure 3 in Long et al.¹¹), the proposed approach leads to a similar estimated weight function for APC but tighter credible intervals as shown in Figures 1, which is partly due to accounting for correlation as well as incorporating additional informative predictors.

5 Discussion

In this article, we investigate a Bayesian modeling approach for assessing associations between a clinical outcome and multiple functional biomarkers that are correlated and subject to measurement error. Our approach accounts for two levels of correlation, i.e., between repeated measurements of a same biomarker and between multiple biomarkers from the same subject. Using the proposed Bayesian inference procedure, it takes 616 seconds of CPU (Intel i7-2600) time to fit the model of interest for the MAP II data and we observe similar run times in the simulation studies. However, as the number of functional biomarkers becomes large, it becomes crucial to incorporate model selection as part of the Bayesian inference procedure in order to avoid the model identifiability issue, which is a topic for our future research. Even in the absence of any model identifiability issue, the fitting procedure can become computationally very intensive as the number of functional biomarkers increases and/or they are measured over time as well as spatially, in which case one could use more efficient algorithms that have been developed in recent years such as the Metropolis adjusted Langevin algorithm²¹ or Riemann manifold Langevin and Hamiltonian Monte Carlo methods²².

While we use the same number of basis functions to model different functional biomarkers, it is straightforward to extend our model so that different numbers of basis functions are used to model different functional biomarkers. In addition, it is of potential interest to investigate more sophisticated statistical techniques for selecting the number of basis

functions and, more importantly, selecting functional biomarkers. The proposed Bayesian approach also allows for prediction for future observations. Specifically, one can obtain samples from the posterior predictive distribution for future observations. For a new subject, let $z_j(t_k)$ be the observed value of biomarker j at t_k ($j = 1, \dots, p$ and $k = 1, \dots, m$); and given $z_j(t_k)$'s, one may be interested in predicting the outcome using the posterior predictive distribution, namely, $f(\tilde{y} | z_j(t_k)$'s). Specifically, one can first draw $\tilde{\mu} = (\tilde{\mu}_1, \dots, \tilde{\mu}_p)$, $\tilde{\Gamma} = (\tilde{\Gamma}_1, \dots, \tilde{\Gamma}_p)$, $\tilde{\Omega} = (\tilde{\Omega}_1, \dots, \tilde{\Omega}_p)$ and $\tilde{\sigma}_b^2$ from their respective conditional posterior distribution; then, draw $\tilde{\gamma} = (\tilde{\gamma}_1, \dots, \tilde{\gamma}_p)$ conditional on $z_j(t_k)$'s, and draw $\tilde{\mu}$, $\tilde{\Gamma}$, $\tilde{\Omega}$ and $\tilde{\sigma}_b^2$ based on Equation (7); then, draw $\tilde{\xi}$ and $\tilde{\phi}$ from their respective conditional posterior distribution of ξ and ϕ , and then, draw $\tilde{y}$ based on Equation (5) and (6) given $\tilde{\gamma}$, $\tilde{\xi}$ and $\tilde{\phi}$. The posterior predictive distribution of $\tilde{y}$ can be further summarized by its mean or mode, which can serve as a point prediction for continuous and binary outcomes, respectively.

An alternative to the proposed Bayesian implementation strategy could be to fit the joint models (3) and (4) in a Frequentist framework; for example, a generalized additive mixed model formulation that can be fitted using the gamm or amer functions in R can be used to model functional biomarkers. However, it is not trivial to perform model-fitting and inference for the joint models in this framework, which is another potential topic for future research.

Acknowledgements

Qi Long and Xiaoxi Zhang contributed equally to this work. This work was supported in part by US NIH PHS Grant UL1 RR025008 from the Clinical and Translational Science Award program and US NIH/NCI grant R01 CA114456.

A Posterior Sampling for Gaussian Outcomes

Without loss of generality, we first consider a special case of Gaussian outcomes and the observed data likelihood is

$$\begin{aligned} \ell(\xi, \phi, \gamma_i, \mathbf{b}_i, \Omega_{ij}, \sigma_b^2, \sigma_y^2; \mathbf{y}, \mathbf{z}) &\propto \prod_{i=1}^n \left[\frac{1}{\sigma_y} \right. \\ &\exp \left\{ -\frac{\left(y_i - \alpha_0 - \alpha_1^T \mathbf{x}_i - \sum_{j=1}^p \eta_j^T \gamma_{ij} \right)^2}{2\sigma_y^2} \right\} \times \prod_{j=1}^p |\Omega_{ij}|^{-\frac{1}{2}} \exp \\ &\left. \left\{ -\frac{1}{2} \left(\mathbf{z}_{ij} - \mathbf{S}_i^T \gamma_{ij} - \mathbf{1}_{m_i} b_i \right)^T \Omega_{ij}^{-1} \left(\mathbf{z}_{ij} - \mathbf{S}_i^T \gamma_{ij} - \mathbf{1}_{m_i} b_i \right) \right\} \times \left(\sigma_b^2 \right)^{-\frac{1}{2}} \exp \left\{ -\frac{b_i^2}{2\sigma_b^2} \right\} \right]. \end{aligned}$$

We specify priors for parameters of interest as follows

$$\begin{aligned}\gamma_{ij} &\sim N(\boldsymbol{\mu}_j, \boldsymbol{\Gamma}_j) \quad (j=1, \dots, p), \\ \boldsymbol{\xi} &\sim N(\boldsymbol{\mu}_\xi, \boldsymbol{\Sigma}_\xi), \\ \sigma_y^2 &\sim \text{inv} - \chi^2(v_0, s_0^2), \\ b_i &\sim N(0, \sigma_b^2).\end{aligned}$$

We then specify priors on hyperparameters $(\boldsymbol{\mu}_j, \boldsymbol{\Gamma}_j)$ as follows

$$\begin{aligned}\boldsymbol{\Gamma}_j^{-1} &\sim \text{Wishart}(\mathbf{S}_\Gamma, \nu_\Gamma), \\ \boldsymbol{\mu}_j | \boldsymbol{\Gamma}_j &\sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0).\end{aligned}$$

A.1 The Case of Balanced Design

In this case, $\Omega_{ij} = \Omega_j$ and $m_i = m$. We further specify prior for $\Omega_j, \Omega_j^{-1} \sim \text{Wishart}(\mathbf{S}_\Omega, \nu_\Omega)$. It follows that the full conditional posterior distributions are

$$\begin{aligned}P(\gamma_{ij} | \text{rest}, \text{Obs}) &\sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ \text{with } \boldsymbol{\mu} &= \boldsymbol{\Sigma} \left\{ \frac{1}{\sigma_y^2} \left(\mathbf{y}_i - \alpha_0 - \boldsymbol{\alpha}_1^T \mathbf{x}_i - \sum_{l=1 \neq j}^p \boldsymbol{\eta}_l^T \gamma_{il} \right) \boldsymbol{\eta}_j + \mathbf{S}_i \boldsymbol{\Omega}_j^{-1} (\mathbf{z}_{ij} - \mathbf{1}_{m_i} \mathbf{b}_i) + \boldsymbol{\Gamma}_j^{-1} \boldsymbol{\mu}_j \right\} \quad \text{and} \quad \boldsymbol{\Sigma} = \left\{ \frac{\boldsymbol{\eta}_j \boldsymbol{\eta}_j^T}{\sigma_y^2} + \mathbf{S}_i \boldsymbol{\Omega}_j^{-1} \mathbf{S}_i^T + \boldsymbol{\Gamma}_j^{-1} \right\}^{-1}.\end{aligned}$$

$$\begin{aligned}P(\boldsymbol{\Gamma}_j^{-1} | \text{rest}, \text{Obs}) &\sim \text{Wishart}(\mathbf{S}, \nu) \\ \text{with } \mathbf{S} &= \left\{ \sum_{i=1}^n (\gamma_{ij} - \boldsymbol{\mu}_j) (\gamma_{ij} - \boldsymbol{\mu}_j)^T + \mathbf{S}_\Gamma^{-1} \right\}^{-1} \quad \text{and} \quad \nu = \nu_\Gamma + n.\end{aligned}$$

$$\begin{aligned}P(\boldsymbol{\mu}_j | \text{rest}, \text{Obs}) &\sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ \text{with } \boldsymbol{\mu} &= \boldsymbol{\Sigma} \left\{ \sum_{i=1}^n \boldsymbol{\Gamma}_j^{-1} \gamma_{ij} + \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\mu}_0 \right\} \quad \text{and} \quad \boldsymbol{\Sigma} = \left\{ n \boldsymbol{\Gamma}_j^{-1} + \boldsymbol{\Sigma}_0^{-1} \right\}^{-1}.\end{aligned}$$

$$\begin{aligned}P(b_i | \text{rest}, \text{Obs}) &\sim N(\mu, \sigma^2) \\ \text{with } \mu &= \sigma^2 T \quad \text{and} \quad \sigma^2 = \left\{ a + \sigma_b^{-2} \right\}^{-1}, \quad \text{where } a = \mathbf{W}_i^T (\boldsymbol{\Omega}_i^*)^{-1} \mathbf{W}_i, \quad \text{and} \quad T = \mathbf{W}_i^T (\boldsymbol{\Omega}_i^*)^{-1} (\mathbf{z}_i - \mathbf{S}_i^* \boldsymbol{\gamma}_i).\end{aligned}$$

$$\begin{aligned}P(\sigma_b^2 | \text{rest}, \text{Obs}) &\sim \text{Inv} - \chi^2(\nu, s^2) \\ \text{with } \nu &= \nu_b + n \quad \text{and} \quad \nu s^2 = \nu_b s_b^2 + \sum_{i=1}^n b_i^2.\end{aligned}$$

$$P(\boldsymbol{\Omega}_j^{-1} | \text{rest}, \text{Obs}) \sim \text{Wishart}(\mathbf{S}, \nu)$$

$$\text{with } \mathbf{S} = \left\{ \sum_{i=1}^n (\mathbf{z}_{ij} - \mathbf{S}_i^T \boldsymbol{\gamma}_{ij} - \mathbf{1}_{m_i} b_i) (\mathbf{z}_{ij} - \mathbf{S}_i^T \boldsymbol{\gamma}_{ij} - \mathbf{1}_{m_i} b_i)^T + \mathbf{S}_{\Omega}^{-1} \right\}^{-1} \quad \text{and} \quad \nu = \nu_{\Omega} + n.$$

$$P(\boldsymbol{\xi} | \text{rest}, \text{Obs}) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\text{with } \boldsymbol{\mu} = \boldsymbol{\Sigma} \left\{ \sum_{i=1}^n \frac{\mathbf{w}_i \mathbf{y}_i}{\sigma_y^2} + \boldsymbol{\Sigma}_{\xi}^{-1} \boldsymbol{\mu}_{\xi} \right\} \quad \text{and} \quad \boldsymbol{\Sigma} = \left\{ \sum_{i=1}^n \frac{\mathbf{w}_i \mathbf{w}_i^T}{\sigma_y^2} + \boldsymbol{\Sigma}_{\xi}^{-1} \right\}^{-1}, \quad \text{where } \mathbf{w}_i = (\mathbf{1}, \mathbf{x}_i^T, \gamma_{i1}^T, \gamma_{i2}^T, \dots, \gamma_{ip}^T)^T.$$

$$P(\sigma_y^2 | \text{rest}, \text{Obs}) \sim \text{Inv} - \chi^2(\nu, s^2)$$

$$\text{with } \nu = \nu_0 + n \quad \text{and} \quad \nu s^2 = \nu_0 s_0^2 + \sum_{i=1}^n (y_i - \mathbf{w}_i^T \boldsymbol{\xi})^2.$$

A.2 The Case of Unbalanced Design

In the case of an unbalanced design, Ω_{ij} is assumed to have an AR(1) correlation structure, that is, the (k_1, k_2) 'th element of Ω_{ij} is $\phi_j \rho_j^{|t_{ik_1} - t_{ik_2}|}$ with $|\rho| < 1$. Let $\Omega_{ij} = \phi_j \mathbf{R}_{ij}$, where $\mathbf{R}_{ij}$ is the correlation matrix. Let $\Delta_i(k) = t_{ik} - t_{i,(k-1)}$ and $\Delta_i^*(k) = t_{i,(k+1)} - t_{i,(k-1)}$ then some algebra can show that

$$|\mathbf{R}_{ij}| = \prod_{k=2}^{m_i} (1 - \rho_j^{2\Delta_i(k)}),$$

and

$$\begin{aligned} (\mathbf{R}_{ij}^{-1})_{11} &= (1 - \rho_j^{2\Delta_i(2)})^{-1}, \\ (\mathbf{R}_{ij}^{-1})_{m_i m_i} &= (1 - \rho_j^{2\Delta_i(m_i)})^{-1}, \\ (\mathbf{R}_{ij}^{-1})_{kk} &= (1 - \rho_j^{2\Delta_i^*(k)}) \left\{ (1 - \rho_j^{2\Delta_i(k)}) (1 - \rho_j^{2\Delta_i(k+1)}) \right\}^{-1} \quad \text{for } 1 < k < m_i, \\ (\mathbf{R}_{ij}^{-1})_{k,(k+1)} &= (\mathbf{R}_{ij}^{-1})_{(k+1),k} = -\rho_j^{\Delta_i(k+1)} (1 - \rho_j^{2\Delta_i(k+1)})^{-1} \quad \text{for } 1 \leq k \leq m_i - 1, \\ (\mathbf{R}_{ij}^{-1})_{k_1 k_2} &= 0 \quad \text{for } |k_1 - k_2| > 1. \end{aligned}$$

Note that $\Omega_{ij}^{-1} = \frac{1}{\phi_j} \mathbf{R}_{ij}^{-1}$ and $|\Omega_{ij}| = \phi_j^{m_i} |\mathbf{R}_{ij}|$.

The observed data likelihood becomes

$$\ell(\boldsymbol{\xi}, \phi, \boldsymbol{\gamma}_i, \mathbf{b}_i, \boldsymbol{\Omega}_{ij}, \sigma_b^2, \sigma_y^2, \mathbf{y}, \mathbf{z}) \propto \prod_{i=1}^n \left[\frac{1}{\sigma_y} \exp \left\{ -\frac{\left(y_i - \alpha_0 - \boldsymbol{\alpha}_1^T \mathbf{x}_i - \sum_{j=1}^p \boldsymbol{\eta}_j^T \boldsymbol{\gamma}_{ij} \right)^2}{2\sigma_y^2} \right\} \times \prod_{j=1}^p \phi_j^{-\frac{m_{ij}}{2}}, |\mathbf{R}_{ij}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \frac{1}{\phi_j} (\mathbf{z}_{ij} - \mathbf{S}_i^T \boldsymbol{\gamma}_{ij})^2 \right\} \right]$$

We further specify flat priors for $\rho_j (|\rho_j| < 1)$ and $\phi_j (\phi_j > 0)$, and let

$w_i = (1, \mathbf{x}_i^T, \gamma_{i1}^T, \gamma_{i2}^T, \dots, \gamma_{ip}^T)^T$. It follows that the full conditional posterior distributions are

$$\begin{aligned} & \mathbf{1.} \\ & P(\boldsymbol{\gamma}_{ij} | rest, Obs) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ & \text{with } \boldsymbol{\mu} = \boldsymbol{\Sigma} \left\{ \frac{1}{\sigma_y^2} \left(\mathbf{y}_i - \alpha_0 - \boldsymbol{\alpha}_1^T \mathbf{x}_i - \sum_{l=1 \neq j}^p \boldsymbol{\eta}_l^T \gamma_{il} \right) \boldsymbol{\eta}_j + \mathbf{S}_i \boldsymbol{\Omega}_j^{-1} (\mathbf{z}_{ij} - \mathbf{1}_{m_i} \mathbf{b}_i) + \boldsymbol{\Gamma}_j^{-1} \boldsymbol{\mu}_j \right\} \quad \text{and} \quad \boldsymbol{\Sigma} = \left\{ \frac{\boldsymbol{\eta}_j \boldsymbol{\eta}_j^T}{\sigma_y^2} + \mathbf{S}_i \boldsymbol{\Omega}_j^{-1} \mathbf{S}_i^T + \boldsymbol{\Gamma}_j^{-1} \right\}^{-1} \end{aligned}$$

$$\begin{aligned} & \mathbf{2.} \\ & P(\boldsymbol{\Gamma}_j^{-1} | rest, Obs) \sim Wishart(\mathbf{S}, \nu) \\ & \text{with } \mathbf{S} = \left\{ \sum_{i=1}^n (\boldsymbol{\gamma}_{ij} - \boldsymbol{\mu}_j) (\boldsymbol{\gamma}_{ij} - \boldsymbol{\mu}_j)^T + \mathbf{S}_j^{-1} \right\}^{-1} \quad \text{and} \quad \nu = \nu_{\boldsymbol{\Gamma}} + n. \end{aligned}$$

$$\begin{aligned} & \mathbf{3.} \\ & P(\boldsymbol{\mu}_j | rest, Obs) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ & \text{with } \boldsymbol{\mu} = \boldsymbol{\Sigma} \left\{ \sum_{i=1}^n \boldsymbol{\Gamma}_j^{-1} \boldsymbol{\gamma}_{ij} + \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\mu}_0 \right\} \quad \text{and} \quad \boldsymbol{\Sigma} = \left\{ n \boldsymbol{\Gamma}_j^{-1} + \boldsymbol{\Sigma}_0^{-1} \right\}^{-1}. \end{aligned}$$

$$\begin{aligned} & \mathbf{4.} \\ & P(b_i | rest, Obs) \sim N(\mu, \sigma^2) \\ & \text{with } \mu = \sigma^2 T \quad \text{and} \quad \sigma^2 = \left\{ a + \sigma_b^{-2} \right\}^{-1}, \quad \text{where } a = \mathbf{W}_i^T (\boldsymbol{\Omega}_i^*)^{-1} \mathbf{W}_i, \quad \text{and} \quad T = \mathbf{W}_i^T (\boldsymbol{\Omega}_i^*)^{-1} (\mathbf{z}_i - \mathbf{S}_i^* \boldsymbol{\gamma}_i). \end{aligned}$$

$$\begin{aligned} & \mathbf{5.} \\ & P(\sigma_b^2 | rest, Obs) \sim Inv - \chi^2(\nu, s^2) \\ & \text{with } \nu = \nu_b + n \quad \text{and} \quad \nu s^2 = \nu_b s_b^2 + \sum_{i=1}^n b_i^2. \end{aligned}$$

$$\begin{aligned} & \mathbf{6.} \\ & (\boldsymbol{\xi} | rest, Obs) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ & \text{with } \boldsymbol{\mu} = \boldsymbol{\Sigma} \left\{ \sum_{i=1}^n \frac{w_i y_i}{\sigma_y^2} + \boldsymbol{\Sigma}_{\xi}^{-1} \boldsymbol{\mu}_{\xi} \right\} \quad \text{and} \quad \boldsymbol{\Sigma} = \left\{ \sum_{i=1}^n \frac{w_i w_i^T}{\sigma_y^2} + \boldsymbol{\Sigma}_{\xi}^{-1} \right\}^{-1}. \end{aligned}$$

$$\begin{aligned} & \mathbf{7.} \\ & P(\sigma_y^2 | rest, Obs) \sim Inv - \chi^2(\nu, s^2) \\ & \text{with } \nu = \nu_0 + n \quad \text{and} \quad \nu s^2 = \left\{ \nu_0 s_0^2 + \sum_{i=1}^n (y_i - w_i^T \boldsymbol{\xi})^2 \right\}. \end{aligned}$$

$$\begin{aligned} & \mathbf{8.} \\ & P(\phi_j | rest, Obs) \sim Inv - \chi^2(\nu, s^2) \\ & \text{with } \nu = \sum_{i=1}^n m_i \quad \text{and} \quad \nu s^2 = \sum_{i=1}^n \left(\mathbf{z}_{ij} - \mathbf{S}_i^T \boldsymbol{\gamma}_{ij} - \mathbf{1}_{m_i} b_i \right)^T (\mathbf{R}_{ij})^{-1} \left(\mathbf{z}_{ij} - \mathbf{S}_i^T \boldsymbol{\gamma}_{ij} - \mathbf{1}_{m_i} b_i \right). \end{aligned}$$

$$\begin{aligned} & \mathbf{9.} \\ & P(\rho_j | rest, Obs) \propto \\ & \prod_{i=1}^n |\mathbf{R}_{ij}(\rho_j)|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\rho_j} \left(\mathbf{z}_{ij} - \mathbf{S}_i^T \boldsymbol{\gamma}_{ij} - \mathbf{1}_{m_i} b_i \right)^T \mathbf{R}_{ij}(\rho_j)^{-1} \left(\mathbf{z}_{ij} - \mathbf{S}_i^T \boldsymbol{\gamma}_{ij} - \mathbf{1}_{m_i} b_i \right) \right\}, \quad \text{which needs to be approximated} \end{aligned}$$

B Posterior Sampling for Binary Outcomes

For binary outcomes, a logistic model is assumed for Equation (6). Then, the observed data likelihood becomes

$$\ell(\boldsymbol{\xi}, \phi, \boldsymbol{\gamma}_i, \mathbf{b}_i, \boldsymbol{\Omega}_{ij}, \sigma_b^2; \mathbf{y}, \mathbf{z}) \propto \prod_{i=1}^n \left[\left\{ 1 + \exp \left(-\alpha_0 - \boldsymbol{\alpha}_1^T \mathbf{x}_i - \sum_{j=1}^p \boldsymbol{\eta}_j^T \boldsymbol{\gamma}_{ij} \right) \right\}^{-y_i} \left\{ 1 + \exp \left(\alpha_0 + \boldsymbol{\alpha}_1^T \mathbf{x}_i + \sum_{j=1}^p \boldsymbol{\eta}_j^T \boldsymbol{\gamma}_{ij} \right) \right\}^{y_i-1} \right] \times \prod_{j=1}^p |\boldsymbol{\Omega}_{ij}|^{-\frac{1}{2}},$$

We specify priors for parameters of interest as follows

$$\begin{aligned} \boldsymbol{\gamma}_{ij} & \sim N(\boldsymbol{\mu}_j, \boldsymbol{\Gamma}_j) \quad (j=1, \dots, p), \\ \boldsymbol{\xi} & \sim N(\boldsymbol{\mu}_\xi, \boldsymbol{\Sigma}_\xi), \\ \sigma_b^2 & \sim Inv - \chi^2(\nu_b, s_b^2). \end{aligned}$$

We also specify priors on hyperparameters $(\boldsymbol{\mu}_j, \boldsymbol{\Gamma}_j)$ as follows

$$\begin{aligned} \boldsymbol{\Gamma}_j^{-1} & \sim \text{Wishart}(\mathbf{S}_\Gamma, \nu_\Gamma), \\ \boldsymbol{\mu}_j | \boldsymbol{\Gamma}_j & \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0). \end{aligned}$$

Here, we only consider the special case of a balanced design, i.e., $m_i = m$ and $\Omega_{ij} = \Omega_j$. We specify prior for Ω_j , $\Omega_j^{-1} \sim \text{Wishart}(\mathbf{S}_\Omega, \nu_\Omega)$ and let $\text{logit}(p_i) = \alpha_0 + \boldsymbol{\alpha}_1^T \mathbf{x}_i + \sum_{j=1}^p \boldsymbol{\eta}_j^T \boldsymbol{\gamma}_{ij}$. Note that p_i is a function of $\boldsymbol{\xi}$ and $\boldsymbol{\gamma}_{ij}$, and hence needs not be drawn. It follows that the full conditional posterior distributions are

$$\begin{aligned} & \mathbf{1.} \\ & P(\boldsymbol{\gamma}_{ij} | rest, Obs) \\ & \propto p_i^{y_i} (1 - p_i)^{1-y_i} \times \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\gamma}_{ij}^T \left(\mathbf{S}_i \boldsymbol{\Omega}_j^{-1} \mathbf{S}_i^T + \boldsymbol{\Gamma}_j^{-1} \right) \boldsymbol{\gamma}_{ij} - 2 \boldsymbol{\gamma}_{ij}^T \left\{ \mathbf{S}_i \boldsymbol{\Omega}_j^{-1} (\mathbf{z}_{ij} - \mathbf{1}_{m_i} b_i) + \boldsymbol{\Gamma}_j^{-1} \boldsymbol{\mu}_j \right\} \right] \right\} \end{aligned}$$

$$\begin{aligned} & \mathbf{2.} \\ & P(\boldsymbol{\Gamma}_j^{-1} | rest, Obs) \sim \text{Wishart}(\mathbf{S}, \nu) \\ & \text{with } \mathbf{S} = \left\{ \sum_{i=1}^n (\boldsymbol{\gamma}_{ij} - \boldsymbol{\mu}_j) (\boldsymbol{\gamma}_{ij} - \boldsymbol{\mu}_j)^T + \mathbf{S}_\Gamma^{-1} \right\}^{-1} \quad \text{and} \quad \nu = \nu_\Gamma + n. \end{aligned}$$

3.

$$P(\boldsymbol{\mu}_j | rest, Obs) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\text{with } \boldsymbol{\mu} = \boldsymbol{\Sigma} \left\{ \sum_{i=1}^n \boldsymbol{\Gamma}_j^{-1} \boldsymbol{\gamma}_{ij} + \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\mu}_0 \right\} \quad \text{and} \quad \boldsymbol{\Sigma} = \left\{ n \boldsymbol{\Gamma}_j^{-1} + \boldsymbol{\Sigma}_0^{-1} \right\}^{-1}.$$

4.

$$P(b_i | rest, Obs) \sim N(\mu, \sigma^2)$$

$$\text{with } \mu = \sigma^2 T \quad \text{and} \quad \sigma^2 = \left\{ a + \sigma_b^{-2} \right\}^{-1}, \quad \text{where } a = \mathbf{W}_i^T (\boldsymbol{\Omega}_i^*)^{-1} \mathbf{W}_i, \quad \text{and} \quad T = \mathbf{W}_i^T (\boldsymbol{\Omega}_i^*)^{-1} (\mathbf{z}_i - \mathbf{S}_i^* \boldsymbol{\gamma}_i).$$

5.

$$P(\sigma_b^2 | rest, Obs) \sim Inv - \chi^2(\nu, s^2)$$

$$\text{with } \nu = \nu_b + n \quad \text{and} \quad \nu s^2 = \nu_b s_b^2 + \sum_{i=1}^n b_i^2.$$

6.

$$P(\boldsymbol{\Omega}_j^{-1} | rest, Obs) \sim Wishart(\mathbf{S}, \nu)$$

$$\text{with } \mathbf{S} = \left\{ \sum_{i=1}^n (\mathbf{z}_{ij} - \mathbf{S}_i^T \boldsymbol{\gamma}_{ij} - \mathbf{1}_{m_i} b_i) (\mathbf{z}_{ij} - \mathbf{S}_i^T \boldsymbol{\gamma}_{ij} - \mathbf{1}_{m_i} b_i)^T + \mathbf{S}_{\Omega}^{-1} \right\}^{-1} \quad \text{and} \quad \nu = \nu_{\Omega} + n.$$

7.

$$P(\boldsymbol{\xi} | rest, Obs) \propto \prod_{i=1}^n \left\{ p_i^{y_i} (1 - p_i)^{1-y_i} \right\} \times \exp \left[-\frac{1}{2} \left(\boldsymbol{\xi}^T \boldsymbol{\Sigma}_{\xi}^{-1} \boldsymbol{\xi} - 2 \boldsymbol{\xi}^T \boldsymbol{\Sigma}_{\xi}^{-1} \boldsymbol{\mu}_{\xi} \right) \right]$$

The conditionals of $P(\boldsymbol{\xi} | rest, Obs)$ and $P(\boldsymbol{\gamma}_{ij} | rest, Obs)$ need to be approximated using an ARMS-algorithm. For an unbalanced design, a similar algorithm can also be developed.

References

1. Daniel CR, Bostick RM, Flanders WD, Long Q, Fedirko V, Sidelnikov E, et al. TGF- α expression as a potential biomarker of risk within the normal-appearing colorectal mucosa of patients with and without incident sporadic adenoma. *Cancer Epidemiology Biomarkers & Prevention*. 2009; 18:65–73.
2. Dash C, Kennedy V, Bostick RM. APC, β -catenin and E-cadherin protein expression in colon crypts in the Markers of Adenomatous Polyps II (MAP II) study. *American Association for Cancer Research Annual Meeting Abstracts*. 2006; 1:164.
3. Long Q, Flanders WD, Fedirko V, Bostick RM. Robust statistical methods for analysis of biomarkers measured with batch/experiment-specific errors. *Statistics in Medicine*. 2010; 29(3): 361–70. [PubMed: 20020422]
4. Fleming, TR.; Harrington, DP. *Counting Processes and Survival Analysis*. Wiley; New York: 1991.
5. Cardot H, Ferraty F, Sarda P. Spline estimators for the functional linear model. *Statistica Sinica*. 2003; 13:571–91.
6. Cardot H, Sarda P. Estimation in generalized linear models for functional data via penalized likelihood. *Journal of Multivariate Analysis*. 2005; 92:24–41.
7. Cardot H, Crambes C, Kneip A, Sarda P. Smoothing spline estimators in functional linear regression with errors-in-variables. *Computational Statistics and Data Analysis*. 2007; 51:4832–48.
8. Li Y, Hsing T. On rates of convergence in functional linear regression. *Journal of Multivariate Analysis*. 2006; 98:1782–1804.
9. James GM. Generalized linear models with functional predictors. *J R Statist Soc B*. 2002; 64:411–23.

10. Müller HG, Stadtmüller U. Generalized functional linear models. *The Annals of Statistics*. 2005; 33:774–805.
11. Long Q, Zhang X, Bostick RM. Semiparametric estimation for joint modeling of colorectal cancer risk and functional biomarkers measured with errors. *Biometrical Journal*. 2011; 53:393–410. [PubMed: 21404316]
12. McCullagh, P.; Nelder, JA. *Generalized Linear Models*. 2nd Edition. Chapman & Hall/CRC; New York: 1989.
13. Morris JS, Vannucci M, Brown PJ, Carroll RJ. Wavelet-Based Nonparametric Modeling of Hierarchical Functions in Colon Carcinogenesis. *Journal of the American Statistical Association*. 2003; 98:573–583.
14. Apanasovich TV, Sheather S, Lupton JR, Popovic N, Turner ND, Chapkin RS, et al. Testing for Spatial Correlation in Nonstationary Binary Data, with Application to Aberrant Crypt Foci in Colon Carcinogenesis. *Biometrics*. 2003; 59:752–761. [PubMed: 14969452]
15. Li Y, Wang N, Hong M, Turner ND, Lupton JR, Carroll RJ. Nonparametric estimation of correlation functions in longitudinal and spatial data, with application to colon carcinogenesis experiments. *Annals of Statistics*. 2007; 35:1608–1643.
16. Baldandayuthapani V, Mallick BK, Hong MY, Lupton JR, Turner ND, Carroll RJ. Bayesian Hierarchically Spatially Correlated Functional Data Analysis with Application to Colon Carcinogenesis. *Biometrics*. 2008; 64:64–73. [PubMed: 17608780]
17. Gilks WR, Best NG, Tan KKC. Adaptive Rejection Metropolis Sampling within Gibbs Sampling. *Applied Statistics*. 1995; 44(4):455–472.
18. Prentice R, Pyke R. Logistic disease incidence models and case-control studies. *Biometrika*. 1979; 66:403–11.
19. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of complexity and fit. *Journal of the Royal Statistical Society, Series B*. 2002; 64:583–639.
20. Gerdes H, Gillin JS, Zimbalist E, Urmacher C, Lipkin M, Winawer SJ. Expansion of the epithelial proliferative compartment and frequency of adenomatous polyps in the colon correlate with the strength of family history of colorectal cancer. *Cancer Research*. 1993; 53:279–82. [PubMed: 8417821]
21. Roberts G, Stramer O. Langevin diffusions and Metropolis-Hastings algorithms. *Methodol Comput Appl Probab*. 2003; 4:337–358.
22. Girolami M, Calderhead B. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society: Series B*. 2011; 73(2):123–214.

Figure 1.
Estimated weight function $\hat{\beta}_1(\cdot)$ (solid line) and its 95% credible interval (CI) (dashed line) for APC.

Figure 2.
Estimated weight function $\hat{\beta}_2(\cdot)$ (solid line) and its 95% credible interval (CI) (dashed line) for TGF- α .

Table 1

Simulation results for Gaussian outcomes using the proposed Bayesian model and the Independence model.

	Bayesian Model					Independence Model				
	Bias	SMSE	pSD	SD	CR (%)	Bias	SMSE	SE	SD	CR (%)
$\max(\Omega_j) = 0.2, \sigma_b^2 = 0.01$										
α_0	0.021	0.050	0.043	0.046	90.7	0.062	0.087	0.040	0.061	58.9
α_1	0.000	0.021	0.019	0.021	93.6	0.000	0.028	0.018	0.028	79.5
η_{11}	0.000	0.020	0.019	0.020	93.4	0.017	0.032	0.018	0.027	72.8
η_{12}	0.001	0.021	0.019	0.021	94.1	0.004	0.028	0.018	0.028	78.3
η_{13}	-0.007	0.022	0.019	0.020	91.3	-0.021	0.035	0.018	0.028	67.3
η_{14}	0.003	0.020	0.019	0.020	92.3	0.006	0.030	0.018	0.029	76.8
η_{15}	0.013	0.023	0.019	0.020	86.7	0.053	0.060	0.017	0.028	24.6
η_{16}	-0.012	0.023	0.019	0.020	90.0	-0.027	0.040	0.018	0.029	60.1
η_{21}	0.003	0.021	0.019	0.021	92.2	0.011	0.031	0.018	0.029	74.6
η_{22}	-0.004	0.021	0.019	0.020	93.3	0.005	0.029	0.018	0.029	77.9
η_{23}	0.000	0.021	0.019	0.021	91.9	0.003	0.029	0.018	0.028	79.1
η_{24}	0.007	0.022	0.019	0.020	91.1	0.002	0.029	0.018	0.029	78.4
η_{25}	-0.002	0.020	0.019	0.020	93.2	-0.004	0.029	0.018	0.029	78.0
η_{26}	0.001	0.020	0.019	0.020	95.0	-0.002	0.029	0.018	0.029	77.5
$\max(\Omega_j) = 0.2, \sigma_b^2 = 0.1$										
α_0	0.020	0.051	0.043	0.047	90.0	0.061	0.087	0.041	0.062	59.1
α_1	-0.001	0.021	0.019	0.021	93.2	0.000	0.028	0.019	0.028	79.7
η_{11}	0.000	0.020	0.019	0.020	92.3	0.027	0.038	0.018	0.027	61.2
η_{12}	0.001	0.021	0.019	0.021	94.2	0.004	0.028	0.018	0.028	79.3
η_{13}	-0.007	0.022	0.019	0.021	91.3	-0.021	0.035	0.018	0.029	67.2
η_{14}	0.003	0.021	0.019	0.020	93.6	0.006	0.030	0.019	0.030	77.1
η_{15}	0.013	0.024	0.019	0.021	86.9	0.053	0.060	0.018	0.028	26.0
η_{16}	-0.011	0.023	0.019	0.020	90.5	-0.027	0.040	0.018	0.029	60.9

	Bayesian Model					Independence Model				
	Bias	SMSE	pSD	SD	CR (%)	Bias	SMSE	SE	SD	CR (%)
η_{21}	0.003	0.021	0.019	0.021	92.7	0.021	0.036	0.018	0.029	65.0
η_{22}	-0.004	0.021	0.019	0.020	93.1	0.005	0.030	0.019	0.029	77.6
η_{23}	0.000	0.021	0.019	0.021	93.1	0.003	0.029	0.018	0.029	78.9
η_{24}	0.007	0.022	0.019	0.020	91.0	0.001	0.029	0.018	0.029	78.4
η_{25}	-0.002	0.020	0.019	0.020	93.6	-0.004	0.029	0.018	0.029	77.9
η_{26}	0.001	0.020	0.019	0.020	94.1	-0.002	0.029	0.019	0.029	77.9
$\max(\Omega_j) = 0.2, \sigma_b^2 = 0.1$										
α_0	0.047	0.074	0.052	0.057	83.8	0.134	0.159	0.048	0.084	30.6
α_1	-0.001	0.025	0.024	0.025	92.4	-0.001	0.039	0.022	0.039	74.0
η_{11}	0.001	0.024	0.024	0.024	93.6	0.053	0.066	0.022	0.038	37.6
η_{12}	0.002	0.025	0.023	0.025	94.7	0.009	0.040	0.022	0.039	71.4
η_{13}	-0.017	0.030	0.023	0.025	87.0	-0.045	0.060	0.022	0.040	46.9
η_{14}	0.006	0.025	0.023	0.025	93.4	0.012	0.043	0.022	0.042	69.3
η_{15}	0.029	0.038	0.023	0.025	73.0	0.114	0.120	0.020	0.037	2.7
η_{16}	-0.027	0.037	0.023	0.025	76.2	-0.059	0.071	0.022	0.040	34.2
η_{21}	0.007	0.027	0.024	0.026	91.3	0.041	0.057	0.021	0.039	49.6
η_{22}	-0.007	0.026	0.024	0.025	92.1	0.015	0.043	0.023	0.041	68.5
η_{23}	0.000	0.026	0.023	0.026	93.0	0.007	0.040	0.022	0.040	72.2
η_{24}	0.017	0.030	0.023	0.025	86.4	0.002	0.041	0.022	0.041	70.9
η_{25}	-0.004	0.025	0.023	0.025	92.5	-0.009	0.042	0.022	0.041	68.8
η_{26}	0.002	0.025	0.024	0.025	93.7	-0.005	0.042	0.022	0.042	69.3

$n = 200$, $m_j = 20$, and $\sigma_{ij}^2 = 0.2$. SMSE, square root of the mean squared error; pSD, mean of the posterior standard deviation for the Bayesian model; SE, mean of the standard error estimate for the independence model; SD, Monte Carlo standard deviation of the posterior mean for the Bayesian model or the parameter estimate for the independence model; CR, Monte Carlo coverage rate of 95% credible interval for the Bayesian model or confidence interval for the independence model.

Table 2
Simulation results for binary outcomes with $n = 100$ and $m_i = 20$ using the proposed Bayesian model.

	Bias	SMSE	pSD	SD	CR (%)
α_0	-0.047	0.248	0.286	0.233	95.9
α_1	0.063	0.232	0.229	0.223	92.6
η_{11}	0.036	0.240	0.244	0.237	93.0
η_{12}	-0.082	0.250	0.244	0.236	92.6
η_{13}	0.079	0.251	0.246	0.238	92.2
η_{14}	-0.055	0.252	0.242	0.246	91.7
η_{15}	-0.065	0.242	0.242	0.233	92.6
η_{16}	0.067	0.251	0.243	0.242	92.1
η_{21}	-0.085	0.253	0.245	0.239	93.6
η_{22}	0.089	0.263	0.240	0.248	91.0
η_{23}	-0.064	0.251	0.242	0.242	91.2
η_{24}	-0.066	0.241	0.246	0.232	93.9
η_{25}	0.061	0.252	0.242	0.245	92.5
η_{26}	0.080	0.238	0.247	0.224	93.7

$\max(\eta_j) = 0.1$, and $\sigma_0^2 = 0.02$. SMSE, square root of the mean squared error; pSD, mean of the posterior standard deviation; SD, Monte Carlo standard deviation of the posterior mean; CR, Monte Carlo coverage rate of 95% credible interval.

Table 3

Data analysis for the MAP II study: parameter estimates and their posterior standard deviations (pSD).

Parameter	Estimate (pSD)
α_0	-4.468 (1.495)
α_1	0.064 (0.028)
η_{11}	-0.500 (0.306)
η_{12}	-0.224 (0.810)
η_{13}	2.266 (0.774)
η_{21}	-0.268 (0.065)
η_{22}	-1.346 (0.482)
η_{23}	-0.852 (0.281)