



Published in final edited form as:

Ann Appl Stat. 2025 March ; 19(1): 505–528. doi:10.1214/24-aos1970.

DYNAMIC PREDICTION WITH MULTIVARIATE LONGITUDINAL OUTCOMES AND LONGITUDINAL MAGNETIC RESONANCE IMAGING DATA

Haotian Zou^{1,a}, Luo Xiao^{2,c}, Donglin Zeng^{3,d}, Sheng Luo^{1,b} For the Alzheimer's Disease Neuroimaging Initiative *

¹Department of Biostatistics and Bioinformatics, Duke University

²Department of Statistics, North Carolina State University

³Department of Biostatistics, University of Michigan

Abstract

Alzheimer's Disease (AD) is a common neurodegenerative disorder impairing multiple domains. Recent AD studies, for example, the Alzheimer's Disease Neuroimaging Initiative (ADNI) study, collect multimodal data to better understand AD severity and progression. To facilitate precision medicine for high-risk individuals, it is essential to develop an AD predictive model that leverages multimodal data and provides accurate personalized predictions of dementia occurrences. In this article we propose a multivariate functional mixed model with longitudinal magnetic resonance imaging data (MFMM-LMRI) that jointly models longitudinal neurological scores, longitudinal voxelwise MRI data, and the survival outcome as dementia onset. We model longitudinal MRI data using the joint and individual variation explained (JIVE) approach. We investigate two functional forms linking the longitudinal and survival processes. We adopt the Markov chain Monte Carlo (MCMC) method to obtain posterior samples. We establish a dynamic prediction framework that predicts longitudinal trajectories and the probability of dementia occurrence. The simulation study with various sample sizes and event rates supports the validity of the method. We apply the MFMM-LMRI to the motivating ADNI study and conclude that additional ApoE- $\epsilon 4$ alleles and a higher latent disease profile are associated with a higher risk of dementia onset. We detect a significant association between the longitudinal MRI data and the survival outcome. The instantaneous model with longitudinal MRI data has the best fitting and predictive performance.

Keywords

Alzheimer's Disease; functional data; longitudinal magnetic resonance imaging data; joint model

^a haotian.zou@duke.edu . ^b sheng.luo@duke.edu . ^c lxiao5@ncsu.edu . ^d dzeng@email.unc.edu .

*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.

1. Introduction.

Alzheimer's Disease (AD) is the most common neurodegenerative disorders impairing multiple domains (e.g., cognition, memory, and behavior) and progresses heterogeneously in time and across domains (Knopman et al. (2021), Gross et al. (2016)). Recent AD studies collect data from multiple sources to better understand AD severity and progression (Milà-Alomà et al. (2020)). For example, the Alzheimer's Disease Neuroimaging Initiative (ADNI) study (Veitch et al. (2022)) collects cross-sectional demographic and genomics data as well as longitudinal clinical data (e.g., neuropsychological tests), biomarkers from biospecimens (e.g., plasma $A\beta_{42}/A\beta_{40}$ ratio and pTau-181), and neuroimaging data (e.g., magnetic resonance imaging, MRI), together generating multimodal data. These multimodal data are all predictive of AD progression and the dementia occurrence. For example, Kong et al. (2015) first treated the time until dementia onset as the survival outcome and used the cognition scores, single-nucleotide polymorphism (SNP) data, and the hippocampal radial distance data to predict the risk of dementia onset. The prediction models were also explored in many recent works (Li et al. (2018, 2019), Jiang, Xie and Colditz (2021)). However, the prediction values of the multimodal data are mostly shown on a population level instead of an individual level (de Wilde et al. (2017), van Maurik et al. (2017, 2019)). Thus, there exists a critical need to develop an AD predictive model that provides accurate personalized predictions of the dementia occurrence so that clinicians may adopt personalized treatments at patient's early disease stages (e.g., mild cognitive impairment, MCI) to slow AD progression (Petersen (2016), Albert et al. (2011)). In this article we aim to develop a predictive model that leverages multimodal data, including multivariate longitudinal clinical outcomes, longitudinal voxelwise MRI data, and the survival outcome as the time until dementia onset.

The longitudinal outcomes collected from AD studies are sparse, nonlinear, and correlated within subjects (Li, Xiao and Luo (2022)). Thus, a flexible nonparametric data-adaptive functional regression approach can be used to account for the heterogeneity across the longitudinal outcomes (Wang, Chiou and Müller (2016), Morris (2015)). To model sparse longitudinal functional data, Yao, Müller and Wang (2005) adopted the functional principal component analysis (FPCA) and developed the conditional expectation approach to estimate the principal component scores. The univariate functional regression framework is further extended to modeling multivariate correlated functional data (i.e., neuropsychological tests). For example, Chiou, Chen and Yang (2014) developed the multivariate functional principal component analysis (mFPCA) method to decompose multivariate functional data into the mean function and linear combinations of orthonormal eigenfunctions. Li, Xiao and Luo (2020) developed the multivariate fast covariance estimation (mFACE) approach to estimate the auto- and cross-covariance matrices for multivariate functional data. The multivariate functional regression framework has been extensively studied (Tang, Tang and Zhu (2017), Hasenstab et al. (2017), Song, Xia and Zhu (2017), Zhu et al. (2017)). Bayesian methods have been widely used in the estimation of functional components in the functional regression model. For example, Morris and Carroll (2006) adopted the Markov chain Monte Carlo (MCMC) algorithm to obtain posterior samples from the wavelet functional mixed

model. Goldsmith and Kitago (2016) adopted the variational Bayes method for a bivariate functional mixed effects model.

In the presence of the survival outcome (i.e., the dementia event in AD), the joint model (JM) framework is a widely used method to model the longitudinal outcomes (e.g., neurological tests) and survival data (Hickey et al. (2018), Papageorgiou et al. (2019)). Yao (2007) proposed a univariate functional joint model, adopted B-spline basis functions to model longitudinal functional data, and linked the correlation with the survival outcome via the Cox model. Recent works extend the univariate joint model to multivariate settings (Proust-Lima, Dartigues and Jacqmin-Gadda (2016), Long and Mills (2018), Mauff et al. (2020)). For example, Rizopoulos and Ghosh (2011) developed a semiparametric multivariate joint model, proposed three functional forms of the latent longitudinal mean, and adopted the Bayesian approach for inference. Additionally, Li, Xiao and Luo (2022) modeled the sparse and nonlinear longitudinal outcomes as functional data, developed a nonparametric multivariate functional mixed model (MFMM), and adopted an expectation-maximization (EM) algorithm for inference. Zou et al. (2023) extended the random effects model in the MFMM to six functional forms to flexibly model the association between multivariate longitudinal outcomes and survival data. One important application of the joint model is in dynamic prediction (Goerdten et al. (2019), Wang, Luo and Li (2017), Li and Luo (2019)). In contrast with static prediction, dynamic prediction updates an individual's survival probability (i.e., the probability of remaining stable MCI) whenever the individual's latest data are obtained, which facilitates clinical decision-making and treatment counseling.

Magnetic resonance imaging (MRI) is a widely used clinical tool to detect brain atrophy in AD patients (Vlaardingerbroek and Boer (2013), Polzehl and Tabelow (2019)). The structural MRI (sMRI) provides a detailed anatomical structure of the whole brain as well as regions of interest (ROI) and has been frequently used in large-scale AD studies, including the ADNI study (Veitch et al. (2022)). Voxel-based morphometry (VBM) is a standard tool to analyze the differences in voxel volumes for multiple groups of individuals (Ashburner and Friston (2000)). For instance, the medial temporal lobe (MTL, the hippocampus region included) atrophy has been detected in subjects with dementia diagnosis when compared to healthy controls (Busatto et al. (2003), Pennanen et al. (2005)). Additionally, recent AD studies collect longitudinal MRI data (e.g., the ADNI study), which provide clinicians with rich information on the subject-specific longitudinal change of the voxel volumes. The voxelwise longitudinal MRI data can be treated as longitudinal functional data, and the functional regression approach can be used for modeling (Li and Luo (2019), Zou et al. (2021)).

To model longitudinal functional data, Di et al. (2009) proposed a two-way functional ANOVA model and decomposed the univariate longitudinal functional data into the mean function, visit-specific shift, subject-specific deviation, and the subject- and visit-specific random deviation functions. Greven et al. (2011) proposed a two-dimensional semiparametric functional regression model, decomposed the longitudinal imaging data into the fixed main effect, random functional intercept and slope, and subject- and visit-specific random deviation, and adopted the FPCA to decompose the random components. Additionally, Goldsmith et al. (2012) developed a parametric linear mixed model, treated

the longitudinal diffusion tensor imaging (DTI) data as a functional predictor, and used the FPCA to decompose the functional data. Moreover, Chen and Müller (2012) proposed a nonparametric function-on-scalar regression model, based on a two-step Karhunen–Loève expansion, and modeled the random effect functions using the time and voxel-dependent eigenfunctions and time-dependent principal component functions. Park and Staicu (2015) proposed a parsimonious nonparametric functional mixed model and decomposed the latent longitudinal functional data, using time-invariant orthogonal basis functions and time-dependent basis coefficients, and adopted a flexible error structure. To provide an accurate prediction of the risk of dementia, the longitudinal functional imaging data are frequently modeled as functional predictors in the survival model by Shi et al. (2021), Jiang et al. (2023), Kong et al. (2018). Recently, Kang and Song (2023) proposed the joint modeling of the functional imaging data and survival data, developed a semiparametric functional regression model, and included the random functional components as functional predictors in the survival model.

In this manuscript we extend the multivariate functional mixed model (MFMM) proposed by Li, Xiao and Luo (2022) to the MFMM-longitudinal MRI (MFMM-LMRI) that jointly models multivariate sparse longitudinal outcomes (i.e., neurological scores), longitudinal neuroimaging data (i.e., voxelwise MRI data), and the survival outcome (i.e., the dementia event) and provides accurate personalized predictions of longitudinal trajectories and the dementia occurrence. To account for the association between the longitudinal clinical data and longitudinal MRI data, we extend Park & Staicu’s model and use the joint and individual variation explained (JIVE) approach to decompose the longitudinal MRI data (Lock et al. (2013)). To the best of our knowledge, no work has been done in leveraging multivariate longitudinal data, longitudinal imaging data, and survival data in the AD prediction model. Comparing with existing literature and our previous work of modeling multivariate longitudinal and survival data (Zou et al. (2023)), our MFMM-LMRI model is novel in three aspects. First, the longitudinal voxelwise MRI data are modeled using a nonparametric functional mixed model, with a decomposition similar in nature to the general JIVE approach. Second, we model the correlation between multivariate longitudinal outcomes and longitudinal MRI data via the random functional joint variation based on the latent functional principal component scores from multivariate longitudinal outcomes. Third, we incorporate the longitudinal outcomes and longitudinal MRI data in the dynamic prediction framework, which provides personalized predictions of longitudinal trajectories and the probability of remaining stable MCI for healthcare professionals and patients. We also assess the added predictive values from the longitudinal MRI data.

The article is organized as follows. In Section 2 we introduce the motivating ADNI study. In Section 3 we propose the MFMM-LMRI model and the estimation approach. In Section 4 we conduct an extensive simulation study to validate our method. In Section 5 we apply our model to the motivating ADNI study. In Section 6 we summarize the findings and discuss the limitations and future directions.

2. Motivation.

The motivation for this article comes from the ongoing Alzheimer’s Disease Neuroimaging Initiative (ADNI) study (Veitch et al. (2022)). Based on a previous finding (Li et al. (2017)), we select five longitudinal outcomes that are predictive of the dementia occurrence: 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) (Rosen, Mohs and Davis (1984), Tombaugh and McIntyre (1992), Hughes et al. (1982)). We include subjects with mild cognitive impairment (MCI) at baseline. We use a multivariate functional mixed model (MFMM) (Li, Xiao and Luo (2022)) that automatically captures the heterogeneity and correlation among multivariate longitudinal outcomes, as displayed in Figure 1. We preprocess the longitudinal T1-weighted structural MRI data using the pipelines detailed in Zou et al. (2021). We use the time from baseline to the first dementia diagnosis as the survival event time. Subjects without dementia diagnosis during the follow-up are treated as right-censored.

Based on the data downloaded on November 19, 2019, we select 1018 subjects with mild cognitive impairment (MCI) at baseline and remove 86 subjects with missing ApoE- $\epsilon 4$ genotype and one subject with missing age. We exclude 189 subjects without valid MRI data. The final analysis dataset consists of 742 subjects among whom 268 had dementia diagnosed during the study. There are 3323 visits with longitudinal cognition scores and 2877 visits with MRI data. The mean follow-up time is 2.70 years with the standard deviation being 2.47 years. The outcomes ADAS-Cog 13, RAVLT-immediate, RAVLT-learning, MMSE, and CDR-SB have 29, 15, 15, 8, and 33 missing values, respectively.

3. Method.

Data used in the preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer’s Disease (AD).

We assume that there are $i = 1, \dots, N$ subjects, $j = 1, \dots, J$ longitudinal outcomes, $k = 1, \dots, K_i$ clinical visits for subject i . Let $Y_{ij}(t_{ik})$ denote the subject i ’s longitudinal outcome j at visit k (or at time t_{ik}). We denote the missing outcome as $\tilde{Y}_{ij}(t_{ik})$, the missing indicator as I_{ijk} (equals to 1 if $Y_{ij}(t_{ik})$ is missing), and we assume missing at random as the missingness mechanism. Let $m_i(v, t_g)$ be subject i ’s MRI data at visit time t_g for voxel v , $g = 1, \dots, G_i$, where G_i is the number of visits with MRI data for subject i and $v = 1, \dots, V$ is the voxel location. We denote V as the number of selected voxels detailed in Section 3.1. For the survival outcome, we denote $T_i = \min(T_i^*, C_i)$ as the event time, $\delta_i = I(T_i^* \leq C_i)$ as the censoring indicator, where T_i^* is the failure time, C_i is the censoring time, and $I(\cdot)$ is an indicator function. Note that for subject i , time t_{iK_i} and t_{G_i} are less than or equal to the event time T_i because of truncation.

3.1. Voxel selection.

The preprocessed sMRI data has a total of $S = 1,843,303$ voxels. To reduce the computation and memory burden, we use the logistic regression classifier (LRC) to select informative voxels (Huang et al. (2017)). We also explore other voxel selection approaches (i.e., voxel averaging, baseline MRI, joint model, and univariate functional principal component analysis), with details in Supplementary Material Section 8 (Zou et al. (2025a)). Let $F_{i,s}(t_g)$ be the volume of voxel s for subject i at visit g , where $s = 1, \dots, S$ and $g = 1, \dots, G_i$. We denote the volume vector as $\mathbf{F}_{i,s} = \{F_{i,s}(t_1), \dots, F_{i,s}(t_{G_i})\}'$ for subject i and voxel s . We fit a logistic regression model to select voxels: $\text{logit}(p_i) = \mathbf{w}_i' \mathbf{F}_{i,s}$, where we use the survival information (whether subjects convert to AD, δ_i) as the outcome and $\mathbf{w}_i = \{w(t_1), \dots, w(t_{G_i})\}'$ is the regression coefficient vector. We use the cubic B-spline function to approximate the weight function: $w(t) \approx \sum_{p=1}^P a_p b_p(t)$, where $b_p(t)$ is the p th prespecified cubic B-spline function evaluated at time t and a_p is the regression coefficient. We set the number of B-spline functions P as 15. The estimated weight vector is denoted as $\widehat{\mathbf{w}}_i = \{\widehat{w}(t_1), \dots, \widehat{w}(t_{G_i})\}'$.

We compute the predicted survival status for subject i and voxel s as $c_{i,s} = 1 / \{1 + \exp(-\widehat{\mathbf{w}}_i' \mathbf{F}_{i,s})\}$. The concordance index is computed as $C_s = \sum_{i=1}^N I\{I(c_{i,s} > 0.5) = \delta_i\} / N$, where $I(\cdot)$ is an indicator function. We select 1000 informative voxels based on the rank of the concordance index. We also conduct a sensitivity analysis using 3000 and 10,000 voxels and display the results in Supplementary Material Table S7 (Zou et al. (2025a)).

Based on one reviewer's comment, we conduct the voxel averaging using the $S = 1,843,303$ voxels from the whole brain. We average the volumes of all voxels within a three-voxel radius, encompassing a $7 * 7 * 7$ cubic (343 voxels) around each target voxel. After voxel averaging, we have a total of 6562 ROIs in the whole brain. We display the results in Supplementary Material Table S7 (Zou et al. (2025a)).

3.2. Model.

We assume the longitudinal outcome $Y_{ij}(t_{ik})$ is measured with a random noise ϵ_{ijk} so that $Y_{ij}(t_{ik}) = X_{ij}(t_{ik}) + \epsilon_{ijk}$, where $X_{ij}(t_{ik})$ is the latent mean of the longitudinal outcome $Y_{ij}(t_{ik})$. We also assume the MRI data $m_i(v, t)$ is a noisy measurement of the latent MRI mean $x_{mi}(v, t)$, with a random noise $\epsilon_{mi}(v, t)$, so that $m_i(v, t) = x_{mi}(v, t) + \epsilon_{mi}(v, t)$. Note that the subscript m denotes the MRI data. In the following context, the random components $U_i(t)$, $W_{ij}(t)$, $U_{mi}(t)$, and $f_{mi}(v, t)$ are separate terms. We denote the baseline covariate vector as \mathbf{Z}_i . We model the longitudinal outcomes, longitudinal MRI data, and the survival process in the multivariate functional mixed model,

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

$$m_i(v, t) = \mu_m(v, t) + \lambda_m(v)U_{mi}(t) + f_m(v, t) + \epsilon_{mi}(v, t),$$

(2)

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

(3)

In Model (1) we decompose the latent longitudinal outcome $X_{ij}(t)$ into the fixed mean function $\mu_j(t)$, the shared random profile $U_i(t)$, and the subject-outcome specific random profile $W_{ij}(t)$, with the association parameter β_j . The mean function $\mu_j(t)$ is assumed to be dependent on time t , and an additive model can be adopted to account for the effect of some baseline covariates. The random profile $U_i(t)$ is assumed to be shared across J multivariate longitudinal outcomes and measures the shared latent disease profile. The subject-outcome specific random profile $W_{ij}(t)$ measures the subject-outcome specific disease profile. The association parameter β_j is the expected change of the outcome j in one unit increase of the shared random profile $U_i(t)$. Higher values of the shared random profile $U_i(t)$ indicate worse cognitive function because we set ADAS-Cog 13 as the first longitudinal outcome and set its association parameter $\beta_1 = 1$ for model identifiability (refer to Supplementary Material Section 2 for a detailed justification) (Zou et al. (2025a)). The random error is assumed to follow $\epsilon_{ijk} \sim N(0, \sigma_j^2)$.

In Model (2) we decompose the MRI data $m_i(v, t)$ into the mean function $\mu_m(v, t)$, the joint variation (the time-dependent MRI component $U_{mi}(t)$ and a location-dependent coefficient function $\lambda_m(v)$), the individual variation (the MRI-specific variation $f_{mi}(v, t)$), and the random noise $\epsilon_{mi}(v, t)$. The mean function $\mu_m(v, t)$ is dependent on the voxel location v and time t , and an additive model can be adopted to incorporate time-varying effects of baseline covariates. The time-dependent MRI component $U_{mi}(t)$ accounts for the correlation between the longitudinal outcomes and MRI data via the latent functional principal component scores ξ_{it} detailed in Section 3.3. The location-dependent coefficient function $\lambda_m(v)$ measures the strength of the correlation between the longitudinal outcomes and MRI data, that is, $\lambda_m(v) = 0$ indicates independence. We assume the random error $\epsilon_{mi}(v, t) \sim N(0, \sigma_m^2)$. We assume $U_{mi}(t)$, $f_{mi}(v, t)$, $W_{ij}(t)$, $\epsilon_{mi}(v, t)$, and ϵ_{ijk} are independent.

In Model (3) $h_i(t)$ is the hazard function, and $h_0(t)$ is the baseline hazard function modeled with a piecewise constant function (Lawless and Zhan (1998), Feng, Wolfe and Port (2005)). The functional form $F(X_i, t)$ is the contribution of J multivariate longitudinal outcomes to the hazard function, which is detailed in Section 3.4. We denote the MRI-specific functional principal component (FPC) score vector as $\boldsymbol{\eta}_{mi}$, derived from the MRI-specific variation $f_{mi}(v, t)$ and is detailed in Section 3.3. The parameter vector $\boldsymbol{\gamma}_m$ denotes the association between the latent MRI-specific FPC score vector $\boldsymbol{\eta}_{mi}$ and the risk of dementia onset.

In summary, the multivariate functional mixed model with longitudinal MRI data (MFMM-LMRI) has four modeling advantages when compared with existing methods. First, the longitudinal submodel (1) captures nonlinear and correlated longitudinal patterns via the mean function $\mu_j(t)$ and random component functions $U_i(t)$ and $W_{ij}(t)$, where the parametric

multivariate joint model with functional data (MJM-FD), proposed by Zou et al. (2021), is sensitive to model misspecification. Second, the longitudinal MRI data is decomposed via the nonparametric joint and individual variation explained (JIVE) approach, which is more robust than parametric modeling approaches (Goldsmith et al. (2012)). Third, we achieve computational efficiency via modeling the latent MRI-specific FPC score vector η_{mi} in the survival model (3) and retain most of the MRI-specific variations. Fourth, the missing data in the longitudinal outcomes are imputed based on the joint model (refer to Section 3.4), which increases the robustness of the modeling framework.

3.3. Modeling of latent random processes.

For the longitudinal outcomes, we approximate the mean function using cubic B-spline functions: $\mu_j(t) \approx \sum_{p=1}^P A_{jp} b_p(t)$, where $b_p(t)$ is the prespecified cubic B-spline function p measured at time t , A_{jp} is the regression coefficient, and P is the number of cubic B-spline functions. We set $P = 5$ to approximate the mean function. We also explore other numbers of B-spline functions: $P = (5, 6, 7, 8, 9)$ and obtain very similar parameter estimates (not shown). The random profile $U_i(t)$ and the subject-outcome specific random profile $W_{ij}(t)$ are assumed to be two zero-mean Gaussian processes with the covariance functions $C_0(t, t')$ and $C_1(t, t')$, respectively, where t and t' are two different time points. For the subject-outcome specific random profile $W_{ij}(t)$, we assume that the covariance function $C_1(t, t')$ is the same for all J longitudinal outcomes to reduce the computational burden and for a flexible interpretation.

We decompose the covariance matrices $C_0(t, t')$ and $C_1(t, t')$ using the Mercer's Theorem: $C_0(t, t') = \sum_{l=1}^{\infty} d_{0l} \phi_l(t) \phi_l(t')$, and $C_1(t, t') = \sum_{l=1}^{\infty} d_{1l} \psi_l(t) \psi_l(t')$, where d_{0l} and d_{1l} are nonincreasing eigenvalues and $\phi_l(t)$ and $\psi_l(t)$ are orthogonal eigenfunctions, that is, $\int \phi_l(t) \phi_{l'}(t) dt = 1 \cdot I(l = l')$. Using the Karhunen-Loève expansion, the random profiles $U_i(t)$ and $W_{ij}(t)$ are expressed as $U_i(t) = \sum_{l=1}^{\infty} \xi_{il} \phi_l(t)$ and $W_{ij}(t) = \sum_{l=1}^{\infty} \zeta_{ijl} \psi_l(t)$, where ξ_{il} is the functional principal component (FPC) score for the shared random profile $U_i(t)$ and ζ_{ijl} is the FPC score for the subject-outcome specific profile $W_{ij}(t)$. We assume the FPC scores $\xi_{il} \sim N(0, d_{0l})$ and $\zeta_{ijl} \sim N(0, d_{1l})$ and are independent across j and l .

We estimate the covariance matrix $C_{jj}(t, t') = \text{cov}(X_{ij}(t), X_{ij}(t'))$ using the multivariate fast covariance estimation (mFACE) approach (Li, Xiao and Luo (2020)) on a refined grid \mathcal{S} , where \mathcal{S} is an equally spaced time grid. In Supplementary Material Section 1 (Zou et al. (2025a)), we provide a detailed estimation procedure to estimate the covariance matrices $C_0(t, t')$ and $C_1(t, t')$. We perform an eigendecomposition on the estimated covariance matrices $\widehat{C}_0(t, t')$ and $\widehat{C}_1(t, t')$ to obtain the estimated eigenfunctions $\widehat{\phi}_l(t)$ and $\widehat{\psi}_l(t)$ and the estimated eigenvalues \widehat{d}_{0l} and \widehat{d}_{1l} . The number of eigenvalues is determined using the proportion of variance explained (PVE) criteria (Di et al. (2009)), where we set L_0 and L_1 as the number of eigenvalues for the shared random profile $U_i(t)$ and the subject-outcome specific random profile $W_{ij}(t)$, respectively.

For the MRI data, we first estimate the mean function $\mu_m(v, t)$, using the sandwich smoother (Xiao, Li and Ruppert (2013)), and obtain the estimated mean function $\widehat{\mu}_m(v, t)$.

We denote the demeaned MRI data as $m_i^*(v, t) = m_i(v, t) - \hat{\mu}_m(v, t)$. We decompose the time-dependent MRI component $U_{mi}(t) = \sum_{i=1}^{\infty} \xi_{il} \phi_{mi}(t)$ using the Karhunen–Loeve’s expansion, where ξ_{il} is the FPC score from the subject-specific disease profile $U_i(t)$ in Model (1) and accounts for the correlation between the longitudinal outcomes and MRI data, and $\phi_{mi}(t)$ are orthogonal eigenfunctions for $U_{mi}(t)$ (refer to Supplementary Material Section 2 for a detailed justification) (Zou et al. (2025a)). We estimate the coefficient function $\hat{\lambda}_m(v)$ and eigenfunctions $\hat{\phi}_{mi}(v)$ detailed in Supplementary Material Section 2 (Zou et al. (2025a)).

Similarly, using the bivariate Karhunen–Loeve’s expansion, we decompose the MRI-specific variation $f_{mi}(v, t) = \sum_{i=1}^{\infty} \xi_{mil}(t) \psi_{mi}(v)$, as in Park and Staicu (2015), where $\psi_{mi}(v)$ are voxel-specific orthogonal eigenfunctions and $\xi_{mil}(t)$ are the corresponding time-dependent coefficient functions, which determine the time-varying effects of the MRI-specific variation. We assume the coefficient function $\xi_{mil}(t)$ is a univariate zero-mean stochastic process with the covariance function $\Sigma_i(t, t') = \text{cov}(\xi_{mil}(t), \xi_{mil}(t'))$, for $l = 1, \dots, \infty$, and is correlated over time within the same subject i , but uncorrelated across subjects. By the Mercer’s Theorem, we decompose $\Sigma_i(t, t') = \sum_{r=1}^{\infty} d_{mlr} \omega_{mlr}(t) \omega_{mlr}(t')$, where $\omega_{mlr}(t)$ are the orthogonal eigenfunctions and d_{mlr} are the nonincreasing eigenvalues. Using the Karhunen–Loeve’s expansion, we express the coefficient function $\xi_{mil}(t) = \sum_{r=1}^{\infty} \eta_{mlr} \omega_{mlr}(t)$, where η_{mlr} are the latent FPC scores for the coefficient function $\xi_{mil}(t)$ and we assume $\eta_{mlr} \sim N(0, d_{mlr})$ and are independent across l and r . Thus, we express the MRI-specific variation $f_{mi}(v, t) = \sum_{l=1}^{\infty} \sum_{r=1}^{\infty} \eta_{mlr} \omega_{mlr}(t) \psi_{mi}(v)$. We estimate the eigenfunctions $\hat{\psi}_{mi}(v)$, $\hat{\omega}_{mlr}(t)$, and eigenvalues \hat{d}_{mlr} detailed in Supplementary Material Section 2 (Zou et al. (2025a)). We estimate the number of eigenfunctions L_m for $f_{mi}(v, t)$ and R_l for $\xi_{mil}(t)$, using the PVE criteria detailed in Supplementary Material Section 2 (Zou et al. (2025a)). Thus, the MRI-specific variation $f_{mi}(v, t)$ is approximated as $f_{mi}(v, t) \approx \sum_{l=1}^{L_m} \sum_{r=1}^{R_l} \eta_{mlr} \omega_{mlr}(t) \psi_{mi}(v)$.

To reduce the dimension of the MRI data, we project the MRI data $m_i(v, t)$ onto L_m eigenfunctions $\psi_{mi}(v)$, and we obtain $m_{il}(t) = \mu_{mi}(t) + \lambda_{mi} U_{mi}(t) + \xi_{mil}(t) + \epsilon_{mi}(t)$, where $\mu_{mi}(t) = \int_V \mu_m(v, t) \psi_{mi}(v) dv$, $\lambda_{mi} = \int_V \lambda_m(v) \psi_{mi}(v) dv$, $\xi_{mil}(t) = \int_V f_{mi}(v, t) \psi_{mi}(v) dv \approx \sum_{r=1}^{R_l} \eta_{mlr} \omega_{mlr}(t)$, and $\epsilon_{mi}(t) = \int_V \epsilon_m(v, t) \psi_{mi}(v) dv$ and is assumed to follow $N(0, \sigma_m^2)$. We denote the latent MRI-specific FPC score vector as $\boldsymbol{\eta}_{mi} = (\boldsymbol{\eta}_{mil})$, for $l = 1, \dots, L_m$, which is used to model the hazard function in Model (3), where $\boldsymbol{\eta}_{mil} = (\eta_{mlr})$, for $r = 1, \dots, R_l$. By projection we transform the high-dimensional longitudinal MRI data $m_i(v, t)$ with the dense voxel grid and the sparse time grid to the low-dimensional MRI-specific FPC scores $\boldsymbol{\eta}_{mi}$, which reduces the computational burden while retaining most of the MRI-specific variations. To facilitate understanding of the model structure, we display Figure 2.

3.4. Functional forms and likelihood expression.

We propose two functional forms of $F(X_i, t)$: Model 1 (the random effects model) and Model 2 (the instantaneous model).

For Model 1 (the random effects model), the functional form

$F(X_i, t) = \sum_{l=1}^{L_0} \gamma_{0l} \xi_{il} + \sum_{j=1}^J \sum_{l=1}^{L_1} \gamma_{1jl} \zeta_{ijl}$. The association between the longitudinal and survival outcome are quantified using the latent FPC scores ξ_{il} from the shared disease profile $U_i(t)$ and ζ_{ijl} from the subject-outcome specific disease profile $W_{ij}(t)$. 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 we denote the number of FPC scores for $U_i(t)$ and $W_{ij}(t)$ as L_0 and L_1 , respectively. We set the vector $\gamma_0 = (\gamma_{0l}), l = 1, \dots, L_0$ and $\gamma_{1j} = (\gamma_{1jl}), l = 1, \dots, L_1$, and $\gamma_1 = (\gamma_{1j}), j = 1, \dots, J$.

For Model 2 (the instantaneous model), the functional form $F(X_i, t) = \gamma_0 U_i(t) + \sum_{j=1}^J \gamma_{1j} W_{ij}(t)$. The association between the longitudinal and survival outcome are quantified using the current values of the shared disease profile $U_i(t)$ and subject-outcome specific pattern $W_{ij}(t)$, for $j = 1, \dots, J$. We set the vector $\gamma_0 = (\gamma_0)$ and $\gamma_1 = (\gamma_{1j}), j = 1, \dots, J$.

We denote the parameter space as $\Theta = (\beta, \sigma, \sigma_m, \alpha, \gamma_0, \gamma_1, \gamma_m, \xi, \zeta, \eta_m)$, where $\beta = (\beta_j)_{j=2, \dots, J}$, $\sigma = (\sigma_j)_{j=1, \dots, J}$, $\sigma_m = (\sigma_{ml})_{l=1, \dots, L_m}$.

Based on Models (1) to (3), the logarithm likelihood is formulated as

$$l(\Theta) = \log \left\{ \prod_{i=1}^N L_i^{\text{long}} L_i^{\text{MRI}} L_i^{\text{surv}} P(\Theta) \right\}, \quad \text{where}$$

$$L_i^{\text{long}} = \prod_{j=1}^J \prod_{k=1}^{K_j} P(Y_{ij}(t_{jk}) | \Theta)^{1 - I_{ijk}} P(\tilde{Y}_{ij}(t_{jk}) | \Theta)^{I_{ijk}}, \quad \text{and}$$

$$L_i^{\text{MRI}} = \prod_{l=1}^{L_m} \prod_{g=1}^{G_l} P(\hat{m}_{il}(t_g) | \Theta), \quad \text{and}$$

$$L_i^{\text{surv}} = h_i(T_i | \Theta)^{\delta_i} S_i(T_i | \Theta),$$
(4)

$$\tilde{Y}_{ij}(t) \sim N \left(\hat{\mu}_j(t) + \beta_j \left[\sum_{l=1}^{L_0} \xi_{il} \hat{\phi}_l(t) + \sum_{l=1}^{L_1} \zeta_{ijl} \hat{\psi}_l(t) \right], \sigma_j^2 \right),$$
(5)

$$\hat{m}_{il}(t) \sim N \left(\hat{\mu}_{ml}(t) + \hat{\lambda}_{ml} \sum_{l=1}^{L_0} \xi_{il} \hat{\phi}_{ml}(t) + \sum_{r=1}^{R_l} \eta_{mlr} \hat{\omega}_{mlr}(t), \sigma_{ml}^2 \right),$$

$$\xi_{il} \sim N(0, \hat{d}_{0l}), \quad \zeta_{ijl} \sim N(0, \hat{d}_{1l}), \quad \eta_{mlr} \sim N(0, \hat{d}_{mlr}),$$
(6)

$$\beta, \alpha, \gamma_0, \gamma_1, \gamma_m \sim N(0, 10^2), \quad \sigma, \sigma_m \sim \text{Inverse-Gamma}(0.1, 0.1).$$
(7)

We denote $P(\Theta)$ as the prior distribution and $S_i(t)$ as the survival probability at time t . We use noninformative priors for all parameters in equation (7). We use the Gaussian quadrature to approximate the survival function $S_i(t)$ (refer to Supplementary Material Section 3 for the detailed derivation) (Zou et al. (2025a)).

We adopt the Markov chain Monte Carlo (MCMC) method based on the No-U-Turn Sampling (NUTS) algorithm implemented in `Stan` to obtain posterior samples (Hoffman and Gelman (2014)). The missing outcomes are treated as unknown parameters and are imputed from the posterior samples based on the distribution in equation (5). We use two MCMC chains, each of which consists of 2000 warm-up iterations and 1000 postwarm-up iterations. The improved scale reduction statistic \widehat{R} is used to diagnose the convergence of the chains (Vehtari et al. (2021)), where an $\widehat{R} < 1.1$ indicates a satisfactory convergence. From a total of 2000 posterior samples, we derive the mean, standard deviation, and 95% credible interval (CI). For easy implementation the `Stan` code for Model 2 (the instantaneous model) with longitudinal MRI data is provided in Supplementary Material Section 9 (Zou et al. (2025a,b)).

We use the two model assessment statistics to determine the model with the best fit: the empirical Akaike information criterion (EAIC) and the empirical Bayesian information criterion (EBIC). Smaller values of the two statistics indicate a better model fit. We first compute the posterior deviation using Monte Carlo approximation: $\bar{D}(\Theta) \approx -2/Q \sum_{q=1}^Q \sum_{i=1}^N l_i(\Theta^{(q)})$, where $\Theta^{(q)}$ is the q th posterior sample of the parameter space Θ . We compute $\text{EAIC} = \bar{D}(\Theta) + 2 \cdot p$ and $\text{EBIC} = \bar{D}(\Theta) + \log(N) \cdot p$, where p is the number of parameters in the parameter space Θ .

3.5. Dynamic prediction.

Suppose a new subject B has longitudinal outcomes and MRI data until the landmark time T ; we predict the longitudinal outcomes and survival probability at a future time $T' = T + \delta_i (T' > T)$. Let $\mathbf{Y}_B(T) = (Y_{B_j}(t_k))_{j=1, \dots, J, k=1, \dots, K_B}$ be a vector of all longitudinal outcomes until the landmark time T , where K_B is the number of clinical visits until time T for subject B . Let $m_B(v, t_g)$ denote the MRI data at voxel v and time t_g for subject B , where $g = 1, \dots, G_B$ and G_B is the number of MRI visits until the landmark time T . Let $\boldsymbol{\eta}_{mB} = (\boldsymbol{\eta}_{mBl})$, for $l = 1, \dots, L_m$, denote the vector of the MRI-specific FPC scores.

We project the MRI data $m_B(v, t)$ onto the eigenfunctions $\widehat{\psi}_{ml}(v)$, for $l = 1, \dots, L_m$. We obtain the projection vector $\widehat{\mathbf{m}}_{Bl} = (\widehat{m}_{Bl}(t_1), \dots, \widehat{m}_{Bl}(t_{G_B}))$, where $m_{Bl}(t_g) = \int_V m_B(v, t_g) \psi_{ml}(v)$, for $l = 1, \dots, L_m$. We use the MCMC method based on the NUTS algorithm to obtain the posterior samples of the FPC scores $\boldsymbol{\xi}_B$, $\boldsymbol{\zeta}_B$, and $\boldsymbol{\eta}_{mB}$ based on the posterior distribution in the following equation:

$$P(\boldsymbol{\xi}_B, \boldsymbol{\zeta}_B, \boldsymbol{\eta}_{mB} \mid \mathbf{Y}_B^{(T)}, \widehat{\mathbf{m}}_{Bl}(t), \delta_B = 0, \widehat{\Theta}) \\ \propto \prod_{j=1}^J \prod_{k=1}^{K_B} P(Y_{B_j}(t_{Bk}) \mid \boldsymbol{\xi}_B, \boldsymbol{\zeta}_B, \widehat{\Theta}) \\ \cdot \prod_{l=1}^{L_m} \prod_{g=1}^{G_B} P(\widehat{m}_{Bl}(t_g) \mid \boldsymbol{\xi}_B, \boldsymbol{\eta}_{mB}, \widehat{\Theta}) S_B(T \mid \mathbf{Z}_B, \boldsymbol{\xi}_B, \boldsymbol{\zeta}_B, \boldsymbol{\eta}_{mB}, \widehat{\Theta}) P(\boldsymbol{\xi}_B, \boldsymbol{\zeta}_B, \boldsymbol{\eta}_{mB} \mid \widehat{\Theta}).$$

where $\widehat{\Theta}$ is the estimated mean of Θ .

We compute the predicted longitudinal outcomes and survival probability detailed in the following equations:

$$E(Y_{Bj}(T') | \xi_B^{(q)}, \zeta_B^{(q)}, \widehat{\Theta}) \approx \widehat{\mu}_j(T') + \widehat{\beta}_j \left[\sum_{l=1}^{L_0} \xi_{Bil}^{(q)} \widehat{\phi}_l(T') + \sum_{l=1}^{L_1} \zeta_{Bil}^{(q)} \widehat{\psi}_l(T') \right],$$

$$S_B(T' | T, Z'_B, \xi_B^{(q)}, \zeta_B^{(q)}, \eta_{mB}^{(q)}, \widehat{\Theta}) \approx \frac{S_B(T' | Z'_B, \xi_B^{(q)}, \zeta_B^{(q)}, \eta_{mB}^{(q)}, \widehat{\Theta})}{S_B(T | Z'_B, \xi_B^{(q)}, \zeta_B^{(q)}, \eta_{mB}^{(q)}, \widehat{\Theta})},$$

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

The prediction performance is measured using the discrimination (how well the model discriminates event-free subjects and subjects with events) and calibration measurements (how well the predicted survival probability agrees with the observed survival status). We use the time-dependent area under the receiver operation curve (AUC) for the discrimination measurement, where the time-dependent AUC is computed using the integration of the time-dependent sensitivity and specificity values (Li, Greene and Hu (2018)). For the calibration measurement, we compute the Brier score (BS) as the average of the weighted squared difference between the predicted survival probability and the observed survival status (Graf et al. (1999)). Higher values of AUC and lower values of BS indicate a better prediction performance. For an illustration purpose, we compute the integrated AUC (iAUC) and integrated BS (iBS). For a fixed landmark time T , we integrate the estimated AUC and BS values over the dense prediction window δ_t using the Simpson's rule.

4. Simulation study.

We conduct a simulation study to validate the proposed method. We simulate $J = 3$ longitudinal outcomes, MRI data, and survival data from the multivariate functional mixed model with longitudinal MRI data (MFMM-LMRI),

$$\begin{aligned} Y_{ij}(t_{ik}) &= \mu_j(t_{ik}) + \beta_j[U_i(t_{ik}) + W_{ij}(t_{ik})] + \epsilon_{ijk}, \\ m_i(v, t) &= \mu_m(v, t) + \lambda_m(v)U_m(t) + f_m(v, t) + \epsilon_{mi}(v, t), \\ h_i(t) &= h_0(t)\exp(Z'_i\alpha + F(X_i, t) + \eta'_{mi}\gamma_m). \end{aligned}$$

We set the functional form as Model 2 (the instantaneous model):

$(X_i, t) = \gamma_0 U_i(t) + \sum_{j=1}^J \gamma_{1j} W_{ij}(t)$, where $\gamma_0 = 0.4$ and $\gamma_1 = (0.35, 0.30, -0.30)$. We explore four settings with different sample sizes and event rates (ER): Setting 1 ($N = 800$ for the training set, $N = 300$ for the testing set, baseline hazard $h_0(t) = \exp(-1.5)$ so ER = 35%); Setting 2 ($N = 600$ for the training set, $N = 200$ for the testing set, $h_0(t) = \exp(-1.5)$); Setting 3 ($N = 800$ for the training set, $N = 300$ for the testing set, $h_0(t) = \exp(-2)$ so ER = 25%); Setting 4 ($N = 600$ for the training set, $N = 200$ for the testing set, $h_0(t) = \exp(-2)$). Note that

Settings 1 and 3 have sample sizes close to the real data with different event rates, and Settings 2 and 4 have smaller sample sizes. For each setting we randomly split subjects into the training and testing sets. We also explore another scenario where we intentionally misspecify the functional form, with details and results in Supplementary Material Section 4 (Zou et al. (2025a)).

For the longitudinal outcomes, we set the true mean function $\mu_i(t)$ equal to the estimated mean function from the real data analysis for the first three longitudinal outcomes. The FPC scores are simulated from $\xi_{it} \sim N(0, d_{0i})$ and $\zeta_{ijt} \sim N(0, d_{1i})$, where $\mathbf{d}_0 = (3, 1)/\sqrt{10}$ and $d_{11} = 1$, for $L_0 = 2$ 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 $\beta = (\beta_2, \beta_3) = (-0.6, -0.4)$ and the standard deviation for random errors $\sigma = (1, 0.6, 0.4)$. For subject i we simulate longitudinal outcomes on a sparse time grid $\mathcal{S}_i = [0, t_i, t_i + 0.1, \dots, t_i + 0.9]$, where t_i is sampled from a multinomial distribution with equal probabilities on $[0.01, 0.02, \dots, 0.1]$. We simulate the missing probability $\text{logit}(p_{ijk}) = 0.5(t_{ik} - 5)$ and the missing status $I_{ijk} \sim \text{Bernoulli}(p_{ijk})$, thus achieving around a 2.5% missing rate for all three longitudinal outcomes. The MAR assumption holds because the longitudinal outcome $Y_{ij}(t_{ik})$ is dependent on time t_{ik} .

We simulate the baseline MRI data $m_i(v, t)$ observed on a dense voxel grid $\mathbf{V} = (1/V, 2/V, \dots, 1)$, where $V = 1000$ voxels. For subject i the MRI data are observed on a sparse time grid $(t_1, t_2, \dots, t_{G_i})$, where $t_1 = 0$, and t_2 is sampled from a multinomial distribution with equal probabilities on $(0.01, 0.02, \dots, 0.10)$, and $t_g = t_2 + 0.1(g - 2)$, for $g = 3, \dots, 11$. The mean function $\mu_m(v, t) = 6 + 3\sin(\pi v/2) + 4\cos(\pi v/2)t + 6t^2$. The coefficient function $\lambda_m(v) = \sin(2\pi v) + \cos(2\pi v)$. The eigenfunctions $\phi_{ml}(t) = \sqrt{2}\sin((2l - 1)\pi t/2)$, 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 MRI-specific variation $f_{mi}(v, t) = \sum_{l=1}^{L_m} \xi_{mil}(t)\psi_{ml}(v)$, where $\xi_{mil}(t) = \sum_{r=1}^{R_l} \eta_{mlr}\omega_{mlr}(t)$. The number of eigenfunctions $R_l = 2$, for $l = 1, \dots, L_m = 5$ and the eigenfunctions $\omega_{mlr}(t) = \sqrt{2}\sin((2r - 1)\pi t/2)$, for $l = 1, \dots, L_m = 5$ and $r = 1, \dots, R_l = 2$. We simulate the FPC scores $\eta_{mlr} \sim N(0, d_{mlr})$, where $d_{mlr} = 0.6^l - 1 \cdot 0.5^{r-1}$, for $l = 1, \dots, L_m = 5$ and $r = 1, \dots, R_l = 2$. We simulate the random error from $\epsilon_{mi}(v, t) \sim N(0, \sigma_m^2)$, where $\sigma_m = 0.2$.

For the survival outcome, we simulate $Z_i \sim \text{Binomial}(2, 0.4)$ to mimic the number of ApoE-ε4 alleles. We set $\alpha = 0.3$ and $\boldsymbol{\gamma}_m = (\boldsymbol{\gamma}_{m1}, \dots, \boldsymbol{\gamma}_{m5})'$, where $\boldsymbol{\gamma}_{m1} = (0.4, 0.2)$, $\boldsymbol{\gamma}_{m2} = (0.30, 0.15)$, $\boldsymbol{\gamma}_{m3} = (-0.4, -0.2)$, $\boldsymbol{\gamma}_{m4} = (-0.6, -0.3)$, $\boldsymbol{\gamma}_{m5} = (-0.3, -0.15)$. The survival probability \mathcal{S}_i is simulated from $U(0, 1)$, and we compute the event time T_i^* using the bisection method by solving the equation $\log(\mathcal{S}_i) = -H_i(T_i^*)$. The censoring time $C_i \sim U(0, 1.6)$.

We repeat the simulation 200 times. We obtain the posterior means, standard errors, and 95% credible intervals for each replication r , where $r = 1, \dots, 200$. From a total of $R = 200$ replications, we compute the bias (the difference between the averaged mean and the true value), standard deviation (SD of R posterior means), standard error (SE, square root of

the averaged variances), and coverage probability (CP, frequency of 95% credible intervals covering the true value). To evaluate the estimation performance for the mean functions and eigenfunctions, we compute the averaged mean-squared error (AMSE), defined as $AMSE = \sum_{r=1}^R \sum_{g=1}^G (\hat{f}_r(g) - f(g))^2 / (RG)$, where g is the grid location, G is the number of grid points on the refined time grid $\mathcal{S} = (0, 0.01, \dots, 1)$ or the dense voxel grid \mathcal{V} , $f(g)$ is the true function, and $\hat{f}_r(g)$ is the estimated function at location g and replication r , for $r = 1, \dots, R$. For dynamic prediction we apply our method to the testing dataset for R replications. We compute the true and estimated iAUC and iBS for landmark times at 0.4, 0.5, 0.55, and 0.6 for each replication r . We set the prediction window $\delta_r = (0.1, 0.11, \dots, 0.25)$. From a total of R replications, we average and compute the true and estimated iAUC and iBS values.

Table 1 (upper panel) summarizes the mean, standard deviation (SD), standard error (SE), coverage probability (CP), and average mean squared error (AMSE) for the simulation Setting 1. Table 1 suggests that most parameters are estimated with small bias and CP being close to the nominal level of 0.95, except for the parameter γ_{m21} . This is likely due to the weakness of the signal (FPC scores) explained by the MRI data. The functions are estimated close to the true functions, indicated by AMSE being close to zero. Table S3 in the Supplementary Material summarizes the mean, SD, SE, CP, and AMSE for simulation Settings 2 to 4 (Zou et al. (2025a)). Table S3 suggests that, for all simulation settings, the parameters and functions are estimated with small bias with CP being close to the nominal level of 0.95.

Table 1 (lower panel) summarizes the true and estimated iAUC and iBS for simulation Setting 1. Table 1 suggests that the estimated iAUC and iBS values are close to their true values. Table S4 in the Supplementary Material summarizes the iAUC and iBS for simulation Settings 2 to 4 (Zou et al. (2025a)), which suggests that the estimated iAUC and iBS are close to their true values for all settings with various sample sizes and event rates.

5. Real data application.

We apply the proposed multivariate functional mixed model–longitudinal MRI (MFMM-LMRI) to the motivating ADNI study. For subject i at time t_{ik} , let $Y_{ij}(t_{ik})$, where $j = 1, \dots, 5$, be ADAS-Cog 13 (higher worse), RAVLT-immediate (higher better), RAVLT-learning (higher better), MMSE (higher better), and CDR-SB (higher worse) scores after the Box–Cox transformation for normality. We use a 95% PVE and determine the number of the functional principal component (FPC) scores for the shared random profile $U_i(t)$ as $L_0 = 2$ and for the outcome-specific random profile $W_{ij}(t)$ as $L_1 = 1$. We select 1000 voxels based on the rank of the concordance index in Section 3.1. We use a 95% PVE and determine the number of FPC scores for the MRI-specific variation $f_m(v, t)$ as $L_m = 5$. The knots for the baseline hazard function $h_0(t)$ are set as (four, six, eight) years. We also explore other knot locations and obtain very similar results (not shown). The baseline covariate vector $Z_i = (\text{Age}_i, \text{Sex}_i, \text{Education}_i, \text{ApoE}_i)'$, where the Male sex is set as the reference, Education_i is the education years for subject i , and ApoE_i is the number of ApoE – $\epsilon 4$ alleles for subject i .

We fit the MFMM-LMRI for the two functional forms: Model 1 (the random effects model) and Model 2 (the instantaneous model). For each functional form, we compare the model assessment statistics for the five sets of models: (1) longitudinal whole-brain: the voxels from the longitudinal whole-brain MRI data are selected using LRC, with $V = 1000$ voxels; (2) baseline whole-brain: the voxels are from the baseline whole-brain MRI data, with the same $V = 1000$ voxels; (3) longitudinal hippocampus: the longitudinal MRI voxels are from the hippocampus region, with $V = 9846$ voxels; (4) baseline hippocampus: the MRI voxels are from the baseline MRI data and hippocampus region; (5) MFMM: the model does not include any MRI data. We compare the whole-brain model vs. the hippocampus model because brain atrophy in the hippocampus region is associated with a higher risk of dementia. We also compare the MFMM-LMRI model with the other four candidate models (refer to Supplementary Material Section 6 for details) (Zou et al. (2025a)): (1) Model 3-2S (the two-stage sequential estimation approach based on Model 2 with longitudinal whole-brain voxels, to illustrate the necessity of joint modeling), (2) Model 3-MJM-MRI (the parametric mixed model that treats the baseline MRI data as functional predictors detailed in Zou et al. (2021), to display the advantage of modeling nonlinear longitudinal trajectories using functional mixed models), (3) Model 3-bCox (the Cox model with baseline covariates and five longitudinal outcomes at baseline, to show the strength of modeling longitudinal trajectories in the survival model), and (4) Model 3-NM (the nonmixed functional model, where the longitudinal model $Y_{ij}(t) = \mu_{ij}(t) + \epsilon_{ij}(t)$ and the survival model $h_i(t) = h_0(t)\exp(\mathbf{Z}_i'\boldsymbol{\alpha})$, to demonstrate the strength of modeling longitudinal outcomes using Model (1)).

Table 2 displays the model assessment statistics for all candidate models. Among the models using the instantaneous model as the functional form (Model 2), the longitudinal whole-brain MRI model has smaller EAIC and EBIC statistics when compared to the baseline whole-brain MRI model, the longitudinal hippocampus model, the baseline hippocampus model, the model without MRI data (MFMM), and the random effects model (Model 1). Table 2 also suggests that the MFMM-LMRI model has substantially better assessment statistics when compared to Model 3-2S (the two-stage approach), Model 3-MJM-MRI (the parametric mixed model), Model 3-bCox (the Cox model with baseline longitudinal outcomes), and Model 3-NM (the nonmixed functional model). The results indicate that Model 2 with longitudinal whole-brain MRI data has the best model fitting performance and it is selected as the final model.

Table 3 summarizes the mean, standard error, and 95% credible intervals from Model 2 (the instantaneous model) with longitudinal whole-brain voxels. For the longitudinal outcomes, one unit increase in ADAS-Cog 13 is associated with a 0.618 unit decrease in RAVLT-immediate (95% CI: $[-0.645, -0.594]$), 0.680 unit decrease in RAVLT-learning (95% CI: $[-0.715, -0.645]$), 0.037 unit decrease in MMSE (95% CI: $[-0.038, -0.035]$), 0.176 unit increase in CDR-SB (95% CI: $[0.167, 0.185]$), after their Box-Cox transformation. Thus, higher scores of ADAS-Cog 13 (worse cognitive functions) are associated with lower scores in RAVLT-immediate, RAVLT-learning, MMSE, and higher scores in CDR-SB (worse memory and functional activities). Table S5 in the Supplementary Material summarizes the estimated variances of the FPCs from Model 2 with longitudinal whole-brain voxels (Zou et al. (2025a)). Figure S1 in the Supplementary Material summarizes the estimated mean and

eigenfunctions for the longitudinal outcomes (Zou et al. (2025a)). For the survival outcome, one additional ApoE- $\epsilon 4$ allele is associated with a 37.3% increase in log hazard (95% CI: [0.176, 0.573]). One unit increase of the shared disease profile $U_j(t)$ (worse cognitive functions) are associated with a 69.7% increase (95% CI: [0.533, 0.861]) in log-hazard (a higher risk of dementia onset). We also conclude that higher values of the subject-outcome specific disease profiles in ADAS-Cog 13 and CDR-SB (worse memory and cognitive functions) are associated with a higher risk of dementia onset, indicated by the significant $\hat{\gamma}_{11}$ (0.266, 95% CI: [0.094, 0.449]) and $\hat{\gamma}_{15}$ (0.292, 95% CI: [0.149, 0.442]).

The parameter γ_m quantifies the association between the l -th latent FPC score vector and the risk of dementia onset. For instance, the estimated $\hat{\gamma}_{m11}$ suggests that a unit increase in the first latent FPC score η_{m11} is associated with a 0.290 unit (95% CI: [-0.050, 0.618]) increase in the logarithm hazard rate of dementia onset. We perform a likelihood ratio test to examine the significance of the association between the longitudinal MRI data and the survival outcome: $H_0: \gamma_m = 0$ and H_1 : at least one γ_m does not equal to zero. The test statistic $T = 22.6 \xrightarrow{d} \chi_6^2$ and the p-value is around 0.001, providing a strong evidence of the significant association between the longitudinal MRI data and the risk of dementia onset.

We apply the dynamic prediction method to all candidate models. We randomly set aside 556 subjects (75%) as the training set and the remaining 186 subjects (25%) as the testing set. The landmark time $T = (2, 2.5, 3, 3.5, 4)$ and the prediction window $\delta_t = (0.50, 0.51, \dots, 1.50)$. For each candidate model, we repeat the dynamic prediction 100 times and compute the mean of the integrated AUC (iAUC) and integrated BS (iBS). Table 4 summarizes the mean of iAUC and iBS for all candidate models. In summary, Model 2 (the instantaneous model) with longitudinal whole-brain voxels has the best overall predictive performance in iAUC and iBS values. Comparing with Model 3-2S (the two-stage approach based on M2-Long-whole), Model 3-MJM-MRI (the multivariate joint model with MRI data), Model 3-bCox (the Cox model with baseline longitudinal outcomes), and Model 3-NM (the nonmixed functional model), Model 2 with longitudinal whole-brain voxels show a substantially better predictive performance, suggesting the improvement in predictive performance using the joint modeling when compared to a two-stage approach, by modeling longitudinal outcomes using functional mixed models, by modeling longitudinal trajectories in the survival model, and by modeling random component functions in the multivariate functional mixed model.

Table S7 in the Supplementary Material presents the iAUC and iBS results from various voxel selection approaches (Zou et al. (2025a)), including voxel averaging, baseline MRI, joint model, and univariate functional principal component analysis, detailed in Supplementary Material Section 8 (Zou et al. (2025a)). A comparison between Table 4 and Table S7 indicates similar predictive performances across all five voxel selection methods. Additionally, we compare the MFMM-MRI model to the tensor regression method (Zhou, Li and Zhu (2013)) and three machine learning algorithms: Random Survival Forest (Ishwaran et al. (2008)), Survival Deep Neural Network (Katzman et al. (2018)), and Bayesian Additive Regression Trees (Sparapani et al. (2016)). The results, detailed in Supplementary Material Table S7 (Zou et al. (2025a)), show that the MFMM-MRI

model consistently achieves higher iAUC values compared to both the tensor regression method and the machine learning algorithms. Supplementary Material Table S8 details the standard deviation of the iAUC and iBS values for all candidate models (Zou et al. (2025a)), providing insight into the variability of these metrics. Furthermore, Supplementary Material Table S9 lists the p-values from the paired t-tests comparing the iAUC and iBS values between the final model, M2-Long-whole, and other models (Zou et al. (2025a)). Due to potential violations of the independence assumption, these p-values should be interpreted with caution.

To illustrate the model's capability of dynamic prediction, we randomly select a subject C with censoring time at 5.98 years. Figure 3 displays the prediction of the five longitudinal outcomes and the survival probability for subject C given landmark times at two or three years. Subject C has good cognitive and memory functions (low ADAS-Cog 13 scores and high MMSE scores before three years) and the survival probability decreases slowly with a high probability of remaining MCI at the censoring time (5.98 years).

6. Discussion.

In this article we propose the multivariate functional mixed model with longitudinal MRI data (MFMM-LMRI) that leverages multimodal data, including the longitudinal clinical data (i.e., neuropsychological test scores), the longitudinal neuroimaging data (i.e., longitudinal voxelwise magnetic resonance imaging data), and the survival data (i.e., time until dementia onset). We use the nonparametric functional mixed effects model that jointly models multivariate longitudinal outcomes and longitudinal MRI data and adopt the decomposition similar in nature to the joint and individual variation explained (JIVE) approach to decompose these two outcomes. We investigate two functional forms (the random effects model and the instantaneous model) linking the longitudinal and survival outcome. We use the Markov chain Monte Carlo (MCMC) method based on the No-U-Turn Sampling (NUTS) algorithm to obtain posterior samples. We also develop the dynamic prediction framework that provides accurate personalized predictions of longitudinal trajectories and the risk of dementia onset based on the subject's latest data. We use the logistic regression classifier (LRC) to select informative voxels from the whole brain and apply the MFMM-LMRI to the motivating ADNI study. The instantaneous model (Model (2) with longitudinal voxels from the whole brain has the best fitting and predictive performance. Comparing Model 2 with or without longitudinal MRI data, we conclude that the longitudinal MRI data from the whole brain significantly adds the prognostic value and improves the model fit.

We conclude that higher scores in ADAS-Cog 13 (worse cognition) are associated with lower RAVLT-immediate, RAVLT-learning, MMSE, and higher CDR-SB scores (worse memory and functional activities). Higher values of the shared disease profile $U_i(t)$ and the subject-outcome specific disease profiles $W_{i,j}(t)$ for ADAS-Cog 13 and CDR-SB are associated with a higher risk of dementia onset. Despite the statistical insignificance of the individual association parameters $\hat{\gamma}_m$, a likelihood ratio test indicates a significant association between the longitudinal MRI data and the survival outcome, with a p-value of 0.001. This suggests that the collective effect of the latent functional principal component (FPC) scores, derived from MRI-specific variations, significantly impacts the risk of dementia onset. We

also explore other voxel selection approaches (i.e., voxel averaging, baseline MRI, joint model, and univariate functional principal component analysis) and obtain very similar predictive performance results. The simulation study with various sample sizes and event rates supports the validity of the method.

There are a few limitations we may address as future directions. First, the longitudinal submodel of the MFMM-LMRI may be extended to a generalized functional mixed model to model multivariate longitudinal outcomes with distributions in the exponential family (i.e., binary or ordinal outcomes). The generalized functional additive model with various functional forms of the transformed mean function may be applied to model non-Gaussian functional data (Scheipl, Gertheiss and Greven (2016)). Second, the MFMM-LMRI may also be extended to incorporate relevant genetic markers from the genome-wide association studies (GWAS). While ApoE- ϵ 4 is a major genetic risk factor for AD (Corder et al. (1993)), GWAS identified more than 40 susceptible loci associated with AD risk (Andrews, Fulton-Howard and Goate (2020)). Thus, the polygenic risk scores (PRS), based on the weighted sum of the risk alleles for each invariant, may be included in the survival model (Escott-Price et al. (2015), Chouraki et al. (2016)). Third, the missing at random (MAR) assumption may not hold in the ADNI study, because subjects with deteriorating cognitive functions are more likely to have missed visits, leading to missing data that is not at random (MNAR). To address this, we may employ model-based imputation techniques, such as selection or pattern-mixture models, which account for potential dependencies between missingness and unobserved variables (Galimard et al. (2016), Hammon and Zinn (2020)). By integrating these models into our MFMM-LMRI framework, we aim to adjust for potential biases arising from MNAR data.

In this article we employ a sequential estimation procedure for statistical inference, wherein the mean and eigenfunctions for the longitudinal outcomes and MRI data are initially estimated independently of the survival data. During the MCMC procedure, we perform posterior sampling of the parameters and functional principal component (FPC) scores using these preestimated functions. This method is designed to balance computational efficiency with feasibility. Although our simulation studies in various settings did not reveal any sizable bias, we recommend that readers exercise caution when applying the provided `stan` code. Developing a fully integrated joint estimation procedure represents a valuable direction for future research.

In this study we employ logistic regression to select informative voxels from MRI data. As alternatives, leveraging the entire brain's volume information through high-dimensional Functional Principal Component Analysis (FPCA) with singular value decomposition could be considered (Zipunnikov et al. (2011)). Additionally, treating MRI data as three-dimensional functional data and using a multivariate FPCA approach offers a promising direction (Katina, Vittert and Bowman (2021)). To control the type-I error rate in voxel selection effectively, methods, such as covariate adaptive techniques, could be utilized to manage the false discovery rate (FDR) or the familywise error rate (FWER) (Zhang, Xia and Zou (2019), Zhou, Zhang and Chen (2021)).

In the ADNI dataset, various neuropsychological outcomes, such as the Trail Making Test (TMT) Parts A and B, which assess executive function, are available. To incorporate scores from distinct domains like executive function into our model, we could extend the longitudinal submodel of the MFMM-LMRI framework to include multiple latent shared variations. This approach aligns with recent advancements, for example, Li and Xiao (2023), where the multivariate functional mixed model is expanded to a latent factor model incorporating multiple independent latent processes $U_{il}(t)$ for $l = 1, \dots, L$. However, this extension significantly increases computational complexity, requiring the development of more efficient algorithms.

Many AD studies, including the ADNI study, collect various types of medical imaging data, that is, the structural MRI data, the diffusion tensor imaging (DTI) data, and the positron emission tomography (PET) scans (Veitch et al. (2022), Alexander et al. (2007), Bailey et al. (2005)). The proposed multivariate functional mixed model with longitudinal MRI data (MFMM-LMRI) may be applied to model the longitudinal high-dimensional DTI or PET data (Keihaninejad et al. (2013), Leuzy et al. (2022)). Moreover, the MFMM-LMRI may be extended to a multivariate longitudinal imaging model that simultaneously models multivariate longitudinal outcomes, multiple types of longitudinal imaging data (i.e., MRI, DTI, and PET), and the survival outcome.

The multiomics data are commonly collected in AD studies to better understand the disease mechanism (Hasin, Seldin and Lusic (2017)). For instance, the ADNI study (Veitch et al. (2022)) collects metabolomics (i.e., the bile acid and purines), lipidomics, and proteomics data, which provide a rich source of data to determine the effects of metabolites and proteins in AD progression. Thus, the MFMM-LMRI may be extended to account for the multiomics data. Specifically, the multiomics factor analysis (MOFA) approach that integrates various types of omics data via factor analysis may be applied (Argelaguet et al. (2018)). Alternatively, a Bayesian latent variable model that clusters multiomics data and reduces the dimension using the latent variables may also be adopted (Mo et al. (2018)).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic

Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Funding.

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

The research of Luo Xiao was supported by National Institute of Neurological Disorders and Stroke (grant numbers: R01NS126449 and R01NS112303).

REFERENCES

- Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ et al. (2011). 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. *Alzheimer's Dement.* 7 270–279. [PubMed: 21514249]
- Alexander AL, Lee JE, Lazar M and Field AS (2007). Diffusion tensor imaging of the brain. *Neurotherapeutics* 4 316–329. [PubMed: 17599699]
- Andrews SJ, Fulton-Howard B and Goate A (2020). Interpretation of risk loci from genome-wide association studies of Alzheimer's disease. *Lancet Neurol.* 19 326–335. [PubMed: 31986256]
- Argelaguet R, Velten B, Arnol D, Dietrich S, Zenz T, Marioni JC, Buettner F, Huber W and Stegle O (2018). Multi-omics factor analysis—a framework for unsupervised integration of multiomics data sets. *Mol. Syst. Biol* 14 e8124. 10.15252/msb.20178124 [PubMed: 29925568]
- Ashburner J and Friston KJ (2000). Voxel-based morphometry—the methods. *NeuroImage* 11 805–821. 10.1006/nimg.2000.0582 [PubMed: 10860804]
- Bailey DL, Maisey MN, Townsend DW and Valk PE (2005). *Positron Emission Tomography 2*. Springer, Berlin.
- Busatto GF, Garrido GE, Almeida OP, Castro CC, Camargo CH, Cid CG, Buchpiguel CA, Furuie S and Bottino CM (2003). A voxel-based morphometry study of temporal lobe gray matter reductions in Alzheimer's disease. *Neurobiol. Aging* 24 221–231. [PubMed: 12498956]
- Chen K and Müller H-G (2012). Modeling repeated functional observations. *J. Amer. Statist. Assoc* 107 1599–1609. MR3036419 10.1080/01621459.2012.734196
- Chiou J-M, Chen Y-T and Yang Y-F (2014). Multivariate functional principal component analysis: A normalization approach. *Statist. Sinica* 24 1571–1596. MR3308652
- Chouraki V, Reitz C, Maury F, Bis JC, Bellenguez C, Yu L, Jakobsdottir J, Mukherjee S, Adams HH et al. (2016). Evaluation of a genetic risk score to improve risk prediction for Alzheimer's disease. *J. Alzheimer's Dis* 53 921–932. [PubMed: 27340842]
- Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL and Pericak-Vance MA (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261 921–923. 10.1126/science.8346443 [PubMed: 8346443]
- de Wilde A, van Maurik IS, Kunneman M, Bouwman F, Zwan M, Willemse EA, Biessels GJ, Minkman M, Pel R et al. (2017). Alzheimer's biomarkers in daily practice (ABIDE) project: Rationale and design. *Alzheimer's Dement.* 6 143–151.
- Di C-Z, Crainiceanu CM, Caffo BS and Punjabi NM (2009). Multilevel functional principal component analysis. *Ann. Appl. Stat* 3 458–488. MR2668715 10.1214/08-AOAS206 [PubMed: 20221415]
- Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, Badarinarayan N, Gerad/Perades, IGAP consortia et al. (2015). Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain* 138 3673–3684. [PubMed: 26490334]
- Feng S, Wolfe RA and Port FK (2005). Frailty survival model analysis of the national deceased donor kidney transplant dataset using Poisson variance structures. *J. Amer. Statist. Assoc* 100 728–735. MR2206989 10.1198/016214505000000123

- Galimard J-E, Chevret S, Protopopescu C and Resche-Rigon M (2016). A multiple imputation approach for MNAR mechanisms compatible with Heckman's model. *Stat. Med* 35 2907–2920. MR3528233 10.1002/sim.6902 [PubMed: 26893215]
- Goerden J, Iki I, Danso SO, Carrière I and Muniz-Terrera G (2019). Statistical methods for dementia risk prediction and recommendations for future work: A systematic review. *Alzheimer's Dement.* 5 563–569.
- Goldsmith J, Crainiceanu CM, Caffo B and Reich D (2012). Longitudinal penalized functional regression for cognitive outcomes on neuronal tract measurements. *J. R. Stat. Soc. Ser. C. Appl. Stat* 61 453–469. MR2914521 10.1111/j.1467-9876.2011.01031.x
- Goldsmith J and Kitago T (2016). Assessing systematic effects of stroke on motor control by using hierarchical function-on-scalar regression. *J. R. Stat. Soc. Ser. C. Appl. Stat* 65 215–236. MR3456686 10.1111/rssc.12115
- Graf E, Schmoor C, Sauerbrei W and Schumacher M (1999). Assessment and comparison of prognostic classification schemes for survival data. *Stat. Med* 18 2529–2545. 10.1002/(sici)1097-0258(19990915/30)18:17/18<2529::aid-sim274>3.0.co;2-5 [PubMed: 10474158]
- Greven S, Crainiceanu C, Caffo B and Reich D (2011). Longitudinal functional principal component analysis. In *Recent Advances in Functional Data Analysis and Related Topics*. *Contrib. Statist* 149–154. Physica-Verlag/Springer, Heidelberg. MR2815575 10.1007/978-3-7908-2736-1_23
- Gross AL, Mungas DM, Leoutsakos J-MS, Albert MS, Jones RN, Alzheimer's Disease Neuroimaging Initiative et al. (2016). Alzheimer's disease severity, objectively determined and measured. *Alzheimer's Dement.* 4 159–168.
- Hammon A and Zinn S (2020). Multiple imputation of binary multilevel missing not at random data. *J. R. Stat. Soc. Ser. C. Appl. Stat* 69 547–564. MR4098961 10.1111/rssc.12401
- Hasenstab K, Scheffler A, Telesca D, Sugar CA, Jeste S, DiStefano C and entürk D (2017). A multi-dimensional functional principal components analysis of EEG data. *Biometrics* 73 999–1009. MR3713133 10.1111/biom.12635 [PubMed: 28072468]
- Hasin Y, Seldin M and Lusis A (2017). Multi-omics approaches to disease. *Genome Biol.* 18 1–15. [PubMed: 28077169]
- Hickey GL, Philipson P, Jorgensen A and Kolamunnage-Dona R (2018). Joint models of longitudinal and time-to-event data with more than one event time outcome: A review. *Int. J. Biostat* 14 20170047. MR3812146 10.1515/ijb-2017-0047
- Hoffman MD and Gelman A (2014). The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res* 15 1593–1623. MR3214779
- Huang M, Yang W, Feng Q and Chen W (2017). Longitudinal measurement and hierarchical classification framework for the prediction of Alzheimer's disease. *Sci. Rep* 7 1–13. [PubMed: 28127051]
- Hughes CP, Berg L, Danziger W, Coben LA and Martin RL (1982). A new clinical scale for the staging of dementia. *Br. J. Psychiatry* 140 566–572. [PubMed: 7104545]
- Ishwaran H, Kogalur UB, Blackstone EH and Lauer MS (2008). Random survival forests. *Ann. Appl. Stat* 2 841–860. MR2516796 10.1214/08-AOAS169
- Jiang S, Cao J, Rosner B and Colditz GA (2023). Supervised two-dimensional functional principal component analysis with time-to-event outcomes and mammogram imaging data. *Biometrics* 79 1359–1369. MR4606358 10.1111/biom.13611 [PubMed: 34854477]
- Jiang S, Xie Y and Colditz GA (2021). Functional ensemble survival tree: Dynamic prediction of Alzheimer's disease progression accommodating multiple time-varying covariates. *J. R. Stat. Soc. Ser. C. Appl. Stat* 70 66–79. MR4204938 10.1111/rssc.12449
- Kang K and Song XY (2023). Joint modeling of longitudinal imaging and survival data. *J. Comput. Graph. Statist* 32 402–412. MR4592919 10.1080/10618600.2022.2102027
- Katina S, Vittert L and Bowman AW (2021). Functional data analysis and visualisation of three-dimensional surface shape. *J. R. Stat. Soc. Ser. C. Appl. Stat* 70 691–713. MR4275842 10.1111/rssc.12482
- Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T and Kluger Y (2018). DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med. Res. Methodol* 18 1–12. [PubMed: 29301497]

- Keihaninejad S, Zhang H, Ryan NS, Malone IB, Modat M, Cardoso MJ, Cash DM, Fox NC and Ourselin S (2013). An unbiased longitudinal analysis framework for tracking white matter changes using diffusion tensor imaging with application to Alzheimer's disease. *NeuroImage* 72 153–163. [PubMed: 23370057]
- Knopman DS, Amieva H, Petersen RC, Chételat G, Holtzman DM, Hyman BT, Nixon RA and Jones DT (2021). Alzheimer disease. *Nat. Rev. Dis. Primers* 7 1–21. [PubMed: 33414454]
- Kong D, Giovanello KS, Wang Y, Lin W, Lee E, Fan Y, Murali Doraiswamy P, Zhu H, Alzheimer's Disease Neuroimaging Initiative et al. (2015). Predicting Alzheimer's disease using combined imaging-whole genome SNP data. *J. Alzheimer's Dis* 46 695–702. [PubMed: 25869783]
- Kong D, Ibrahim JG, Lee E and Zhu H (2018). FLCRM: Functional linear Cox regression model. *Biometrics* 74 109–117. MR3777931 10.1111/biom.12748 [PubMed: 28863246]
- Lawless J and Zhan M (1998). Analysis of interval-grouped recurrent-event data using piecewise constant rate functions. *Canad. J. Statist* 26 549–565.
- Leuzy A, Smith R, Cullen NC, Strandberg O, Vogel JW, Binette AP, Borroni E, Janelidze S, Ohlsson T et al. (2022). Biomarker-based prediction of longitudinal Tau positron emission tomography in Alzheimer disease. *JAMA Neurol.* 79 149–158. [PubMed: 34928318]
- Li C, Xiao L and Luo S (2020). Fast covariance estimation for multivariate sparse functional data. *Stat* 9 e245. MR4116315 [PubMed: 34262756]
- Li C, Xiao L and Luo S (2022). Joint model for survival and multivariate sparse functional data with application to a study of Alzheimer's Disease. *Biometrics* 78 435–447. MR4450566 10.1111/biom.13427 [PubMed: 33501651]
- Li H, Habes M, Wolk DA, Fan Y, Alzheimer's Disease Neuroimaging Initiative et al. (2019). A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal magnetic resonance imaging data. *Alzheimer's Dement.* 15 1059–1070. [PubMed: 31201098]
- Li K, Chan W, Doody RS, Quinn J, Luo S and Alzheimer's Disease Neuroimaging Initiative (2017). Prediction of conversion to Alzheimer's Disease with longitudinal measures and time-to-event data. *J. Alzheimer's Dis* 58 361–371. 10.3233/JAD-161201 [PubMed: 28436391]
- Li K and Luo S (2019). Dynamic predictions in Bayesian functional joint models for longitudinal and time-to-event data: An application to Alzheimer's disease. *Stat. Methods Med. Res* 28 327–342. MR3903744 10.1177/0962280217722177 [PubMed: 28750578]
- Li K, O'Brien R, Lutz M, Luo S, Initiative ADN et al. (2018). A prognostic model of Alzheimer's disease relying on multiple longitudinal measures and time-to-event data. *Alzheimer's Dement.* 14 644–651. [PubMed: 29306668]
- Li L, Greene T and Hu B (2018). 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* 27 2264–2278. MR3825906 10.1177/0962280216680239 [PubMed: 27895266]
- Li R and Xiao L (2023). Latent factor model for multivariate functional data. *Biometrics* 79 3307–3318. MR4680723 10.1111/biom.13924 [PubMed: 37661821]
- Lock EF, Hoadley KA, Marron JS and Nobel AB (2013). Joint and individual variation explained (JIVE) for integrated analysis of multiple data types. *Ann. Appl. Stat* 7 523–542. MR3086429 10.1214/12-AOAS597 [PubMed: 23745156]
- Long JD and Mills JA (2018). Joint modeling of multivariate longitudinal data and survival data in several observational studies of Huntington's disease. *BMC Med. Res. Methodol* 18 1–15. [PubMed: 29301497]
- Mauff K, Steyerberg E, Kardys I, Boersma E and Rizopoulos D (2020). Joint models with multiple longitudinal outcomes and a time-to-event outcome: A corrected two-stage approach. *Stat. Comput* 30 999–1014. MR4108688 10.1007/s11222-020-09927-9
- Milà-Alomà M, Salvadó G, Gispert JD, Vilor-Tejedor N, Grau-Rivera O, Sala-Vila A, Sánchez-Benavides G, Arenaza-Urquijo EM, Crous-Bou M et al. (2020). Amyloid beta, Tau, synaptic, neurodegeneration, and glial biomarkers in the preclinical stage of the Alzheimer's continuum. *Alzheimer's Dement.* 16 1358–1371. [PubMed: 32573951]
- Mo Q, Shen R, Guo C, Vannucci M, Chan KS and Hilsenbeck SG (2018). A fully Bayesian latent variable model for integrative clustering analysis of multi-type omics data. *Biostatistics* 19 71–86. MR3799604 10.1093/biostatistics/kxx017 [PubMed: 28541380]

- Morris JS (2015). Functional regression. *Annu. Rev. Stat. Appl* 2 321–359.
- Morris JS and Carroll RJ (2006). Wavelet-based functional mixed models. *J. R. Stat. Soc. Ser. B. Stat. Methodol* 68 179–199. MR2188981 10.1111/j.1467-9868.2006.00539.x
- Papageorgiou G, Mauff K, Tomer A and Rizopoulos D (2019). An overview of joint modeling of time-to-event and longitudinal outcomes. *Annu. Rev. Stat. Appl* 6 223–240. MR3939519 10.1146/annurev-statistics-030718-105048
- Park SY and Staicu A-M (2015). Longitudinal functional data analysis. *Stat* 4 212–226. MR3405402 10.1002/sta4.89 [PubMed: 26594358]
- Pennanen C, Testa C, Laakso M, Hallikainen M, Helkala E, Hänninen T, Kivipelto M, Könönen M, Nissinen A et al. (2005). A voxel based morphometry study on mild cognitive impairment. *J. Neurol. Neurosurg. Psychiatry* 76 11–14. [PubMed: 15607988]
- Petersen RC (2016). Mild cognitive impairment. *Continuum (Minneapolis, Minn.)* 22 404–418. 10.1212/CON.0000000000000313 [PubMed: 27042901]
- Polzehl J and Tabelow K (2019). *Magnetic Resonance Brain Imaging*. Springer, Berlin.
- Proust-Lima C, Dartigues J-F and Jacqmin-Gadda H (2016). Joint modeling of repeated multivariate cognitive measures and competing risks of dementia and death: A latent process and latent class approach. *Stat. Med* 35 382–398. MR3455508 10.1002/sim.6731 [PubMed: 26376900]
- Rizopoulos D and Ghosh P (2011). A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. *Stat. Med* 30 1366–1380. MR2828959 10.1002/sim.4205 [PubMed: 21337596]
- Rosen WG, Mohs RC and Davis KL (1984). A new rating scale for Alzheimer's disease. *Am. J. Psychiatr* 141 1356–1364. 10.1176/ajp.141.11.1356 [PubMed: 6496779]
- Scheipl F, Gertheiss J and Greven S (2016). Generalized functional additive mixed models. *Electron. J. Stat* 10 1455–1492. MR3507370 10.1214/16-EJS1145
- Shi H, Ma D, Nie Y, Beg MF, Pei J, Cao J, Alzheimer's Disease Neuroimaging Initiative et al. (2021). Early diagnosis of Alzheimer's disease on ADNI data using novel longitudinal score based on functional principal component analysis. *J. Med. Imag* 8 024502–024502.
- Song X, Xia Y and Zhu H (2017). Hidden Markov latent variable models with multivariate longitudinal data. *Biometrics* 73 313–323. MR3632377 10.1111/biom.12536 [PubMed: 27148857]
- Sparapani RA, Logan BR, McCulloch RE and Laud PW (2016). Nonparametric survival analysis using Bayesian Additive Regression Trees (BART). *Stat. Med* 35 2741–2753. MR3513715 10.1002/sim.6893 [PubMed: 26854022]
- Tang A-M, Tang N-S and Zhu H (2017). Influence analysis for skew-normal semiparametric joint models of multivariate longitudinal and multivariate survival data. *Stat. Med* 36 1476–1490. MR3631973 10.1002/sim.7211 [PubMed: 28070895]
- Tombaugh TN and McIntyre NJ (1992). The mini-mental state examination: A comprehensive review. *J. Amer. Geriatr. Soc* 40 922–935. [PubMed: 1512391]
- van Maurik IS, Vos SJ, Bos I, Bouwman FH, Teunissen CE, Scheltens P, Barkhof F, Frolich L, Kornhuber J et al. (2019). Biomarker-based prognosis for people with mild cognitive impairment (ABIDE): A modelling study. *Lancet Neurol.* 18 1034–1044. [PubMed: 31526625]
- van Maurik IS, Zwan MD, Tijms BM, Bouwman FH, Teunissen CE, Scheltens P, Wattjes MP, Barkhof F, Barkhof J et al. (2017). Interpreting biomarker results in individual patients with mild cognitive impairment in the Alzheimer's biomarkers in daily practice (ABIDE) project. *JAMA Neurol.* 74 1481–1491. [PubMed: 29049480]
- Vehtari A, Gelman A, Simpson D, Carpenter B and Bürkner P-C (2021). Rank-normalization, folding, and localization: An improved \hat{R} for assessing convergence of MCMC (with discussion). *Bayesian Anal.* 16 667–718. MR4298989 10.1214/20-ba1221
- Veitch DP, Weiner MW, Aisen PS, Beckett LA, DeCarli C, Green RC, Harvey D, Jack CR Jr, Jagust W et al. (2022). Using the Alzheimer's Disease Neuroimaging Initiative to improve early detection, diagnosis, and treatment of Alzheimer's disease. *Alzheimer's Dement.* 18 824–857. [PubMed: 34581485]
- Vlaardingerbroek MT and Boer JA (2013). *Magnetic Resonance Imaging: Theory and Practice*. Springer, Berlin.

- Wang J, Luo S and Li L (2017). Dynamic prediction for multiple repeated measures and event time data: An application to Parkinson's disease. *Ann. Appl. Stat* 11 1787–1809. MR3709578 10.1214/17-AOAS1059 [PubMed: 29081873]
- Wang J-L, Chiou J-M and Müller H-G (2016). Functional data analysis. *Annu. Rev. Stat. Appl* 3 257–295.
- Xiao L, Li Y and Ruppert D (2013). Fast bivariate P -splines: The sandwich smoother. *J. R. Stat. Soc. Ser. B. Stat. Methodol* 75 577–599. MR3065480 10.1111/rssb.12007
- Yao F. (2007). Functional principal component analysis for longitudinal and survival data. *Statist. Sinica* 17 965–983. MR2408647
- Yao F, Müller H-G and Wang J-L (2005). Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc* 100 577–590. MR2160561 10.1198/016214504000001745
- Zhang MJ, Xia F and Zou J (2019). Fast and covariate-adaptive method amplifies detection power in large-scale multiple hypothesis testing. *Nat. Commun* 10 1–11. [PubMed: 30602773]
- Zhou H, Li L and Zhu H (2013). Tensor regression with applications in neuroimaging data analysis. *J. Amer. Statist. Assoc* 108 540–552. MR3174640 10.1080/01621459.2013.776499
- Zhou H, Zhang X and Chen J (2021). Covariate adaptive familywise error rate control for genome-wide association studies. *Biometrika* 108 915–931. MR4341359 10.1093/biomet/asaa098 [PubMed: 34803516]
- Zhu H, Morris JS, Wei F and Cox DD (2017). Multivariate functional response regression, with application to fluorescence spectroscopy in a cervical pre-cancer study. *Comput. Statist. Data Anal* 111 88–101. MR3630220 10.1016/j.csda.2017.02.004
- Zipunnikov V, Caffo B, Yousem DM, Davatzikos C, Schwartz BS and Crainiceanu C (2011). Functional principal component model for high-dimensional brain imaging. *NeuroImage* 58 772–784. 10.1016/j.neuroimage.2011.05.085 [PubMed: 21798354]
- Zou H, Li K, Zeng D, Luo S and Alzheimer's Disease Neuroimaging Initiative (2021). Bayesian inference and dynamic prediction of multivariate joint model with functional data: An application to Alzheimer's disease. *Stat. Med* 40 6855–6872. MR4352771 10.1002/sim.9214 [PubMed: 34649301]
- Zou H, Xiao L, Zeng D and Luo S (2025a). Supplementary Material for “Dynamic Prediction with Multivariate Longitudinal Outcomes and Longitudinal Magnetic Resonance Imaging Data.” 10.1214/24-AOAS1970SUPPA.
- Zou H, Xiao L, Zeng D and Luo S (2025b). Code for “Dynamic Prediction with Multivariate Longitudinal Outcomes and Longitudinal Magnetic Resonance Imaging Data.” 10.1214/24-AOAS1970SUPPB.
- Zou H, Zeng D, Xiao L and Luo S (2023). Bayesian inference and dynamic prediction for multivariate longitudinal and survival data. *Ann. Appl. Stat* 17 2574–2595. MR4637681 10.1214/23-aos1733 [PubMed: 37719893]

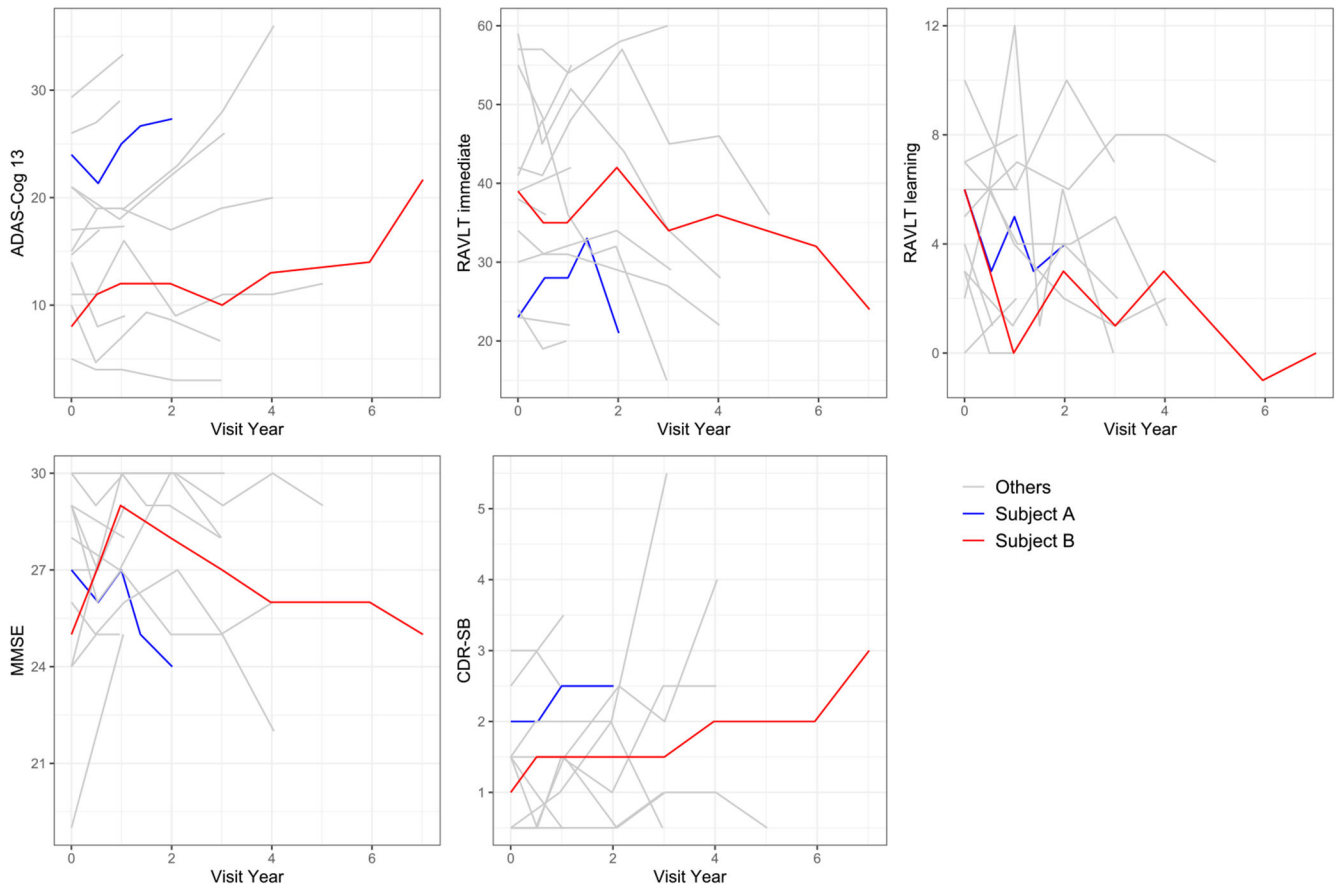


Fig. 1. Spaghetti plot of ADAS-Cog 13, RAVLT-immediate, RAVLT-learning, MMSE, and CDR-SB scores for randomly selected 15 subjects with MCI at baseline.

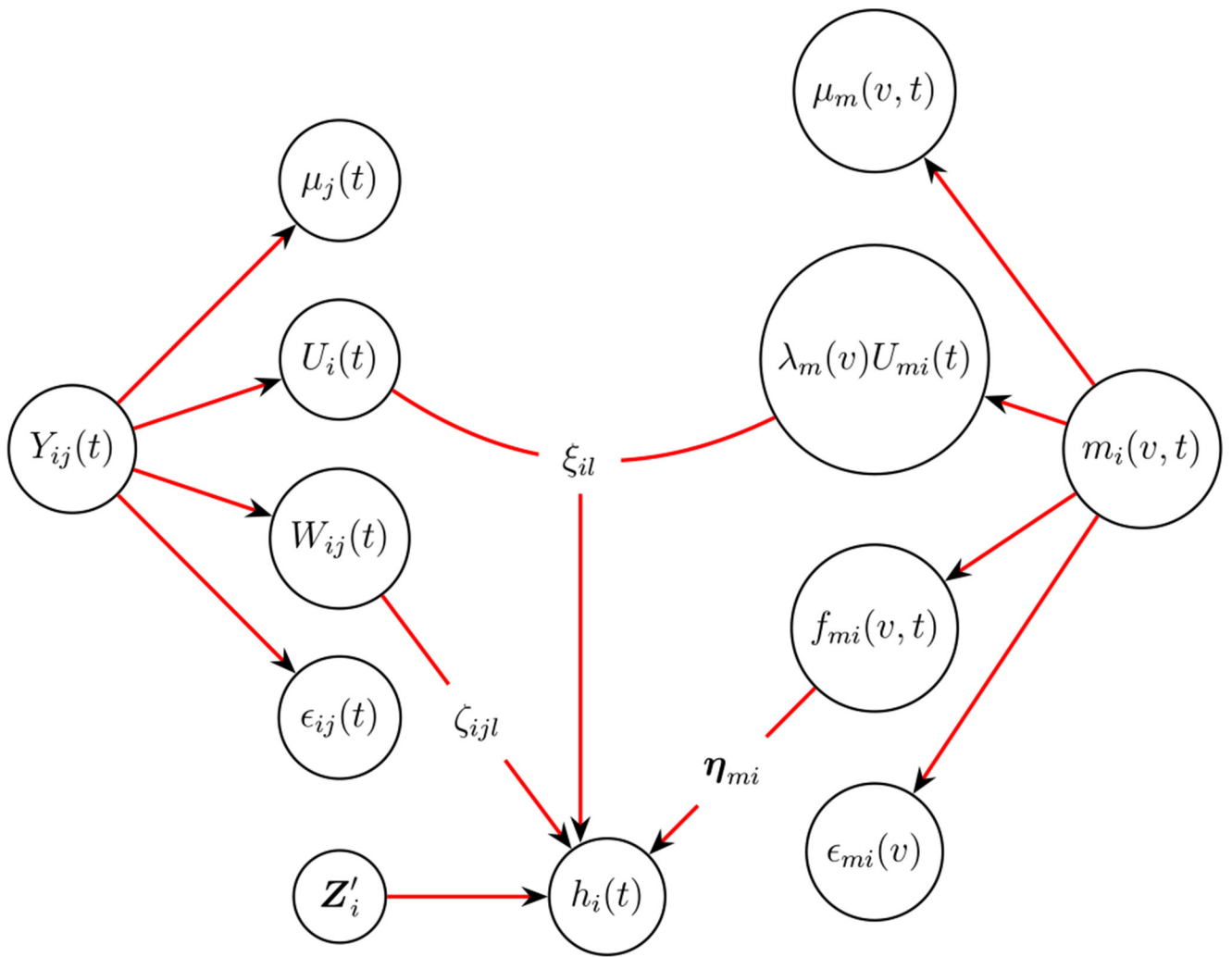


Fig. 2. A path diagram showing the structure of the multivariate functional mixed model with longitudinal MRI data.

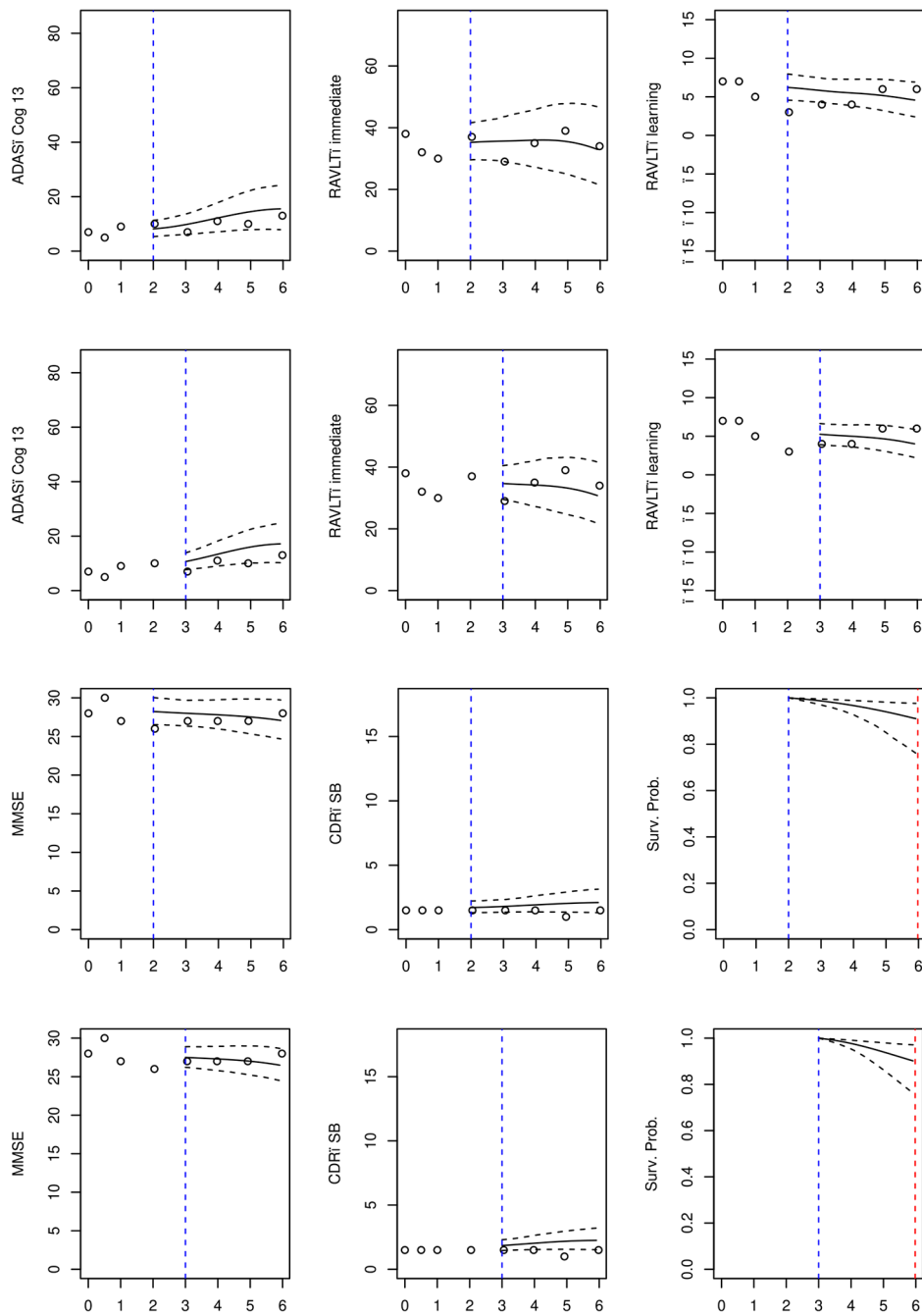


Fig. 3. Prediction of ADAS-Cog 13, RAVLT-immediate, RAVLT-learning, MMSE, CDR-SB scores, and the probability of remaining stable MCI for a new subject C at the landmark times of two years (rows 1 and 3) and three years (rows 2 and 4). The dotted circles are the observed neurological scores. The blue-dotted and red-dashed vertical lines present the landmark time (two or three years) and the censoring time (5.98 years), respectively. The solid lines

represent the predicted longitudinal trajectories or survival probabilities, with the dashed lines being 95% pointwise credible intervals.

TABLE 1

True values of parameters, bias, standard deviation (SD), standard error (SE), coverage probability (CP), averaged mean-squared error (AMSE), and true and estimated iAUC and iBS from 200 simulation replications based on Setting 1 ($N = 800$ for the training set, $N = 300$ for the testing set, with event rate around 35%)

Parameters	Bias	SD	SE	CP	AMSE
Longitudinal outcomes					
$\beta_2 = -0.6$	-0.002	0.011	0.012	0.955	$\mu_1(t)$ 0.011
$\beta_3 = 0.4$	-0.001	0.016	0.017	0.965	$\mu_2(t)$ 0.003
$\sigma_1 = 1$	0.010	0.010	0.012	0.890	$\mu_3(t)$ 0.005
$\sigma_2 = 0.6$	0.002	0.006	0.007	0.980	$\phi_1(t)$ 0.007
$\sigma_3 = 1.4$	0.001	0.012	0.015	1.000	$\phi_2(t)$ 0.020
Survival outcome					
$\alpha = 0.3$	-0.001	0.106	0.108	0.945	$\psi_1(t)$ 0.005
$\gamma_0 = 0.4$	0.023	0.130	0.090	0.820	$\mu_m(v, t)$ 0.044
$\gamma_{11} = 0.35$	-0.032	0.087	0.083	0.930	$\lambda_m(v)$ 0.015
$\gamma_{12} = 0.3$	0.002	0.093	0.086	0.955	$\phi_{m1}(t)$ 0.006
$\gamma_{13} = -0.3$	-0.035	0.140	0.133	0.930	$\phi_{m2}(t)$ 0.020
$\gamma_{m11} = 0.4$	-0.025	0.103	0.095	0.915	$\Psi_{m1}(v)$ 0.013
$\gamma_{m12} = 0.2$	-0.080	0.128	0.121	0.895	$\Psi_{m2}(v)$ 0.019
$\gamma_{m21} = 0.3$	-0.099	0.199	0.134	0.785	$\Psi_{m3}(v)$ 0.016
$\gamma_{m22} = 0.15$	-0.068	0.153	0.138	0.885	$\Psi_{m4}(v)$ 0.017
$\gamma_{m31} = -0.4$	0.049	0.150	0.155	0.945	$\Psi_{m5}(v)$ 0.008
$\gamma_{m32} = -0.2$	-0.017	0.184	0.204	0.970	$\omega_{m11}(t)$ 0.007
$\gamma_{m41} = -0.6$	0.127	0.233	0.210	0.870	$\omega_{m12}(t)$ 0.010
$\gamma_{m42} = -0.3$	0.107	0.239	0.254	0.950	$\omega_{m21}(t)$ 0.009
$\gamma_{m51} = -0.3$	-0.004	0.250	0.260	0.950	$\omega_{m22}(t)$ 0.011
$\gamma_{m52} = -0.15$	-0.028	0.313	0.336	0.950	$\omega_{m31}(t)$ 0.007
					$\omega_{m32}(t)$ 0.009
					$\omega_{m41}(t)$ 0.007
					$\omega_{m42}(t)$ 0.009
					$\omega_{m51}(t)$ 0.007
					$\omega_{m52}(t)$ 0.009
Landmark time	True iAUC	Est. iAUC	True iBS	Est. iBS	
T = 0.4	0.788	0.741	0.061	0.064	
T = 0.5	0.757	0.718	0.057	0.059	
T = 0.55	0.739	0.705	0.055	0.056	
T = 0.6	0.712	0.678	0.053	0.054	

Notations and abbreviations: iAUC: integrated area under receiver operating characteristic curve; iBS: integrated Brier score.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE 2

The model performance statistics (A total of 1000 voxels are selected via the logistic regression classifier)

	EAIC	EBIC	N. Parameters	Computation time (hrs)
Model 1—Longitudinal whole-brain	38,414	83,038	9681	0.7
Model 1—Baseline whole-brain	39,510	84,128	9680	0.4
Model 1—Longitudinal hippocampus	41,927	93,399	11,167	3.1
Model 1—Baseline hippocampus	42,484	93,943	11,164	0.3
Model 1—MFMM	39,502	84,097	9675	0.4
Model 2—Longitudinal whole-brain	38,411	83,030	9680	1.9
Model 2—Baseline whole-brain	39,533	84,147	9679	1.7
Model 2—Longitudinal hippocampus	41,490	93,408	11,166	2.1
Model 2—Baseline hippocampus	42,519	93,973	11,163	1.4
Model 2—MFMM	39,534	84,125	9674	1.7
Model 3-2S	41,611	86,229	9680	2.6
Model 3-MJM-MRI	43,548	88,543	9762	11.4
Model 3-bCox	39,554	84,122	9669	5.8
Model 3-NM	58,293	102,837	9664	1.0

Notations and abbreviations: EAIC: empirical Akaike information criterion; EBIC: empirical Bayesian information criterion. Model 1: the random effects model; Model 2: the instantaneous model; MFMM: the multivariate functional mixed model; Model 3-2S: the two-stage approach; Model 3-MJM-MRI: the parametric multivariate joint model with MRI data; Model 3-bCox: the Cox model with baseline covariates and five baseline longitudinal outcomes; Model 3-NM: the nonmixed functional model. The computation was conducted on a 13th Intel i7-13,700 2.10 GHz CPU.

TABLE 3

Mean, standard error, and 95% credible intervals from Model 2 (the instantaneous model) with longitudinal whole-brain voxels

	Mean	SE	2.5%	97.5%
Longitudinal outcomes				
β_2 : RAVLT-immediate	-0.618	0.013	-0.645	-0.594
β_3 : RAVLT-learning	-0.680	0.017	-0.715	-0.645
β_4 : MMSE	-0.037	0.001	-0.038	-0.035
β_5 : CDR-SB	0.176	0.005	0.167	0.185
Survival outcome				
Age	-0.004	0.010	-0.023	0.014
Female sex	0.158	0.155	-0.144	0.448
Education years	0.006	0.023	-0.041	0.051
ApoE- $\epsilon 4$ alleles	0.373	0.101	0.176	0.573
γ_0	0.697	0.084	0.533	0.861
γ_{11} : ADAS-Cog 13	0.266	0.090	0.094	0.449
γ_{12} : RAVLT-immediate	0.111	0.095	-0.076	0.296
γ_{13} : RAVLT-learning	0.024	0.092	-0.142	0.204
γ_{14} : MMSE	-0.172	0.063	-0.301	-0.057
γ_{15} : CDR-SB	0.292	0.075	0.149	0.442
γ_{m11}	0.290	0.172	-0.050	0.618
γ_{m21}	0.013	0.139	-0.257	0.288
γ_{m31}	0.107	0.119	-0.115	0.348
γ_{m41}	-0.050	0.205	-0.436	0.354
γ_{m42}	0.812	0.471	-0.196	1.710
γ_{m51}	0.093	0.176	-0.235	0.433

Notations and abbreviations: ADAS-Cog: Alzheimer’s Disease assessment scale—cognitive subscale; RAVLT: Rey auditory verbal learning test; MMSE: minimal state examination; CDR-SB: clinical dementia rating scale—sum of boxes. 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 survival outcome, for $j = 1, \dots, J = 5$. The parameter vector $\gamma_m = \gamma_{ml}$ for $l = 1, \dots, L_m = 5$, is the association parameter between the MRI-specific FPC score vector η_{mi} and the survival outcome, where $\gamma_{ml} = (\gamma_{mlr})$ for $r = 1, \dots, R_l$

TABLE 4

The estimated iAUC and iBS (via the logistic regression classifier (we selected 1000 voxels))

Landmark	<u>M1-Long-whole</u>		<u>M1-Base-whole</u>		<u>M1-Long-Hippo</u>		<u>M1-Base-Hippo</u>		<u>M1-MFMM</u>	
	iAUC	iBS	iAUC	iBS	iAUC	iBS	iAUC	iBS	iAUC	iBS
T = 2	0.886	0.086	0.887	0.086	0.879	0.088	0.884	0.086	0.877	0.088
T = 2.5	0.864	0.080	0.855	0.081	0.842	0.082	0.844	0.081	0.872	0.083
T = 3	0.854	0.095	0.845	0.095	0.833	0.097	0.836	0.095	0.840	0.101
T = 3.5	0.838	0.088	0.804	0.090	0.808	0.089	0.784	0.090	0.751	0.096
T = 4	0.849	0.080	0.799	0.083	0.840	0.079	0.797	0.083	0.774	0.084
Landmark	<u>M2-Long-whole</u>		<u>M2-Base-whole</u>		<u>M2-Long-Hippo</u>		<u>M2-Base-Hippo</u>		<u>M2-MFMM</u>	
	iAUC	iBS	iAUC	iBS	iAUC	iBS	iAUC	iBS	iAUC	iBS
T = 2	0.882	0.087	0.884	0.088	0.876	0.087	0.881	0.088	0.873	0.088
T = 2.5	0.872	0.079	0.878	0.078	0.859	0.078	0.866	0.077	0.871	0.079
T = 3	0.858	0.095	0.860	0.094	0.848	0.094	0.848	0.093	0.834	0.096
T = 3.5	0.834	0.088	0.832	0.086	0.807	0.088	0.815	0.086	0.767	0.094
T = 4	0.861	0.079	0.847	0.082	0.858	0.077	0.843	0.081	0.804	0.079
Landmark	<u>M3-2S</u>		<u>M3-MJM-MRI</u>		<u>M3-bCox</u>		<u>M3-NM</u>			
	iAUC	iBS	iAUC	iBS	iAUC	iBS	iAUC	iBS		
T = 2	0.854	0.091	0.870	0.097	0.863	0.087	0.697	0.102		
T = 2.5	0.843	0.078	0.867	0.089	0.810	0.086	0.675	0.092		
T = 3	0.826	0.092	0.843	0.107	0.799	0.096	0.682	0.107		
T = 3.5	0.807	0.085	0.795	0.101	0.757	0.088	0.565	0.102		
T = 4	0.833	0.078	0.829	0.088	0.762	0.083	0.577	0.097		

Notations and abbreviations: iAUC: integrated area under the receiver operation curve; iBS: integrated Brier score. M1: the random effects model; M2: the instantaneous model; Long-whole: longitudinal whole-brain MRI data; Base-whole: baseline whole-brain MRI data; Long-Hippo: longitudinal hippocampus MRI data; Base-Hippo: baseline hippocampus MRI data; MFMM: the multivariate functional mixed model; M3-2S: the two-stage approach; M3-MJM-MRI: the parametric multivariate joint model with MRI data; M3-bCox: the Cox model with baseline covariates and five baseline longitudinal outcomes; M3-NM: the nonmixed functional model. M2-Long-whole is formatted bold indicating the final model