Received: 12 April 2022 | Revised: 9 November 2022 | Accepted: 26 January 2023

DOI: 10.1002/sim.9683

RESEARCH ARTICLE

Statistics in Medicine WILEY

Multivariate functional mixed model with MRI data: An application to Alzheimer's disease

Haotian Zou[1] | **Luo Xiao**[2] | **Donglin Zeng**[1] | **Sheng Luo**[3] | **for the Alzheimer's Disease Neuroimaging Initiative**[1]

[1]Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina

[2]Department of Statistics, North Carolina State University, Raleigh, North Carolina

[3]Department of Biostatistics and Bioinformatics, Duke University, Durham, North Carolina

Correspondence
Sheng Luo, Department of Biostatistics and Bioinformatics, Duke University, Durham, NC, USA.
Email: sheng.luo@duke.edu

Funding information
National Institute on Aging, Grant/Award Numbers: P30AG028716, P30AG072958, R01AG064803

Summary

Alzheimer's Disease (AD) is the leading cause of dementia and impairment in various domains. Recent AD studies, (ie, Alzheimer's Disease Neuroimaging Initiative (ADNI) study), collect multimodal data, including longitudinal neurological assessments and magnetic resonance imaging (MRI) data, to better study the disease progression. Adopting early interventions is essential to slow AD progression for subjects with mild cognitive impairment (MCI). It is of particular interest to develop an AD predictive model that leverages multimodal data and provides accurate personalized predictions. In this article, we propose a multivariate functional mixed model with MRI data (MFMM-MRI) that simultaneously models longitudinal neurological assessments, baseline MRI data, and the survival outcome (ie, dementia onset) for subjects with MCI at baseline. Two functional forms (the random-effects model and instantaneous model) linking the longitudinal and survival process are investigated. We use Markov Chain Monte Carlo (MCMC) method based on No-U-Turn Sampling (NUTS) algorithm to obtain posterior samples. We develop a dynamic prediction framework that provides accurate personalized predictions of longitudinal trajectories and survival probability. We apply MFMM-MRI to the ADNI study and identify significant associations among longitudinal outcomes, MRI data, and the risk of dementia onset. The instantaneous model with voxels from the whole brain has the best prediction performance among all candidate models. The simulation study supports the validity of the estimation and dynamic prediction method.

KEYWORDS

Alzheimer's disease, Bayesian method, dynamic prediction, functional regression, magnetic resonance imaging

1 | INTRODUCTION

Alzheimer's Disease (AD) is the leading cause of dementia and impairment in various domains (eg, cognition, memory, and daily activities).[1] In US, the predicted number of subjects with AD reach 6.5 million in 2022, imposing a significant health burden.[2] Recent AD studies, that is, Alzheimer's Disease Neuroimaging Initiative (ADNI) and National Alzheimer's Coordinating Center (NACC), collect multimodal data, including longitudinal neurological assessments,

biomarkers, medical imaging, and genomics data to better understand disease severity and progression pattern.[3,4] Mild cognitive impairment (MCI) is an important clinical progression stage of AD.[5] Cognition decline and impairment of one or more cognitive domains are presented in subjects with MCI due to AD. Adopting early personalized interventions in the MCI stage is essential to delay the time until dementia onset.[6] Hence, it is important to develop an AD predictive model that leverages multimodal data and provides accurate personalized predictions of dementia onset.

The joint model (JM) is a broadly applicable approach to jointly analyze longitudinal repeated measurements and survival outcome.[7,8] Recently, extensive research works applied the JM framework to investigate AD progression pattern and predict the risk of dementia onset.[9-11] In the JM framework, the longitudinal and survival outcomes are commonly modelled using generalized mixed models and Cox proportional hazards models, respectively.[8] Multiple functional forms have been investigated to link the association between the latent longitudinal mean and hazard function.[12,13] Recently, the extension of JM to modelling multivariate longitudinal processes is investigated in multiple literature, while complex functional forms are incorporated to model the association between these two processes.[14-16] Additionally, Bayesian shrinkage priors are adopted to determine the most appropriate functional form.[17] Dynamic prediction is an important application of JM, which provides personalized predictions of longitudinal trajectories and the probability of event occurrence (eg, dementia onset) based on subject's latest data,[8,18,19] and facilitates clinical decision making, treatment selection, and counseling the disease progress with patients.

Magnetic resonance imaging (MRI) is an widely used imaging technique that provides detailed anatomical structures of the whole brain.[20,21] A common MRI tool that detects volume change in regions of interest (ROI) for patients with AD is the structural MRI (sMRI). Voxel-based morphometry (VBM) is a standard analytical method that performs voxel-wise comparison of the ROI volumes among multiple groups of subjects.[22,23] For example, recent literature reported significant volume atrophy in hippocampus, cortical, and brainstem regions for subjects with early AD as compared with healthy controls using the VBM analysis.[24,25] Alternatively, the volumetric sMRI data can be treated as a functional outcome with millions of voxels observed on a dense voxel grid, and the functional regression approach can be adopted to handle the high-dimensional sMRI data as in previous literature.[9,11]

Functional regression is commonly used to model functional outcomes observed on a discrete grid.[26,27] To model the univariate longitudinal variable as a sparse functional outcome, Guo[28] proposed a semi-parametric functional mixed-effects model (FMEM) that decomposed the latent longitudinal mean into covariate-specific fixed functions and random effects functions. Yao[29] extended the FMEM to the shared latent functional principal components (FPC) model to jointly analyze the univariate functional longitudinal outcome and the survival process. To model the multivariate longitudinal outcomes, Li et al[10] proposed the non-parametric multivariate functional mixed model (MFMM) that captured nonlinear longitudinal trajectories and simultaneously modelled the survival outcome, linked by the FPC scores that are shared among multiple longitudinal outcomes and the outcome-specific FPC scores. On the other hand, it is essential to model voxel-wise sMRI data at baseline as a functional covariate because brain atrophy in some voxels negatively impact cognitive functions.[9,11] To this end, Zou et al[11] used parametric mixed models for multivariate longitudinal outcomes while modelling the baseline sMRI data as functional covariates. However, the model did not capture complex nonlinear longitudinal trajectories and was sensitive to model misspecification.

In this article, we propose to extend the flexible non-parametric MFMM proposed by Li et al,[10] to a multivariate functional mixed model with MRI data (MFMM-MRI) that jointly models multivariate longitudinal outcomes (ie, neurological scores), baseline MRI data, and the survival outcome (ie, dementia onset). We decompose the functional volumetric sMRI data in a similar fashion to the joint and individual variation explained (JIVE) approach.[30] Specifically, the MRI data are decomposed into the joint variation (the function determines the shared variation between longitudinal outcomes and MRI data) and the individual variation (the function determines the MRI-specific varying pattern). We also develop dynamic prediction methods that provide personalized predictions of dementia onset for subjects with MCI at baseline. Comparing with existing literature, our method is novel in three aspects. First, we use a non-parametric functional mixed model to model multivariate longitudinal outcomes and investigate two functional forms of the association between the longitudinal and survival process, which increase the flexibility in modelling and interpretation. Second, we include the baseline MRI data as a functional outcome and use a general and non-parametric decomposition similar in nature with JIVE. Third, we provide personalized predictions of longitudinal outcomes and the probability of remaining stable MCI, which facilitates clinical prognosis and early personalized interventions. The multivariate functional mixed model with MRI data (MFMM-MRI) provides personalized predictions of neurological scores and the risk of dementia onset using longitudinal outcomes, baseline MRI data, and the survival outcome. Additionally, we are able to determine the associations among those three types of outcomes.

This article is structured as follows. In Section 2, we give a brief introduction to the motivating ADNI study. In Section 3, we illustrate the MFMM-MRI model, its estimation, and the dynamic prediction method. In Section 4, the MFMM-MRI is applied to the motivating ADNI study. In Section 5, a simulation study is conducted to support the validity of the estimation and prediction method. In Section 6, we summarize our findings and discuss the limitations and future directions.

2 | MOTIVATION

The development of MFMM-MRI is motivated from the Alzheimer's Disease Neuroimaging Initiative (ADNI) study with completed ADNI-1/GO/2 stages and the ongoing ADNI-3 stage. The primary objective of the ADNI study is to determine and predict cognition decline using biomarkers, neurological assessments, medical imaging data, and genetic data for subjects with different stages of AD.[3] Based on a previous finding,[31] we select five longitudinal neurological measurements that are predictive of dementia onset: Alzheimer's Disease assessment scale - cognitive subscale, 13 items (ADAS-Cog 13), Rey auditory verbal learning test, immediate recall (RAVLT-immediate), RAVLT learning curve (RAVLT-learning), Mini-mental state examination (MMSE), and clinical dementia rating scale - sum of boxes (CDR-SB).[32-35] Higher values in ADAS-Cog 13 and CDR-SB and lower values in the other outcomes indicate worse cognitive and memory functions and are associated with faster AD progression. Figure 1 displays a spaghetti plot of these five selected longitudinal outcomes for randomly selected 15 subjects with MCI at baseline, suggesting the sparsity of longitudinal outcomes. Additionally, both subjects A and B show an increasing (clinically deteriorating) pattern for ADAS-Cog 13 and CDR-SB scores and a decreasing (clinically declining) trend for RAVLT and MMSE scores, suggesting that multivariate longitudinal outcomes are correlated within the same subject. Thus, we adopt the multivariate functional mixed model to model the five selected longitudinal outcomes observed on a sparse time grid.[10]

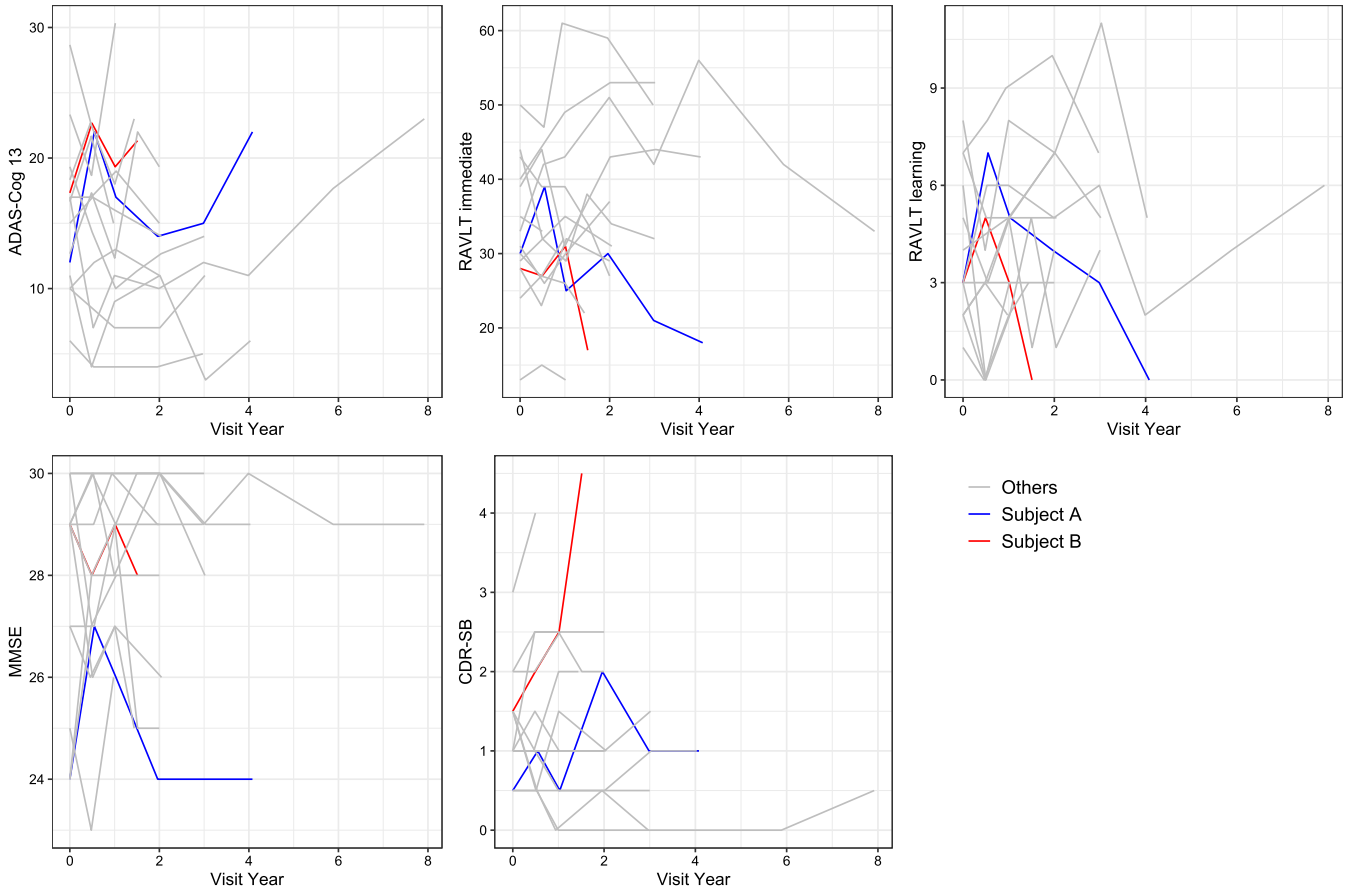

FIGURE 1 Spaghetti plot of ADAS-Cog 13, RAVLT-immediate, RAVLT-learning, MMSE, and CDR-SB scores for randomly selected 15 subjects with MCI at baseline

Based on the data downloaded on January 15, 2020, we include 1018 subjects with MCI at baseline from the ADNI-1/GO/2/3 stages. We exclude 86 subjects with missing ApoE- $\epsilon 4$ alleles because the ApoE- $\epsilon 4$ allele is associated with an increased risk of dementia occurrence.[36] Moreover, we exclude one subject with missing education years and 189 subjects with invalid or no MRI data at baseline. The survival time is the time from baseline to the first dementia diagnosis. Subjects without dementia diagnosis during the follow-up period are treated as right-censored and the event time is set as the last visit time. The final analysis dataset consists of 742 subjects and 3323 visits, with 268 dementia occurrences. The outcomes ADAS-Cog 13, RAVLT-immediate, RAVLT-learning, MMSE, CDR-SB have 29, 15, 15, 8, 35 missing values, respectively. Subjects have a mean follow-up period of 2.70 years, with SD being 2.47 years. The minimum and maximum number of visits are 1 visit and 13 visits, with median being 4 visits.

3 | METHODS

Suppose there are N subjects (denoted as $i = 1, 2, \ldots, N$) and J longitudinal outcomes (denoted as $j = 1, 2, \ldots, J$), measured at time $(t_{i1}, t_{i2}, \ldots, t_{iK_i})$, where K_i is the number of visits for subject i. Let $Y_{ij}(t_{ik})$ be the observed j-th longitudinal outcome for subject i and visit k, for $k = 1, \ldots, K_i$, with a random noise ϵ_{ijk}. Let $\tilde{Y}_{ij}(t_{ik})$ denote the missing longitudinal outcome and I_{ijk} be the missing status indicator (1 if outcome $Y_{ij}(t_{ik})$ is missing and 0 otherwise) and we assume missing at random (MAR). For subject i, We denote the baseline MRI data as $m_i(v)$ at voxel location v observed on a one-dimensional grid $\boldsymbol{V}$, which is the vector of voxels of interest with length V, where $V = 9846$ if hippocampal voxels are of clinical interest and $V = 642$ after voxel-selection from the whole brain. The event time and censoring status for subject i are denoted as $T_i = \min(T_i^*, C_i)$ and $\delta_i = I(T_i^* \leq C_i)$, respectively, where T_i^* and C_i are the failure time and censoring time, respectively, and $I()$ is an indicator function.

3.1 | MRI preprocess pipelines and voxel selection steps

The MRI data are T1-weighted structural MRI data, which undergone gradient warping, B1 non-uniformity correction, and N3 correction. The MRI preprocess pipelines and voxel selection steps are detailed in Zou et al[11] In brief, we register the neck-skull-stripped MRI data to the `JHU-MNI-ss (Eve)` Template and the registered MRI data are standardized and merged with the white matter parcellation map (WMPM Type II) of the `Eve` Template, with $S = 1\,843\,303$ voxels.

For the voxel selection steps, we independently fit Cox proportional hazards models for S voxels from the whole brain: $h_i(t) = h_0(t)\exp(\boldsymbol{Z}_i'\boldsymbol{\alpha} + \gamma_s \mathrm{Vol}_{i,s})$, where $\boldsymbol{Z}_i'$ is the covariate vector and $\mathrm{Vol}_{i,s}$ is the voxel volume for voxel s and subject i. We adjust the p-value of the regression coefficient γ_s using the false discovery rate (FDR) and the p-value threshold is set as 0.01.[11] Similar ideas of voxel selection were adopted in recent literature. For example, Huang et al[37] used the logistic regression classifier to identify significant voxels, where they utilized the survival information (whether patients converted to AD). Additionally, Petrone et al[38] used the Jacobian determinant maps for feature selection and they also classified patients into three groups and used the F-test to identify significant voxels.

3.2 | Model

The longitudinal outcome $Y_{ij}(t_{ik})$ is assumed to be a noisy measurement of the latent process $X_{ij}(t_{ik})$ that $Y_{ij}(t_{ik}) = X_{ij}(t_{ik}) + \epsilon_{ijk}$. The MRI data $m_i(v)$ is also assumed to have a random noise $\epsilon_{mi}(v)$. We denote subject i's baseline covariate vector as $\boldsymbol{Z}_i'$. We model the longitudinal outcome $Y_{ij}(t_{ik})$, baseline MRI data $m_i(v)$, and the survival process in the multivariate functional mixed model with MRI (MFMM-MRI):

$$Y_{ij}(t_{ik}) = X_{ij}(t_{ik}) + \epsilon_{ijk} = \mu_j(t_{ik}) + \beta_j(U_i(t_{ik}) + W_{ij}(t_{ik})) + \epsilon_{ijk}, \tag{1}$$

$$m_i(v) = \mu_m(v) + \beta_m u_{mi}(v) + f_{mi}(v) + \epsilon_{mi}(v), \tag{2}$$

$$h_i(t) = h_0(t)\exp(\boldsymbol{Z}_i'\boldsymbol{\alpha} + F(X_i, t) + \boldsymbol{g}_{mi}'\boldsymbol{\gamma}_m). \tag{3}$$

In Model (1), we decompose the latent longitudinal process $X_{ij}(t_{ik})$ into three components: the outcome-specific mean function $\mu_j(t_{ik})$, the shared random profile $U_i(t_{ik})$, and the subject-outcome specific random profile $W_{ij}(t_{ik})$, with an

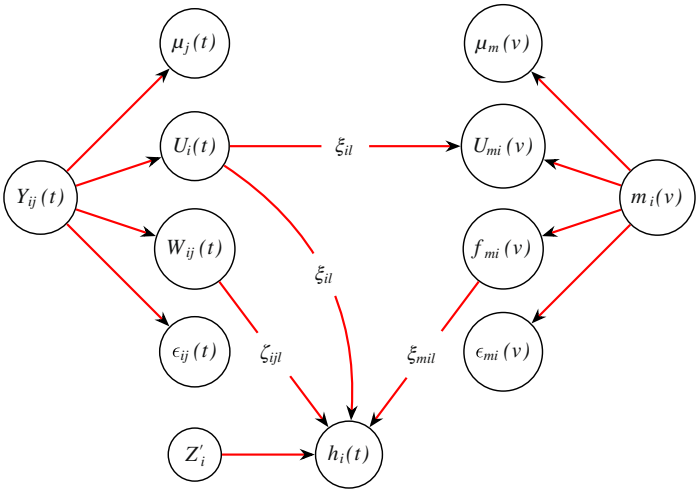

FIGURE 2 Directed acyclic graph showing the latent structural equations based on models (1) to (3)

association parameter β_j. The mean function $\mu_j(t)$ is time dependent and an additive model can be adopted to incorporate the time-varying effects of some baseline covariates. The shared random profile $U_i(t)$ quantifies the subject-specific random deviation from the mean function and is shared across all longitudinal outcomes. The subject-outcome specific random profile $W_{ij}(t)$ quantifies subject-outcome specific disease progression pattern and we assume random profiles $U_i(t)$, $W_{ij}(t)$, and random error ϵ_{ijk} are independent. Higher values of random profiles $U_i(t)$ and $W_{ij}(t)$ imply worse cognitive functions because we set $\beta_1 = 1$ for the outcome ADAS-Cog 13 to achieve model identifiability. The association parameter β_j is the expected increase of the outcome j in one unit increase of the random profiles $U_i(t)$ and $W_{ij}(t)$. The random error is assumed to follow $\epsilon_{ijk} \sim N(0, \sigma_j^2)$.

In Model (2), the subscript m is to denote the MRI data. We decompose the MRI data into three components, similar in nature with the JIVE approach,[30] (1) mean function $\mu_m(v)$; (2) joint variation (the MRI component $u_{mi}(v)$ shared with the multivariate longitudinal outcomes via the latent functional principal component scores ξ_{il} detailed in Section 3.3, and being multiplied by a scale parameter β_m); and (3) the individual variation (the MRI-specific varying pattern $f_{mi}(v)$), while the random noise $\epsilon_{mi}(v) \sim N(0, \sigma_m^2)$. The scale parameter β_m measures the strength of the correlation between the MRI data and longitudinal outcomes (eg, $\beta_m = 0$ indicates independence). We assume that the MRI components $u_{mi}(v)$, $f_{mi}(v)$, and random error $\epsilon_{mi}(v)$ are independent.

In the survival model (3), $h_0(t)$ is the baseline hazard function and $\boldsymbol{\alpha}$ is the regression coefficient vector. The functional form $F(X_i, t)$ quantifies the contribution of the multivariate longitudinal outcomes towards the hazard function, whose expressions are detailed in Section 3.3. The parameter vector $\boldsymbol{\gamma}_m$ is the contribution of the MRI-specific pattern $\boldsymbol{g}'_{mi}$ to the survival outcome, where $\boldsymbol{g}'_{mi}$ is the projection of the individual variation $f_{mi}(v)$ to its eigenfunctions (see details in Section 3.4). Figure 2 displays a directed acyclic graph showing the latent structural equations based on Models (1) to (3).

3.3 | Functional components, functional forms and Bayesian inference

In the longitudinal Model (1), we approximate the mean function $\mu_j(t)$ using cubic B-spline functions with P degrees of freedom, so that $\mu_j(t) \approx \sum_{j=1}^{P} b_p(t) A_{jp}$, where $b_p(t)$ is the p-th cubic B-spline function at time t and A_{jp} is the regression coefficient. We set P as 9 because the sensitivity analysis displayed in Supplementary Table S1 suggests that the model assessment statistics are very close for different numbers of B-spline functions: $P = (7, 8, 9, 10, 11)$. The random profiles $U_i(t)$ and $W_{ij}(t)$ are assumed to be two zero-mean stochastic processes with covariance functions $C_0(t, t')$ and $C_1(t, t')$, respectively, where t and t' are two different time points. By Mercer's Theorem, we decompose covariance functions $C_0(t, t') = \sum_{l=1}^{\infty} d_{0l} \phi_l(t) \phi_l(t')$ and $C_1(t, t') = \sum_{l=1}^{\infty} d_{1l} \psi_l(t) \psi_l(t')$, where d_{0l} and d_{1l} are non-increasing eigenvalues, and $\phi_l(t)$ and $\psi_l(t)$ are orthonormal eigenfunctions of $U_i(t)$ and $W_{ij}(t)$, respectively, that is, $\int \phi_l(t)^2 dt = 1$, and $\int \phi_l(t) \phi_{l'}(t) dt = 0$ for $l \neq l'$. By Karhunen-Loeve's expansion, we express random profiles $U_i(t) = \sum_{l=1}^{\infty} \xi_{il} \phi_l(t)$ and $W_{ij}(t) = \sum_{l=1}^{\infty} \zeta_{ijl} \psi_l(t)$, where ξ_{il} and ζ_{ijl} are independent functional principal component (FPC) scores, which are assumed to follow $\xi_{il} \sim N(0, d_{0l})$ and $\zeta_{ijl} \sim N(0, d_{1l})$.

In the MRI Model (2), we approximate the mean function $\mu_m(v)$ using the sample mean $\hat{\mu}_m(v)$ because the dense sMRI data provide good estimates of the mean function, and we smooth the mean function using cubic B-spline functions. We decompose the shared MRI component as $u_{mi}(v) = \sum_{i=1}^{\infty}\xi_{il}\phi_{ml}(v)$, and we approximate it as $u_{mi}(v) \approx \sum_{i=1}^{L_0}\xi_{il}\phi_{ml}(v)$, where $\phi_{ml}(v)$ is $u_{mi}(v)$'s orthonormal eigenfunctions and L_0 is the number of FPC scores for random profile $U_i(t)$ and its specification is detailed in Section 3.4. The FPC scores ξ_{il} are from the shared random profile $U_i(t)$ in Model (1) and they model the correlation between the MRI data and multivariate longitudinal outcomes. The MRI-specific varying pattern $f_{mi}(v)$ is assumed to be a zero-mean stochastic process with the covariance function $C_{mm}(v, v')$, where v and v' are two different voxels. Similar as above, we decompose $f_{mi}(v) = \sum_{l=1}^{\infty}\xi_{mil}\psi_{ml}(v)$, where ξ_{mil} is the MRI-specific FPC score, and $\psi_{ml}(v)$ is $f_{mi}(v)$'s orthonormal eigenfunctions, and we assume FPC scores $\xi_{mil} \sim N(0, d_{ml})$, where d_{ml} is the non-increasing eigenvalue of the MRI-specific FPC score ξ_{mil}. We approximate the MRI-specific varying pattern as $f_{mi}(v) \approx \sum_{l=1}^{L_m}\xi_{mil}\psi_{ml}(v)$, where the number of eigenfunctions L_m is determined by the proportion of variance explained (PVE) criteria and is detailed in Supplementary Section 2.[39]

We project the MRI data $m_i(v)$ onto L_m dimensional orthonormal eigenfunctions $\psi_{mj}(v)$, for $j = 1, \dots, L_m$. We obtain $m_{ij} \approx \mu_{mj} + \beta_m\sum_{l=1}^{L_0}\xi_{il}f_{lj} + \xi_{mil} + \epsilon_{mij}$, where $m_{ij} = \int_v m_i(v)\psi_{mj}(v)dv$, $\mu_{mj} = \int_v \mu_m(v)\psi_{mj}(v)dv$, $f_{lj} = \int_v \phi_{ml}(v)\psi_{mj}(v)dv$, and $\epsilon_{mij} = \int_v \epsilon_{mi}(v)\psi_{mj}(v)dv$. We denote the MRI-specific pattern as $\boldsymbol{g}'_{mi} = (\xi_{mi1}, \dots, \xi_{miL_m})$, which are used to model the hazard function in Model (3). We do not include FPC scores ξ_{il} in the MRI-specific pattern $\boldsymbol{g}'_{mi}$ because the function form $F(X_i, t)$ contains information of the FPC scores ξ_{il} and is detailed in the following context. By projection, we transform the high-dimensional MRI data (V voxels) to low-dimensional latent variables (L_m FPC scores), which facilitates computational efficiency while retaining most MRI-specific variation.

To investigate the association between multivariate longitudinal outcomes and survival process, we propose two functional forms of $F(X_i, t)$: the random-effects model and instantaneous model.

Model 1 (random-effects model), $F(X_i, t) = \sum_{l=1}^{\infty}\gamma_{0l}\xi_{il} + \sum_{j=1}^{J}\sum_{l=1}^{\infty}\gamma_{1jl}\zeta_{ijl}$. The association is quantified via the shared latent features (FPC scores) ξ_{il} and subject-outcome specific latent features ζ_{ijl}. We estimate the number of FPC scores in Section 3.4 and we approximate $F(X_i, t) \approx \sum_{l=1}^{L_0}\gamma_{0l}\xi_{il} + \sum_{j=1}^{J}\sum_{l=1}^{L_1}\gamma_{1jl}\zeta_{ijl}$, where L_0 is the number of FPC scores for the shared random profile $U_i(t)$ and L_1 the number of FPC scores for the subject-specific random profile $W_{ij}(t)$. We set $\boldsymbol{\gamma_0} = (\gamma_{01}, \dots, \gamma_{0L_0})$ and $\boldsymbol{\gamma_{1j}} = (\gamma_{1j1}, \dots, \gamma_{1jL_1})$.

Model 2 (instantaneous model), $F(X_i, t) = \gamma_0 U_i(t) + \sum_{j=1}^{J}\gamma_{1j}W_{ij}(t)$, which quantifies the association via the current values of the shared random profile $U_i(t)$ and subject-outcome specific random profile $W_{ij}(t)$ at time t. We set $\boldsymbol{\gamma_0} = \gamma_0$ and $\boldsymbol{\gamma_{1j}} = \gamma_{1j}$.

These two functional forms are widely used to investigate the association between the longitudinal and survival process for various types of diseases (eg, cardiovascular diseases,[13] primary biliary cirrhosis,[17] and Alzheimer's Disease[10]). For Model 1 (random-effects model), because the eigenfunctions $\phi_l(t)$ and $\psi_l(t)$ are forced to be positive for easy interpretation, higher values of the shared FPC scores ξ_{il} and ζ_{ijl} suggest worse cognitive behavior and faster dementia conversion. For Model 2 (instantaneous model), higher values of the shared random profile $U_i(t)$ and outcome-specific progression patterns $W_{ij}(t)$ are associated with a worse cognitive behavior and faster dementia conversion.

We denote the parameter space as $\boldsymbol{\Theta} = (\boldsymbol{A}, \boldsymbol{\beta}, \boldsymbol{d_0}, \boldsymbol{d_1}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}, \boldsymbol{\gamma}_m)$, where $\boldsymbol{A} = (\boldsymbol{A_1}, \dots, \boldsymbol{A_j}, \dots, \boldsymbol{A_J})$, $\boldsymbol{A_j} = (A_{j1}, \dots, A_{jP})$, $\boldsymbol{\beta} = (\beta_2, \dots, \beta_J)$, $\boldsymbol{d_0} = (d_{01}, \dots, d_{0L_0})$, $\boldsymbol{d_1} = (d_{11}, \dots, d_{1L_1})$, and $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_J)$. The log-likelihood is expressed as:

$$l(\boldsymbol{Y}|\boldsymbol{\Theta}) = \sum_{i=1}^{N}\log(L_i^{\text{long}}L_i^{\text{surv}}p(\boldsymbol{\xi}_i, \boldsymbol{\zeta}_i|\boldsymbol{d_0}, \boldsymbol{d_1})p(\boldsymbol{\xi}_{mi}|\boldsymbol{d_m})), \tag{4}$$

$$\text{where } L_i^{\text{long}} = \prod_{j=1}^{J}\prod_{k=1}^{K_i}p(Y_{ij}(t_{ik})|\xi_{il}, \zeta_{ijl})^{1-I_{ijk}}p(\tilde{Y}_{ij}(t_{ik})|\xi_{il}, \zeta_{ijl})^{I_{ijk}},$$

$$\text{and } L_i^{\text{surv}} = h_i(T_i|\boldsymbol{\xi}_i, \boldsymbol{\zeta}_i, \boldsymbol{g}_{mi})^{\delta_i}S_i(T_i|\boldsymbol{\xi}_i, \boldsymbol{\zeta}_i, \boldsymbol{g}_{mi}). \tag{5}$$

Equation (4) does not include the likelihood of the MRI data because the MRI data $m_i(v)$ is projected onto eigenfunctions $\psi_{ml}(v)$ and the MRI-specific FPC scores ξ_{mil} are estimated using their expectations detailed in Section 3.4. We denote the FPC score vector $\boldsymbol{\xi}_i = (\xi_{i1}, \dots, \xi_{iL_0})$, $\boldsymbol{\zeta}_i = (\boldsymbol{\zeta}_{i1}, \dots, \boldsymbol{\zeta}_{ij}, \dots, \boldsymbol{\zeta}_{iJ})$, $\boldsymbol{\zeta}_{ij} = (\zeta_{ij1}, \dots, \zeta_{ijL_1})$, and $\boldsymbol{\xi}_{mi} = (\xi_{mi1}, \dots, \xi_{miL_m})$. We denote $S_i(t)$ as the survival probability at time t for subject i. We use a piecewise constant function to model the baseline hazard function $h_0(t)$.[40,41] We use the Gaussian quadrature to approximate the survival function $S_i(t)$, whose formulation is detailed in Supplementary Section 3. The choice of priors and derivation of posterior distribution are detailed as follows:

$$\boldsymbol{A}, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}, \boldsymbol{\gamma}_m \sim N(0, 10^2),$$
$$\boldsymbol{d_0}, \boldsymbol{d_1}, \boldsymbol{\sigma} \sim \text{Inv\_Gamma}(0.1, 0.1),$$
$$p(\boldsymbol{\Theta}|\boldsymbol{Y}) \propto \prod_{i=1}^{N} L_i^{\text{long}} L_i^{\text{surv}} p(\boldsymbol{\xi}_i, \boldsymbol{\zeta}_i|\boldsymbol{d_0}, \boldsymbol{d_1}) p(\boldsymbol{\xi}_{mi}|\boldsymbol{d_m}) p(\boldsymbol{\Theta}),$$

where Inv_Gamma is the inverse gamma distribution and $p(\boldsymbol{\Theta})$ is the prior distribution of the parameter space $\boldsymbol{\Theta}$.

We adopt the Markov Chain Monte Carlo (MCMC) method to obtain posterior samples from the likelihood based on No-U-Turn Sampling (NUTS) algorithm implemented in `Stan`.[42] The missing outcomes $\tilde{Y}_{ij}(t_{ik})$ are treated as unknown parameters and are imputed in the MCMC procedure. The scale reduction statistic $\hat{R}$ is used to assess the convergence of the model, where $\hat{R} < 1.1$ indicates satisfied convergence.[42] For the real data analysis and simulation study, we use two chains with 2000 warm-up iterations and 1000 post burn-in iterations for each MCMC chain. From a total of 2000 posterior samples, we obtain the estimated mean, standard deviation, and 95% credible intervals. To facilitate easy implementation, the `Stan` code is given in Supplementary Section 9. Refer to https://github.com/zouhaotian/MFMM_MRI for detailed R codes in real data analysis and simulation study.

3.4 | Estimation of functional components and model assessment criteria

Let $C_{jj'}(t, t')$ be the covariance function between the latent processes $X_{ij}(t)$ and $X_{ij'}(t')$. We use the multivariate fast covariance estimation (mFACE) method to estimate the covariance function $C_{jj'}(t, t')$ on a refined grid $\boldsymbol{S}$, where $\boldsymbol{S}$ is an equally spaced time grid.[43] In Supplementary Section 1, we provide the detailed estimation procedure to obtain the estimated covariance functions $\hat{C}_0(t, t')$ and $\hat{C}_1(t, t')$. We use the eigendecomposition to obtain the estimated eigenfunctions $\hat{\phi}_l(t)$ and $\hat{\psi}_l(t)$ and eigenvalues $\hat{d}_{0l}$ and $\hat{d}_{1l}$. Using the PVE criteria, we identify the number of FPC scores as L_0 for the shared random profile $U_i(t)$, and L_1 for the subject-outcome specific random profile $W_{ij}(t)$. And we approximate the random profiles $U_i(t) \approx \sum_{l=1}^{L_0} \xi_{il}\phi_l(t)$ and $W_{ij}(t) \approx \sum_{l=1}^{L_1} \zeta_{ijl}\psi_l(t)$.

We denote the covariance function between the longitudinal outcome $Y_{ij}(t)$ and MRI data $m_i(v)$ as $C_{jm}(t, v) = \text{cov}(Y_{ij}(t), m_i(v))$. We denote the covariance function between the MRI data $m_i(v)$ and $m_i(v')$ as $C_{mm}(v, v') = \text{cov}(m_i(v), m_i(v'))$. We express the covariance functions using the following equations (refer to Supplementary Section 2 for the detailed derivation):

$$C_{jm}(t, v) = \beta_j \beta_m \sum_{l=1}^{\infty} d_{0l}\phi_l(t)\phi_{ml}(v), \tag{6}$$

$$C_{mm}(v, v') = \beta_m^2 \sum_{l=1}^{\infty} d_{0l}\phi_{ml}(v)\phi_{ml}(v') + \sum_{l=1}^{\infty} d_{ml}\psi_{ml}(v)\psi_{ml}(v') + I(v = v')\sigma_m^2. \tag{7}$$

We denote the demeaned MRI data as: $m_i^*(v) = m_i(v) - \hat{\mu}_m(v)$. From the multivariate longitudinal model (1), we obtain the number of FPC scores L_0, estimated FPC scores $\hat{\xi}_{il}$, and the variances of FPCs $\hat{d}_{0l}$, for $l = 1, \ldots, L_0$. More explicitly, we denote the FPC score vector $\hat{\boldsymbol{\xi}}_i = (\hat{\xi}_{i1}, \ldots, \hat{\xi}_{iL_0})$, and $\hat{\boldsymbol{D}}_0$ is a diagonal matrix with its diagonal elements being the eigenvalue $\hat{d}_{0l}$. To estimate eigenfunctions $\phi_{ml}(v)$, $\psi_{ml}(v)$, the scale parameter β_m, and the standard deviation σ_m, we adopt an iterative search procedure detailed in Supplementary Section 2. To facilitate easy implementation, we display an algorithm graph 1.

After we obtain the estimated eigenfunctions $\hat{\phi}_{ml}(v)$ and $\hat{\psi}_{ml}(v)$, we use the Riemann's sum to approximate m_{ij}, so $\hat{m}_{ij} \approx \int_V m_i(v)\hat{\psi}_{mj}(v)dv$, $\hat{\mu}_{mj} \approx \int_V \hat{\mu}_m(v)\hat{\psi}_{mj}(v)dv$, $\hat{f}_{lj} \approx \int_V \hat{\phi}_{ml}(v)\hat{\psi}_{mj}(v)dv$. Thus, we estimate the FPC score ξ_{mij} using its expectation: $E(\hat{\xi}_{mij}) = \hat{m}_{ij} - (\hat{\mu}_{mj} + \hat{\beta}_m \sum_{l=1}^{L_0} \xi_{il}\hat{f}_{lj})$ because $E(\hat{\epsilon}_{mij}) = E(\int_v \epsilon_{mi}(v)\hat{\psi}_{mj}(v)dv) = 0$, and the MRI-specific pattern $\boldsymbol{g}'_{mi} = (\xi_{mi1}, \ldots, \xi_{miL_m})$. The eigenvalue d_{ml} is estimated based on the variance of the FPC scores ξ_{mil}.

We assess the model fitting performance using empirical Akaike information criterion (EAIC), empirical Bayesian information criterion (EBIC), leave-one-out information criterion (LOOIC), and widely applicable information criterion (WAIC).[44,45] Lower values of the above statistics suggest a better model fitting performance. Specifically, we first compute the posterior mean deviation $\overline{D}(\boldsymbol{\Theta}) = -2E(\sum_{i=1}^{N} l_i(\boldsymbol{\Theta}))$ using Monte Carlo approximation: $\overline{D}(\boldsymbol{\Theta}) \approx -2/Q \sum_{q=1}^{Q} \sum_{i=1}^{N} l_i(\boldsymbol{\Theta}^{(q)})$, where $l_i(\boldsymbol{\Theta})$ is the log-likelihood for subject i, Q is the number of posterior samples, and the superscript (q) denotes the q-th posterior sample. The EAIC and EBIC statistics are computed as: $\text{EAIC} = \overline{D}(\boldsymbol{\Theta}) + 2p$, and $\text{EBIC} = \overline{D}(\boldsymbol{\Theta}) + p \log N$,

Algorithm 1. Estimation algorithm

Require: The estimated FPC scores $\hat{\xi}_{il}$ and variance $\hat{d}_{0l}$ via mFACE method.
Input: $\hat{\xi}_{il}$, $\hat{d}_{0l}$, and MRI data $m_i(v)$.
Output: Eigenfunctions $\phi_{ml}(v)$, $\psi_{ml}(v)$, scale parameter β_m, and SD σ_m.

1: **Initialization** Smooth $\hat{C}^*_M(v,v') = \text{cov}(m_i(v), m_i(v'))$. Estimate $\hat{\sigma}_m = 1/\text{V}\sum_{v=1}^{V} C^*_M(v,v) - \hat{C}^*_M(v,v)$. Eigendecomposition on $\hat{C}^*_M(v,v')$ for L_0 eigenvalues $\hat{d}^*_{0l}$ and eigenfunctions $\phi^{(0)}_{ml}(v)$.
2: **Initialization** Regress $\hat{d}^*_{0l}$ with $\hat{d}_{0l}$ for β^2_m. Regress $\hat{C}_{jm}(t,v) = \text{cov}(Y_{ij}(t), m_i(v))$ with $\hat{\beta}_j \sum_{l=1}^{L_0} \hat{d}_{0l}\hat{\phi}_l(t)\phi^{(0)}_{ml}(v)$ to determine sign of $\beta^{(0)}_m$.
3: **while** $|\beta^{(t+1)}_m - \beta^{(t)}_m| \geq \epsilon = 1e-5$ **do**
4: Compute $M^*_1(v) = M^*(v) - \beta^{(0)}_m \hat{\boldsymbol{\xi}}\Phi^{(0)}_m$. Smooth $\hat{C}^*_{M1}(v,v') = \text{cov}(M^*_1(v), M^*_1(v'))$. Eigendecomposition on $\hat{C}^*_{M1}(v,v')$ for first L_m (determined by PVE) eigenfunctions $\psi^{(1)}_{ml}(v)$ and eigenvalues $d^{(1)}_{ml}$.
5: Smooth $\hat{C}^*_{M2}(v,v') \approx \hat{C}^*_M(v,v') - \sum_{l=1}^{L_m} d^{(1)}_{ml}\psi^{(1)}_{ml}(v)\psi^{(1)}_{ml}(v')$. Eigendecomposition on $\hat{C}^*_{M2}(v,v')$ for first L_0 eigenfunctions $\phi^{(1)}_{ml}(v)$ and eigenvalues $\hat{d}^*_{0l}$. Estimate $\beta^{(1)}_m$ by regressing $\hat{d}^*_{0l}$ and $\hat{d}_{0l}$ and taking squared root.
6: **end while**
7: **Finalize** Estimate SD and 95% CI for $\hat{\beta}_m$ and SD σ_m from B bootstrap replications.

where p is the number of parameters in the parameter space $\boldsymbol{\Theta}$. The LOOIC and WAIC statistics are computed based on the log pointwise predictive density implemented in the `loo` package.[45]

3.5 | Dynamic prediction

Suppose there is a new subject B with multivariate longitudinal outcomes observed at K_B visits and the baseline MRI data. For a future time T' ($T' > T$), we predict subject B's longitudinal outcomes and the event-free probability at time T'. Let $\boldsymbol{Y}^{(T)}_B = (Y_{Bj}(t_{Bk}))_{j=1,\dots,J,k=1,\dots,K_B}$ be all longitudinal outcomes until time T, where $Y_{Bj}(t_{Bk})$ is j-th longitudinal outcome for subject B at visit k. We denote the FPC score vector for longitudinal outcomes as $\boldsymbol{\xi}_B$ and $\boldsymbol{\zeta}_B$ for subject B. We denote the covariate vector as $\boldsymbol{Z}'_B$ and the FPC score vector for the MRI specific-varying pattern as $\boldsymbol{\xi}_{mB}$. Let $\boldsymbol{g}'_{mB} = (\xi_{mB1}, \dots, \xi_{mBL_m})$ denote the MRI-specific pattern as in Section 3.4. We adopt the MCMC method to obtain samples from the posterior distribution of the score vectors $\boldsymbol{\xi}_B$, $\boldsymbol{\zeta}_B$, and $\boldsymbol{\xi}_{mB}$ as in Equation (8). We use two MCMC chains, each of which consists of 2000 warmup iterations and 1000 post burn-in iterations, with a total of $Q = 2000$ posterior samples.

$$P(\boldsymbol{\xi}_B, \boldsymbol{\zeta}_B, \boldsymbol{\xi}_{mB}|\boldsymbol{Y}^{(T)}_B, T, \delta_B = 0, \hat{\boldsymbol{\Theta}}) \propto \prod_{j=1}^{J}\prod_{k=1}^{K_B} p(Y_{Bj}(t_{Bk})|\boldsymbol{\xi}_B, \boldsymbol{\zeta}_B, \hat{\boldsymbol{\Theta}}) S_B(T|\boldsymbol{Z}'_B, \boldsymbol{\xi}_B, \boldsymbol{\zeta}_B, \boldsymbol{g}_{mB}, \hat{\boldsymbol{\Theta}}) p(\boldsymbol{\xi}_B, \boldsymbol{\zeta}_B, \boldsymbol{\xi}_{mB}|\hat{\boldsymbol{\Theta}}), \tag{8}$$

where $\hat{\boldsymbol{\Theta}}$ is the estimated mean of $\boldsymbol{\Theta}$. Equation (8) does not include the likelihood of the MRI data because the MRI-specific FPC scores are estimated via projection detailed in Section 3.4.

We compute the predicted longitudinal outcomes and survival probability at time T' from Q posterior samples as in Equations (9) and (10), respectively.

$$E(Y_{Bj}(T')|\boldsymbol{\xi}^{(q)}_B, \boldsymbol{\zeta}^{(q)}_B, \hat{\boldsymbol{\Theta}}) \approx \hat{\mu}_j(T') + \hat{\beta}_j\left(\sum_{l=1}^{L_0}\xi^{(q)}_{Bl}\hat{\phi}_l(T') + \sum_{l=1}^{L_1}\zeta^{(q)}_{Bjl}\hat{\psi}_l(T')\right), \tag{9}$$

$$S_B(T'|T, \boldsymbol{Z}'_B, \boldsymbol{\xi}^{(q)}_B, \boldsymbol{\zeta}^{(q)}_B, \boldsymbol{g}^{(q)}_{mB}, \hat{\boldsymbol{\Theta}}) \approx S_B(T'|\boldsymbol{Z}'_B, \boldsymbol{\xi}^{(q)}_B, \boldsymbol{\zeta}^{(q)}_B, \boldsymbol{g}^{(q)}_{mB}, \hat{\boldsymbol{\Theta}})/S_B(T|\boldsymbol{Z}'_B, \boldsymbol{\xi}^{(q)}_B, \boldsymbol{\zeta}^{(q)}_B, \boldsymbol{g}^{(q)}_{mB}, \hat{\boldsymbol{\Theta}}), \tag{10}$$

where the superscript (q) denotes the q-th posterior sample. We compute the mean of predicted longitudinal outcomes and event-free survival probability based on Q predicted values.

To assess the model predictive performance, we use discrimination (how well the model discriminates subjects with or without event) and calibration measurements (how well the model agrees with the observed survival status). Specifically, the area under the time-dependent receiver operating characteristic curve (AUC) is adopted to measure the discrimination, based on the trapezoidal integration of the time-dependent sensitivity and specificity.[46] For the

calibration measurement, we use the Brier score (BS) based on the average of the weighted squared differences between the predicted and observed survival status.[47] Higher values of AUC and lower values of BS suggest a better predictive performance. For the illustration purpose, we integrate the estimated AUC and BS values over a dense prediction window δ_t given a landmark time T and we obtain the integrated AUC (iAUC) and integrated BS (iBS) using the Simpson's Rule.

4 | REAL DATA APPLICATION

For subject i at visit k, let $Y_{i1}(t_{ik}), \ldots, Y_{i5}(t_{ik})$ denote the observed ADAS-Cog 13, RAVLT-immediate, RAVLT-learning, MMSE, and CDR-SB scores after Box-Cox transformation (for normality). We determine the number of FPC scores of the shared random profile $U_i(t)$ and subject-outcome random profile $W_{ij}(t)$ using 95% PVE threshold and we obtain $L_0 = 2$ and $L_1 = 1$. We use a piecewise constant function to model the baseline hazard function with knots at 4, 6, and 8 years. The knots are so selected because the estimated cumulative hazard function displayed in Supplementary Figure S1 suggests that its slope is slightly different among 0 to 4 years, 4 to 6 years, 6 to 8 years, and 8 years later. We explored other knot selections and obtained very similar results (not shown). The covariate vector $\boldsymbol{Z}_i' = (\text{Age}_i, \text{Sex}_i, \text{Education}_i, \text{ApoE}_i)$, where we set Male as the reference sex, education as the number of education years, and ApoE as the number of ApoE- $\epsilon 4$ alleles.

We compare the model fitting performance for the three sets of models: (1) MFMM-whole: the MRI data is from whole-brain voxels after voxel selection, with $V = 642$ voxels and $L_m = 10$ using 95% PVE; (2) MFMM-hippo: the MRI data is from hippocampal voxels only, with $V = 9846$ voxels and $L_m = 13$ using 95% PVE; (3) MFMM: the model is without MRI data. The reason of comparing MFMM-hippo vs. MFMM is that hippocampal volume shrinkage is an important risk factor of dementia onset.[24] Similarly, we compare two functional forms of $F(X_i, t)$ as in Section 3.3, that is, Model 1 (random effects model) and Model 2 (instantaneous model), rendering a total of six models for comparison. We also compare the MFMM-MRI model with other four candidate models: Model 3 - 2S (a two-stage sequential estimation approach based on Model 2); Model 3 - MJM-MRI (the multivariate joint model with MRI data,[11] with details available in Supplementary Section 5); Model 3 - bCox (the Cox model with baseline longitudinal outcomes); Model 3 - NM (the non-mixed functional model that the longitudinal model is $Y_{ij}(t_{ik}) = \mu_j(t_{ik}) + \epsilon_{ijk}$ and the Cox model includes baseline covariates). We compare MFMM-MRI with four candidate models to show the advantages of the joint modelling vs. the two-stage approach, modeling nonlinear longitudinal outcomes and MRI data using functional mixed models, modelling longitudinal trajectories in the survival model, and modelling the random component functions in the longitudinal model.

Table 1 compares model fitting performance of candidate models. Comparing the three sets of models in Model 1, Model 1 with whole-brain voxels (M1-whole) has the smallest LOOIC and WAIC statistics, suggesting a better model

TABLE 1 Model comparison statistics for candidate models: Model 1-whole (random-effects model with whole-brain voxels); Model 1-hippo (random-effects model with hippocampal voxels); Model 1 (random-effects model without MRI data); Model 2 (instantaneous model); Model 3 - 2S (two-stage approach); Model 3 - MJM-MRI (multivariate joint model with MRI data); Model 3 - bCox (the Cox model with baseline longitudinal outcomes); Model 3 - NM (non-mixed functional model)

Model	EAIC	EBIC	LOOIC	WAIC	Time (h)
Model 1-whole	19 917	20 295	23 284	22 570	3.3
Model 1-hippo	19 950	20 342	23 334	22 618	2.6
Model 1	19 915	20 247	23 314	22 583	2.5
Model 2-whole	**19 887**	**20 260**	**23 283**	**22 560**	5.9
Model 2-hippo	19 922	20 309	23 277	22 596	8.0
Model 2	19 892	20 219	23 303	22 584	6.7
Model 3 - 2S	20 998	21 325	24 311	23 662	5.0
Model 3 - MJM-MRI	26 645	27 019	28 104	28 005	11.4
Model 3 - bCox	20 011	20 333	23 340	22 661	5.8
Model 3 - NM	40 317	40 585	40 460	40 476	1.0

Note: The computation is run a four-core 2.50 GHz Intel processor.
Abbreviations: EAIC, empirical akaike information criterion; EBIC, empirical Bayesian information criterion; LOOIC, leave-one-out information criterion; WAIC, widely application information criterion.

performance. Model 1 with hippocampal voxels (M1-hippo) and without MRI data (M1) show similar model assessment statistics. Table 1 suggests that Model 2 with whole-brain voxels (M2-whole) has a better performance (smaller EAIC and WAIC and similar LOOIC) as compared with Model 2 with hippocampal voxels (M2-hippo) and without MRI data (M2). Since EBIC statistics give a large penalty to the number of parameters, Model 2 with whole-brain voxels (M2-whole) show slightly larger EBIC statistics as compared to Model 2. Model 2 with whole-brain voxels (M2-whole) has smaller assessment statistics as compared with M1-whole, suggesting that M2-whole has the best performance among all candidate models. Additionally, Model 2 with whole-brain voxels (M2-whole) show substantially better performance as compared to Model 3 - 2S, Model 3 - MJM-MRI, Model 3 - bCox, and Model 3 - NM, indicating the advantages of joint modelling vs. the two-stage approach, non-parametric mixed functional models vs parametric linear models, modelling longitudinal trajectories in the survival model, and modelling random component functions in the longitudinal model. Thus, we select Model 2 with whole-brain voxels (M2-whole) as the final model. Specifically in the survival model (3), the covariate vector $\boldsymbol{Z}_{\boldsymbol{i}}' = (\text{Age}_i, \text{Sex}_i, \text{Education}_i, \text{ApoE}_i)$, the functional form $F(X_i, t) = \gamma_0 U_i(t) + \sum_{j=1}^{J=5} \gamma_{1j} W_{ij}(t)$, and the MRI-specific pattern $\boldsymbol{g}_{mi}' = (\xi_{mi1}, \ldots, \xi_{mi10})$, where $U_i(t)$ is the shared random profile, $W_{ij}(t)$ is the subject-outcome specific random profile, and ξ_{mil} is the MRI-specific FPC score, for $l = 1, \ldots, L_m = 10$.

Table 2 summarizes the mean, SE, and 95% credible intervals from Model 2 with whole-brain voxels. For the multivariate longitudinal outcomes, the estimated association parameters $\hat{\beta}_2$ to $\hat{\beta}_5$ are all significant, suggesting that one unit increase of the shared random profile $U_i(t)$ and subject-outcome specific random profile $W_{ij}(t)$ are associated with 0.540 unit decrease in RAVLT-immediate (95% CI: [− 0.567, − 0.514]), 0.600 unit decrease in RAVLT-learning (95% CI: [− 0.636, − 0.563]), 0.033 unit decrease in MMSE (95% CI: [− 0.035, − 0.031]), and 0.159 unit increase in CDR-SB (95% CI: [0.150, 0.169]), after their Box-Cox transformation (refer to Supplementary Table S2 for power parameters used in Box-Cox transformation). Results suggest that higher values of ADAS-Cog 13 (worse cognitive functions) are associated with lower values of RAVLT-immediate, RAVLT-learning, MMSE, and higher CDR-SB scores (worse functional activities and memory functions). The estimated variances of FPC scores, mean functions, and eigenfunctions are displayed in Supplementary Table S2 and Figure S2. For the MRI components, the estimated scale parameter $\hat{\beta}_m$ is − 0.053 (95% bootstrap CI: [− 0.072, − 0.035] based on 120 bootstrap replications), suggesting that one unit increase in $u_{mi}(v)$ (the MRI component shared with longitudinal outcomes) is associated with 0.053 unit decrease in the volume for voxel v. Results suggest that higher values of FPC scores ξ_{il} (worse cognitive functions) are associated with a greater decrease of the voxel volumes in MRI data. Figure 3 displays the decomposition of the MRI data for a randomly selected subject in the brain heat map (in sagittal view) using BrainNet Viewer.[48]

For the survival outcome as dementia onset, results suggest that one additional ApoE- $\epsilon 4$ allele is associated with a 27.9% (95% CI: [0.078, 0.472]) increase in log hazard. The estimated association parameter $\hat{\gamma}_0$ is significant, suggesting that one unit increase of the shared random profile $U_i(t)$ (worse cognitive behaviors) is associated with a 60.7% (95% CI: [0.522, 0.702]) increase in log hazard (higher risk of dementia onset). The estimated association parameters $\hat{\gamma}_{11}$, $\hat{\gamma}_{12}$, and $\hat{\gamma}_{15}$ are marginally significant, suggesting higher values of ADAS-Cog 13, lower values of RAVLT-immediate, and higher values of CDR-SB are associated with a higher risk of dementia onset. Most of the estimated association parameters $\hat{\gamma}_{ml}$, for $l = 1, \ldots, 10$, are negative and statistically significant, suggesting that voxel volume shrinkage in MRI data is associated with a higher risk of dementia onset. We use the likelihood ratio test (LRT) to test the hypothesis that $H_0 : \gamma_{m1} = \cdots = \gamma_{m10} = 0$ vs. $H_1 :$ otherwise.[49] The test statistic $T = 39.89 \xrightarrow{d} \chi^2_{10}$, and p-value is less than 0.001, providing sufficient evidence of the significant association between the MRI-specific varying pattern and dementia onset.

We randomly split the dataset into the training and testing set with sample sizes being 556 and 186, respectively (3:1 ratio). We repeat the dynamic prediction for 100 times and compute the mean of iAUC and iBS. Table 3 compares the predictive performance for all candidate models using iAUC and iBS measurements at landmark times 2, 2.5, 3, 3.5, and 4 years. The prediction window $\delta_t = (0.5, 0.51, \ldots, 1.5)$. Comparing the three sets of models in Model 1, results suggest that Model 1 with whole-brain voxels (M1-whole) shows highest iAUC and similar iBS values as compared with Model 1 with hippocampal voxels (M1-hippo) and without MRI data (M1), for all landmark times. Results also suggest that Model 2 with whole-brain voxels (M2-whole) has the best predictive performance as compared with Model 2 with hippocampal voxels (M2-hippo) and without MRI data (M2), and Model 1 with whole-brain voxels (M1-whole). Comparing Model 2 with whole-brain voxels (M2-whole) with the other four candidate models (Model 3 - 2S, Model 3 - MJM-MRI, Model 3 - bCox, and Model 3 - NM), the results suggest that Model 2 has higher iAUC values for most landmark times. Thus, the proposed model show improved predictive performance as compared to the models using the two-stage approach, parametric linear mixed models, without longitudinal trajectories in the survival model, and without random component functions in the longitudinal model. In summary, Model 2 with whole-brain voxels (M2-whole) has the best fitting and predictive performance among all candidate models.

TABLE 2 Mean, standard error, and 95% credible intervals for parameters from model 2 (instantaneous model) with whole-brain voxels (M2-whole)

	Mean	SE	2.5%	97.5%
Longitudinal outcomes and MRI data				
β_2: RAVLT-immediate	− 0.540	0.014	− 0.567	− 0.514
β_3: RAVLT-learning	− 0.600	0.019	− 0.636	− 0.563
β_4: MMSE	− 0.033	0.001	− 0.035	− 0.031
β_5: CDR-SB	0.159	0.005	0.150	0.169
σ_1: ADAS-Cog 13	1.065	0.016	1.035	1.096
σ_2: RAVLT-immediate	0.612	0.008	0.596	0.629
σ_3: RAVLT-learning	1.334	0.018	1.298	1.370
σ_4: MMSE	0.059	0.001	0.057	0.060
σ_5: CDR-SB	0.263	0.004	0.256	0.271
β_m	− 0.053	0.009	− 0.072	− 0.035
σ_m	0.404	0.007	0.392	0.417
Survival outcome				
Age	− 0.022	0.009	− 0.040	− 0.004
Female sex	0.134	0.145	− 0.148	0.423
Education years	− 0.006	0.024	− 0.055	0.041
ApoE- ϵ4 alleles	0.279	0.100	0.078	0.472
γ_0	0.607	0.045	0.522	0.702
γ_{11}: ADAS-Cog 13	0.169	0.099	− 0.012	0.364
γ_{12}: RAVLT-immediate	0.164	0.092	− 0.010	0.336
γ_{13}: RAVLT-learning	0.006	0.094	− 0.174	0.200
γ_{14}: MMSE	− 0.132	0.081	− 0.288	0.029
γ_{15}: CDR-SB	0.279	0.083	0.120	0.446
γ_{m1}	− 0.752	0.230	− 1.199	− 0.304
γ_{m2}	− 1.194	0.306	− 1.792	− 0.594
γ_{m3}	− 0.358	0.419	− 1.155	0.486
γ_{m4}	0.922	0.593	− 0.229	2.046
γ_{m5}	− 1.765	0.601	− 2.952	− 0.573
γ_{m6}	− 1.387	0.913	− 3.228	0.495
γ_{m7}	1.739	0.954	− 0.108	3.572
γ_{m8}	− 3.129	1.060	− 5.237	− 1.086
γ_{m9}	− 0.167	1.416	− 2.985	2.567
γ_{m10}	0.707	1.452	− 2.174	3.540

Note: The SE and 95% CI for the scale parameter β_m and standard deviation σ_m is derived using bootstrap method as in Section 3.4. The parameters γ_0 and γ_{1j} are the association parameters for the shared random profile $U_i(t)$ and subject-outcome specific random profile $W_{ij}(t)$ with the risk of dementia onset, respectively, for $j = 1, \ldots, J = 5$. The parameter γ_{ml} is the association parameter between the MRI-specific varying pattern ξ_{mil} and the risk of dementia onset, for $l = 1, \ldots, L_m = 10$.

Abbreviations: ADAS-Cog, the Alzheimer's disease assessment scale - cognitive subscale; CDR-SB, clinical dementia rating scale - sum of boxes; MMSE, mini-mental state examination; RAVLT, Rey auditory verbal learning test.

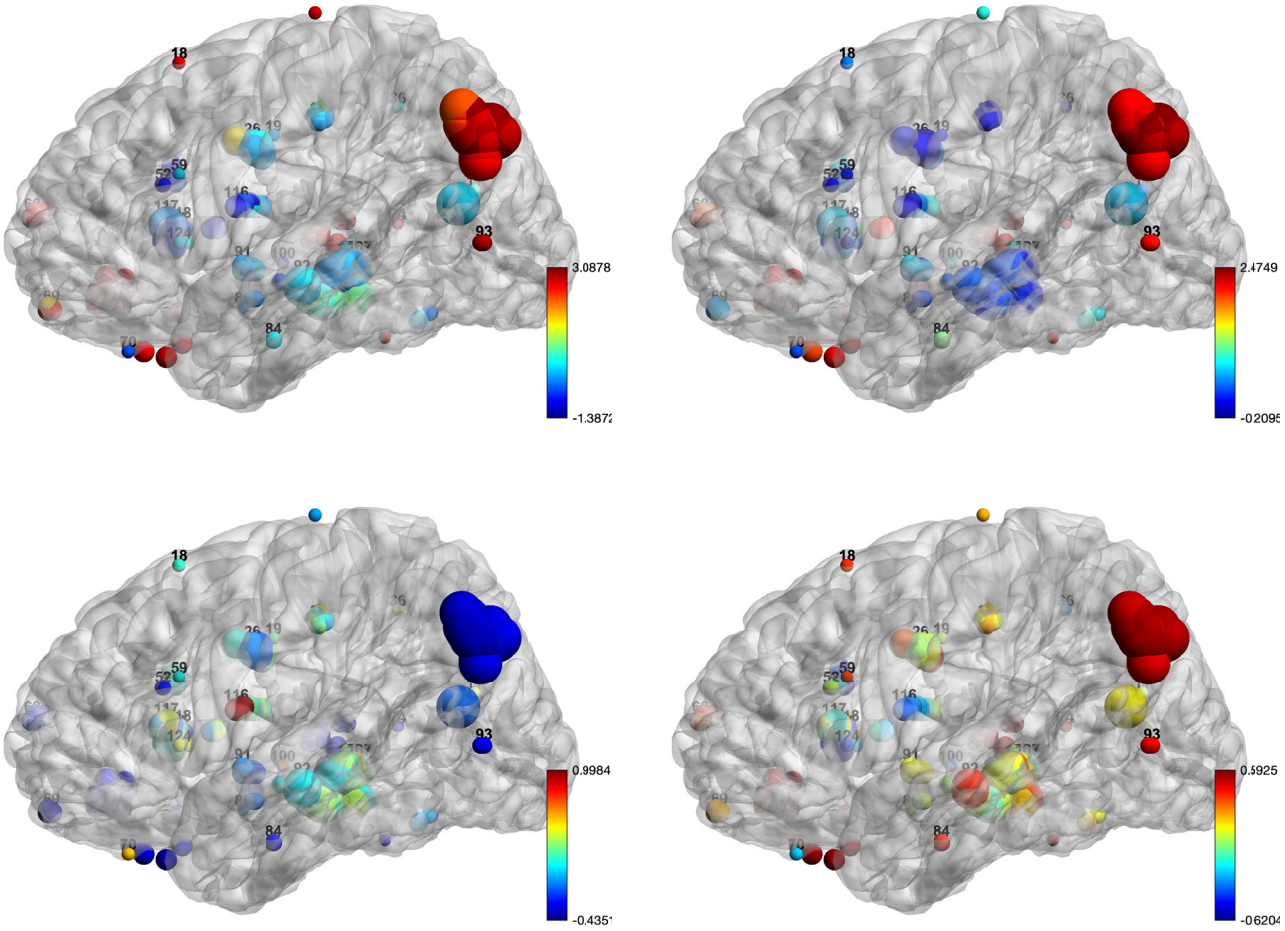

FIGURE 3 The brain heat map (in sagittal view) that expresses the decomposition of the MRI data for a randomly selected subject i as: $m_i(v) \approx \hat{\mu}_m(v) + \hat{\beta}_m \hat{u}_{mi}(v) + \hat{f}_{mi}(v)$. The four plots (from top left in clockwise order) represent the observed voxel volumes $m_i(v)$, the estimated mean function $\hat{\mu}_m(v)$, the MRI-specific varying pattern $\hat{f}_{mi}(v)$, and the MRI component shared with longitudinal outcomes $\hat{u}_{mi}(v)$. The colored dots are the voxels, with colors being volumes of voxels. The labels for regions of interest (ROI) are displayed as numbers (refer to Supplementary Tables S10 and S11). The voxel sizes are proportional to the number of voxels in each ROI

Figure 4 displays the prediction of neurological scores and probability of remaining stable MCI for a randomly selected subject A at the landmark times of 2 and 3 years, based on Model 2 with whole-brain voxels (M2-whole). Figure 4 suggests that subject A has a relatively mild disease status (low ADAS-Cog 13 scores before 2 years) and good functional activities (low CDR-SB scores before 2 years). The predicted survival probabilities decrease slowly, with a high chance of remaining MCI at the censoring time (7.98 years).

5 | SIMULATION STUDY

5.1 | Simulation settings

A simulation study is conducted to validate the proposed estimation and prediction method. The true values of parameters are close to the estimates from the real data analysis based on Model 2 (the instantaneous model) with whole-brain voxels. To check the robustness of the method, we explore three simulation scenarios detailed in Supplementary Section 7. For simulation Scenario 1 (we simulate data from Model 2), we consider four simulation settings with different sample sizes and event rates (ER) detailed in Supplementary Section 7. In the following context, we only display the results from

TABLE 3 Predictive performance for candidate models at landmark times 2, 2.5, 3, 3.5, and 4 years

	M1-whole		M1-hippo		M1		M2-whole		M2-hippo		M2	
Landmark (T)	**iAUC**	**iBS**	**iAUC**	**iBS**	**iAUC**	**iBS**	**iAUC**	**iBS**	**iAUC**	**iBS**	**iAUC**	**iBS**
T = 2	0.879	0.091	0.876	0.090	0.877	0.088	0.884	0.087	0.875	0.090	0.873	0.088
T = 2.5	0.886	0.087	0.866	0.083	0.872	0.083	0.899	0.081	0.867	0.081	0.871	0.079
T = 3	0.854	0.102	0.835	0.096	0.840	0.101	0.866	0.095	0.834	0.098	0.834	0.096
T = 3.5	0.783	0.095	0.765	0.089	0.751	0.096	0.802	0.089	0.769	0.090	0.767	0.094
T = 4	0.801	0.084	0.780	0.079	0.774	0.084	0.834	0.077	0.811	0.079	0.804	0.079
	M3 - 2S		**M3 - MJM-MRI**		**M3 - bCox**		**M3 - NM**					
Landmark (T)	**iAUC**	**iBS**	**iAUC**	**iBS**	**iAUC**	**iBS**	**iAUC**	**iBS**				
T = 2	0.850	0.097	0.808	0.116	0.864	0.087	0.696	0.103				
T = 2.5	0.843	0.080	0.838	0.100	0.810	0.086	0.675	0.092				
T = 3	0.816	0.093	0.829	0.120	0.798	0.096	0.681	0.107				
T = 3.5	0.777	0.088	0.817	0.110	0.757	0.087	0.564	0.102				
T = 4	0.798	0.083	0.821	0.100	0.760	0.083	0.577	0.097				

Abbreviations: iAUC, integrated area under receiver operating characteristic curve; iBS, integrated Brier score; M1, Model 1, random-effects model; M2, Model 2, instantaneous model; M3 - 2S, (two-stage sequential estimation approach); M3 - MJM-MRI, (multivariate joint model with MRI data; M3 - bCox, (the Cox model with baseline longitudinal outcomes); M3 - NM, (non-mixed functional model).

simulation Scenario 1 and Setting 1 ($N = 800$ for training set, $N = 300$ for testing set, $h_0(t) = \exp(-1.5)$). The results for Settings 2 to 4 are displayed in Supplementary Tables S3, S4, and S5. The results for Scenarios 2 and 3 are displayed in Supplementary Tables S8 and S9. We also compare the model with four candidate models as in Section 4 and detailed in Supplementary Section 7 to show the strengths of the joint modelling vs the two-stage approach, modeling longitudinal outcomes and MRI data using functional mixed models, modelling longitudinal trajectories in the survival model, and modelling the random component functions in the longitudinal model. For the four candidate models, we simulate data from Scenario 1 and Setting 1 and estimate the parameters using the candidate models. The results are summarized in Supplementary Tables S6 and S7.

For subject i ($i = 1, \dots, N$), we simulate $J = 3$ longitudinal outcomes from Model (1) on a sparse time grid $\boldsymbol{S}_i = [0, t_i, t_i + 0.1, \dots, t_i + 0.9]$ so the time grid has 11 points for each outcome, where t_i is sampled from a multinomial distribution with equal probabilities on [0.01, 0.02, …, 0.1]. The true mean functions $\mu_j(t)$ are equal to the estimated mean functions from real data analysis. The FPC scores are simulated from $\xi_{il} \sim N(0, d_{0l})$ and $\zeta_{ijl} \sim N(0, d_{1l})$, where $\boldsymbol{d_0} = (2, 1)$ and $d_{11} = 1$, with $L_0 = 2$ and $L_1 = 1$. The eigenfunctions $\phi_1(t) = \sqrt{2}\sin(\pi t)$, $\phi_2(t) = \sqrt{2}\cos(\pi t)$, $\psi_1(t) = \sqrt{2}\cos(2\pi t)$. We set the association parameter $\boldsymbol{\beta} = (\beta_2, \beta_3) = (-0.55, -0.60)$ and the standard deviation for random errors $\boldsymbol{\sigma} = (1.10, 0.60, 1.36)$. We simulate missing outcomes with the missingness mechanism being missing at random (MAR), and the missing probability $\text{logit}(p_{ijk}) = 0.5(t_{ik} - 5)$, and $I_{ijk} \sim \text{Ber}(p_{ijk})$, thus achieving around 2.5% missing rate for all three longitudinal outcomes because we assume t_{ik} is the only predictor for the longitudinal outcomes, where Ber is the Bernoulli distribution.

We simulate the baseline MRI data $m_i(v)$ from Model (2) observed on a dense voxel grid $\boldsymbol{V} = (1/V, 2/V, \dots, 1)$, where $V = 642$ voxels, and the scale parameter $\beta_m = -0.05$. The FPC scores for the MRI-specific varying pattern ξ_{mil} are simulated from $\xi_{mil} \sim N(0, d_{ml})$, for $l = 1, \dots, L_m = 5$, and the variance $\boldsymbol{d}_m = (0.10, 0.07, 0.05, 0.03, 0.01)$. The eigenfunctions $\phi_{ml}(v) = \sqrt{2}\sin((2l-1)\pi v/2)$, for $l = 1, \dots, L_0 = 2$, and $\psi_{ml}(v) = \sqrt{2}\cos(\pi(2l-1)v/2)$, for $l = 1, \dots, L_m = 5$. We simulate the random error from $\epsilon_{mi}(v) \sim N(0, \sigma_m^2)$, where the standard deviation $\sigma_m = 0.4$.

To simulate the survival outcome from Model (3), we set the baseline hazard function $h_0(t) = \exp(-1.5)$, $Z_i \sim \text{Binom}(2, 0.4)$, the regression coefficient $\alpha = 0.3$, the association parameters $\gamma_0 = 0.6$, $\boldsymbol{\gamma}_1 = (0.35, 0.3, 0.1)$, and $\boldsymbol{\gamma}_m = (-0.6, -1.1, -0.5, 0.7, -1.8)$, where Binom is the Binomial distribution. The survival probability S_i is simulated from $U(0, 1)$ and we approximate the event time T_i^* using the bisection method by solving the equation $\log(S_i) = -H_i(T_i^*)$. The censoring time is generated from $C_i \sim U(0, 1.6)$. After censoring, we have around 6.5 observations per subject.

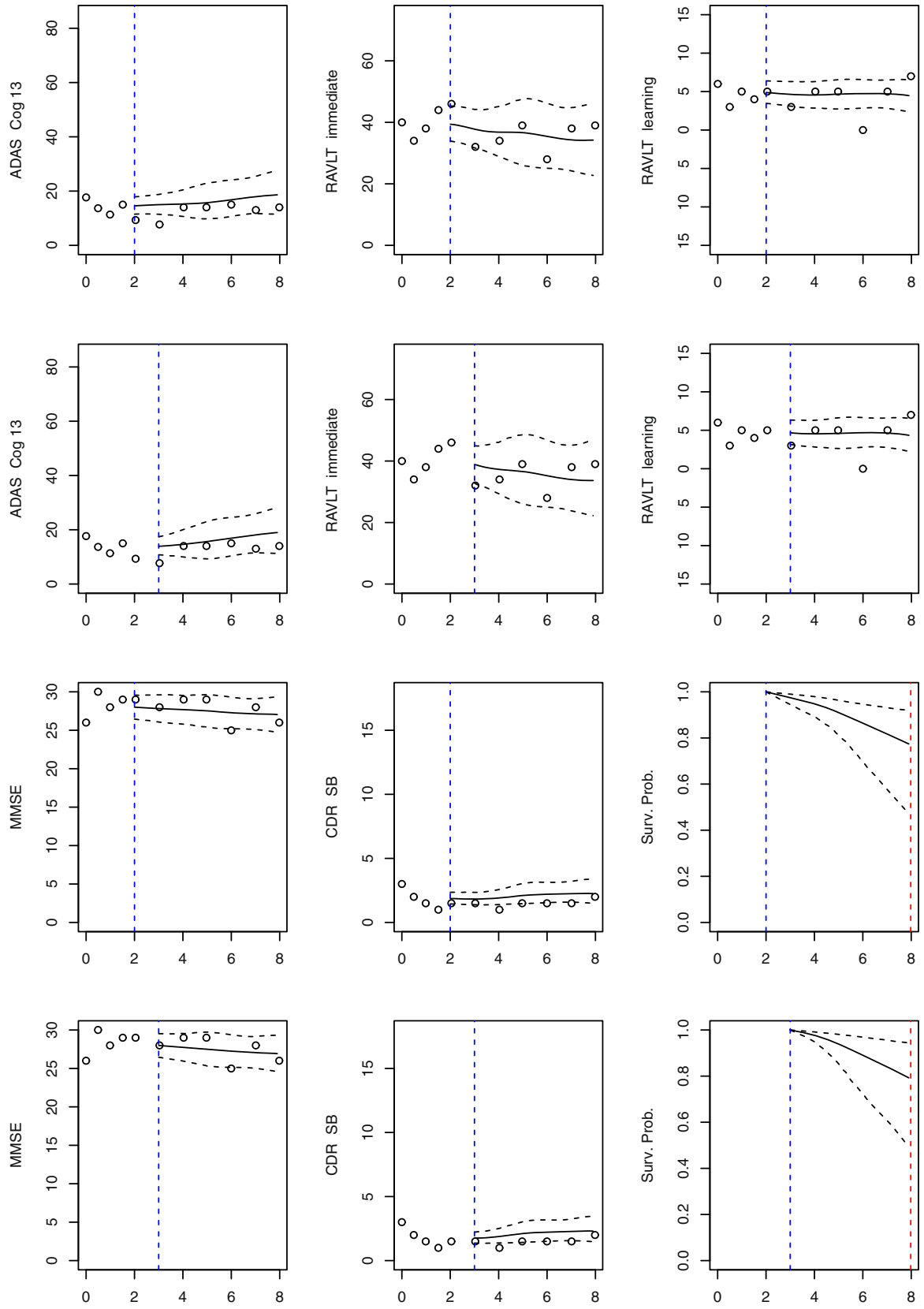

FIGURE 4 Prediction of ADAS-Cog 13, RAVLT-immediate, RAVLT-learning, MMSE, CDR-SB scores, and the probability of remaining stable MCI for a new subject A given the landmark times at 2 years (rows 1 and 3) and 3 years (rows 2 and 4). The dotted circles are observed neurological scores. The blue and red vertical lines represent the landmark time (2 or 3 years) and the censoring time (7.98 years), respectively. The solid lines represent the predicted longitudinal trajectories or survival probabilities, with dashed lines being 95% credible intervals

We repeat the simulation study for $R = 100$ replications. From each replication r, we obtain the posterior mean, standard errors, and 95% credible intervals. From a total of R replications, we compute the bias (difference between the averaged posterior mean and true values of parameters), standard deviation (SD of R posterior means), standard error (SE, square root of averaged variances), and coverage probability (CP, frequency of 95% CI covering true values of parameters). The standard error and coverage probability for parameters d_m and β_m are not presented because those parameters are not in the parameter space as in Section 3.3. To evaluate the estimation performance for mean functions and eigenfunctions, we compute the mean-squared error (MSE), defined as $\text{MSE} = 1/G \sum_{\boldsymbol{G}} (\hat{f}(g) - f(g))^2$, where G is the number of grid points for the refined time grid $\boldsymbol{S} \in [0, 0.01, \ldots, 1]$ or the dense voxel grid $\boldsymbol{V}$, $f(g)$ is the true function, and the mean of estimated functions $\hat{f}(g) = 1/R \sum_{r=1}^{R} \hat{f}_r(g)$. We apply the dynamic prediction method to the testing dataset for R replications. From each replication r, we obtain the true and estimated iAUC and iBS for landmark times at 0.4, 0.5, 0.55, and 0.6. The prediction window is set as $\delta_t = (0.1, 0.11, \ldots, 0.25)$. From R replications, we average and compare the true and estimated iAUC and iBS values.

5.2 | Simulation results

Table 4 summarizes the bias, SD, SE, and CP for parameters from 100 replications. Results suggest that for most parameters, biases are close to zero, with SD being close to SE, and CP being close to the nominal level 0.95 (except for the parameter d_{11}). The results for simulation Setting 2 (the sample size is 1200) displayed in Table S3 suggest that we have much smaller bias, SD, and SE in the parameters γ_{m3}, γ_{m4}, and γ_{m5} as compared with simulation Setting 1 (the sample size is 800). Table 4 suggests that the SD and SE for $\boldsymbol{\gamma}_m$ are close, which indicates the accuracy of the estimation algorithm. Table 4 also suggests that the mean functions and eigenfunctions are estimated close to the true functions with mean-squared error being close to zero. Table 5 summarizes the mean of true and estimated iAUC and iBS from 100 replications. Results suggest that the estimated iAUC and iBS values are close to the true iAUC and iBS values, respectively, for all landmark times. Tables 4 and 5 support the validity of the estimation and dynamic prediction method.

The results for simulation Settings 1 to 4 presented in Supplementary Tables S3 to S5 suggest that in all settings, the parameters are estimated with small bias and the coverage probabilities are close to the nominal value of 0.95, and the estimated iAUC and iBS are close to their true values. For Model 3 - 2S (two-stage approach based on Model 2, the instantaneous model), the results presented in Supplementary Table S6 suggests that the estimated parameters have large bias and low coverage probabilities (i.e., parameters β_2, β_3, σ_1, σ_2, σ_3, and $\boldsymbol{\gamma}_m$). The prediction results for the other four candidate models are presented in Supplementary Table S7. The results suggest that the estimated iAUC and iBS values are close to their true values for Model 3 - 2S (two-stage approach) and Model 3 - MJM-MRI (multivariate joint model with MRI data), while Model 3 - bCox (Cox model using baseline longitudinal outcomes) and Model 3 - NM (non-mixed functional model) have poor prediction performance. Comparing with Model 3 - MJM-MRI, our methods use non-parametric functional mixed models and are more robust to model sparse and nonlinear longitudinal outcomes.

The results for simulation Scenarios 2 and 3 are presented in Supplementary Tables S8 and S9. The results suggest that for Scenario 2 (no correlation between longitudinal and MRI data), the estimated parameters, functions, and iAUC and iBS values are close to their true values, with the coverage probabilities close to the nominal level 0.95. For Scenario 3 (misspecification of the functional form $F(X_i, t)$), we observe large bias in parameters γ_{ml}, and the iAUC and iBS values are not close to their true values. The results suggest that the method is not sensitive to the correlation between the MRI and longitudinal outcomes, and is sensitive to the misspecification of the functional form $F(X_i, t)$.

6 | DISCUSSION

In this article, we propose the multivariate functional mixed model with MRI data (MFMM-MRI) that simultaneously models multivariate longitudinal outcomes (ie, neurological scores), baseline magnetic resonance imaging (MRI) data, and the survival process (ie, dementia onset) for subjects with mild cognitive impairment (MCI) at baseline. We use the non-parametric functional mixed model to decompose the longitudinal and MRI data into the joint and individual variation. We adopt an iterative search procedure to estimate the eigenfunctions and scale parameter for the MRI data. The Markov Chain Monte Carlo (MCMC) approach based on the No-U-Turn Sampling (NUTS) algorithm is adopted to obtain posterior samples from the likelihood. We also develop a dynamic prediction method that provides accurate personalized predictions of longitudinal trajectories and risk of dementia onset, which facilitates clinical decision making

TABLE 4 True values of parameters, bias, standard deviation (SD), standard error (SE), coverage probability (CP), and mean-squared error (MSE) from 100 simulation replications for simulation Scenario 1, Setting 1 (based on Model 2)

Parameters	Bias	SD	SE	CP	MSE	
Longitudinal outcomes						
d_{01}=2	− 0.055	0.103	0.132	0.980	$\mu_1(t)$	0.000
d_{02}=1	− 0.029	0.065	0.065	0.940	$\mu_2(t)$	0.000
d_{11}=1	− 0.026	0.065	0.045	0.810	$\mu_3(t)$	0.000
$\beta_2 = -0.55$	− 0.001	0.008	0.008	0.960	$\phi_1(t)$	0.002
$\beta_3 = -0.6$	− 0.001	0.014	0.013	0.930	$\phi_2(t)$	0.001
σ_1=1.1	0.001	0.010	0.013	0.990	$\psi_1(t)$	0.001
σ_2=0.6	0.002	0.006	0.007	1.000		
σ_3=1.36	0.002	0.012	0.015	0.970		
MRI outcomes						
d_{m1}=0.1	0.001	0.003			$\mu_m(t)$	0.000
d_{m2}=0.07	0.002	0.002			$\phi_{m1}(t)$	0.064
d_{m3}=0.05	0.001	0.002			$\phi_{m2}(t)$	0.090
d_{m4}=0.03	0.000	0.001			$\psi_{m1}(t)$	0.003
d_{m5}=0.01	0.000	0.000			$\psi_{m2}(t)$	0.003
$\beta_m = -0.05$	0.002	0.009			$\psi_{m3}(t)$	0.002
σ_m=0.4	0.000	0.000			$\psi_{m4}(t)$	0.000
Survival outcome						
α=0.3	0.009	0.105	0.102	0.950	$\psi_{m5}(t)$	0.000
γ_0=0.6	− 0.007	0.052	0.055	0.960		
γ_{11}=0.35	0.005	0.083	0.090	0.980		
γ_{12}=0.3	− 0.013	0.088	0.091	0.970		
$\gamma_{13} = 0.1$	0.000	0.095	0.109	0.980		
$\gamma_{m1} = -0.6$	0.076	0.232	0.225	0.940		
$\gamma_{m2} = -1.1$	0.019	0.291	0.270	0.960		
$\gamma_{m3} = -0.5$	− 0.108	0.395	0.316	0.870		
γ_{m4}=0.7	0.109	0.463	0.407	0.900		
$\gamma_{m5} = -1.8$	0.087	0.670	0.701	0.960		

Note: the standard error and coverage probability for parameters d_m and β_m are not presented because those parameters are not in the parameter space as in Section 3.3.

TABLE 5 True and estimated iAUC and iBS from 100 simulation replications for simulation Scenario 1, Setting 1 (based on Model 2)

Landmark time	True iAUC	Estimated iAUC	True iBS	Estimated iBS
T = 0.4	0.818	0.799	0.063	0.065
T = 0.5	0.805	0.786	0.060	0.063
T = 0.55	0.797	0.774	0.058	0.059
T = 0.6	0.785	0.755	0.056	0.058

Abbreviations: iAUC, integrated area under receiver operating characteristic curve; iBS, integrated Brier score.

and treatment selection. We apply the proposed MFMM-MRI to the motivating Alzheimer's Disease Neuroimaging Initiative (ADNI) study. We found that higher ADAS-Cog 13 scores (worse cognitive functions) are associated with lower RAVLT-immediate, RAVLT-learning, MMSE scores, and higher CDR-SB scores (worse memory functions and functional activities). We discover that higher functional principal component (FPC) scores (worse cognitive functions) are associated with a greater decrease in voxel volumes in MRI data. For the survival outcome, we conclude that additional ApoE-ϵ4 alleles, higher ADAS-Cog 13 and CDR-SB scores, and lower RAVLT-immediate scores are associated with a higher dementia risk. Moreover, the MRI-specific varying pattern is significantly associated with the risk of dementia onset based on the likelihood ratio test. The dynamic prediction results show that the proposed instantaneous model with whole-brain voxels has the best prediction performance in terms of calibration and discrimination measurements among all candidate models. The simulation study supports the validity and accuracy of the estimation and dynamic prediction method.

The multivariate functional mixed model with MRI data (MFMM-MRI) leverages multimodal data, which are predictive of dementia occurrence and AD progression. Comparing with existing methods (ie, parametric linear models[11] and the model without MRI data[10]), our approach provides more accurate prediction of AD dementia and facilitates treatment selection for health care professionals and patients at early disease stages (ie, mild cognitive impairment). Additionally, the MFMM-MRI model provides the association of the neurological scores, MRI data, and the survival outcome so that clinicians may determine how each diagnostic test may change this individual's prognosis.

There are several limitations to address as future directions. First, the proposed MFMM-MRI models continuous longitudinal outcomes with normality assumption. The longitudinal submodel of MFMM-MRI may be extended to a generalized functional mixed model to model outcomes with distributions in the exponential family, for example, ordinal, binary, or count.[50] Second, as stated in Zou et al,[11] our voxel selection approach is an ad hoc procedure to reduce the memory burden in the functional principal component analysis (FPCA). The high-dimensional FPCA using the singular value decomposition (SVD) for the partitioned MRI data may be adopted to reduce the computational and memory burden.[51] Alternatively, the MRI data can be treated as a spatial region embedded in three-dimensional space and multivariate FPCA may be applied.[52] Third, the MRI submodel of MFMM-MRI handles baseline MRI data only, while longitudinal sMRI data are available for some subjects. Thus, the functional mixed model may be extended to model two-dimensional MRI data observed on the sparse time and dense voxel grid. Additionally, for the survival submodel of MFMM-MRI, the accelerated failure time (AFT) model is another alternative and the Bayesian assessment criteria may be adopted to select the model with the best fit.[53] Fourth, in the longitudinal and MRI models (1) and (2), we incorporate the correlation between the MRI data and longitudinal outcomes via shared FPC scores, but the baseline MRI data may also change longitudinal trajectories. Thus, the functional form $F(X_i, t)$ in the survival model (3) may be extended to account for the interaction between the longitudinal outcomes and MRI data. Specifically, the interaction between the FPC scores of the longitudinal outcomes and MRI data may be included in both the random-effects model and instantaneous model.

As stated in Section 1, recent AD studies collect multi-omics data, for example, the ADNI study collects genome, transcriptome, and proteome data, and so forth, which provide unique insights to determine disease mechanism together with neurological assessments and imaging data.[54] One recent research article investigated the convolutional neural network (CNN) that integrated multi-omics data into kernel matrices and identified significant associations between single nucleotide variants (SNVs) and quantitative AD traits, including biospecimen samples and MRI data.[55] Mo et al proposed a Bayesian latent variable model that clustered multi-omics data and used the Bayesian variable selection approach to reduce the high-dimensionality and determine significant latent variables.[56] As a future direction, the extension of the above methods to modelling multimodal data, including multi-omics data, longitudinal neurological data, and MRI data, and efficient estimation algorithms are worth of further investigation.

ACKNOWLEDGEMENTS

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

FUNDING INFORMATION

The research of Sheng Luo was supported by National Institute on Aging (grant numbers: R01AG064803, P30AG072958, and P30AG028716).

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflict of interest.

DATA AVAILABILITY STATEMENT

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

ORCID

Haotian Zou https://orcid.org/0000-0002-3595-8716
Luo Xiao https://orcid.org/0000-0001-8707-0914
Sheng Luo https://orcid.org/0000-0003-4214-5809

REFERENCES

1. Knopman DS, Amieva H, Petersen RC, et al. Alzheimer disease. *Nat Rev Dis Primers*. 2021;7(1):1-21.
2. Rajan KB, Weuve J, Barnes LL, McAninch EA, Wilson RS, Evans DA. Population estimate of people with clinical Alzheimer's disease and mild cognitive impairment in the United States (2020–2060). *Alzheimers Dement*. 2021;17(12):1966-1975.
3. Weiner MW, Veitch DP, Aisen PS, et al. 2014 Update of the Alzheimer's disease neuroimaging initiative: a review of papers published since its inception. *Alzheimers Dement*. 2015;11(6):e1-e120.
4. Besser L, Kukull W, Knopman DS, et al. Version 3 of the national Alzheimer's coordinating center's uniform data set. *Alzheimer Dis Assoc Disord*. 2018;32(4):351-358.
5. Petersen RC. Mild cognitive impairment. *Continuum (Minneap Minn)*. 2016;22(2):404.
6. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the national institute on aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):270-279.
7. Hickey GL, Philipson P, Jorgensen A, Kolamunnage-Dona R. Joint models of longitudinal and time-to-event data with more than one event time outcome: a review. *Int J Biostat*. 2018;14(1):1-19.
8. Papageorgiou G, Mauff K, Tomer A, Rizopoulos D. An overview of joint modeling of time-to-event and longitudinal outcomes. *Annu Rev Stat Appl*. 2019;6:223-240.
9. Li K, Luo S. Dynamic predictions in Bayesian functional joint models for longitudinal and time-to-event data: an application to Alzheimer's disease. *Stat Methods Med Res*. 2019;28(2):327-342.
10. Li C, Xiao L, Luo S. Joint model for survival and multivariate sparse functional data with application to a study of Alzheimer's disease. *Biometrics*. 2021;78(2):435-447.
11. Zou H, Li K, Zeng D, Luo S, Initiative ADN. Bayesian inference and dynamic prediction of multivariate joint model with functional data: an application to Alzheimer's disease. *Stat Med*. 2021;40(30):6855-6872.
12. Brown ER. Assessing the association between trends in a biomarker and risk of event with an application in pediatric HIV/AIDS. *Ann Appl Stat*. 2009;3(3):1163.
13. Rizopoulos D, Hatfield LA, Carlin BP, Takkenberg JJ. Combining dynamic predictions from joint models for longitudinal and time-to-event data using Bayesian model averaging. *J Am Stat Assoc*. 2014;109(508):1385-1397.
14. Rizopoulos D, Ghosh P. A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. *Stat Med*. 2011;30(12):1366-1380.
15. Long JD, Mills JA. Joint modeling of multivariate longitudinal data and survival data in several observational studies of Huntington's disease. *BMC Med Res Methodol*. 2018;18(1):1-15.
16. Mauff K, Steyerberg E, Kardys I, Boersma E, Rizopoulos D. Joint models with multiple longitudinal outcomes and a time-to-event outcome: a corrected two-stage approach. *Stat Comput*. 2020;30(4):999-1014.
17. Andrinopoulou ER, Rizopoulos D. Bayesian shrinkage approach for a joint model of longitudinal and survival outcomes assuming different association structures. *Stat Med*. 2016;35(26):4813-4823.
18. Rizopoulos D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics*. 2011;67(3):819-829.
19. Sène M, Taylor JM, Dignam JJ, Jacqmin-Gadda H, Proust-Lima C. Individualized dynamic prediction of prostate cancer recurrence with and without the initiation of a second treatment: Development and validation. *Stat Methods Med Res*. 2016;25(6):2972-2991.
20. Vlaardingerbroek MT, Boer JA. *Magnetic Resonance Imaging: Theory and Practice*. Berlin: Springer Science & Business Media; 2013.
21. Polzehl J, Tabelow K. *Magnetic Resonance Brain Imaging*. Cham: Springer; 2019.
22. Ashburner J, Friston KJ. Voxel-based morphometry—the methods. *Neuroimage*. 2000;11(6):805-821.

23. Whitwell JL. Voxel-based morphometry: an automated technique for assessing structural changes in the brain. *J Neurosci*. 2009;29(31):9661-9664.
24. Gupta Y, Lee KH, Choi KY, et al. Early diagnosis of Alzheimer's disease using combined features from voxel-based morphometry and cortical, subcortical, and hippocampus regions of MRI T1 brain images. *PLoS One*. 2019;14(10):e0222446.
25. Ji X, Wang H, Zhu M, et al. Brainstem atrophy in the early stage of Alzheimer's disease: a voxel-based morphometry study. *Brain Imaging Behav*. 2021;15(1):49-59.
26. Wang JL, Chiou JM, Müller HG. Functional data analysis. *Annu Rev Stat Appl*. 2016;3:257-295.
27. Morris JS. Functional regression. *Annu Rev Stat Appl*. 2015;2:321-359.
28. Guo W. Functional mixed effects models. *Biometrics*. 2002;58(1):121-128.
29. Yao F. Functional principal component analysis for longitudinal and survival data. *Stat Sin*. 2007;17(3):965-983.
30. Lock EF, Hoadley KA, Marron JS, Nobel AB. Joint and individual variation explained (JIVE) for integrated analysis of multiple data types. *Ann Appl Stat*. 2013;7(1):523.
31. Li K, Chan W, Doody RS, Quinn J, Luo S. Prediction of conversion to Alzheimer's disease with longitudinal measures and time-to-event data. *J Alzheimers Dis*. 2017;58(2):361-371.
32. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry*. 1984;141(11):1356-1364.
33. Rey A. *L'examen Clinique en Psychologie*. France: Presses Universitaries; 1958.
34. Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. *J Am Geriatr Soc*. 1992;40(9):922-935.
35. Hughes CP, Berg L, Danziger W, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *Br J Psychiatry*. 1982;140(6):566-572.
36. Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nat Rev Neurol*. 2013;9(2):106-118.
37. Huang M, Yang W, Feng Q, Chen W. Longitudinal measurement and hierarchical classification framework for the prediction of Alzheimer's disease. *Sci Rep*. 2017;7(1):1-13.
38. Petrone PM, Casamitjana A, Falcon C, et al. Prediction of amyloid pathology in cognitively unimpaired individuals using voxel-wise analysis of longitudinal structural brain MRI. *Alzheimer's Res Therapy*. 2019;11(1):1-13.
39. Di CZ, Crainiceanu CM, Caffo BS, Punjabi NM. Multilevel functional principal component analysis. *Ann Appl Stat*. 2009;3(1):458.
40. Lawless J, Zhan M. Analysis of interval-grouped recurrent-event data using piecewise constant rate functions. *Can J Stat*. 1998;26(4):549-565.
41. Feng S, Wolfe RA, Port FK. Frailty survival model analysis of the national deceased donor kidney transplant dataset using Poisson variance structures. *J Am Stat Assoc*. 2005;100(471):728-735.
42. Hoffman MD, Gelman A. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J Mach Learn Res*. 2014;15(1):1593-1623.
43. Li C, Xiao L, Luo S. Fast covariance estimation for multivariate sparse functional data. *Stat*. 2020;9(1):e245.
44. Wang J, Luo S. Bayesian multivariate augmented Beta rectangular regression models for patient-reported outcomes and survival data. *Stat Methods Med Res*. 2017;26(4):1684-1699.
45. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Comput*. 2017;27(5):1413-1432.
46. Li L, Greene T, Hu B. A simple method to estimate the time-dependent receiver operating characteristic curve and the area under the curve with right censored data. *Stat Methods Med Res*. 2018;27(8):2264-2278.
47. Graf E, Schmoor C, Sauerbrei W, Schumacher M. Assessment and comparison of prognostic classification schemes for survival data. *Stat Med*. 1999;18(17-18):2529-2545.
48. Xia M, Wang J, He Y. BrainNet Viewer: a network visualization tool for human brain connectomics. *PloS one*. 2013;8(7):e68910.
49. Kong D, Staicu AM, Maity A. Classical testing in functional linear models. *J Nonparametr Stat*. 2016;28(4):813-838.
50. Scheipl F, Gertheiss J, Greven S. Generalized functional additive mixed models. *Electron J Stat*. 2016;10(1):1455-1492.
51. Zipunnikov V, Caffo B, Yousem DM, Davatzikos C, Schwartz BS, Crainiceanu C. Functional principal component model for high-dimensional brain imaging. *Neuroimage*. 2011;58(3):772-784.
52. Katina S, Vittert L, Bowman A W. Functional data analysis and visualisation of three-dimensional surface shape. *J R Stat Soc Ser C Appl Stat*. 2021;70(3):691-713.
53. Tseng YK, Hsieh F, Wang JL. Joint modelling of accelerated failure time and longitudinal data. *Biometrika*. 2005;92(3):587-603.
54. Saykin AJ, Shen L, Yao X, et al. Genetic studies of quantitative MCI and AD phenotypes in ADNI: progress, opportunities, and plans. *Alzheimers Dement*. 2015;11(7):792-814.
55. Konigorski S, Khorasani S, Lippert C. Integrating omics and MRI data with kernel-based tests and CNNs to identify rare genetic markers for Alzheimer's disease. *arXiv preprint arXiv:1812.00448*; 2018.
56. Mo Q, Shen R, Guo C, Vannucci M, Chan KS, Hilsenbeck SG. A fully Bayesian latent variable model for integrative clustering analysis of multi-type omics data. *Biostat*. 2018;19(1):71-86.

SUPPORTING INFORMATION

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

How to cite this article: Zou H, Xiao L, Zeng D, Luo S, for the Alzheimer's Disease Neuroimaging Initiative. Multivariate functional mixed model with MRI data: An application to Alzheimer's disease. *Statistics in Medicine*. 2023;1-20. doi: 10.1002/sim.9683