Statistics in Medicine

WILEY

Statistics
in Medicine

RESEARCH ARTICLE

A New Algorithm for Sampling Parameters in a Structured Correlation Matrix With Application to Estimating Optimal Combinations of Muscles to Quantify Progression in Duchenne Muscular Dystrophy

Michael K. Kim[1] | Michael J. Daniels[1] | William D. Rooney[2] | Rebecca J. Willcocks[3] | Glenn A. Walter[4] | Krista H. Vandenborne[3]

[1]Department of Statistics, University of Florida, Gainesville, Florida, USA | [2]Advanced Imaging Research Center, Oregon Health and Science University, Portland, Oregon, USA | [3]Department of Physical Therapy, University of Florida, Gainesville, Florida, USA | [4]Department of Physiology and Aging, University of Florida, Gainesville, Florida, USA

Correspondence: Michael J. Daniels (daniels@ufl.edu)

Received: 30 October 2024 | **Revised:** 15 August 2025 | **Accepted:** 25 August 2025

Funding: This work was supported by the National Institutes of Health (Grant Nos. AR056973, HL158963).

Keywords: biomarkers | multivariate longitudinal data | positive definite matrix | structured correlation matrix

ABSTRACT

The goal of this paper is to estimate an optimal combination of biomarkers for individuals with Duchenne muscular dystrophy (DMD), which provides the most sensitive combinations of biomarkers to assess disease progression (in this case, optimal with respect to standardized response mean (SRM) for 4 muscle biomarkers). The biomarker data is incomplete (missing and irregular) multivariate longitudinal data. We propose a normal model with structured covariance designed for our setting. To sample from the posterior distribution of parameters, we develop a Markov Chain Monte Carlo (MCMC) algorithm to address the positive definiteness constraint on the structured correlation matrix. In particular, we propose a novel approach to compute the support of the parameters in the structured correlation matrix; we modify the approach from [1] on the set of the largest possible submatrices of the correlation matrix, where the correlation parameter is a unique element. For each posterior sample, we compute the optimal weights of our construct. We conduct data analysis and simulation studies to evaluate the algorithm and the frequentist properties of the posteriors of correlations and weights. We found that the lower extremities are the most responsive muscles at the early and late ambulatory disease stages, and the biceps brachii is the most responsive at the nonambulatory disease stage.

1 | Introduction

For Bayesian inference using multivariate normal (MVN) models with a structured correlation matrix, posterior computation can be challenging due to positive definiteness restrictions. We introduce an approach based on [1] to address this.

We use this model to address an important problem in Duchenne muscular dystrophy (DMD)—finding an optimal combination of biomarkers. The data are annualized changes of fat fraction (FF) of muscles (which are the biomarkers). We want to compute an optimal combination of these biomarkers to more precisely evaluate disease progression and potentially have more sensitive endpoints for clinical trials. The data available to do this is multivariate longitudinal at irregular time points with substantial missingness. We will use an MVN model with structured correlation to model this data and address the missingness. Previous unpublished work only allowed for complete data.

We first review relevant literature on covariance/correlation matrices. To ensure positive definiteness of the covariance matrix, multiple works have relied on unconstrained parameterizations of either the covariance or correlation matrix. Some earlier works regarding the covariance matrix include the Cholesky decomposition [2], the matrix logarithm [3, 4], or Givens angles [5, 6]. These techniques focus on estimation and do not allow interpretation of new parameters in terms of original variances and correlations. Later works try to address this issue, and one of the most popular examples is the modified Cholesky decomposition [7]. Their utility is limited by the ordering of variables, which is typically only natural for univariate longitudinal data. Other papers that exploit the ordering are [8–11].

There have also been works on unconstrained parameterization methods for the correlation matrix, which are of interest in Bayesian settings that deal with correlations and variances separately. A correlation matrix has the additional constraint of diagonal elements fixed at one. Pinheiro and Bates [12] introduce the spherical parameterization, which is conducted by first performing the Cholesky decomposition on the correlation matrix and then expressing the elements of the Cholesky factor as angles. More recent works expanded on this [13, 14]. Due to the Cholesky decomposition, the angles are dependent on ordering. The angles are defined on the support $(0, \pi]$ to ensure uniqueness and identifiability. Namely, there is a monotone relationship between the angles and the correlations—weaker correlations imply larger angles. For longitudinal data, this implies angles are increasing with time lag.

Zhang et al. [13] illustrate the applicability of the spherical parameterization for unbalanced longitudinal data by proposing a joint mean-variance-correlation generalized linear model. As spherical parameterization is dependent on the ordering of variables, their model only permits popular correlation structures for longitudinal data, such as compound symmetry or AR(1). This is problematic for complex correlated data with partial ordering information, including longitudinal data that has multiple outcomes measured repeatedly from the same subject (i.e., multivariate longitudinal data). Tsay and Pourahmadi [14] address this by showing that positive definiteness is guaranteed for structured correlation matrices using spherical parameterization alongside pivotal angles. By knowing the locations and estimates of pivotal angles, one can obtain the unique correlations and then the implied ("nonpivotal") angles row-by-row. Both papers use maximum likelihood estimation, as it achieves consistency and asymptotic normality. However, if we were to apply Bayesian computation on the angles, we would get different posteriors based on how the data was ordered, as typical priors on the angles are not invariant to ordering. Ghosh et al. [15] introduce shrinkage and selection priors for the angles, which are the inverse cosine of semi-partial correlations, but the priors are only applicable for ordered data. Similar to [13], Ghosh et al. [15] only consider a limited set of correlation structures: AR(1), banded, and block common.

Another unconstrained parameterization method for the correlation matrix uses the partial autocorrelations adjusted for intervening variables [16, 17]. Similar to the angles of hyperspherical parameterization, partial autocorrelations impose an order on the variables and have a one-to-one and recursive relationship with the correlations. This results in similar interpretability issues as before, but the authors do introduce priors for the correlation matrix that are independent of the order of indices. For application, they explore parsimonious modeling for balanced longitudinal (ordered) data with a focus on lags. More recent works take a different approach and apply the matrix logarithm on the correlation matrix [18, 19]. The transformation is one-to-one, invariant to ordering, and offers flexibility for parsimonious modeling and prior specification of unbalanced data. However, the interpretability of the new parameters is not intuitive.

A notable constrained approach for the covariance matrix is the linear covariance model [20], where the covariance matrix is the linear combination of known symmetric matrices and unknown coefficients. Constraints are put on the unknown coefficients, and the approach is applicable for any element-wise estimation of the covariance matrix. On the other hand, a constrained approach that directly models the correlation elements avoids issues with both interpretability and the unity diagonal constraint of the correlation matrix, both posing challenges for many unconstrained parameterization methods. The constrained Bayesian approach puts the positive definiteness constraint in each sampling step. Some earlier works are [1, 21, 22]. Barnard et al. [1] note that a correlation matrix stays positive definite (PD) if one were to replace any unique correlation element inside of it from an interval calculated using the values of all other correlation elements. They demonstrate the effectiveness of this approach with order-invariant priors. They use independent log-normal priors for standard deviations and either the marginally uniform priors for correlations or the jointly uniform prior for the correlation matrix. Wong et al. [21] estimate the covariance matrix of normal data by identifying zeroes in its inverse (see covariance selection models [23]), separating the inverse into a product of inverse partial variances and the matrix of partial correlations, and using "covariance selection" priors that allow zeroes in the precision matrix. There is a constraint on the prior for the matrix of partial correlations. The efficiency of their method and any identification of structure rely on the sparsity of the true precision matrix. Pitt et al. [24] is an extension, and Carter et al. [25] is a generalization of [21].

Similar to [1], Liechty et al. [22] put the positive definiteness constraint on the prior of the correlation matrix and compute intervals to sample correlation elements. They propose prior probability models that group marginal correlations into clusters based on similarities among correlations or variables. This results in group-structured correlation matrices, that is, block common. The "common correlation" prior allows shrinkage toward a diagonal correlation matrix. The "grouped correlation" and "grouped variables" mixture priors have flexibility in shrinkage toward target structures by using a point mass at zero or a small-variance distribution for a term in the mixture. Zhang et al. [26] is a recent work in a constrained Bayesian approach for unstructured and unordered correlation matrices that are functions of individual-level covariates, where each correlation element is specified by a linear model. They focus on how positive definiteness is ensured at different values of covariates. They intersect variations of the intervals defined in [1] to address the positive

definiteness constraint of the correlation matrices. We also intersect "Barnard" intervals, but based on submatrices in the setting of a structured correlation matrix.

We provide a flexible approach to modeling correlation matrices by generalizing the interval approach of [1] for a structured correlation matrix. In our application, it is a correlation matrix of multiple outcomes measured at each time point in an unbalanced longitudinal dataset with missingness. Our structure for the application assumes exchangeable time points, but the approach can be used to model any correlation structure without requiring an ordering of variables.

The paper is organized as follows: In Section 2, we introduce the DMD data and explain the clinical motivation of an optimal combination of biomarkers. In Section 3, we introduce the model and the structure of the correlation matrix. We also introduce the objective function to optimally choose weights of the construct. In Section 4, we provide details on the posterior computation. In particular, details on how to compute a "tight" PD interval of the candidate distribution of a correlation element by applying the interval approach of [1] on the set of the largest possible submatrices of the correlation matrix. In the simulation study of Section 5, we compare the acceptance and positive definiteness rates of correlations drawn from candidate distributions that use our generated PD intervals versus $(-1, 1)$ support. Section 5 also compares the computational performance and frequentist operating characteristics via simulations. Section 6 presents our findings on the DMD data, both clinically and algorithmically. Section 7 provides conclusions and extensions.

2 | Data and Motivation

The muscles of individuals with DMD are progressively replaced by fat, but at different rates, and these rates vary by individual, age, disease stage, and other factors. The disease stages we consider are defined by functional ability. In the early ambulatory stage, individuals can walk and get up from the floor without assistance. In the late ambulatory stage, individuals can walk but can no longer get up from the floor. In the nonambulatory stage, individuals cannot walk. The muscles of interest are two lower extremity muscles, soleus (SOL) and vastus lateralis (VL), and two upper extremity muscles, biceps brachii (BB) and deltoid (DEL). We use magnetic resonance spectroscopy FF measures of the muscles [27]. The specific measurement we use is the annualized change in FF from the current and previous visit of a subject, where visits are somewhere between 6 to 24 months apart. It is calculated as (FF at current visit − FF at previous visit)/(age at current visit − age at previous visit). For brevity, we use "muscle" to describe FF changes in a muscle between visits. Also, we use "time point" interchangeably with "measurement time."

The numbers of subjects (160 unique subjects) and visits vary by ambulatory disease stage. In total, early ambulatory data has 140 subjects and 419 measurement times, late ambulatory data has 74 subjects and 128 measurement times, and nonambulatory data has 51 subjects and 115 measurement times. The data exhibits missingness, as not all muscles are measured for each measurement time. Tables SB2, SB3, and SB4 in the Supporting Information provides details: Table SB2 provides the missingness by muscle, revealing significantly more missingness in the upper extremities than the lower extremities. Table SB3 provides the distribution of the number of missing muscle measurements at a given measurement time. Typically two muscles are missing at a given measurement time, particularly in earlier ambulatory stages. Table SB4 provides the distribution of the number of subject measurement times for the unbalanced DMD data. There is a maximum of 8, 6, and 7 subject measurement times for early, late, and nonambulatory data, respectively.

To account for the heterogeneity in the fat replacement rate of different muscles, we want a sensitive measure of overall muscle quality across different disease stages. We will construct an optimal combination of biomarkers that produces the most clinically meaningful and sensitive combination of FF measures of different muscles with varying weighting coefficients across a wide spectrum of disease stages. The inclusion of nonambulant subjects here is important. The majority of individuals with DMD are nonambulatory but are excluded from most clinical trials due to the inapplicability of traditional functional outcome measures [28].

3 | Model and Objective Function

3.1 | Model

For an individual with muscular dystrophy in a certain ambulatory disease stage, we denote the data as $y_{ij\ell}$, indicating the FF of subject i, measurement time j, and muscle ℓ. In total, we have N subjects, J measurement times, and L muscles. Let $p = J \times L$ be the total number of outcomes for a subject. The model below is for unbalanced longitudinal data, but for simplicity of notation, we will use J and p rather than J_i and p_i.

We model the data by an MVN distribution and assume exchangeable time points for a subject in each ambulatory disease stage. We employ the separation strategy on the covariance matrix [1]:

$$\boldsymbol{Y}_i \overset{\text{ind}}{\sim} \text{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
$$\boldsymbol{\Sigma} = \boldsymbol{SRS}$$

Due to the assumption of exchangeable time points, we put structure on the mean vector, the diagonal standard deviation matrix, and the correlation matrix as follows:

- $\boldsymbol{\mu}_{p\times 1} = (\tilde{\boldsymbol{\mu}}'_{L\times 1}, \ldots, \tilde{\boldsymbol{\mu}}'_{L\times 1})'$ is a mean vector with unique elements $\tilde{\boldsymbol{\mu}} = (\mu_1, \ldots, \mu_L)'$ for each measurement time j, where μ_ℓ is the mean of muscle ℓ, that is, $\text{E}[Y_{ij\ell}] = \mu_\ell$.
- $\boldsymbol{S}_{p\times p} = \text{diag}(\boldsymbol{s}'_{L\times 1}, \ldots, \boldsymbol{s}'_{L\times 1})$ is a diagonal standard deviation matrix with unique elements $\boldsymbol{s} = (s_1, \ldots, s_L)'$ for each measurement time j, where s_ℓ is the standard deviation of muscle ℓ, that is, $\text{SD}[Y_{ij\ell}] = s_\ell$.
- $\boldsymbol{R}_{p\times p}$ is a structured correlation matrix with unique elements

$$\boldsymbol{r} = (\underbrace{\eta_{12}, \ldots, \eta_{L-1,L}}_{\binom{L}{2}}, \underbrace{\rho_{(1)}, \ldots, \rho_{(L)}}_{L}, \gamma)' = (r_1, \ldots, r_q)', \text{ where}$$

- $\eta_{\ell\ell'}$ is the correlation between two different muscles ℓ and ℓ' observed at the same measurement time, that is, $\text{Corr}[Y_{ij\ell}, Y_{ij\ell'}] = \eta_{\ell\ell'}$. There are $\binom{L}{2}$ $\eta_{\ell\ell'}$s, and for each $\eta_{\ell\ell'}$, there are J instances of it in the upper-triangular portion of $\boldsymbol{R}$, so in total there are $JL(L-1)$ $\eta_{\ell\ell'}$s in $\boldsymbol{R}$.
- $\rho_{(\ell)}$ is the correlation between any two different measurement times for muscle ℓ, that is, $\text{Corr}[Y_{ij\ell}, Y_{ij'\ell}] = \rho_{(\ell)}$. There are L $\rho_{(\ell)}$s, and for each $\rho_{(\ell)}$, there are $\binom{J}{2}$ instances of it in the upper-triangular portion of $\boldsymbol{R}$, so in total there are $JL(J-1)$ $\rho_{(\ell)}$s in $\boldsymbol{R}$.
- γ is the correlation between two different muscles at two different measurement times, that is, $\text{Corr}[Y_{ij\ell}, Y_{ij'\ell'}] = \gamma$. There are $L^2J^2 - 2J\binom{L}{2} - 2L\binom{J}{2} - LJ$ or $JL(J-1)(L-1)$ γs in $\boldsymbol{R}$.
- The matrix has $q = \binom{L}{2} + L + 1 = 11$ unique parameters.

The structured correlation matrix for $L = 4$ muscles is as follows:

$$\boldsymbol{R} = \begin{bmatrix}
1 & \eta_{12} & \eta_{13} & \eta_{14} & \rho_{(1)} & \gamma & \gamma & \gamma & \rho_{(1)} & \gamma & \gamma & \gamma & \cdots \\
 & 1 & \eta_{23} & \eta_{24} & \gamma & \rho_{(2)} & \gamma & \gamma & \gamma & \rho_{(2)} & \gamma & \gamma & \cdots \\
 & & 1 & \eta_{34} & \gamma & \gamma & \rho_{(3)} & \gamma & \gamma & \gamma & \rho_{(3)} & \gamma & \cdots \\
 & & & 1 & \gamma & \gamma & \gamma & \rho_{(4)} & \gamma & \gamma & \gamma & \rho_{(4)} & \cdots \\
 & & & & 1 & \eta_{12} & \eta_{13} & \eta_{14} & \rho_{(1)} & \gamma & \gamma & \gamma & \cdots \\
 & & & & & 1 & \eta_{23} & \eta_{24} & \gamma & \rho_{(2)} & \gamma & \gamma & \cdots \\
 & & & & & & 1 & \eta_{34} & \gamma & \gamma & \rho_{(3)} & \gamma & \cdots \\
 & & & & & & & 1 & \gamma & \gamma & \gamma & \rho_{(4)} & \cdots \\
 & & & & & & & & 1 & \eta_{12} & \eta_{13} & \eta_{14} & \cdots \\
 & & & & & & & & & 1 & \eta_{23} & \eta_{24} & \cdots \\
 & & & & & & & & & & 1 & \eta_{34} & \cdots \\
 & & & & & & & & & & & 1 & \cdots \\
\vdots & & & & & & & & & & & &
\end{bmatrix}$$

Given the objective of accounting for correlation within disease stage windows, we assume exchangeable time points for simplicity. But we could easily accommodate time-series structures, including autoregressive structures.

3.2 | Optimal Construct

For an individual with muscular dystrophy at a certain ambulatory disease stage, we compute an optimal convex combination $\tilde{y}_{ij} = \sum_{\ell=1}^{L} w_\ell y_{ij\ell} = \boldsymbol{w}'\boldsymbol{y}_{ij}$, where $\sum_{\ell=1}^{L} w_\ell = 1$, $w_\ell \geq 0$ $\forall \ell$. The weights $\boldsymbol{w}$ are chosen to maximize the standardized response mean (SRM) of the construct, where $\boldsymbol{\Sigma}_\eta = \text{Var}[\boldsymbol{Y}_{ij}]$ is the $L \times L$ covariance matrix containing the correlation parameters, $\eta_{\ell\ell'}$s:

$$\boldsymbol{w} = \underset{\boldsymbol{w}}{\text{argmax}} \text{SRM}[\tilde{Y}_{ij}] = \underset{\boldsymbol{w}}{\text{argmax}} \frac{\text{E}[\tilde{Y}_{ij}]}{\text{SD}[\tilde{Y}_{ij}]} = \underset{\boldsymbol{w}}{\text{argmax}} \frac{\boldsymbol{w}'\tilde{\boldsymbol{\mu}}}{\sqrt{\boldsymbol{w}'\boldsymbol{\Sigma}_\eta \boldsymbol{w}}} \tag{1}$$

We note that the standard deviation of the construct does not vary across time, but the weights are optimized for each ambulatory disease stage. These weights provide the largest mean of convex combinations of 1-year changes in FF of muscles with respect to variability in the changes for each disease stage.

4 | Posterior Computation

Full details of the Markov Chain Monte Carlo (MCMC) algorithm for posterior computation are presented in Section SA.1 of the Supporting Information. Here, we focus on our new approach for sampling the correlations and computing the optimal biomarker constructs.

4.1 | Sampling the Structured Correlation Parameters

In a Metropolis-Hastings (M-H) step, when the correlation matrix has dimension greater than 2, a correlation candidate may produce a non-PD correlation matrix if we were to sample from a candidate distribution with a $(-1, 1)$ support. To mitigate this, the support of the candidate distribution should be shrunk to a tighter interval. Barnard et al. [1] introduced an approach that computes an interval for a unique correlation element inside of a correlation matrix by solving a quadratic equation using determinants of augmented correlation matrices (details in Section SA.4 of the Supporting Information). The correlation matrix remains PD as long as the value of the correlation element lies within the interval. Their approach works for any element inside of an unstructured correlation matrix. We will refer to their approach as the "Barnard approach" and the interval computed from the approach as the "PD interval."

We adapt the Barnard approach on a structured correlation matrix $\boldsymbol{R}$ with unique elements $(r_1, \ldots, r_q)'$, which are fewer than those in an unstructured correlation matrix. We first create the set of all the largest possible submatrices of $\boldsymbol{R}$, where each of these submatrices contains the kth correlation $r_k \in (r_1, \ldots, r_q)'$ only once, and the rest of the submatrix is filled with current values of correlation elements other than the kth correlation $\boldsymbol{r}_{(-k)}^{(t-1)}$. After computing the PD interval for each submatrix, we define the support of the candidate distribution of $r_k^{(t)}$ as the intersection of all the PD intervals. If there is only one largest possible submatrix, then we use its PD interval as the support. We denote the support as $(L_k^{(t)}, U_k^{(t)})$. Our approach improves the rate at which a correlation candidate results in a PD $\boldsymbol{R}$.

To find the largest possible submatrix, the elements of the correlation matrix need to be reordered. In particular, we can build a largest possible submatrix for r_k with respect to *any* structured correlation matrix by selecting a subset of the original outcomes $\hat{\boldsymbol{y}} \subset \boldsymbol{y} = (y_1, \ldots, y_p)'$ using Algorithm SA1 in the Supporting Information, which is summarized below:

1. The first two outcomes in the subset should correspond to correlation r_k.
2. Under the condition that r_k remains as a unique element inside the submatrix, check one-by-one whether the remaining outcomes can be added to the subset.

To compute the tightest (intersected) PD interval, we use all combinations of the largest possible submatrix for r_k. We can apply Algorithm SA1 in the Supporting Information on all permutations of the ordering of outcomes. This can be computationally expensive and can result in duplicate PD intervals.

To save computational time for our structure specified in Section 3.1, we devised individualized algorithms for $r_1, \ldots, r_q$, efficiently generating all combinations of the largest possible submatrix for each r_k (Section SA.5 of the Supporting Information). An alternative to further reduce computational cost is to compute the PD interval of only one submatrix for each r_k, particularly when the submatrix has a dimension close to that of the correlation matrix. We explore this option in Section 5.

For the candidate distribution, we recommend two choices. First, a uniform distribution on the derived PD interval, $(L_k^{(t)}, U_k^{(t)})$,

$$r_k^{(t)} \sim \text{Unif}(L_k^{(t)}, U_k^{(t)}) \tag{2}$$

Second, a beta distribution, shifted and scaled to the PD interval, $(L_k^{(t)}, U_k^{(t)})$, whose mode is fixed at the value of the previous iteration, $r_k^{(t-1)}$,

$$\begin{aligned} r_k^{(t)} &\sim \text{R-Beta}(\alpha_k^{(t)}, \beta_k^{(t)}, L_k^{(t)}, U_k^{(t)}), \text{ where} \\ \alpha_k^{(t)} &= \frac{(\kappa_k - 1)L_k^{(t)} + (2 - \kappa_k)r_k^{(t-1)} - U_k^{(t)}}{L_k^{(t)} - U_k^{(t)}} \\ \beta_k^{(t)} &= \frac{(\kappa_k - 1)U_k^{(t)} + (2 - \kappa_k)r_k^{(t-1)} - L_k^{(t)}}{U_k^{(t)} - L_k^{(t)}} = \kappa_k - \alpha_k^{(t)} \end{aligned} \tag{3}$$

and with concentration (and tuning) parameter $\kappa_k > 2$. We will refer to this distribution as the reparameterized-beta distribution.

If the reparameterized-beta distribution is used, we can tighten the interval we draw $r_k^{(t)}$ from by setting κ_k to a higher value, which makes the candidate distribution be more centered around the previous value $r_k^{(t-1)}$. Doing this will increase the percentage of times that $\boldsymbol{R}$ is PD throughout the M-H for r_k, since we would be reducing the possible candidates that would be far away from $r_k^{(t-1)}$ in which it is ensured to produce a PD $\boldsymbol{R}$. In turn, the acceptance rate for r_k will also increase, since the prior forces the M-H algorithm to automatically reject an $r_k^{(t)}$ that would produce a non-PD $\boldsymbol{R}$. Candidate distributions like the reparameterized-beta distribution allow the researcher to achieve the standard 25% M-H acceptance rate for r_k by adjusting κ_k [29]. However, there will be more potential autocorrelation in a chain given the centering on the previous iteration value. Candidate distributions like the uniform distribution allow for a more even exploration of correlation candidates on $(L_k^{(t)}, U_k^{(t)})$ that would not be directly influenced by $r_k^{(t-1)}$.

4.2 | Computing the Posterior Distribution of the Weights

The SRM of $\tilde{Y}_{ij}$ is a function of parameters $\tilde{\boldsymbol{\mu}}$ and $\boldsymbol{\Sigma}_\eta$ in Equation (1). Therefore, we compute the posterior sample of the weights by maximizing the SRM for each posterior draw of the model parameters defined in Section 3.1. A standard approach to compute a posterior point estimate for weights on a simplex is to take the element-wise posterior mean. However, the posterior distributions of the weights were occasionally skewed and/or bimodal with peaks at zero. To address this, we defined our posterior point estimate as the set of weights that maximizes the SRM when plugging in the posterior medians of the model parameters from Section 3.1. This summary satisfies the simplex constraint and avoids both disproportionate influence from sharp modal peaks and the loss of substantial posterior mass (e.g., maximum a posteriori estimates).

5 | Simulations

We conduct simulations to assess frequentist operating characteristics of the optimal weights and SRM. In particular, we evaluate the coverages of 95% credible intervals, biases, and mean squared errors (MSEs) of the weights and SRM using simulated data.

We also do simulations to examine the acceptance and positive definiteness rates of correlations drawn using our proposed algorithm.

To generate the simulated data for each ambulatory disease stage, we use the posterior medians of the model parameters for the DMD data (Table SB18 in the Supporting Information) as the truth. We generate 500 simulated datasets for $N = 100$ subjects and $J = 4$ time points. We consider four true distributions: normal (Section 3.1), t_{10} [30], t_3, and skew-normal with shape (skewness) parameter of 0.1 [31]. We also generate simulated data with missingness using the normal model. We emulate the column-wise missingness and row-wise missingness patterns of the DMD data (Tables SB2 and SB3 in the Supporting Information) by using $(0.05, 0.05, 0.75, 0.75)'$ and $(0.1, 0.6, 0.1, 0.2)'$, respectively.

For each simulated dataset, we run 50000 iterations with 1000 burn-in. Unless otherwise noted, posterior results assume that correlations were sampled using an M-H algorithm with candidate distribution, reparameterized-beta on the PD interval defined in Equation (3). We will refer to this distribution as R-Beta(L, U). We will refer to Equation (2) as Unif(L, U).

5.1 | Coverages, Biases, and MSEs of Weights and SRM

Coverages of 95% credible intervals of weights and SRM are given in Table 1. The coverages for normal (with missingness) are close to those for normal. For t_{10}, the coverages of weights are close to those for normal, but there is some undercoverage with SRM. For t_3, there is considerable undercoverage with the weights due to the distribution's heavy tails, but there is overcoverage for true weights equal to zero. For skew-normal, there is mostly undercoverage with the weights and severe undercoverage with SRM.

Biases of weights and SRM are given in Table 2. The magnitudes of the biases are relatively small with respect to the weights $\in (0, 1)$ and SRM $\in (0.8, 1.4)$. Biases of weights slightly increase from normal to t_{10}, moderately increase from t_{10} to t_3, and considerably increase from t_3 to skew-normal. Biases of SRM slightly increases from normal to t_{10}, then considerably increases from t_{10} to t_3 and from t_3 to skew-normal.

TABLE 1 | Coverages of 95% credible intervals of weights and SRM.

Coverage	w_1	w_2	w_3	w_4	SRM
Normal – no miss – early ambulatory	95.2%	96.2%	95.6%	100%	94.8%
Normal – no miss – late ambulatory	95.6%	95.4%	95.4%	95.0%	94.8%
Normal – no miss – nonambulatory	95.6%	96.8%	96.0%	99.6%	95.0%
t_{10} – no miss – early ambulatory	91.6%	93.4%	94.2%	100%	92.6%
t_{10} – no miss – late ambulatory	93.8%	93.4%	93.0%	95.2%	90.6%
t_{10} – no miss – nonambulatory	94.0%	94.2%	93.8%	99.4%	91.6%
t_3 – no miss – early ambulatory	77.2%	76.2%	78.2%	99.2%	56.6%
t_3 – no miss – late ambulatory	74.6%	72.8%	78.6%	78.6%	55.0%
t_3 – no miss – nonambulatory	75.8%	76.8%	75.0%	96.2%	50.8%
Skew-normal – no miss – early ambulatory	88.6%	73.8%	93.4%	99.8%	19.0%
Skew-normal – no miss – late ambulatory	89.2%	94.8%	95.0%	90.6%	20.0%
Skew-normal – no miss – nonambulatory	96.0%	93.8%	82.2%	97.2%	39.2%
Normal – yes miss – early ambulatory	95.8%	95.8%	95.2%	100%	95.2%
Normal – yes miss – late ambulatory	94.0%	94.8%	93.8%	94.6%	94.2%
Normal – yes miss – nonambulatory	95.2%	96.4%	94.8%	99.0%	94.2%

TABLE 2 | Biases of weights and SRM.

Bias	w_1	w_2	w_3	w_4	SRM
Normal – no miss – early ambulatory	0.00	0.00	−0.00	0.00	0.01
Normal – no miss – late ambulatory	−0.00	−0.00	−0.00	0.01	0.00
Normal – no miss – nonambulatory	−0.00	−0.00	−0.00	0.01	−0.00
t_{10} – no miss – early ambulatory	0.01	−0.01	−0.00	0.00	0.01
t_{10} – no miss – late ambulatory	0.00	−0.00	−0.01	0.01	0.00
t_{10} – no miss – nonambulatory	−0.00	0.00	−0.01	0.01	−0.01
t_3 – no miss – early ambulatory	−0.01	0.01	−0.00	0.01	0.08
t_3 – no miss – late ambulatory	−0.00	−0.00	−0.01	0.02	0.08
t_3 – no miss – nonambulatory	−0.02	0.01	−0.01	0.02	0.09
Skew-normal – no miss – early ambulatory	0.06	−0.08	0.02	0.00	0.25
Skew-normal – no miss – late ambulatory	−0.03	−0.01	0.00	0.04	0.21
Skew-normal – no miss – nonambulatory	−0.00	0.02	−0.04	0.02	0.15
Normal – yes miss – early ambulatory	−0.00	0.00	−0.01	0.01	0.01
Normal – yes miss – late ambulatory	−0.01	−0.00	−0.01	0.03	0.01
Normal – yes miss – nonambulatory	−0.01	0.00	−0.02	0.03	0.00

Root MSEs of weights and SRM are given in Table SB5 in the Supporting Information. Root MSEs of weights generally have moderate values for normal, t_{10}, and skew-normal. MSEs of weights for t_3 and MSEs of SRM for t_3 and skew-normal are large. Some of the smallest MSEs are for true weights equal to zero. It is expected that biases and MSEs for simulated data with missingness are slightly larger than those for simulated data without missingness. For SRM, there is a noticeable decrease in coverage and increase in bias and MSE from t_{10} to t_3 and from t_3 to skew-normal. The poorer frequentist operating characteristics are expected for a mis-specified observed data model.

5.2 | Acceptance and Positive Definiteness Rates of Drawn Correlations

This study is conducted with the normal distribution scenario and no missingness. It is meant to evaluate the proposed approach for sampling correlation parameters of the structured correlation matrix specified in Section 3.1. We compare the M-H acceptance and positive definiteness rates of correlations drawn from R-Beta(L, U), R-Beta(L_1, U_1), R-Beta$(-1, 1)$, Unif(L, U), Unif(L_1, U_1), and Unif$(-1, 1)$, where (L_1, U_1) is the PD interval for one randomly selected submatrix. Note that

$(L, U) = (L_1, U_1)$ for $\rho_{(\ell)}$, $\ell = 1, \ldots, L$ since there is only one variation of the largest possible submatrix for $\rho_{(\ell)}$ (explanation in Section SA.5 of the Supporting Information). We average the M-H acceptance rates across the 500 datasets.

The M-H acceptance rates of correlations drawn from different candidate distributions are given in Tables SB6, SB7, and SB8 in the Supporting Information. For the correlations drawn from a reparameterized-beta distribution, acceptance rates are between 25% to 26% after adjusting the tuning parameters. For correlations drawn from Unif(L, U), the acceptance rates are between 5.4% and 8.0% for the ηs, between 4.5% and 7.6% for the ρs, and less than or equal to 2.3% for γ. For correlations drawn from Unif(L_1, U_1), the acceptance rates are between 5.2% and 7.9% for the ηs and 2.0% for γ. For the correlations drawn from Unif$(-1, 1)$, the acceptance rates are between 4.0% and 6.4% for the ηs, between 3.6% and 5.8% for the ρs, and less than or equal to 1.9% for γ. These results are as expected.

Next, we consider the percentage of times a correlation candidate resulted in a PD correlation matrix. We average the positive definiteness rates across the 500 datasets. The positive definiteness rates of correlations drawn from different candidate distributions are given in Tables SB9, SB10, and SB11 in the Supporting Information. In terms of positive definiteness rates, correlations drawn from R-Beta(L, U) are the highest and Unif$(-1, 1)$ the lowest as expected. When ηs are drawn from a reparameterized-beta distribution or when ρs are drawn from Unif(L, U), their positive definiteness rates are consistent across the ambulatory disease stages.

It is apparent from Tables SB9, SB10, and SB11 that tuning the reparameterized-beta distribution to achieve 25% acceptance rates in the M-H has a noticeable impact on the positive definiteness rates of the correlations. This is evident by the smaller differences in the positive definiteness rates between R-Beta(L, U) and R-Beta$(-1, 1)$, compared to the corresponding differences between Unif(L, U) and Unif$(-1, 1)$. Using one randomly selected submatrix to compute the PD interval (L_1, U_1) is computationally efficient and yields competitive but smaller positive definiteness rates than the intersected PD interval (L, U). Unlike the uniform distribution, the reparameterized-beta distribution is centered on the previous iteration value of a correlation element, a value shaped by the prior to guarantee a PD correlation matrix. Therefore, tightening the support of the reparameterized-beta distribution typically increases autocorrelation between samples across iterations.

We note that a resulting structured correlation matrix from a correlation candidate is more likely to be PD if the correlation has a lower count of instances in the matrix. For example, in our structured correlation matrix $\boldsymbol{R}$, there are $2J$ instances of an $\eta_{\ell\ell'}$, $J(J-1)$ instances of a $\rho_{(\ell)}$, and $JL(J-1)(L-1)$ instances of γ. So if we were to draw correlations from Unif$(-1, 1)$, that is, without influence from PD intervals or tuning parameters, we expect γ to have the lowest rate of producing a PD $\boldsymbol{R}$, and for $J > 3$, we expect ηs to have higher rates than ρs. This intuition holds in Tables SB9, SB10, and SB11 when correlations are drawn from a uniform distribution. In the case of the reparameterized-beta distribution, the positive definiteness rates of some ρs are slightly smaller than that of γ due to tuning parameters.

Some further intuition concerning the positive definiteness rates of correlations drawn from candidate distributions supported on the PD interval is that the dimension of the largest possible submatrix for $\rho_{(\ell)}$ is 1 greater than that for $\eta_{\ell\ell'}$, but the submatrix for $\rho_{(\ell)}$ has no other variations, whereas the submatrix for $\eta_{\ell\ell'}$ has J variations and thus J PD intervals to intersect for the support of the $\eta_{\ell\ell'}$ candidate. And while the dimension of the largest possible submatrix for γ is only 3, there are $2\binom{L}{2}$ PD intervals to intersect for the support of the γ candidate.

6 | Analysis of DMD Biomarkers

We use our model to make inference on optimal combination of biomarkers in the DMD data with respect to SRM. Recall from Section 2, we have four muscles of interest, and we measure their annualized FF change between visits. We focus on two lower extremity muscles, SOL and VL, and two upper extremity muscles, BB and DEL. The data is stratified into three ambulatory disease stages.

6.1 | Computations

We run 4 MCMC chains, each with 75000 iterations and 1000 burn-in. The M-H acceptances rates of correlations drawn from different candidate distributions are given in Tables SB12, SB13, and SB14 in the Supporting Information. Across ambulatory disease stages, there is more variation among the acceptance rates of correlations here than those from the simulations (Section 5.2). This is most likely because the simulated data is balanced without missingness and has a relatively small number of time points $J = 4$. The DMD data is unbalanced with missingness and has varying distributions of the number of subject measurement times for different ambulatory disease stages (Tables SB2, SB3, and SB4 in the Supporting Information).

Next, the positive definiteness rates of correlations drawn from different candidate distributions are given in Tables SB15, SB16, and SB17 in the Supporting Information. As explained in Section 5.2, when drawing from Unif$(-1, 1)$, we expect γ to have the lowest rate of producing a PD $\boldsymbol{R}$, and for $J > 3$, we expect ηs to have higher rates than ρs. These expectations are largely supported by the results in Tables SB15, SB16, and SB17, including for correlations drawn from R-Beta(L, U) or Unif(L, U). The only exceptions occur in the late ambulatory data, where not all positive definiteness rates of ηs exceed those of ρs under Unif$(-1, 1)$ (and R-Beta(L, U) to a much lesser extent). Differences in positive definiteness rates between ηs and ρs may be less pronounced in the late ambulatory group because 54.1% of subjects had only one total measurement time, compared to 25.7% for early ambulatory and 39.2% for nonambulatory subjects (Table SB4).

The positive definiteness rates are lower here than those from the simulations (Section 5.2). While the simulated data has a relatively small number of time points $J = 4$, the maximum of subject measurement times for the DMD data is $J = 8, 6, 7$ (early, late, non, respectively). The DMD data is also incomplete as opposed to the simulated data (Tables SB2, SB3, and SB4).

6.2 | Posteriors of Model Parameters

Posterior distributions of model parameters $\tilde{\boldsymbol{\mu}}$, $\boldsymbol{s}$, and $\boldsymbol{r}$ are summarized using 95% credible intervals and posterior medians. These are given in Tables SB19, SB20, SB21, SB22, SB23, SB24, SB25, SB26, and SB27 in the Supporting Information.

The means of lower extremities increase from early to late ambulatory and then decrease from late to nonambulatory, whereas the means of upper extremities increase from early to nonambulatory. Note however that DEL shows little to no change from late to nonambulatory. The means of upper extremities become greater than the means of lower extremities by the nonambulatory disease stage. These patterns suggest that DMD affects the lower extremities of individuals more (larger average annualized changes of FF) for early and late ambulatory disease stages. By the time they cannot walk, DMD affects the upper extremities more, particularly BB.

The standard deviations increase from early to nonambulatory, that is, there is more variability in the annualized changes of FF in the muscles for later ambulatory disease stages. From early to nonambulatory, VL has the highest standard deviation, and SOL has the lowest standard deviation.

Most of the correlations are positive, and they decrease as the individuals' ability to walk deteriorates. This is particularly apparent when comparing just the early and nonambulatory disease stages. Generally, there are less relations between two muscles or two measurement times at later ambulatory disease stages in terms of annualized changes of FF in the muscles. Note that $\rho_{(1)}$, $\rho_{(2)}$, $\rho_{(3)}$, γ have near 0 correlations at later ambulatory disease stages, but $\rho_{(4)}$ keeps a high correlation relative to other temporal correlations at later ambulatory disease stages. Between different measurement times, the annualized changes of FF become more similar for a lower extremity by the late ambulatory disease stage and for BB by the nonambulatory disease stage. A possible explanation is that by the late ambulatory disease stage, DMD has progressed in the lower extremities to the point that there is a similar amount of FF replacement between different measurement times, and likewise for BB by the nonambulatory disease stage. This is consistent with DMD targeting the lower extremities first.

6.3 | Posteriors of Weights and SRM

Table 3 provides the point estimate of the weights, which is the set of weights that optimize the SRM, by plugging in the posterior medians of the model parameters (Table SB18 in the Supporting Information). Posterior density is given in Figure SB1 in the Supporting Information. Credible intervals are given in Tables SB28, SB29, and SB30 in the Supporting Information. The estimates of the individual muscle SRMs, the SRM using optimal weights (SRM_{opt}), and the SRM using equal weights $\boldsymbol{w}_{\text{equal}} = (0.25, 0.25, 0.25, 0.25)'$ ($\text{SRM}_{\text{equal}}$) are also given in Table 3. $\text{SRM}_{\text{equal}}$ uses the same model parameters as SRM_{opt}. Note that SRM_{opt} and $\text{SRM}_{\text{equal}}$ account for correlations between muscles unlike the individual muscle SRMs.

SOL weight increases from early to late ambulatory and then decreases from late to nonambulatory. VL weight decreases from early to nonambulatory. BB weight very slightly decreases from early to late ambulatory and then greatly increases from late to nonambulatory. DEL weight is zero or close to zero across the ambulatory disease stages which may be explained by its relatively high correlation with other muscles, particularly SOL and BB (Table SB18). Among the four muscles, the lower extremities and BB are the most responsive muscles to FF replacement across all ambulatory disease stages. As the individuals lose their ability to walk, BB becomes increasingly more responsive and eventually contributes the most weight by the nonambulatory disease stage. At early ambulatory, VL contributes more weight than SOL, but this eventually shifts by late ambulatory and persists to nonambulatory.

TABLE 3 | Point estimate of weights, individual muscle SRMs, SRM from optimal weights, SRM from equal weights.

	Early ambulatory	Late ambulatory	Nonambulatory
w_1[SOL]	0.290	0.459	0.358
w_2[VL]	0.465	0.218	0.051
w_3[BB]	0.245	0.235	0.591
w_4[DEL]	0	0.088	0
SRM_{SOL}	0.706	0.957	0.706
SRM_{VL}	0.913	0.804	0.335
SRM_{BB}	0.563	0.687	1.010
SRM_{DEL}	0.368	0.770	0.567
SRM_{opt}	0.985	1.183	1.144
$\text{SRM}_{\text{equal}}$	0.901	1.137	0.940

TABLE 4 | Point estimate of weights, SRM from optimal weights, SRM from equal weights (for model with only SOL and VL).

	Early ambulatory	Late ambulatory	Nonambulatory
w_1[SOL]	0.392	0.665	0.936
w_2[VL]	0.608	0.335	0.064
SRM_{opt}	0.935	1.098	0.708
$\text{SRM}_{\text{equal}}$	0.932	1.065	0.579

The trends for individual muscle SRMs across the ambulatory disease stages are similar to the observed trends from before for the point estimate of the weights. As expected, SRM_{opt} is greater than any individual muscle SRM and $\text{SRM}_{\text{equal}}$. We note that VL SRM is greater than $\text{SRM}_{\text{equal}}$ at early ambulatory, and BB SRM is greater than $\text{SRM}_{\text{equal}}$ at nonambulatory.

Table 4 provides corresponding estimates as in Table 3 if the model only considers SOL and VL. Individual muscle SRMs for SOL and VL are the same as in Table 3. We find similar trends for the point estimate of the weights with the exception that SOL shows only an increasing trend from early to nonambulatory since no upper extremities are present. The SRM from optimal weights are smaller across the ambulatory disease stages when the model only considers lower extremities, particularly at nonambulatory.

We visualize the joint posterior distribution of the weights that live on the simplex by plotting each set of weights as a point on a tetrahedron and then using a three-dimensional version of the boxplot called `gemPlot` (`R` package) [32]. See Section SB.3.1 of the Supporting Information for further details and figures.

7 | Discussion

Our objective was to estimate an optimal combination of biomarkers in order to assess disease progression of individuals with DMD. To do this, we modeled the data using an MVN distribution with a structured correlation matrix to account for the high rate of missingness. For Bayesian inference, we addressed the positive definiteness constraint of a structured correlation matrix by proposing a generalization of the interval approach in [1]. In particular, we adapted the Barnard approach to the set of the largest possible submatrices of a structured correlation matrix, where a target correlation parameter is a unique element inside each submatrix. This procedure computes a "tight" interval for the support of a correlation parameter. We provided detailed rationale and specific algorithms on how to build the largest possible submatrices in the context of our parameterization. We note that our approach can be used for any correlation structure without necessitating an ordering of variables including partial time series structures. For example, to replace the exchangeable structure with an autoregressive order one, we would replace γ with $\gamma^{|\text{lag}|}$

Our approach for the correlation structure in the MVN model could also be used for a t or skew-normal model. This was demonstrated in the simulations where we fit the MVN model for datasets generated by t_{10}, t_3, and skew-normal distributions. MSEs and coverages of credible intervals were as expected and magnitudes of biases were small for the weights and SRM of the construct. The simulations also showed that the positive definiteness and acceptance rates of the correlation parameters in MCMC were improved when using our approach. We compared different candidate distributions with tightened supports for the correlations parameters against Unif$(-1, 1)$.

For each posterior sample of the model parameters, we computed optimal weights for the construct to evaluate DMD progression across different ambulatory disease stages. We also demonstrated how to visualize the joint posterior of the weights on the simplex. We found that at the early and late ambulatory disease stages, the lower extremities were the most responsive muscles. After the individuals lost their ability to walk, biceps brachii became the most responsive muscle. The posterior means and SRMs of individual muscles also support our findings. As the disease progressed, variability of individual muscles increased, whereas correlations between muscles or measurement times generally decreased. We note that the deltoid muscle had a weight close to zero and has a relatively high correlation with other muscles across all disease stages. FF of the deltoid is difficult to measure using magnetic resonance spectroscopy due to heterogeneous distribution of fat across the muscle, making it difficult to capture a representative sample in a rectangular voxel.

We note that the constructs can vary with other disease characteristics besides ambulatory status, including other functional milestones or continuous outcomes, for example, 6MWD (6-minute walking distance); we are currently working on this extension. We will also explore alternative objective functions besides the SRM to optimize the weights. In addition, our approach to compute "almost" PD intervals can be used for concentration matrices. Finally, we could consider a "composite" proposal, that first samples from the reparameterized-beta and then samples from the uniform based on our intervals to increase the acceptance rate and allow for a more efficient exploration of the posterior.

Acknowledgments

All co-authors were partially supported by either NIH R01 AR056973 and/or Wellstone grant P50 (P50AR052646). Daniels was also partially supported by NIH R01 HL158963. MRI data was collected in the McKnight Brain Institute at the National High Magnetic Field Laboratory's Advanced Magnetic Resonance Imaging and Spectroscopy (AMRIS) Facility, which is supported by National Science Foundation Cooperative Agreement No. DMR-1644779 and DMR-1157490 and the State of Florida, and in OHSU's Advanced Imaging Research Center, supported by NIH S10OD021701 for the 3T Siemens Prisma MRI instrument. Two instrument grants at OHSU also supported the data collection: NIH S10-OD018224 and S10-OD016356.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Research data are not shared.

References

1. J. Barnard, R. McCulloch, and X. L. Meng, "Modeling Covariance Matrices in Terms of Standard Deviations and Correlations, With Application to Shrinkage," *Statistica Sinica* 10 (2000): 1281–1311.

2. M. J. Lindstrom and D. M. Bates, "Newtonâeuro Raphson and EM Algorithms for Linear Mixed-Effects Models for Repeated-Measures Data," *Journal of the American Statistical Association* 83, no. 404 (1988): 1014–1022.

3. T. Leonard and J. S. Hsu, "Bayesian Inference for a Covariance Matrix," *Annals of Statistics* 20, no. 4 (1992): 1669–1696.

4. T. Y. Chiu, T. Leonard, and K. W. Tsui, "The Matrix-Logarithmic Covariance Model," *Journal of the American Statistical Association* 91 (1996): 198–210.

5. R. Yang and J. O. Berger, "Estimation of a Covariance Matrix Using the Reference Prior," *Annals of Statistics* 22 (1994): 1195–1211.

6. M. J. Daniels and R. E. Kass, "Non-Conjugate Bayesian Estimation of Covariance Matrices and Its Use in Hierarchical Models," *Journal of the American Statistical Association* 94, no. 448 (1999): 1254–1263.

7. M. Pourahmadi, "Joint Mean-Covariance Models With Applications to Longitudinal Data: Unconstrained Parameterisation," *Biometrika* 86, no. 3 (1999): 677–690.

8. J. Pan and G. Mackenzie, "On Modelling Mean-Covariance Structures in Longitudinal Studies," *Biometrika* 90, no. 1 (2003): 239–244.

9. J. Pan and G. MacKenzie, "Regression Models for Covariance Structures in Longitudinal Studies," *Statistical Modelling* 6, no. 1 (2006): 43–57.

10. C. Leng, W. Zhang, and J. Pan, "Semiparametric Mean–Covariance Regression Analysis for Longitudinal Data," *Journal of the American Statistical Association* 105, no. 489 (2010): 181–193.

11. W. Zhang and C. Leng, "A Moving Average Cholesky Factor Model in Covariance Modelling for Longitudinal Data," *Biometrika* 99, no. 1 (2012): 141–150.

12. J. C. Pinheiro and D. M. Bates, "Unconstrained Parametrizations for Variance-Covariance Matrices," *Statistics and Computing* 6 (1996): 289–296.

13. W. Zhang, C. Leng, and C. Y. Tang, "A Joint Modelling Approach for Longitudinal Studies," *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 77, no. 1 (2015): 219–238.

14. R. S. Tsay and M. Pourahmadi, "Modelling Structured Correlation Matrices," *Biometrika* 104, no. 1 (2017): 237–242.

15. R. P. Ghosh, B. Mallick, and M. Pourahmadi, "Bayesian Estimation of Correlation Matrices of Longitudinal Data," *Bayesian Analysis* 16, no. 3 (2021): 1039–1058.

16. M. J. Daniels and M. Pourahmadi, "Modeling covariance matrices via partial autocorrelations," *Journal of Multivariate Analysis* 100, no. 10 (2009): 2352–2363.

17. Y. Wang and M. J. Daniels, "Bayesian Modeling of the Dependence in Longitudinal Data via Partial Autocorrelations and Marginal Variances," *Journal of Multivariate Analysis* 116 (2013): 130–140.

18. I. Archakov and P. R. Hansen, "A New Parametrization of Correlation Matrices," *Econometrica* 89, no. 4 (2021): 1699–1715.

19. J. Hu, Y. Chen, C. Leng, and C. Y. Tang, "Regression Analysis of Correlations for Correlated Data," arXiv preprint arXiv:2109.05861 (2021).

20. T. W. Anderson, "Asymptotically Efficient Estimation of Covariance Matrices With Linear Structure," *Annals of Statistics* 1, no. 1 (1973): 135–141.

21. F. Wong, C. K. Carter, and R. Kohn, "Efficient Estimation of Covariance Selection Models," *Biometrika* 90, no. 4 (2003): 809–830.

22. J. C. Liechty, M. W. Liechty, and P. Müller, "Bayesian Correlation Estimation," *Biometrika* 91, no. 1 (2004): 1–14.

23. A. P. Dempster, "Covariance Selection," *Biometrics* 28 (1972): 157–175.

24. M. Pitt, D. Chan, and R. Kohn, "Efficient Bayesian Inference for Gaussian Copula Regression Models," *Biometrika* 93, no. 3 (2006): 537–554.

25. C. K. Carter, F. Wong, and R. Kohn, "Constructing Priors Based on Model Size for Nondecomposable Gaussian Graphical Models: A Simulation Based Approach," *Journal of Multivariate Analysis* 102, no. 5 (2011): 871–883.

26. S. Zhang, J. Kuha, and F. Steele, "Modelling Correlation Matrices in Multivariate Dyadic Data: Latent Variable Models for Intergenerational Exchanges of Family Support," arXiv preprint arXiv:2210.14751 (2022).

27. S. C. Forbes, H. Arora, R. J. Willcocks, et al., "Upper and Lower Extremities in Duchenne Muscular Dystrophy Evaluated With Quantitative MRI and Proton MR Spectroscopy in a Multicenter Cohort," *Radiology* 295, no. 3 (2020): 616–625.

28. I. E. Verhaart and A. Aartsma-Rus, "Therapeutic Developments for Duchenne Muscular Dystrophy," *Nature Reviews Neurology* 15, no. 7 (2019): 373–386.

29. A. Gelman, W. R. Gilks, and G. O. Roberts, "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms," *Annals of Applied Probability* 7, no. 1 (1997): 110–120.

30. S. Kotz and S. Nadarajah, *Multivariate t-Distributions and Their Applications* (Cambridge University Press, 2004).

31. A. Azzalini and A. Capitanio, "Statistical Applications of the Multivariate Skew Normal Distribution," *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 61, no. 3 (1999): 579–602.

32. J. Kruppa and K. Jung, "Automated Multigroup Outlier Identification in Molecular High-Throughput Data Using Bagplots and Gemplots," *BMC Bioinformatics* 18, no. 1 (2017): 1–10.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1**: Supporting Information.