Check for updates

Received: 4 April 2021 | Revised: 30 January 2022 | Accepted: 5 March 2022

DOI: 10.1002/sim.9421

RESEARCH ARTICLE

Statistics in Medicine WILEY

Fixed-effects inference and tests of correlation for longitudinal functional data

Ruonan Li[1] | **Luo Xiao**[1] | **Ekaterina Smirnova**[2] | **Erjia Cui**[3] | **Andrew Leroux**[4] | **Ciprian M. Crainiceanu**[3]

[1]Department of Statistics, North Carolina State University, Raleigh, North Carolina, USA

[2]Department of Biostatistics, Virginia Commonwealth University, Richmond, Virginia, USA

[3]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

[4]Department of Biostatistics and Informatics, Colorado School of Public Health, Aurora, Colorado, USA

Correspondence
Luo Xiao, Department of Statistics, North Carolina State University, 2311 Stinson Drive, Raleigh, NC 27695, USA.
Email: lxiao5@ncsu.edu

Funding information
NIH, Grant/Award Numbers: R01 AG064803, R01 NS112303, R56 AG064803

Abstract

We propose an inferential framework for fixed effects in longitudinal functional models and introduce tests for the correlation structures induced by the longitudinal sampling procedure. The framework provides a natural extension of standard longitudinal correlation models for scalar observations to functional observations. Using simulation studies, we compare fixed effects estimation under correctly and incorrectly specified correlation structures and also test the longitudinal correlation structure. Finally, we apply the proposed methods to a longitudinal functional dataset on physical activity. The computer code for the proposed method is available at https://github.com/rli20ST758/FILF.

KEYWORDS

accelerometry data, covariance function, hypothesis test, mixed effects model

1 | INTRODUCTION

Due to technological advancement, longitudinal functional data have become common in scientific studies.[1] For example, objective physical activity measurements[2-5] obtained from wearable accelerometers usually consists of multiple daily profiles measured at every minute of the day for hundreds of study participants at a number of visits. This type of data structure gives rise to longitudinal functional data, which resembles traditional longitudinal data, but has a function instead of a scalar as the basic unit of observation. The average activity measured at multiple visits gives rise to standard longitudinal data, whereas minute-by-minute activity profiles observed at multiple visits give rise to longitudinal functional data. Another example is diffusion tensor imaging (DTI),[6-10] where imaging data along brain tracts were collected for patients of multiple sclerosis and healthy controls at a number of visits. The physical activity measurements and the DTI data are usually collected at sparse longitudinal visits with densely observed functions, the data structure the article focuses on. This new and practical data structure requires specialized methodological tools to address scientific problems using the original structure of the data.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2022 The Authors. *Statistics in Medicine* published by John Wiley & Sons Ltd.

Longitudinal functional data often exhibit complex within- and between-visit correlations. To accommodate such correlations, various functional mixed effects models[6,8,9,11-16] have been proposed. These approaches, after data transformation or projections, allow inference on the fixed effects parameters using the linear mixed effects (LMEs) inferential machinery. Alternatively, bootstrap of study participants[9,17,18] can be used to construct confidence bands and conduct hypothesis tests for fixed effect functions. In spite of the growing interest in this area of research, modeling longitudinal functional data continues to be daunting. On the one hand, it is often difficult to specify within-visit correlation structures which account for the continuity of the functional domain with dense observations. On the other hand, the between-visit correlations may be modeled flexibly using structured specifications that represent the known and usually relatively sparse longitudinal sampling structure (eg, nesting of functions within individuals).

With the above considerations, we consider a unified inferential model framework that builds on the marginal decomposition model proposed by Park and Staicu,[13] which focused on modeling the covariance of longitudinal functional data, but did not consider fixed effects inference or tests of correlations. The marginal decomposition model used longitudinal time-invariant and data-adaptive eigenfunctions to model within-visit correlations and a sequence of bivariate covariance functions (either parametric or nonparametric) to model the longitudinal correlations, and is computationally simpler than the approach in Chen and Müller,[12] which modeled longitudinal functional data with a four-dimensional nonparametric covariance function. For longitudinal functional data with sparse longitudinal visits and densely observed functions, it might be sufficient and preferred to model the longitudinal correlation with a simple parametric specification. More importantly, the separation of within-subject correlation and longitudinal correlation in the marginal decomposition model not only allows us to compare fixed-effects inference under different longitudinal correlation structures but also enables us to compare different longitudinal correlation structures via hypothesis testing, the major contributions of this article. For the latter, we shall conduct multiple tests on the correlation structure of the sequence of bivariate covariance functions for modeling the longitudinal correlations in the above marginal decomposition model. For each test, we use the bootstrap-based approach for covariance testing proposed by Chen et al,[19] which compared parametric covariance functions vs nonparametric covariance function for modeling traditional longitudinal data. To our best knowledge, we have not been aware of any existing test of the longitudinal correlations developed for longitudinal functional data. Moreover, the test in Chen et al[19] is the only available test with an unspecified nonparametric alternative and can be applied to longitudinal data that are observed at irregular time points; see our discussion in Section 3 on alternative tests that are applicable only to limited scenarios.

In terms of fixed-effects inference, the proposed framework may compare favorably against the functional additive model proposed by Scheipl et al,[8] which only considered multilevel functional models with multiple covariance functions. It is worth noting that the aforementioned model can be computationally challenging to fit when using its spline-based formulation of functional random effects. Bootstrapped approaches such as in Park et al[18] are usually computationally demanding, even for small datasets. In addition, as shall be shown in our simulations, simple bootstrapping of subjects may give inefficient inference when the subjects have multiple visits exhibiting within-visit correlations.

In summary, the contributions of this article are to, for longitudinal functional data: (a) generalize standard longitudinal correlation models; (b) introduce testing for longitudinal correlation structures; and (c) study the effect of correctly and incorrectly specified correlation structures on the fixed-effects inference.

The article is organized as follows. In Section 2, we introduce the general functional mixed effects model and the specific longitudinal correlation structures for modeling longitudinal functional data. In Section 3, we address the problem of testing longitudinal correlation structure. Section 4 assesses the properties of the statistical methods proposed using simulations. Section 5 applies our methods to real data. We conclude the article with a discussion in Section 6. The computer code for the proposed method and an associated RMD file with demonstration details are available at https://github.com/rli20ST758/FILF.

2 | FUNCTIONAL MIXED EFFECTS MODELS

Let $\mathcal{T} \subset \mathbb{R}$ be a compact interval and the index $s \in \mathcal{T}$ be called the functional index. Let $y_{ij}(s)$ denote the observed response variable at s for subject i, $1 \le i \le n$, at the jth, $1 \le j \le n_i$, visit that occurred at time T_{ij}. Denote by $\mathbf{x}_{ij} = [x_{ij1}, \dots, x_{ijp}]^T$ the vector of covariates for subject i at visit j, and $\mathbf{x}_i = [\mathbf{x}_{i1}, \dots, \mathbf{x}_{in_i}]^T$ the $n_i \times p$ matrix of covariates for subject i. Some of the covariates may depend on the visit number and some may not, but we use the index j for all variables to simplify

notation. For one data application on physical activity, the covariates are age, body mass index (BMI), sex, and weekend, and thus $p = 4$. Let $\mathbf{y}_{ij} = [y_{ij}(s_1), \dots, y_{ij}(s_m)]^T$, where $\{s_1, \dots, s_m\}$ is a fixed grid in $\mathcal{T}$. Let $\mathbf{y}_i = [\mathbf{y}_{i1}^T, \dots, \mathbf{y}_{in_i}^T]^T$ be the mn_i-dimensional column vector obtained by vectorizing all observations for subject i.

We denote by $\mu_{ij}(\cdot)$ the conditional mean of the functional process $y_{ij}(\cdot)$ given the covariates $\mathbf{x}_{ij}$ and assume that

$$\mu_{ij}(s) = \beta_0(s) + \sum_{r=1}^{p} x_{ijr}\beta_r(s), \tag{1}$$

where $\beta_r(\cdot)$ are smooth, unknown functions of s. Using the approach in Park and Staicu,[13] we decompose the functional residuals, $y_{ij}(s) - \mu_{ij}(s) = W_{ij}(s) + \epsilon_{ij}(s)$, where

$$W_{ij}(s) = \sum_{k\geq 1} \xi_{ijk}\phi_k(s), \tag{2}$$

and $\epsilon_{ij}(s)$ is white noise with zero mean and variance σ_ϵ^2. The functions $\{\phi_1(\cdot), \phi_2(\cdot), \dots\}$ form an orthonormal basis and ξ_{ijk} are associated random coefficients. We assume that ξ_{ijk} are normally distributed with zero mean and covariance $G_k(T_{ij_1}, T_{ij_2}) = \mathrm{Cov}\{\xi_{ij_1k}, \xi_{ij_2k} | T_{ij_1}, T_{ij_2}\}$ and are uncorrelated over i and k. Then the covariance of the process is a four-dimensional operator with the following structure

$$C(s_1, T_{ij_1}, s_2, T_{ij_2}) = \mathrm{Cov}\left\{W_{ij_1}(s_1), W_{ij_2}(s_2) | T_{ij_1}, T_{ij_2}\right\} = \sum_{k\geq 1} G_k(T_{ij_1}, T_{ij_2})\phi_k(s_1)\phi_k(s_2).$$

We also assume that the sampling visits T_{ij} are independent over i and j and have a smooth probability density function, $g(\cdot)$. Then the basis functions $\phi_k(\cdot)$ are eigenfunctions of the marginal covariance operator

$$C_m(s_1, s_2) = \int C(s_1, T, s_2, T)g(T)\mathrm{d}T = \sum_{k\geq 1} \lambda_k\phi_k(s_1)\phi_k(s_2),$$

where $\lambda_k = \int G_k(T, T)g(T)\mathrm{d}T$ are eigenvalues and are assumed to be ordered decreasingly, that is, $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$, and the random coefficients ξ_{ijk} satisfy $\xi_{ijk} = \int W_{ij}(s)\phi_k(s)\mathrm{d}s$.

Standard longitudinal models using working covariance often assume either independence, exchangeability, or leave them unspecified. A natural generalization of these assumptions to longitudinal functional data is via the specification of the working covariance, $G_k(\cdot, \cdot)$, for every eigenfunction, $\phi_k(\cdot)$, indexed by k. The continuity of the functional domain is accounted by within-visit correlation structures and the longitudinal sampling structure is represented by between-visit correlations. We propose to model the within-visit covariance nonparametrically and the between-visit covariance using approaches inspired by standard and functional longitudinal data analysis. The rationale is that the functions are often densely observed and exhibit complex correlation structure, for which a nonparametric approach would provide the needed flexibility, while the longitudinal sampling visits are often sparse, for which a simpler correlation structure might be sufficient and also easy to interpret.

2.1 | Longitudinal working covariances

We consider three common correlation models including two parametric correlation structures, but the proposed method can easily be extended to other parametric correlation structures.

The first is the working independence model, where $W_{ij}(s)$ in Equation (2) are assumed to be independent and identically distributed for all i and j. Under this framework the random scores $\xi_{ijk} = \zeta_{ijk}$, where ζ_{ijk} are uncorrelated across i and j for each k and $\mathrm{Var}(\zeta_{ijk}) = \sigma_{e,k}^2$. Then, we have $G_k(T_{ij_1}, T_{ij_2}) = \sigma_{e,k}^2 1_{\{j_1=j_2\}}$.

The second is the exchangeable correlation model, that is $\xi_{ijk} = \xi_{ik} + \zeta_{ijk}$, where ξ_{ik} are uncorrelated with all ζ_{ijk} and ζ_{ijk} are mutually uncorrelated for all i, k, and j. Denote $\sigma_{0,k}^2 = \mathrm{Var}(\xi_{ik})$ and $\sigma_{e,k}^2 = \mathrm{Var}(\zeta_{ijk})$, and then $G_k(T_{ij_1}, T_{ij_2}) = \sigma_{0,k}^2 + \sigma_{e,k}^2 1_{\{j_1=j_2\}}$. The exchangeable model for longitudinal functional data is a particular case of the multilevel functional principal component model.[11] Indeed, we may write $W_{ij}(s) = U_i(s) + V_{ij}(s)$ with $U_i(s) = \sum_{k\geq 1} \xi_{ik}\phi_k(s)$ and $V_{ij}(s) = \sum_{k\geq 1} \zeta_{ijk}\phi_k(s)$.

The third is the unspecified smooth correlation model, where $G_k(\cdot,\cdot)$ is modeled nonparametrically. In this case $\xi_{ijk} = \xi_{ik}(T_{ij}) + \zeta_{ijk}$, where $\xi_{ik}(T_{ij})$ is a zero-mean smooth process with covariance function $\mathrm{Cov}\left\{\xi_{ik}(T_{ij_1}), \xi_{ik}(T_{ij_2})\right\} = G_{k0}(T_{ij_1}, T_{ij_2})$, and ζ_{ijk} are zero mean random errors with $\mathrm{Var}(\zeta_{ijk}) = \sigma^2_{e,k}$. In particular, ζ_{ijk} are mutually uncorrelated for all i, j, and k. The induced covariance function is $G_k(T_{ij_1}, T_{ij_2}) = G_{k0}(T_{ij_1}, T_{ij_2}) + \sigma^2_{e,k} 1_{\{j_1=j_2\}}$. We further decompose the covariance function $G_{k0}(T_{ij_1}, T_{ij_2})$ as $G_{k0}(T_{ij_1}, T_{ij_2}) = \sum_{\ell=1}^{\infty} \lambda_{k\ell}\psi_{k\ell}(T_{ij_1})\psi_{k\ell}(T_{ij_2})$, where $\lambda_{k1} \geq \lambda_{k2} \geq \cdots \geq 0$ and $\psi_{k\ell}(\cdot)$ are orthonormal eigenfunctions that satisfy $\int_T \psi_{k\ell_1}(T)\psi_{k\ell_2}(T)\mathrm{d}T = 1_{\{\ell_1=\ell_2\}}$. It follows that $\xi_{ijk} = \sum_{\ell=1}^{\infty} \eta_{ik\ell}\psi_{k\ell}(T_{ij}) + \zeta_{ijk}$, where $\eta_{ik\ell}$ are uncorrelated random scores with zero mean and variance $\lambda_{k\ell}$.

The unspecified smooth covariance model contains the independent and exchangeable models as special cases. Indeed, the working independence model corresponds to $G_{k0}(T_{ij_1}, T_{ij_2}) = 0$ while the exchangeable correlation model corresponds to $G_{k0}(T_{ij_1}, T_{ij_2}) = \sigma^2_{0,k}$. Thus, for model estimation below, we will focus on the unspecified covariance model as the other cases can be treated as simpler cases.

2.2 | Model estimation

We first consider estimation of $\mu_{ij}(\cdot)$ using B-spline expansion of the coefficient functions $\beta_r(\cdot)$. Specifically, let $\mathbf{B}(s) = [B_1(s), \ldots, B_c(s)]^T \in \mathbb{R}^c$ be a collection of B-spline basis functions in the unit interval with c being the number of basis functions. Then $\beta_r(s) = \sum_{d=1}^{c} \alpha_{rd} B_d(s)$, where $\boldsymbol{\alpha}_r = [\alpha_{r1}, \ldots, \alpha_{rc}]^T$ is the unknown coefficient vector. We use cubic B-splines and a relatively large number of equally spaced knots, that is, a large c, so that the approximation bias is negligible. Let $\mathbf{B} = [\mathbf{B}(s_1), \ldots, \mathbf{B}(s_m)]^T \in \mathbb{R}^{m\times c}$, then $\boldsymbol{\mu}_{ij} = [\mu_{ij}(s_1), \ldots, \mu_{ij}(s_m)]^T = \mathbf{B}\boldsymbol{\alpha}_0 + \sum_{r=1}^{p} x_{ijr}\mathbf{B}\boldsymbol{\alpha}_r$. Denote $\mathbf{X}_{ij} = [\mathbf{B}, x_{ij1}\mathbf{B}, \ldots, x_{ijp}\mathbf{B}] \in \mathbb{R}^{m\times((p+1)c)}$, $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_0^T, \ldots, \boldsymbol{\alpha}_p^T]^T \in \mathbb{R}^{(p+1)c}$. Then, we have $\boldsymbol{\mu}_{ij} = \mathbf{X}_{ij}\boldsymbol{\alpha}$. Let $\mathbf{X}_i = [\mathbf{X}_{i1}^T, \ldots, \mathbf{X}_{in_i}^T]^T \in \mathbb{R}^{(mn_i)\times((p+1)c)}$ and $\boldsymbol{\mu}_i = [\boldsymbol{\mu}_{i1}^T, \ldots, \boldsymbol{\mu}_{in_i}^T]^T \in \mathbb{R}^{mn_i}$. Then, we have $\boldsymbol{\mu}_i = \mathbf{X}_i\boldsymbol{\alpha}$.

Now we consider $W_{ij}(s) = \sum_{k\geq 1} \xi_{ijk}\phi_k(s)$ in Equation (2), where for each $k \geq 1$, $\xi_{ijk} = \sum_{\ell\geq 1} \eta_{ik\ell}\psi_{k\ell}(T_{ij}) + \zeta_{ijk}$. As both terms involve infinite sums, we propose to use a number of eigenfunctions $\phi_k(\cdot)$ that explains at least a given percentage of the variance explained (PVE) in the marginal covariance function, C_m. Denote the selected number of $\phi_k(\cdot)$ by K. Next we determine the approximation of ξ_{ijk} for all $1 \leq k \leq K$ simultaneously by ranking $\lambda_{k\ell}$, the variances of scores $\eta_{ik\ell}$. We retain the $\lambda_{k\ell}$ that are larger than λ^*, where λ^* is the largest value that satisfies

$$\frac{\sum_{k=1}^{K}\left\{\sigma^2_{e,k} + \sum_{\ell\geq 1}\lambda_{k\ell}1_{\{\lambda_{k\ell}\geq\lambda^*\}}\right\}}{\sum_{k=1}^{K}\left\{\sigma^2_{e,k} + \sum_{\ell\geq 1}\lambda_{k\ell}\right\}} > \mathrm{PVE}. \tag{3}$$

This will discard $\eta_{ik\ell}$ with small variances. According to our simulation studies, discarding small variances leads to improved estimation. Let L_k be the number of eigenfunctions selected for ξ_{ijk}. Then approximately,

$$W_{ij}(s) = \sum_{1\leq k\leq K}\left\{\sum_{1\leq\ell\leq L_k}\eta_{ik\ell}\psi_{k\ell}(T_{ij}) + \zeta_{ijk}\right\}\phi_k(s).$$

Let $\boldsymbol{W}_{ij} = [W_{ij}(s_1), \ldots, W_{ij}(s_m)]^T \in \mathbb{R}^m$, $\boldsymbol{\phi}_k = [\phi_k(s_1), \ldots, \phi_k(s_m)]^T \in \mathbb{R}^m$, $\boldsymbol{\Psi}_{ijk} = [\psi_{k1}(T_{ij}), \ldots, \psi_{kL_k}(T_{ij})]^T \in \mathbb{R}^{L_k}$, and $\boldsymbol{\eta}_{ik} = [\eta_{ik1}, \ldots, \eta_{ikL_k}]^T \in \mathbb{R}^{L_k}$. Then

$$\boldsymbol{W}_{ij} = \sum_{1\leq k\leq K}\left(\boldsymbol{\Psi}_{ijk}^T\boldsymbol{\eta}_{ik} + \zeta_{ijk}\right)\boldsymbol{\phi}_k.$$

Let $\boldsymbol{\Psi}_{ij} = \mathrm{blockdiag}(\boldsymbol{\Psi}_{ij1}^T, \ldots, \boldsymbol{\Psi}_{ijK}^T) \in \mathbb{R}^{K\times L}$, where $L = \sum_{1\leq k\leq K} L_k$, and $\boldsymbol{\eta}_i = [\boldsymbol{\eta}_{i1}^T, \ldots, \boldsymbol{\eta}_{iK}^T]^T \in \mathbb{R}^L$. Note that $\boldsymbol{\eta}_{ik}$ has a diagonal covariance matrix $\boldsymbol{\Lambda}_k = \mathrm{diag}(\lambda_{k1}, \ldots, \lambda_{kL_k})$ and $\boldsymbol{\eta}_i$ has a diagonal covariance matrix $\boldsymbol{\Lambda} = \mathrm{blockdiag}(\boldsymbol{\Lambda}_1, \ldots, \boldsymbol{\Lambda}_K)$. Let $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_K] \in \mathbb{R}^{m\times K}$ and $\boldsymbol{\zeta}_{ij} = [\zeta_{ij1}, \ldots, \zeta_{ijK}]^T \in \mathbb{R}^K$. Then

$$\boldsymbol{W}_{ij} = \boldsymbol{\Phi}\left(\boldsymbol{\Psi}_{ij}\boldsymbol{\eta}_i + \boldsymbol{\zeta}_{ij}\right).$$

Let $\boldsymbol{W}_i = [\boldsymbol{W}_{i1}^T, \ldots, \boldsymbol{W}_{in_i}^T]^T \in \mathbb{R}^{mn_i}$, $\boldsymbol{\Psi}_i = [\boldsymbol{\Psi}_{i1}^T, \ldots, \boldsymbol{\Psi}_{in_i}^T]^T \in \mathbb{R}^{(Kn_i)\times L}$, and $\boldsymbol{\zeta}_i = [\boldsymbol{\zeta}_{i1}^T, \ldots, \boldsymbol{\zeta}_{n_i i}^T]^T \in \mathbb{R}^{Kn_i}$. Then

$$\boldsymbol{W}_i = (\mathbf{I}_{n_i} \otimes \boldsymbol{\Phi}) \left(\boldsymbol{\Psi}_i \boldsymbol{\eta}_i + \boldsymbol{\zeta}_i\right).$$

It is easy to show that $\boldsymbol{\zeta}_{ij}$ has a diagonal covariance matrix $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_{e,1}^2, \ldots, \sigma_{e,K}^2)$ and that $\boldsymbol{\zeta}_i$ has a diagonal covariance matrix $\mathbf{I}_{n_i} \otimes \boldsymbol{\Sigma}$. For the random error term $\epsilon_{ij}(s)$, let $\boldsymbol{\epsilon}_{ij} = [\epsilon_{ij}(s_1), \ldots, \epsilon_{ij}(s_m)]^T \in \mathbb{R}^m$ and $\boldsymbol{\epsilon}_i = [\boldsymbol{\epsilon}_{i1}^T, \ldots, \boldsymbol{\epsilon}_{in_i}^T]^T \in \mathbb{R}^{mn_i}$.

Finally, we obtain the mixed model representation

$$\begin{aligned} \mathbf{y}_i &= \mathbf{X}_i \boldsymbol{\alpha} + (\mathbf{I}_{n_i} \otimes \boldsymbol{\Phi}) \left(\boldsymbol{\Psi}_i \boldsymbol{\eta}_i + \boldsymbol{\zeta}_i\right) + \boldsymbol{\epsilon}_i, \\ \begin{pmatrix} \boldsymbol{\eta}_i \\ \boldsymbol{\zeta}_i \\ \boldsymbol{\epsilon}_i \end{pmatrix} &\sim \mathcal{N}\left[\begin{pmatrix} \mathbf{0}_L \\ \mathbf{0}_{Kn_i} \\ \mathbf{0}_{mn_i} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Lambda} & 0 & 0 \\ 0 & \mathbf{I}_{n_i} \otimes \boldsymbol{\Sigma} & 0 \\ 0 & 0 & \sigma_\epsilon^2 \mathbf{I}_{mn_i} \end{pmatrix}\right], \end{aligned} \tag{4}$$

where $[\boldsymbol{\eta}_i^T, \boldsymbol{\zeta}_i^T, \boldsymbol{\epsilon}_i^T]^T$ are independent across subjects. Therefore, given the eigenfunctions $\boldsymbol{\Psi}_i$ and $\boldsymbol{\Phi}$, the fixed effects $\boldsymbol{\alpha}$ and variance components $\lambda_{k\ell}$, $\sigma_{e,k}^2$, and σ_ϵ^2 can be estimated using standard mixed effects software. To avoid overfitting, we impose a quadratic penalty on each $\boldsymbol{\alpha}_r$. Let $\mathbf{P}$ be a second order difference penalty matrix.[20] To be specific we use the penalty $\sum_{r=0}^p \tau_r \boldsymbol{\alpha}_r^T \mathbf{P} \boldsymbol{\alpha}_r$, where τ_r is the smoothing parameter for the rth function $\beta_r(\cdot)$. The model can be fitted using the *gam* function[21] or *bam* function[22] in the R package *mgcv*[23] and we use generalized cross validation to select the smoothing parameters.

2.3 | Longitudinal covariance estimation and model estimation algorithm

We first obtain an initial estimate of $\boldsymbol{\alpha}$ for the fixed effects, by fitting model (4) without including the term $(\mathbf{I}_{n_i} \otimes \boldsymbol{\Phi}) \left(\boldsymbol{\Psi}_i \boldsymbol{\eta}_i + \boldsymbol{\zeta}_i\right)$. This is equivalent to assume that all observations $y_{ij}(s)$ are mutually independent. A small simulation study shows that the proposed method is not sensitive to the initial estimate. Indeed, see Section S.1 of the supplementary material for alternative methods of the initial estimate for $\boldsymbol{\alpha}$.

With an initial estimate of $\boldsymbol{\alpha}$, the residuals $r_{ij}(s) = y_{ij}(s) - \mu_{ij}(s)$ can be calculated and used for estimating the marginal eigenfunctions $\phi_k(\cdot)$ and the longitudinal covariances. First, the eigenfunctions $\phi_k(\cdot)$ can be estimated by using the fast *face* algorithm[24] of the sandwich smoother[25] on the set of residuals. Specifically, let $\boldsymbol{R}_{ij} = [r_{ij}(s_1), \ldots, r_{ij}(s_m)]^T \in \mathbb{R}^m$, $\boldsymbol{R}_i = [\boldsymbol{R}_{i1}, \ldots, \boldsymbol{R}_{in_i}] \in \mathbb{R}^{m\times n_i}$, and $\boldsymbol{R} = [\boldsymbol{R}_1, \ldots, \boldsymbol{R}_n] \in \mathbb{R}^{m\times N}$, where $N = \sum_{i=1}^n n_i$. Then we apply the *face* algorithm to $N^{-1}\boldsymbol{R}\boldsymbol{R}^T$. Then, given $\phi_k(\cdot)$, we calculate $\int r_{ij}(s)\phi_k(s)\mathrm{d}s$ as the prediction for ξ_{ijk}. For the working independence model, λ_k can be estimated along with $\phi_k(\cdot)$ directly from the *face* algorithm. For the exchangeable model, we have $\xi_{ijk} = \xi_{ik} + \zeta_{ijk}$ and a one-way nested ANOVA model can be used to estimate $\sigma_{0,k}^2$ and $\sigma_{e,k}^2$. Finally, when $G_k(\cdot,\cdot)$ is unspecified, we approximate it using tensor-product B-splines and apply the fast covariance estimation method for sparse functional data[26] to the dataset $\{\xi_{ijk}, T_{ijk}, 1 \le j \le n_i, 1 \le i \le n\}$ to obtain estimates of $G_{k0}(\cdot,\cdot)$ and $\sigma_{e,k}^2$. We use 10 marginal cubic B-splines with equally spaced knots. Then, an eigendecomposition of $G_{k0}(\cdot,\cdot)$ leads to estimates of eigenfunctions $\psi_{k\ell}(\cdot)$ and associated eigenvalues $\lambda_{k\ell}$. We retain all eigenfunctions with associated positive eigenvalues and then use Equation (3) to discard small eigenvalues.

Once we have got estimates of the marginal eigenfunctions and the longitudinal covariances, we refine the estimates of fixed effects by fitting the mixed model (4) to obtain estimates of $\boldsymbol{\alpha}$, denoted by $\hat{\boldsymbol{\alpha}}$. Given the estimated variance parameters and treating the smoothness penalty on the fixed effects function as prior distributions, $\hat{\boldsymbol{\alpha}}$ has a posterior multivariate normal distribution, which can be used to construct either point-wise or simultaneous confidence bands for the fixed effects functions. We omit the details and refer to the textbook by Wood.[27]

We summarize the steps of the proposed method with unspecified smooth covariance in Algorithm 1. The estimation algorithm with independent and exchangeable covariance is available in Section S.2 in the supplementary material. Notice that the fixed effect is actually estimated twice (initial estimate and refined estimate) but not iterative since we have found that more iterations do not contribute to a better estimation. The reason for this is that as long as a good initial estimate of fixed effects is obtained, subsequent estimation of covariances via residuals is already sufficiently accurate and does not have substantial improvement using residuals with refined estimates. As shall be shown in our simulation study, our proposed initial estimates are often comparable with the proposed refined estimates. Notice that we still use

Algorithm 1. Estimate fixed effects assuming unspecified smooth covariance

1. Estimate the initial fixed effects as $\mu_{ij}^0(s)$ by existed methods assuming that all observations $y_{ij}(s)$ are mutually independent, and subtract the initial mean functions from observed data to get residuals as $r_{ij}(s) = y_{ij}(s) - \mu_{ij}^0(s)$.
2. Apply the *face* algorithm to $r_{ij}(s)$ to extract eigenfunctions $\phi_k(s)$ and principal scores ξ_{ijk}.
3. For each layer k, apply the sparse *face* algorithm[25] to ξ_{ijk} to extract eigenfunctions $\psi_{k\ell}(s)$.
4. With covariates x_{ijp} and estimated eigenfunctions $\phi_k(s)$ and $\psi_{k\ell}(s)$, fit the mixed model (4) using the *gam* or *bam* function to obtain the final and refined fixed effects as $\mu_{ij}(s)$.

the refined estimates as our final estimates because inference would not be valid with the initial estimates. Finally, a small simulation for the iterative procedure between estimating the mean and covariance structure is shown in Section S.1 in the supplementary material.

3 | TESTS OF LONGITUDINAL COVARIANCES

We use the goodness of fit test in Chen et al[19] to test the structure of the covariance $G_{k0}(\cdot,\cdot)$ in the following scenarios: (1) test the independent model against the unspecified covariance model; and (2) test the exchangeable model against the unspecified covariance model. The only other available method for carrying out the above tests is the multivariate test in Zhong et al,[28] which can only be applied to data with a common longitudinal design and assumes an unstructured covariance that does not account for smoothness in the alternative covariance. Thus when applicable, the multivariate test tends to be less powerful than the above goodness of fit test. It is also worth mentioning that if the alternative is also parametric, then likelihood-based tests such as the one in Crainiceanu and Ruppert[29] may instead be used.

We will consider two types of tests: test of an individual covariance or joint test of multiple covariances. To illustrate the idea, consider the case when the null hypothesis assumes independence. For the individual test, we may test $H_{0k} : G_{k0}(\cdot,\cdot) = 0$ against $H_{1k} : G_{k0}(\cdot,\cdot)$ is unspecified for each k. For the joint test, we may test $H_0 : G_{k0}(\cdot,\cdot) = 0$ for all $k \geq 1$ against $H_1 : G_{k0}(\cdot,\cdot)$ is unspecified for at least one $k \geq 1$. The null model for the joint test corresponds to the assumption that the longitudinal functions $W_{ij}(\cdot)$ are uncorrelated across visits, j, within study participants.

The individual longitudinal covariance test can be written as

$$H_{0k} : G_{k0}(T,T') = \tilde{G}_{k0}(T,T') \text{ vs } H_{1k} : G_{k0}(T,T') \neq \tilde{G}_{k0}(T,T'), \tag{5}$$

where $\tilde{G}_{k0}(T,T')$ is a parametric covariance function with a finite number of unknown parameters. For example, to test the null hypothesis of an exchangeable covariance model, we have $G_k(T_{ij_1},T_{ij_2}) = \tilde{G}_{k0}(T_{ij_1},T_{ij_2}) + 1_{\{j_1=j_2\}}\sigma_{e,k}^2$, where $\tilde{G}_{k0}(T_{ij_1},T_{ij_2}) = \sigma_{0,k}^2$. For testing the null hypothesis of working independence, everything stays the same, except that $\tilde{G}_{k0}(T_{ij_1},T_{ij_2}) = 0$.

We now briefly describe the test proposed in Chen et al[19] Suppose that the null covariance is G_0 and the alternative is G_A. First, we estimate the unknown parameters under the null hypothesis and denote the resulting estimate by $\widehat{G}_0$. Under the alternative hypothesis, we estimate G_A nonparametrically using, for example, smoothing based on tensor product of B-splines. If the estimator under the alternative hypothesis is $\widehat{G}_A$ then the test statistic is the Hilbert-Schmidt norm distance

$$T_n = \left\| \widehat{G}_A - \mathcal{K}\widehat{G}_0 \right\|_{\mathrm{HS}},$$

where $\|f\|_{\mathrm{HS}} = \sqrt{\iint f(T,T')^2 \mathrm{d}T\mathrm{d}T'}$ and $\mathcal{K}\widehat{G}_0$ is the smoothed null covariance estimate using again tensor-product of B-splines. The rationale for smoothing $\widehat{G}_0$ is to remove the potential bias induced by smoothing procedure used for obtaining $\widehat{G}_A$. Larger values of the test statistic, T_n, will provide evidence against the null hypothesis. The null distribution of T_n is obtained via wild bootstrap,[30] that is, the estimated error variance under the alternative hypothesis is used to generate bootstrap samples.

To explore the overall structure of longitudinal covariances for modeling longitudinal functional data, we propose a joint test of $G_{k0}(\cdot,\cdot)$ for all $1 \le k \le K$. As we do not expect a large K (only a small number of multiple comparisons), the approach we propose is straightforward: conduct individual covariance tests and use a Bonferroni correction to control the family wise error rate. While not explored, one may also use more refined multiple testing methods such as the Benjamini-Hochberg method.[31]

4 | SIMULATION STUDIES

We use simulations to compare the performance of inferential approaches for the longitudinal functional models using independent, exchangeable, and unspecified covariance assumptions. We will focus on estimation and inference for fixed effects as well as size and power of the covariance tests introduced in Section 3. The longitudinal functional models are compared with the standard functional data method implemented in the *pffr* function[8,32,33] of the R package *refund*.[34] This approach was designed for functional data without residual correlation and will be referred to as the ind_obs method, in the sense that the method assumes independence in longitudinal observations. The methods implemented in the *pffr* function work very well in practice when residuals are not correlated. We also compare methods with the bootstrap method[18] to obtain the confidence bands for the coefficient functions. The point estimators for the bootstrap approach are obtained using the *pffr* function.

4.1 | Simulation settings

Data are generated from the model $y_{ij}(s) = \mu_{ij}(s) + W_{ij}(s) + \epsilon_{ij}(s)$, $s \in [0,1]$. The functional mean is $\mu_{ij}(s) = \beta_0(s) + x_{i1}\beta_1(s) + x_{ij2}\beta_2(s)$, where the number of covariates is $p = 2$, $x_{i1} \overset{\text{iid}}{\sim} N(0,1)$, and $x_{ij2} = aT_{ij} + e_{ij}$, where $e_{ij} \sim N(\rho e_{i(j-1)}, 1)$ with $e_{i0} = 0$. We let $a = 1$ and $\rho = 0.7$. The coefficient functions are $\beta_0(s) = 1 + \sqrt{3}\cos(3\pi s)$, $\beta_1(s) = 2 + \cos(2\pi s)$ and $\beta_2(s) = 2 + \sin(\pi s)$. The number of visits per subject, n_i, will be specified later while the visit times, T_{ij}, are generated uniformly from the unit interval and then ordered. Using the orthonormal functions $\phi_1(s) = 1$ and $\phi_2(s) = \sqrt{2}\sin(2\pi s)$, $W_{ij}(\cdot)$ is generated from one of the three models below:

1. Independent covariance: $W_{ij}(s) = \sum_{k=1}^{2} \zeta_{ijk}\phi_k(s)$.
2. Exchangeable covariance: $W_{ij}(s) = \sum_{k=1}^{2} (\xi_{ik} + \zeta_{ijk})\phi_k(s)$.
3. Unspecified covariance: $W_{ij}(s) = \sum_{k=1}^{2} \left\{ \sum_{\ell=1}^{2} \eta_{ik\ell}\psi_{k\ell}(T_{ij}) + \zeta_{ijk} \right\} \phi_k(s)$.

For the independent covariance, $\zeta_{ij1} \overset{\text{iid}}{\sim} N(0, 4.5)$ and $\zeta_{ij2} \overset{\text{iid}}{\sim} N(0,3)$. For the exchangeable covariance, $\xi_{i1} \overset{\text{iid}}{\sim} N(0,3)$, $\xi_{i2} \overset{\text{iid}}{\sim} N(0,2)$, $\zeta_{ij1} \overset{\text{iid}}{\sim} N(0,1.5)$, and $\zeta_{ij2} \overset{\text{iid}}{\sim} N(0,1)$. For the unspecified covariance, $\psi_{11}(T) = \sqrt{2}\cos(2\pi T)$, $\psi_{12}(T) = \sqrt{2}\sin(2\pi T)$, $\psi_{21}(T) = \sqrt{2}\cos(4\pi T)$, and $\psi_{22}(T) = \sqrt{2}\sin(4\pi T)$. We generate the random terms as $\eta_{i11} \overset{\text{iid}}{\sim} N(0,2)$, $\eta_{i12} \overset{\text{iid}}{\sim} N(0,1.5)$, $\eta_{i21} \overset{\text{iid}}{\sim} N(0,1.5)$, $\eta_{i22} \overset{\text{iid}}{\sim} N(0,1)$, $\zeta_{ij1} \overset{\text{iid}}{\sim} N(0,1)$, and $\zeta_{ij2} \overset{\text{iid}}{\sim} N(0,0.5)$. The functional arguments $\{s_1, \ldots, s_m\}$ form a grid of 101 equidistant points in the unit interval. The mutually independent random errors $\epsilon_{ij}(s)$ are generated from $N(0, \sigma_\epsilon^2)$. The value of σ_ϵ^2 is 1.5, which is determined by a signal-to-noise ratio (SNR) with a value of 5. For the independent and exchangeable covariances, we define SNR in the functional data as $\sum_{k\ge1}(\sigma_{0,k}^2 + \sigma_{e,k}^2)/\sigma_\epsilon^2$; for the unspecified covariance, we define SNR as $\sum_{k\ge1}(\sum_{\ell\ge1}\lambda_{k\ell} + \sigma_{e,k}^2)/\sigma_\epsilon^2$.

For simulations, we use a factorial design with three factors: (a) the number of subjects is either 100 or 200; (b) the number of visits n_i are generated from either Unif[3, … , 6] or Unif[8, … , 12]; and (c) the covariance model is either independent, exchangeable or unspecified. There are 12 model conditions. For each condition, 500 replications are conducted on an Intel Core i7-8700 with 32GB RAM using one core per simulation.

For tests of longitudinal covariances, data are generated similarly. However, under the alternative we add terms to ensure that the null model is misspecified. Specifically, we consider

1. Deviation from independent covariance: $\xi_{ijk} \to \xi_{ijk} + \Delta z_{ik}(T_{ij})$.
2. Deviation from exchangeable covariance: $\xi_{ik} + \zeta_{ijk} \to \xi_{ik} + \zeta_{ijk} + \Delta z_{ik}(T_{ij})$.

The scalar Δ controls the magnitude of the deviation from the null model. The scores ξ_{ik} and ζ_{ijk} are independent from the nonlinear random function $z_{ik}(\cdot)$, where $z_{ik}(T) = \sum_{\ell=1}^{2} \eta_{ik\ell}\psi_{k\ell}(T)$. The terms $\eta_{ik\ell}$ and $\psi_{k\ell}(T)$ are specified as before.

In addition to the above simulation settings, we have also conducted simulations when the number of covariates is four and the number of marginal eigenfunctions is also four. The details of the simulation settings and results can be found in Section S.6 of the supplementary material.

4.2 | Model estimation and evaluation criteria

The initial mean estimation is conducted using the *pffr* function using 10 cubic B-splines bases and the second order difference penalty. The marginal eigenfunctions $\phi_k(s)$ are estimated by fitting the functional residuals via the *fpca.face* function[24] in the R package *refund* with 20 knots. The unspecified covariance model components $\psi_{k\ell}(T)$ and $\xi_{ik\ell}$ are estimated using the *face.sparse* function in the R package *face* with 15 knots. The truncation parameters K and L_k (for the unspecified covariance model) are all determined using the prespecified level PVE with a value of 0.95. The final mixed model is fitted using *bam* function[22] in the R package *mgcv*.[23]

The estimation accuracy for functional parameters is assessed using the mean integrated squared errors (MISE) and the properties of the pointwise confidence intervals. Specifically, for a function β with estimate $\hat{\beta}$, $\text{MISE}(\hat{\beta}) = \frac{1}{N_{\text{sim}}}\sum_{i_{\text{sim}}=1}^{N_{\text{sim}}} \int \{\hat{\beta}^{i_{\text{sim}}}(s) - \beta(s)\}^2 ds$. For the pointwise confidence intervals, we report the integrated actual pointwise coverage (IAC) and integrated actual width (IAW), where IAC is defined as $\frac{1}{N_{\text{sim}}}\sum_{i_{\text{sim}}=1}^{N_{\text{sim}}} \int I\{\hat{\beta}^{i_{\text{sim}}}(s) \in \text{CI}_p(s)\} ds$ and $\text{CI}_p(s)$ is the pointwise approximate confidence interval for the functional parameter $\beta(s)$. For example, an approximate 95% pointwise confidence interval $\text{CI}_p(s_0)$ for $\beta_1(s_0)$ can be constructed as $\beta_1(s_0) \pm 1.96 \times \hat{\text{SD}}\{\hat{\beta}_1(s_0)\}$. The length of the confidence interval at s_0 is $3.92 \times \hat{\text{SD}}\{\hat{\beta}_1(s_0)\}$ and $\text{IAW} = \frac{1}{N_{\text{sim}}}\sum_{i_{\text{sim}}=1}^{N_{\text{sim}}} \int 3.92 \times \hat{\text{SD}}\{\hat{\beta}^{i_{\text{sim}}}(s)\} ds$.

4.3 | Simulation results for estimation

Tables 1 to 3 display the results for estimating the coefficient functions when the data generating mechanism uses independent, exchangeable, or unspecified covariances, respectively. For each table, we compare results using five methods. The ind_obs method refers to the method implemented in the *pffr* function; we use the bootstrap method[18] to obtain the confidence bands for the coefficient functions; and the independent, exchangeable and unspecified methods refer to the proposed methods with independent, exchangeable and unspecified covariances, respectively. We first focus on the point estimation, which is partially captured by the MISE in the three tables. Results indicate that as the number of subjects and visits increase, MISEs decrease for all methods; see the columns labeled $\sqrt{\text{MISE}}$ in Tables 1 to 3. Moreover, the MISE obtains the smallest value when the correlation is correctly specified. In particular, the reduction in MISE is substantial for data with exchangeable and unspecified covariance structures when the correct model is used compared to one that does not account for correlation. For example, in Table 2 which corresponds to data simulated from an exchangeable correlation structure, for $n = 200$ and $n_i \sim \text{Unif}[8, 12]$, the $\sqrt{\text{MISE}}$ is 0.028 for $\beta_2(s)$ for the method which accounts for exchangeable correlation, while it is 0.066 for the bootstrap method which does not. Finally, the unspecified correlation model consistently provides good estimators, with the corresponding MISE comparable to that obtains from the correctly specified model. See Section S.3 in the supplementary material for additional simulation results.

For confidence bands, all methods outperform the ind_obs method in terms of coverage. The confidence bands from the ind_obs method are on average too narrow and are far below the nominal coverage level. The correctly specified and the unspecified covariance models provide confidence bands that are close to the nominal, though, slightly conservative sometimes. The bootstrap method provides confidence bands that close to the nominal, though they require more computational times and have higher average width for estimating slope functions and lower average width for estimating intercept functions. When models are misspecified (eg, exchangeable model for unspecified covariance data), the average coverage of the confidence bands tends to be close to the nominal, but at the expense of much wider confidence bands.

In terms of run times per simulation averaged by 500 replications, the bootstrap method is the slowest and the ind_obs method is the fastest. Computational time increases with more complex longitudinal correlation structures. However, when the dataset has a large number of subjects and a large number of visits per subject ($n = 200$ and $n_i \sim \text{Unif}[8, 12]$),

TABLE 1 $\sqrt{\text{MISE}}$, IAC for 95% pointwise confidence interval and IAW for estimating $\beta_0(s)$, $\beta_1(s)$, and $\beta_2(s)$ based on independent covariance data using five methods across 500 simulations

		$\beta_0(s)$			$\beta_1(s)$			$\beta_2(s)$		
Method	**Time**	$\sqrt{\text{MISE}}$	**IAW**	**IAC**	$\sqrt{\text{MISE}}$	**IAW**	**IAC**	$\sqrt{\text{MISE}}$	**IAW**	**IAC**
	$n = 100$ and $n_i \sim$ Unif[3, 6]									
ind_obs	0 (0.0)	0.127	0.23	0.58	0.119	0.15	0.46	0.082	0.09	0.40
Bootstrap	72 (0.1)	0.127	0.54	0.93	0.119	0.51	0.93	0.082	0.35	0.93
Independent	3 (0.0)	0.125	0.80	0.98	0.115	0.51	0.95	0.073	0.33	0.95
Exchangeable	7 (0.1)	0.125	0.82	0.98	0.115	0.52	0.96	0.073	0.34	0.95
Unspecified	8 (0.0)	0.125	0.80	0.98	0.115	0.51	0.95	0.073	0.34	0.95
	$n = 100$ and $n_i \sim$ Unif[8, 12]									
ind_obs	1 (0.0)	0.086	0.16	0.58	0.079	0.10	0.46	0.052	0.07	0.45
Bootstrap	150 (0.1)	0.086	0.36	0.94	0.079	0.34	0.94	0.052	0.23	0.95
Independent	32 (0.1)	0.085	0.53	0.96	0.078	0.34	0.94	0.048	0.23	0.96
Exchangeable	49 (0.4)	0.085	0.54	0.97	0.078	0.35	0.95	0.048	0.23	0.96
Unspecified	37 (0.1)	0.085	0.53	0.96	0.078	0.34	0.94	0.048	0.23	0.96
	$n = 200$ and $n_i \sim$ Unif[3, 6]									
ind_obs	1 (0.0)	0.086	0.16	0.61	0.084	0.11	0.44	0.056	0.07	0.42
Bootstrap	138 (0.1)	0.086	0.38	0.95	0.084	0.36	0.94	0.056	0.25	0.96
Independent	24 (0.1)	0.085	0.57	0.97	0.082	0.36	0.95	0.051	0.24	0.97
Exchangeable	49 (0.4)	0.085	0.57	0.97	0.082	0.36	0.96	0.051	0.24	0.97
Unspecified	33 (0.1)	0.085	0.57	0.97	0.082	0.36	0.95	0.051	0.24	0.97
	$n = 200$ and $n_i \sim$ Unif[8, 12]									
ind_obs	1 (0.0)	0.059	0.11	0.61	0.057	0.07	0.46	0.038	0.05	0.44
Bootstrap	288 (0.4)	0.059	0.25	0.95	0.057	0.24	0.93	0.038	0.17	0.94
Independent	264 (0.9)	0.058	0.38	0.97	0.056	0.24	0.94	0.036	0.16	0.95
Exchangeable	381 (3.0)	0.058	0.39	0.98	0.056	0.24	0.94	0.036	0.17	0.96
Unspecified	275 (0.9)	0.058	0.38	0.97	0.056	0.24	0.94	0.036	0.16	0.95

Note: Time (standard error) with second as a unit is the run time per simulation averaged by 500 replications.

the computation time for fitting the proposed methods can be comparable to or longer than that of the bootstrap method. The proposed method may be sped up via parallel computing in the *bam* function, but in our simulation we only use one core.

In summary, the unspecified covariance models perform well in terms of inference for the fixed effects in a variety of scenarios. However, in practice one may still want to know whether a simpler correlation structure is reasonable for the data, as the correct model slightly outperforms the unspecified covariance model. In the next section, we provide results for testing the longitudinal covariance.

4.4 | Simulation results for longitudinal covariance tests

For the covariance tests introduced in Section 3, we report the empirical type I error rate (estimated size) for nominal levels $\alpha = 0.05$ and 0.10 based on 5000 simulated datasets, and the power at the $\alpha = 0.05$ level with 1000 simulated datasets. For the joint test, we use the Bonferroni correction to account for the multiplicity of hypotheses tests. For the individual covariance tests, the power curves are presented as functions of the deviation from the null, defined as

TABLE 2 $\sqrt{\text{MISE}}$, IAC for 95% pointwise confidence interval and IAW for estimating $\beta_0(s)$, $\beta_1(s)$, and $\beta_2(s)$ based on exchangeable covariance data using five methods across 500 simulations

		$\beta_0(s)$			$\beta_1(s)$			$\beta_2(s)$		
Method	**Time**	$\sqrt{\text{MISE}}$	**IAW**	**IAC**	$\sqrt{\text{MISE}}$	**IAW**	**IAC**	$\sqrt{\text{MISE}}$	**IAW**	**IAC**
	$n = 100$ and $n_i \sim$ Unif[3, 6]									
ind_obs	0 (0.0)	0.225	0.22	0.34	0.226	0.15	0.24	0.119	0.09	0.28
Bootstrap	74 (0.1)	0.225	0.97	0.94	0.226	0.94	0.94	0.119	0.51	0.94
Independent	3 (0.0)	0.223	0.80	0.86	0.220	0.50	0.69	0.107	0.33	0.82
Exchangeable	6 (0.0)	0.215	1.44	0.97	0.202	0.90	0.95	0.063	0.29	0.95
Unspecified	11 (0.0)	0.215	1.42	0.99	0.208	0.87	0.93	0.064	0.29	0.94
	$n = 100$ and $n_i \sim$ Unif[8, 12]									
ind_obs	1 (0.0)	0.206	0.15	0.27	0.205	0.10	0.17	0.092	0.07	0.26
Bootstrap	152 (0.1)	0.206	0.91	0.95	0.205	0.90	0.95	0.092	0.39	0.94
Independent	32 (0.2)	0.205	0.54	0.74	0.202	0.34	0.54	0.086	0.23	0.75
Exchangeable	46 (0.2)	0.203	1.39	0.97	0.187	0.87	0.97	0.037	0.17	0.96
Unspecified	51 (0.2)	0.205	1.36	0.98	0.190	0.84	0.96	0.037	0.17	0.95
	$n = 200$ and $n_i \sim$ Unif[3, 6]									
ind_obs	1 (0.0)	0.162	0.16	0.34	0.149	0.11	0.26	0.085	0.07	0.29
Bootstrap	137 (0.2)	0.162	0.69	0.94	0.149	0.66	0.95	0.085	0.36	0.95
Independent	23 (0.1)	0.160	0.57	0.85	0.147	0.35	0.72	0.079	0.24	0.82
Exchangeable	44 (0.2)	0.151	1.02	0.97	0.141	0.64	0.95	0.047	0.21	0.95
Unspecified	54 (0.2)	0.152	1.01	0.99	0.142	0.63	0.95	0.047	0.21	0.95
	$n = 200$ and $n_i \sim$ Unif[8, 12]									
ind_obs	1 (0.0)	0.147	0.11	0.28	0.140	0.07	0.20	0.066	0.05	0.25
Bootstrap	288 (0.3)	0.147	0.65	0.95	0.140	0.63	0.95	0.066	0.28	0.95
Independent	261 (0.9)	0.147	0.38	0.73	0.139	0.24	0.58	0.063	0.16	0.76
Exchangeable	373 (1.5)	0.144	0.97	0.97	0.131	0.62	0.96	0.028	0.12	0.95
Unspecified	383 (1.5)	0.147	0.98	0.98	0.134	0.61	0.95	0.029	0.12	0.95

Note: Time (standard error) with second as a unit is the run time per simulation averaged by 500 replications.

$$\frac{\Delta^2 \int \text{Var}\{z_{ik}(T)\}dT}{\int \text{Var}\{\xi_{ik}(T)\}dT + \Delta^2 \int \text{Var}\{z_{ik}(T)\}dT}$$

for each k. For the joint test, the deviation from the null is defined as

$$\frac{\Delta^2 \sum_k \int \text{Var}\{z_{ik}(T)\}dT}{\sum_k \int \text{Var}\{\xi_{ik}(T)\}dT + \Delta^2 \sum_k \int \text{Var}\{z_{ik}(T)\}dT}.$$

The deviation from the null hypothesis is a number between 0 and 1.

Table 4 reports the empirical type I errors of the individual and joint tests, indicating that all tests have empirical levels close to their nominal levels. Tests can be conservative for exchangeable covariance with a larger number of visits per subject. We have also evaluated the dependency of the testing statistics from different layers for a few simulation settings; see Section S.4 of the supplementary material for details. While we have found significant positive correlations among the testing statistics for most scenarios, the correlations seem small. Nevertheless, for the joint test where the Bonferroni correction is used, multiple testing methods that account for positive dependencies such as the Simes method[35] might instead be employed.

TABLE 3 $\sqrt{\text{MISE}}$, IAC for 95% pointwise confidence interval and IAW for estimating $\beta_0(s)$, $\beta_1(s)$, and $\beta_2(s)$ based on unspecified covariance data using five methods across 500 simulations

		$\beta_0(s)$			$\beta_1(s)$			$\beta_2(s)$		
Method	**Time**	$\sqrt{\text{MISE}}$	**IAW**	**IAC**	$\sqrt{\text{MISE}}$	**IAW**	**IAC**	$\sqrt{\text{MISE}}$	**IAW**	**IAC**
	$n = 100$ and $n_i \sim$ Unif[3, 6]									
ind_obs	0 (0.0)	0.127	0.23	0.58	0.122	0.15	0.42	0.086	0.10	0.37
Bootstrap	74 (0.1)	0.127	0.53	0.93	0.122	0.50	0.93	0.086	0.37	0.94
Independent	3 (0.0)	0.125	0.81	0.97	0.118	0.51	0.95	0.076	0.34	0.95
Exchangeable	7 (0.1)	0.125	0.82	0.97	0.118	0.52	0.96	0.076	0.34	0.95
Unspecified	19 (0.1)	0.081	0.62	1.00	0.077	0.31	0.93	0.052	0.22	0.94
	$n = 100$ and $n_i \sim$ Unif[8, 12]									
ind_obs	1 (0.0)	0.081	0.16	0.61	0.070	0.10	0.49	0.063	0.07	0.36
Bootstrap	152 (0.1)	0.081	0.34	0.94	0.070	0.31	0.94	0.063	0.28	0.95
Independent	32 (0.1)	0.079	0.55	0.97	0.069	0.34	0.97	0.059	0.23	0.91
Exchangeable	43 (0.3)	0.079	0.55	0.97	0.069	0.34	0.97	0.059	0.23	0.91
Unspecified	77 (0.4)	0.045	0.37	1.00	0.041	0.18	0.95	0.032	0.13	0.94
	$n = 200$ and $n_i \sim$ Unif[3, 6]									
ind_obs	1 (0.0)	0.085	0.16	0.60	0.079	0.11	0.47	0.059	0.07	0.40
Bootstrap	138 (0.1)	0.085	0.37	0.95	0.079	0.35	0.95	0.059	0.26	0.94
Independent	24 (0.1)	0.084	0.56	0.97	0.077	0.36	0.96	0.055	0.24	0.94
Exchangeable	45 (0.3)	0.084	0.57	0.97	0.077	0.36	0.96	0.055	0.24	0.94
Unspecified	116 (0.6)	0.054	0.43	1.00	0.050	0.21	0.94	0.036	0.15	0.95
	$n = 200$ and $n_i \sim$ Unif[8, 12]									
ind_obs	1 (0.0)	0.056	0.11	0.64	0.052	0.07	0.49	0.044	0.05	0.37
Bootstrap	288 (0.4)	0.056	0.24	0.94	0.052	0.22	0.94	0.044	0.20	0.95
Independent	262 (0.8)	0.055	0.38	0.97	0.051	0.24	0.96	0.043	0.16	0.91
Exchangeable	339 (1.5)	0.055	0.38	0.97	0.051	0.24	0.97	0.043	0.16	0.90
Unspecified	585 (3.6)	0.031	0.25	1.00	0.029	0.12	0.94	0.022	0.09	0.95

Note: Time (standard error) with second as a unit is the run time per simulation averaged by 500 replications.

Figure 1 presents the power curves for the proposed tests, indicating similar patterns for the two individual and the joint tests. The test for independent covariance has higher power than the one for exchangeable covariance data. For both types of the null hypothesis, power increases with the number of subjects and observation per subject. Power seems to be particularly sensitive to the number of visits, n_i, per curve. Indeed, with 100 study participants and exchangeable covariance, for all three tests, the power for $n_i \sim$ Unif[8, 12] is close to one under a deviation around 0.25. In contrast, for $n_i \sim$ Unif[3, 6], the deviation needs to be closer to 0.7 to reach the same power. As expected, power depends on how closely the true model matches the specified alternative assumed by the test.

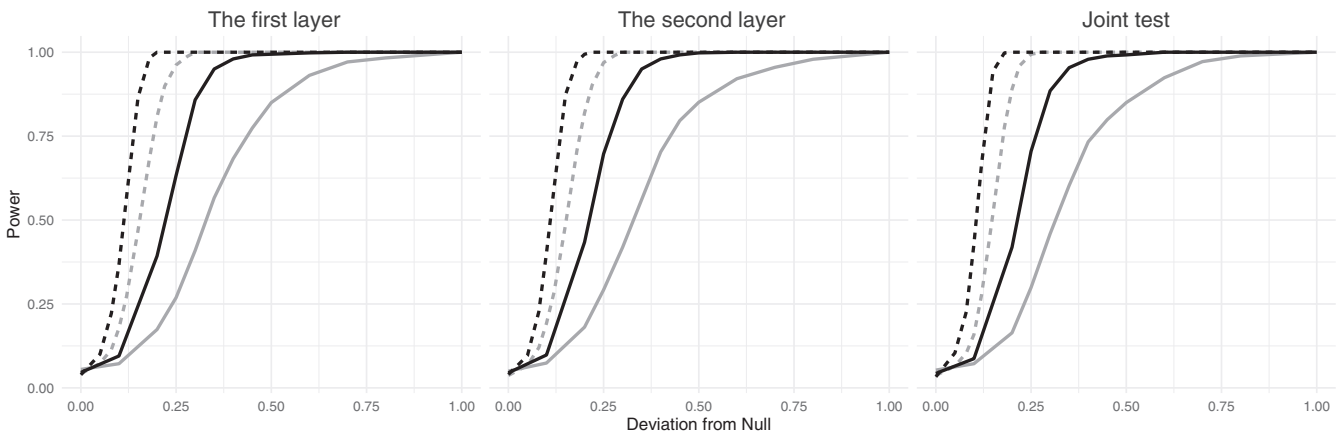
In summary, in all settings, the proposed tests maintain proper size and have the power to detect alternative longitudinal covariance both when the null hypothesis assumes independent covariance and exchangeable covariance.

5 | DATA APPLICATION

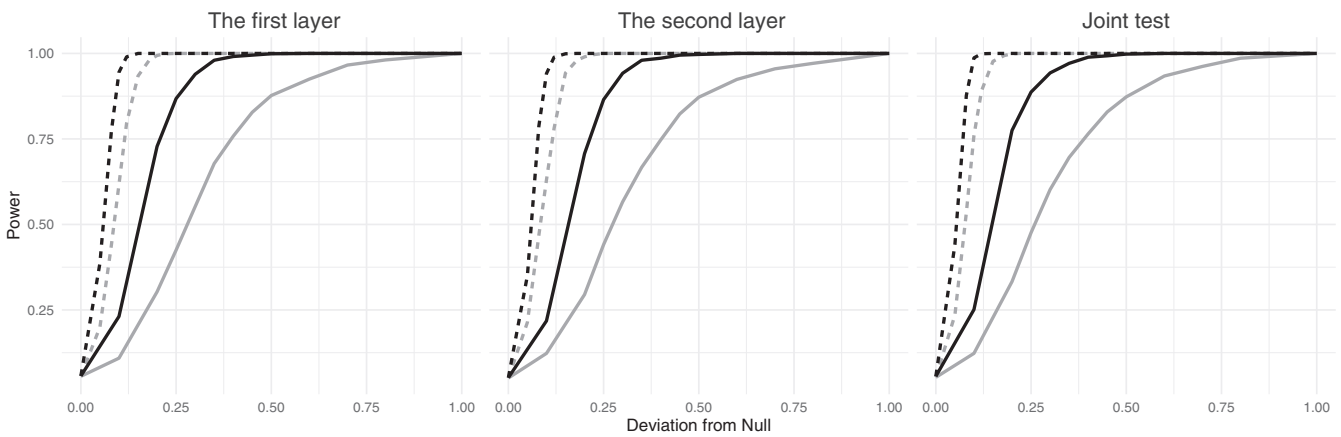
Wearable devices such as accelerometers provide objective and detailed measurements of physical activity and enable researchers to study how physical activity changes with age, personal characteristics, and health status. In earlier studies

TABLE 4 Empirical type I error of covariance tests at the nominal $\alpha = 0.05$ and 0.10 levels based on 5000 datasets, by sample size (n) and observations per subject (n_i)

			$n = 100$		$n = 200$	
Null	n_i	**Layer**	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.10$
Exchangeable	Unif[3, 6]	First layer	0.055	0.105	0.046	0.098
		Second layer	0.050	0.099	0.045	0.093
		Joint test	0.053	0.103	0.043	0.090
	Unif[8, 12]	First layer	0.042	0.094	0.039	0.085
		Second layer	0.037	0.080	0.040	0.092
		Joint test	0.037	0.078	0.033	0.078
Independent	Unif[3, 6]	First layer	0.056	0.115	0.057	0.110
		Second layer	0.051	0.108	0.052	0.103
		Joint test	0.053	0.105	0.056	0.106
	Unif[8, 12]	First layer	0.053	0.103	0.057	0.108
		Second layer	0.054	0.103	0.055	0.102
		Joint test	0.056	0.104	0.058	0.108



(A) Null: exchangeable covariance.



(B) Null: independent covariance.

FIGURE 1 Power curves (type I error $\alpha = 0.05$) for individual and joint tests of covariances, under deviation from the null. Shown are: $n_i \sim$ Unif[3, 6] (solid line) and $n_i \sim$ Unif[8, 12] (dashed line), for $n = 100$ (gray) and $n = 200$ (black)

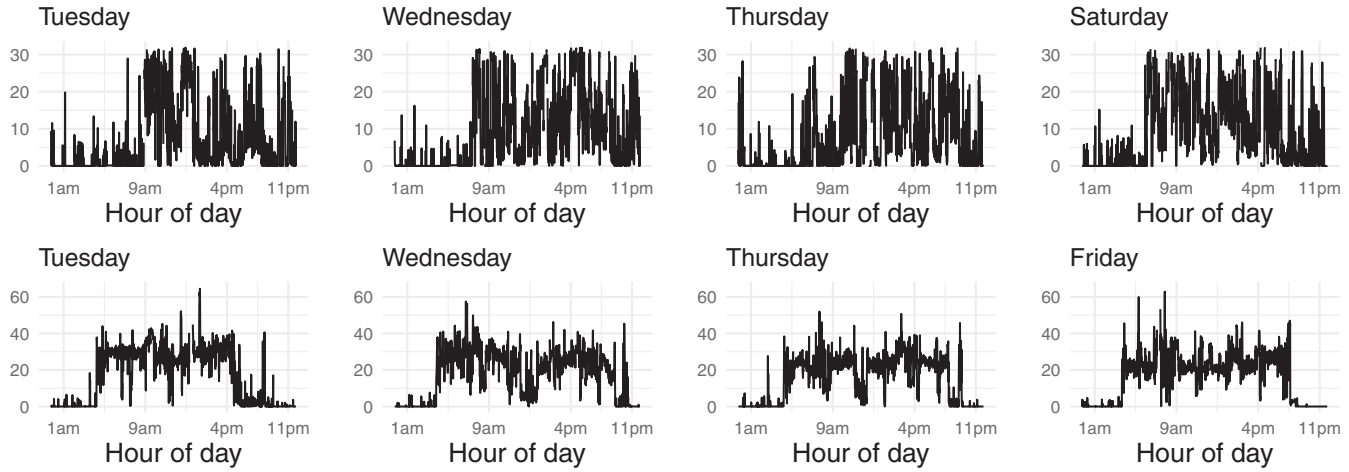


FIGURE 2 Physical activity profiles of two NHANES study participants over available days. Panels in the top row are from a male with age 54 and BMI 20.5. Panels in the bottom row are from a female with age 57 and BMI 31.54. Each panel displays MIMS of 1 day from midnight to midnight, titled by day of the week

based on cross-sectional data, the systematic and random circadian rhythms of physical activity was analyzed as functions of time of day and age.[2,3]

In this analysis, our objective is to quantify the circadian rhythm of physical activity over multiple days and how they relate to age, sex, and BMI and provide associated inference. Data were collected from the National Health and Nutrition Examination Survey (NHANES), a large cohort study conducted in 2-year waves by the US Centers for Disease Control and Prevention (CDC) to assess the health and nutritional status of US population. We focus on the NHANES 2011-2012 and 2013-2014 data, where each study participant was asked to wear the wrist-worn device all the time for seven consecutive days. The physical activity data were collected and summarized in minute-level Monitor-Independent Movement Summary (MIMS) units, a physical activity intensity unit optimized to capture normal human motion.[36] For each day of the study participant, our initial data are MIMS measured at 1440 minutes from midnight to midnight. The final dataset consists of 513 study participants (208 men and 305 women) with an average age of 57 years and an average BMI value of 28. Figure 2 displays the physical activity profiles of two randomly selected NHANES study participants. For each study participant, we further subset the days with at least 95% estimated wear time, leading to a total of 3187 days of wear with an average of 6.21 days of wear per study participant. For subject i at day j, the average MIMS at hour s of a day is our longitudinal functional outcome, denoted by $y_{ij}(s)$.

Activity profiles often vary for weekends (Saturday/Sunday) and weekdays (Monday-Friday). Thus, we create a binary variable which is 1 for weekends and 0 for weekdays. Then, we include sex (denoted by sex_i), age (age_i), BMI (BMI_i), and weekend (weekend_i) as the covariates in this analysis. The mean function is $\mu_{ij}(s) = \beta_0(s) + \text{sex}_i\beta_{\text{sex}}(s) + \text{age}_i\beta_{\text{age}}(s) + \text{BMI}_i\beta_{\text{BMI}}(s) + \text{weekend}_i\beta_{\text{weekend}}(s)$. The variable sex_i is 0 for males and 1 for females. Both age and BMI are centered so that $\beta_0(s)$ is the population mean MIMS corresponding to males of age 57 and with BMI of 28. The longitudinal covariate T_{ij} is the day of wear. Because physical activity profiles are periodic, that is, physical activity counts are the same at the start of the day and at the end of the day, the coefficient function should also be cyclic and we use cyclic P-splines. To investigate the longitudinal correlation structure of the data, we first estimate the marginal covariance function via the *fpca.face* R function using cubic B-spline basis functions with 10 knots. Using the procedure described in Section 2.2, we identify that eight eigenfunctions explain 90% of the marginal variance. Using these eight eigenfunctions, we conduct a joint test for the null hypothesis that all longitudinal covariances $G_{k0}(\cdot,\cdot)$ can be induced by an exchangeable covariance between daily profiles within each subject. The P-values are 0.04, 0.25, 0.06, 0.49, 0.62, 0.49, 0.09 and 0.19, respectively. Thus, with the Bonferroni correction, we do not reject the null hypothesis at the 0.05 level. The final model fit is obtained by using exchangeable covariances for all longitudinal covariances.

The estimated coefficient functions are displayed in Figure 3, along with the 95% pointwise confidence bands. The top left panel in Figure 3 displays the estimated mean activity profile $\beta_0(s)$ for males with age 57 and BMI 28. The shape is consistent with that of published average activity plots over the course of the day. On average, individuals are less active during night. Then, their activities increase sharply between 5 AM and 11 AM, sustain between 11 AM and 7 PM, and

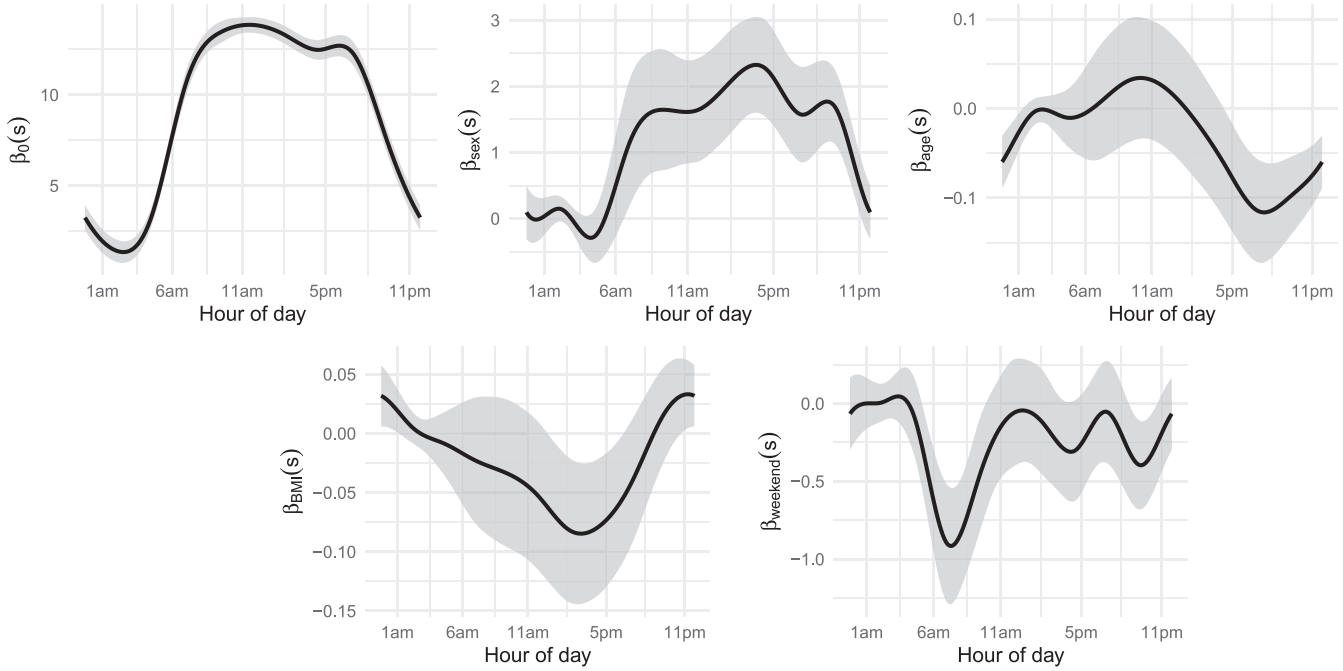


FIGURE 3 Estimated coefficient functions for the NHANES dataset along with pointwise 95% confidence bands under exchangeable longitudinal covariance structure

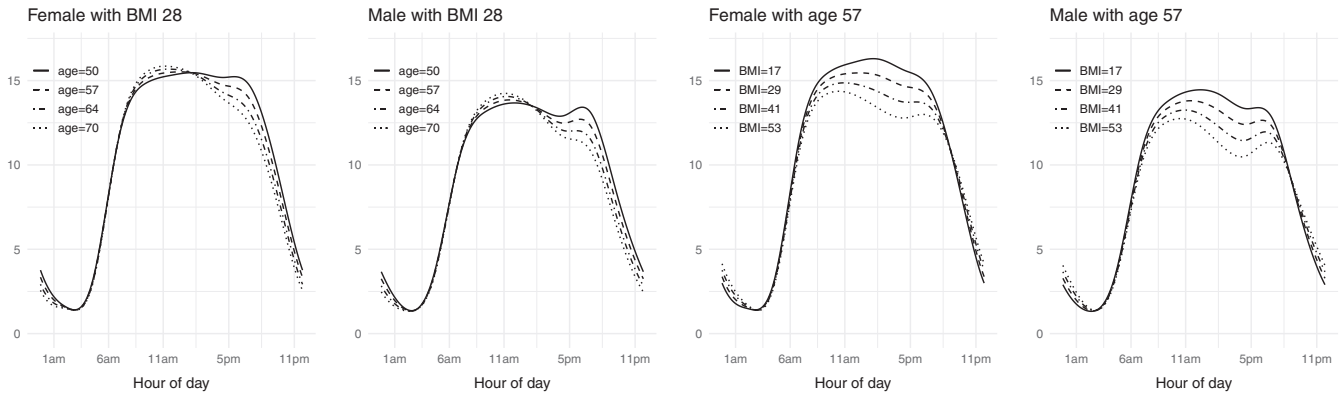


FIGURE 4 Left two panels: Estimated mean activity profiles of females and males with BMI 28 at four different ages. Right two panels: Estimated mean activity profiles of females and males with age 57 and with four different BMI values

decrease rapidly after 7 PM. The top middle panel indicates that females in this age group (50-70) have significantly more activities than males throughout most of the day except for midnight. The top right panel displays the estimated coefficient function for age, indicating a loss of activity for older individuals from afternoon to 1 AM. There is no substantial age effect in activity in the morning while the most considerable age effect happens in the early night. These findings are consistent with previous findings that older individuals sleep early at night and wake up early in the morning, while younger individuals are more energetic at night. The bottom left panel indicates that activity patterns tend to decrease with increasing BMI, especially in the daytime. Compared to weekdays, weekends correspond to lower levels of activity in the whole day, significantly in the morning from 5 AM to 11 AM. These results are consistent with common sense that individuals are less active on weekends morning.

Figure 4 displays the estimated mean activity profiles for females and males, with different age and BMI combinations. From the left two panels, we see that for a given BMI value, there is a pronounced decrease in activity as a function of age, both for females and males. Moreover, younger individuals have on average two activity peaks, while older individuals tend to lose the second activity peak. It indicates that the change of physical activity patterns with age is due to less activity

in the afternoon and at night. The right two panels indicate that the average activity profile decreases with increasing BMI. The negative effect of age only works after noon while individuals with larger BMI have less activity during all daytime. Moreover, the effect of two units on the BMI scale is roughly equivalent to two years of age in terms of overall activity loss. Finally, the estimated covariances of physical activity profiles within each day and across days, respectively, and the top three estimated eigenfunctions $\phi_k(s)$ are shown in Section S.5 in supplementary material.

6 | DISCUSSION

We introduced an inferential framework for estimating fixed effects functions for longitudinal functional data with complex correlation structures. Methods are natural extensions of standard longitudinal data methods with scalar observations and use the flexible marginal decomposition model framework in Park and Staicu.[13]

Simulation studies indicate that the estimation procedure works remarkably well even when the longitudinal correlations are misspecified. However, the confidence bands obtained under misspecified correlation models may have smaller than nominal coverage or are wider than necessary. Our results indicate that using unspecified correlation matrices is the safer approach when it is computationally feasible. More work will be necessary to improve the computational feasibility of these methods for larger sample sizes in terms of the number of study participants, visits, and observations per curve.

We also introduced a testing framework for longitudinal correlations for longitudinal functional data, which extended the testing method in Chen et al.[19] The proposed tests provide insight into what correlation structures are appropriate and may help reduce the computational cost when a simpler longitudinal correlation is chosen.

This article only considered scalar covariates and it would be of interest to extend the proposed framework to include functional covariates.

ACKNOWLEDGEMENTS

The authors thank the reviewers for their constructive and helpful comments, which greatly improved the article. This work was partially supported by the NIH grants R56 AG064803 (LX), R01 AG064803 (RL and LX), and R01 NS112303 (RL and LX). This work represents the opinions of the researchers and not necessarily that of the granting organizations.

DATA AVAILABILITY STATEMENT

The relevant R code to implement the proposed method and an RMD file with details of this article are freely available at https://github.com/rli20ST758/FILF.

ORCID

Ruonan Li https://orcid.org/0000-0001-7293-9779
Luo Xiao https://orcid.org/0000-0001-8707-0914

REFERENCES

1. Ramsay JO, Silverman BW. *Functional Data Analysis*. New York, NY: Springer; 2005.
2. Xiao L, Huang L, Schrack J, Ferrucci L, Zipunnikov V, Crainiceanu CM. Quantifying the lifetime circadian rhythm of physical activity: a covariate-dependent functional approach. *Biostatistics*. 2015;16(2):352-367.
3. Goldsmith J, Zipunnikov V, Schrack J. Generalized multilevel function-on-scalar regression and principal component analysis. *Biometrics*. 2015;71(2):344-353.
4. Leroux A, Di J, Smirnova E, et al. Organizing and analyzing the activity data in NHANES. *Stat Biosci*. 2019;11(2):262-287.
5. Smirnova E, Leroux A, Cao Q, et al. The predictive performance of objective measures of physical activity derived from accelerometry data for 5-year all-cause mortality in older adults: national health. *J Gerontol A*. 2020;75(9):1779-1785.
6. Greven S, Crainiceanu CM, Caffo B, Reich D. Longitudinal functional principal component analysis. *Electron J Stat*. 2010;4:1022-1054.
7. Goldsmith J, Bobb J, Crainiceanu CM, Caffo B, Reich D. Penalized functional regression. *J Comput Graph Stat*. 2011;20(4):830-851.
8. Scheipl F, Staicu AM, Greven S. Functional additive mixed models. *J Comput Graph Stat*. 2015;24(2):477-501.
9. Zhu H, Chen K, Luo X, Yuan Y, Wang JL. FMEM: functional mixed effects models for longitudinal functional responses. *Stat Sin*. 2019;29(4):2007-2033.
10. Cui E, Crainiceanu CM, Leroux A. Additive functional Cox model. *J Comput Graph Stat*. 2021;30(3):780-793.
11. Di CZ, Crainiceanu CM, Caffo BS, Punjabi NM. Multilevel functional principal component analysis. *Ann Appl Stat*. 2009;3(1):458-488.
12. Chen K, Müller HG. Modeling repeated functional observations. *J Am Stat Assoc*. 2012;107(500):1599-1609.
13. Park SY, Staicu AM. Longitudinal functional data analysis. *Stat*. 2015;4(1):212-226.

14. Zhang L, Baladandayuthapani V, Zhu H, et al. Functional CAR models for large spatially correlated functional datasets. *J Am Stat Assoc*. 2016;111(514):772-786.
15. Chen K, Delicado P, Müller HG. Modelling function-valued stochastic processes, with applications to fertility dynamics. *J Royal Stat Soc B*. 2017;79(1):177-196.
16. Lee W, Miranda MF, Rausch P, et al. Bayesian semiparametric functional mixed models for serially correlated functional data, with application to glaucoma data. *J Am Stat Assoc*. 2019;114(526):495-513.
17. Crainiceanu CM, Staicu AM, Ray S, Punjabi N. Bootstrap-based inference on the difference in the means of two correlated functional processes. *Stat Med*. 2012;31(26):3223-3240.
18. Park SY, Staicu AM, Xiao L, Crainiceanu CM. Simple fixed-effects inference for complex functional models. *Biostatistics*. 2018;19(2):137-152.
19. Chen S, Xiao L, Staicu AM. A smoothing-based goodness-of-fit test of covariance for functional data. *Biometrics*. 2019;75(2):562-571.
20. Eilers PHC, Marx BD. Flexible smoothing with B-splines and penalties. *Stat Sci*. 1996;11(2):89-121.
21. Wood SN, Pya N, Safken B. Smoothing parameter and model selection for general smooth models. *J Am Stat Assoc*. 2016;111(516):1548-1575.
22. Wood SN, Goude Y, Shaw S. Generalized additive models for large data sets. *J Royal Stat Soc C*. 2015;64(1):139-155.
23. Wood SN. mgcv: mixed GAM computation vehicle with automatic smoothness estimation; 2019.
24. Xiao L, Zipunnikov V, Ruppert D, Crainiceanu CM. Fast covariance estimation for high-dimensional functional data. *Stat Comput*. 2016;26(1):409-421.
25. Xiao L, Li Y, Ruppert D. Fast bivariate P-splines: the sandwich smoother. *J Royal Stat Soc B*. 2013;75(3):577-599.
26. Xiao L, Li C, Checkley W, Crainiceanu CM. Fast covariance estimation for sparse functional data. *Stat Comput*. 2018;28(3):511-522.
27. Wood SN. *Generalized Additive Models: An Introduction with R*. Boca Raton, FL: Chapman & Hall/CRC Press; 2006.
28. Zhong X, Shi H, Ming Q, et al. Whole-brain resting-state functional connectivity identified major depressive disorder: a multivariate pattern analysis in two independent samples. *J Affect Disord*. 2017;218:346-352.
29. Crainiceanu CM, Ruppert D. Likelihood ratio tests in linear mixed models with one variance component. *J Royal Stat Soc B*. 2004;66(1):165-185.
30. Wu CFJ. Jackknife, bootstrap and other resampling methods in regression analysis. *Ann Stat*. 1986;14(4):1261-1295.
31. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc B*. 1995;57(1):289-300.
32. Ivanescu AE, Staicu AM, Scheipl F, Greven S. Penalized function-on-function regression. *Comput Stat*. 2015;30(2):539-568.
33. Scheipl F, Gertheiss J, Greven S. Generalized functional additive mixed models. *Electron J Stat*. 2016;10(1):1455-1492.
34. Goldsmith J, Scheipl F, Huang L, et al. Refund: regression with functional data; 2018.
35. Sarkar SK, Chang CK. The Simes method for multiple hypothesis testing with positively dependent test statistics. *J Am Stat Assoc*. 1997;92(440):1601-1608.
36. John D, Tang Q, Albinali F, Intille S. An open-source monitor-independent movement summary for accelerometer data processing. *J Meas Phys Behav*. 2019;2(4):268-281.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Li R, Xiao L, Smirnova E, Cui E, Leroux A, Crainiceanu CM. Fixed-effects inference and tests of correlation for longitudinal functional data. *Statistics in Medicine*. 2022;1-16. doi: 10.1002/sim.9421