

Dynamic prediction of Alzheimer's disease progression using features of multiple longitudinal outcomes and time-to-event data

Kan Li¹  | Sheng Luo² 

¹Merck Research Lab, Merck & Co, North Wales, Pennsylvania

²Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, North Carolina

Correspondence

Sheng Luo, Department of Biostatistics and Bioinformatics, Duke University Medical Center, 2424 Erwin Road, Suite 11082, Durham, NC 27705.
 Email: sheng.luo@duke.edu

Funding information

National Institute of Neurological Disorders and Stroke, Grant/Award Number: R01NS091307; National Institute of Aging, Grant/Award Number: R56AG062302

This paper is motivated by combining serial neurocognitive assessments and other clinical variables for monitoring the progression of Alzheimer's disease (AD). We propose a novel framework for the use of multiple longitudinal neurocognitive markers to predict the progression of AD. The conventional joint modeling longitudinal and survival data approach is not applicable when there is a large number of longitudinal outcomes. We introduce various approaches based on functional principal component for dimension reduction and feature extraction from multiple longitudinal outcomes. We use these features to extrapolate the health outcome trajectories and use scores on these features as predictors in a Cox proportional hazards model to conduct predictions over time. We propose a personalized dynamic prediction framework that can be updated as new observations collected to reflect the patient's latest prognosis, and thus intervention could be initiated in a timely manner. Simulation studies and application to the Alzheimer's Disease Neuroimaging Initiative dataset demonstrate the robustness of the method for the prediction of future health outcomes and risks of target events under various scenarios.

KEY WORDS

AUC, functional data analysis, multivariate longitudinal data, neuroimaging, two stage

1 | INTRODUCTION

Alzheimer's disease (AD) is a progressive disorder that causes brain cells to degenerate and die and is the most common cause of dementia among older adults.¹ Symptoms of AD include memory loss and cognitive decline. Since mild cognitive impairment (MCI) is often considered as a transitional stage to AD, MCI patients are usually enrolled as the target population for early prognosis and evaluating interventions.² In the Alzheimer's Disease Neuroimaging Initiative (ADNI) study (<http://adni.loni.usc.edu/>), researchers collect a multitude of longitudinal outcomes including neurocognitive tests, neuroimaging (eg, magnetic resonance imaging, MRI), genetic, and other clinical variables to measure the progression of AD. Given the lack of disease-modifying treatments for AD, the growing public threat of AD has raised the urgency to utilize these multimodal (neurocognitive tests, MRI, and genetic) data to develop robust prediction models of AD conversion, which help in the early diagnosis and timely intervention of this condition.³ This article is motivated by the joint analysis of the rich multimodal data from 355 MCI patients in the ADNI study. We are particularly interested in addressing two questions: (1) propose an effective way to model the association between multiple longitudinal health outcomes and the time to event of interest such as AD conversion, and (2) develop a personalized dynamic prediction framework for the AD conversion from MCI that best utilizes the subject-specific longitudinal profiles.

Joint modeling (JM) of longitudinal and survival data is usually a popular framework to appropriately analyze datasets with repeated measurements and time-to-event outcomes.^{4,5} Li et al⁶ adopted a JM approach to investigate the utility of each longitudinal outcome independently in determining the risk of AD conversion. In this study, they discovered that longitudinal Alzheimer's Disease Assessment Scale-Cognitive 13 items (ADAS-Cog 13) and longitudinal hippocampal volume (HV) are two of the strongest predictors of AD conversion from MCI among neurocognitive and neuroimaging markers. A practical use of joint models is to obtain dynamic personalized prediction of future longitudinal outcome trajectories and risks of survival event. For example, Li and Luo⁷ proposed a functional joint model that incorporates baseline surface-based hippocampus imaging as a functional covariate in the joint model framework. They demonstrated that such surface-based morphology measure contains more information and has higher predictive power than volume-based measure.⁸ However, the JM approach and the dynamic prediction framework only investigated the association between a single longitudinal outcome and a time-to-event outcome. One challenge of extending the JM for multivariate longitudinal outcomes is the difficulty of identifying a satisfactory parametric family to model the multiple longitudinal trajectories. Moreover, as most extensions to multiple longitudinal outcomes are mainly based on the shared random effects approach and the Bayesian framework,^{9–11} the JM approaches are computationally prohibitive when the number of longitudinal outcomes is large. Proust-Lima et al¹² proposed a joint model for multiple longitudinal outcomes and time-to-event data based on a latent class approach. However, the model must be repeatedly estimated with different numbers of classes and different initial values so that it is more computationally intensive and hard to implement. Thus, multivariate JM approaches have not yet been well implemented in available software.¹³

An alternative statistical framework that is flexible and well suited to model sparsely sampled longitudinal data and time-to-event outcome was provided by the work on functional data analysis.^{14–16} Yao¹⁵ considered the longitudinal outcome as functional data and explored the effect of longitudinal data on survival time. Holte et al¹⁷ extended the model and proposed a two-stage approach where the association between the longitudinal process and the survival endpoint was built upon the latent functional principal component (FPC) scores derived in a separate step. They revealed that it is often the changing patterns of the longitudinal outcome, rather than the actual value at the moment of event, that affects a patient's survival risk. In the following studies, the model was further extended for partial follow-up studies,¹⁸ interval-censored time,¹⁹ and prediction.^{20,21} However, these models focus on a single longitudinal outcome and are not applicable to the studies of neurodegenerative diseases, such as AD, which collect multiple longitudinal measures with more complexity. For example, Figure 1 provides the longitudinal trajectories of ADAS-Cog 13 (panel a) and Functional Activities Questionnaire (FAQ; panel b) for 50 randomly selected MCI patients in the ADNI study. Some of the important characteristics of the data are the following: (1) There are multiple neurocognitive markers that are correlated because they measure different perspectives of the cognitive impairment on the same patient. (2) Data are observed longitudinally across patients and trajectories are highly heterogeneous both within and between patients. (3) These longitudinal

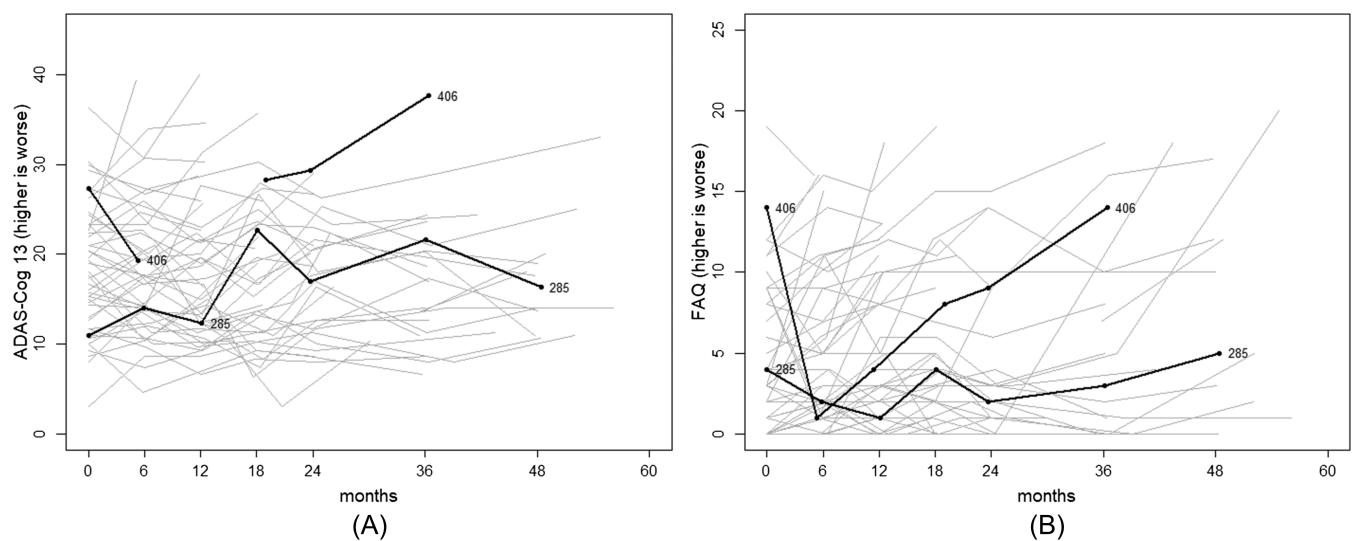


FIGURE 1 Fifty randomly selected mild cognitive impairment (MCI) patients (with two patients highlighted) from the Alzheimer's Disease Neuroimaging Initiative (ADNI) study. A, Longitudinal trajectories of Alzheimer's Disease Assessment Scale-Cognitive 13 items (ADAS-Cog 13); B, Longitudinal trajectories of Functional Activities Questionnaire (FAQ)

outcomes are associated with the time-to-event outcome, which is AD conversion from MCI in our motivating study, eg, participants 285 and 406 had AD diagnosis at month 48 and 36, respectively, as in panel a. (4) Entire visits are missing (in panel a, participant 406 had a missed visit at month 12). To our knowledge, no prior study has leveraged multiple longitudinal outcomes and time-to-event information jointly under this framework to predict the progression of AD.

To answer questions (1) and (2) and address some of the challenges discussed above, we develop a two-stage approach with multivariate FPC analysis (MFPCA)²² and Cox regression (referred to as MFPCCox) framework. In the first stage, we consider the health outcome trajectories as stochastic functions over the longitudinal visit times, and use MFPCA to extract the changing patterns (features) of multiple health outcome trajectories. In the second stage, we use scores (representing the variations of subject-specific health outcome trajectories) on these features as predictors in a Cox regression model to characterize the relationships between survival and multiple longitudinal outcomes. We propose a novel dynamic prediction framework to extrapolate the health outcome trajectories and predict the future event risk for a specific subject. Compared with the existing literature, we make several major contributions as follows: (1) MFPC-Cox provides a comprehensive and robust framework for integrating a large number of longitudinal scalar outcomes in survival analysis. (2) We use the feature information extracted from multiple longitudinal outcomes to make personalized dynamic predictions. We appropriately account for the correlation among outcomes and allow the prediction to be dynamically updated as new outcome measurements are obtained. (3) We evaluate the predictive performance of MFPCCox by using both simulation studies and the application to the ADNI study data. Although two-stage model may yield parameter estimates that are somewhat biased,⁵ our numerical results suggest that the MFPCCox performs just as well, if not better, than the parametric multivariate joint model approach in terms of prediction. The MFPC-Cox is more robust than parametric multivariate JM model if the longitudinal submodels are misspecified. (4) The implementation of the proposed MFPCCox framework and dynamic prediction can be achieved via existing software, which is much less computationally intensive than the multivariate JM approach, such as the model in the work of Rizopoulos and Ghosh.⁹

The article is organized as follows. In Section 2, we introduce the modeling framework of MFPCCox. We describe the method to extract features from multiple longitudinal outcomes and then propose a Cox regression model that uses extracted features as predictors. A dynamic prediction framework is also developed to assess the subject-specific health outcomes and disease risk. In Section 3, we apply the proposed MFPCCox to the ADNI study. In Section 4, simulation studies are conducted to examine the performance of the proposed methods. Concluding remarks and discussion are presented in Section 5.

2 | METHODS

In this section, we describe the data structure and then introduce the approaches to extract features from longitudinal outcomes. We use the features to predict future health outcomes and risk of disease progression.

2.1 | Data structure

We consider a study that recruits I subjects. Each subject i ($i = 1, \dots, I$) is followed up from the study onset until observed event time $T_i^* = \min\{T_i, C_i\}$, where T_i is the true event time of interest, C_i is the independent censoring time, and $\delta_i = I(T_i^* \leq C_i)$ is the censoring indicator. In addition to a set of P baseline or time invariant clinical covariates, denoted by vector $Z_i = \{Z_{ip}\}_{p=1,\dots,P}$, the study also repeatedly collects multiple health outcomes to measure the disease progression. For each subject i at visit j ($j = 1, \dots, J_i$), we observe data $\{Y_{ijq}\}_{q=1,\dots,Q}$, where Y_{ijq} is the q th health outcome observed at time t_{ij} ($t_{ij} \leq T_i^*$) and Q is the total number of outcomes. For notational convenience, we let $Y_{ijq} = Y_{iq}(t_{ij})$ and use these two notations interchangeably. To model the continuous longitudinal outcomes, we assume that the observed data Y_{ijq} is a noisy measurement of the latent outcome process $X_{iq}(t)$, where time $t \in \mathcal{T} = [0, \tau]$ and $\tau = \max\{T_i^* : i = 1, \dots, I\}$. It gives $Y_{ijq} = X_{iq}(t_{ij}) + \varepsilon_{ijq}$, where ε_{ijq} are independent measurement errors with mean zero and variance $\sigma_{\varepsilon_q}^2$.

2.2 | Functional principal component analysis (FPCA) for multivariate longitudinal outcomes

Let $\mathbf{X}_i(t) = \{X_{iq}(t)\}_{q=1,\dots,Q}$ be the multivariate longitudinal processes of the i th subject. We illustrate how to model $\mathbf{X}_i(t)$ via MFPCA, which is formulated as a two-step procedure. In Step 1, we let $\mu_q(t)$ be the unknown smoothed

mean function of $X_{iq}(t)$ and $\Sigma_q(t, t') = \text{cov}\{X_{iq}(t), X_{iq}(t')\}$ be the covariance function that models the correlation of outcome q 's trajectory between time points t and t' . The spectral decomposition of the covariance function is given by $\Sigma_q(t, t') = \sum_{l=1}^{\infty} \lambda_{ql} \phi_{ql}(t) \phi_{ql}(t')$, where $\{\lambda_{ql}\}_{l=1, \dots, \infty}$ are nonincreasing eigenvalues and $\{\phi_{ql}(t)\}_{l=1, \dots, \infty}$ are the corresponding orthonormal eigenfunctions. The Karhunen-Loéve expansion of $X_{iq}(t)$ is given by

$$X_{iq}(t) = \mu_q(t) + \sum_{l=1}^{\infty} \xi_{iql} \phi_{ql}(t), \quad (1)$$

where the FPC scores $\{\xi_{iql}\}_{l=1, \dots, \infty}$ are uncorrelated random variables with mean zero and variance λ_{ql} . We can view $\phi_{ql}(t)$ as a changing pattern of the longitudinal process, and ξ_{iql} describe how strongly the data from subject i follow this pattern. We assume that the longitudinal process can be adequately approximated by the first L_q eigenfunctions, and adopt a truncated version of the individual trajectory represented as $X_{iq}(t) \approx \mu_q(t) + \sum_{l=1}^{L_q} \xi_{iql} \phi_{ql}(t)$. The number of components L_q can be determined based on prespecified percentage of variance explained (PVE). Specifically, L_q may be chosen as the minimum integer such that $\sum_{l=1}^{L_q} \lambda_{ql} / \sum_{l=1}^{\infty} \lambda_{ql} \geq \text{PVE}$, eg, PVE = 80%, 90%, or 95%.

In practice, observed outcomes Y_{ijq} are available only at discrete random times t_{ij} and missing for $t > T_i^*$. The FPCA is conducted via principal analysis by conditional estimation (PACE) algorithm,²³ which has been shown to be versatile and powerful when applied to sparse and irregularly measured longitudinal data contaminated with measurement errors. The PACE algorithm is applied to the entire set of observed data for the q th outcome, generating estimated mean function $\hat{\mu}_q(t)$, error variance $\hat{\sigma}_{\epsilon_q}$, covariance function $\hat{\Sigma}_q(t, t')$, eigenvalues $\hat{\lambda}_{ql}$, and eigenfunctions $\hat{\phi}_{ql}(t)$. The estimated FPC scores of subject i are given by

$$\hat{\xi}_{iql} = \hat{\lambda}_{ql} (\hat{\phi}_{iql})^\top \hat{\Sigma}_{Y_{iq}}^{-1} (Y_{iq} - \hat{\mu}_{iq}), \quad (2)$$

where vectors $Y_{iq} = \{Y_{ijq}\}_{j=1, \dots, J_i}$, $\hat{\mu}_{iq} = \{\hat{\mu}_q(t_{ij})\}_{j=1, \dots, J_i}$, $\hat{\phi}_{iql} = \{\hat{\phi}_{ql}(t_{ij})\}_{j=1, \dots, J_i}$, and $\hat{\Sigma}_{Y_{iq}}$ is a $J_i \times J_i$ matrix with the (j, j') entry $(\hat{\Sigma}_{Y_{iq}})_{j, j'} = \hat{\Sigma}_q(t_{ij}, t_{ij'}) + \hat{\sigma}_{\epsilon_q}^2 \delta_{jj'}$, $\delta_{jj'} = 1$ if $j = j'$ and 0 if $j \neq j'$. We apply the PACE algorithm on all Q longitudinal outcomes and estimate the eigenfunctions $\hat{\phi}_q(t) = \{\hat{\phi}_{ql}(t)\}_{l=1, \dots, L_q}$ and the FPC scores $\hat{\xi}_{iq} = \{\hat{\xi}_{iql}\}_{l=1, \dots, L_q}$, with a suitably chosen truncation L_q for the q th outcome. The subject i 's vector of estimated FPC scores across all outcomes is denoted by $\hat{\xi}_i = \{\hat{\xi}_{iq}\}_{q=1, \dots, Q}$, which is of length $L_+ = \sum_{q=1}^Q L_q$.

Because each longitudinal outcome is associated with the disease progression, there may exist nonnegligible correlation among the FPC scores derived from multiple longitudinal outcomes. In Step 2, MFPCA indirectly models the correlations among outcomes via the correlations among the FPC scores. Specifically, let Θ be an $I \times L_+$ matrix, whose i th row is $\hat{\xi}_i^\top$. A matrix eigenanalysis is performed on the $L_+ \times L_+$ matrix $H = (n - 1)^{-1} \Theta^\top \Theta$ resulting into estimated eigenvalues $\{\hat{\nu}_k\}_{k=1, \dots, L_+}$ and orthonormal eigenvectors $\{\hat{\mathbf{c}}_k\}_{k=1, \dots, L_+}$. Estimates for the multivariate eigenfunctions for the q th outcome are given by $\hat{\psi}_{qk}(t) = \sum_{l=1}^{L_q} [\hat{\mathbf{c}}_k]_l^{(q)} \hat{\phi}_{ql}(t)$, where $[\hat{\mathbf{c}}_k]_l^{(q)}$ denoted the q th block of the orthonormal eigenvector $\hat{\mathbf{c}}_k$. The set of multivariate eigenfunctions $\hat{\psi}_k = \{\hat{\psi}_{qk}(t)\}_{k=1, \dots, L_+}$ characterizes the k th changing pattern of the multivariate longitudinal processes $X_i(t)$. Estimates for the MFPC scores of subject i can be calculated via

$$\hat{\rho}_{ik} = \sum_{q=1}^Q \sum_{l=1}^{L_q} [\hat{\mathbf{c}}_k]_l^{(q)} \hat{\xi}_{iql}. \quad (3)$$

An optimal number $D \leq L_+$ of multivariate FPC can be chosen based on PVE, or some information criterion such as Akaike information criterion. Then, the first D MFPC scores $\hat{\rho}_i = \{\hat{\rho}_{ik}\}_{k=1, \dots, D}$, which quantify the essential features extracted from the multivariate trajectories, are used as the predictors in a Cox model. The underlying trajectories of the q th outcome can be approximated by

$$E(Y_{iq}(t)) = \hat{X}_{iq}(t) \approx \hat{\mu}_q(t) + \sum_{k=1}^D \hat{\rho}_{ik} \hat{\psi}_{qk}(t). \quad (4)$$

2.3 | Survival analysis

The primary aim of our work is to fit the survival model using all follow-up data in the cohort and, subsequently, to make risk predictions at any given landmark time s for individuals still at risk (ie, $T_i^* > s$). Therefore, the estimated MFPC

scores $\hat{\rho}_i$, which is computed using all the follow-up data, can be used as predictors in modeling the relations between the survival time and the patterns of longitudinal outcomes. With preparation in Section 2.2, for each subject i , we use a Cox proportional hazards model that specifies the hazard function for T_i as

$$h_i(t) = h_0(t) \exp \left\{ \mathbf{Z}_i^\top \boldsymbol{\gamma} + \hat{\rho}_i^\top \boldsymbol{\beta} \right\}, \quad (5)$$

where $h_0(t)$ is an unspecified baseline hazard function that can be approximated by a piecewise-constant function or spline functions, $\boldsymbol{\gamma} = \{\gamma_p\}_{p=1,\dots,p}$ is the vector of regression coefficients for the time-independent covariates \mathbf{Z}_i , and $\boldsymbol{\beta} = \{\beta_k\}_{k=1,\dots,D}$ is the vector of regression coefficients for the multivariate longitudinal predictors through the Estimated MFPC scores. This framework combines MFPCA and Cox model and we refer to it as MFPCox.

2.4 | Personalized dynamic predictions

To apply the above framework for dynamic prediction, we fit the model in the training dataset and use the estimated parameters to provide dynamic prediction for a new subject. We first obtain estimated parameters in functional data analysis and scores $\hat{\rho}_i$ using all the longitudinal measurements from the training dataset. Then, we estimate $\hat{\gamma}$, $\hat{\beta}$, and $\hat{h}_0(t)$ by fitting model (5) using the estimated MFPC scores, time-independent covariates, and survival information of subjects in training dataset. Given a new subject N who is event-free till a landmark time s , the observations for the scalar outcomes are $\mathbf{Y}_{Nq} = \{Y_{Nq}(t_{Nj})\}_{j=1,\dots,j'}^*$ at time points $t_{N1}, \dots, t_{Nj'} \leq s \leq \tau$. We compute the FPC scores of the q th outcome, based on the observations up to time s , for this subject using model (2) $\hat{\xi}_{Nql} = \hat{\lambda}_{ql}(\hat{\phi}_{Nql})^\top (\hat{\Sigma}_{Y_{Nq}})^{-1}(\mathbf{Y}_{Nq} - \hat{\mu}_{Nq})$, $l = 1, \dots, L_q$, where $\hat{\phi}_{Nql}$, $\hat{\Sigma}_{Y_{Nq}}$, and $\hat{\mu}_{Nq}$ are constructed similarly as $\hat{\phi}_{iql}$, $\hat{\Sigma}_{Y_{iq}}$, and $\hat{\mu}_{iq}$ estimated in the training dataset. Based on the estimated FPC scores derived from Q longitudinal outcomes, we calculate the MFPC score $\hat{\rho}_N$ using model (3).

Based on the estimated MFPC scores $\hat{\rho}_N$ from the subject-specific profiles of multiple health outcomes, we predict the future outcome trajectories at a future time point $s' = (s + \Delta t) \in \mathcal{T}$ for subject N , where Δt denotes a fixed window of prediction, whereas the varying landmark time s denotes the time at which prediction is made given the subject-specific history. The latent longitudinal process of the q th outcome for subject N can be reconstructed by model (4) such that $\hat{X}_{Nq}(s') = \hat{\mu}_q(s') + \sum_{k=1}^D \hat{\rho}_{Nk} \hat{\psi}_{qk}(s')$, which utilizes correlations across multiple outcomes. The latent longitudinal processes $\hat{X}_{Nq}(s')$ can be used either to predict the future health outcome at time s' or to impute missing values occurred before time s for this new subject (eg, the missed visit from subject 406 as illustrated in Figure 1).

To predict the risk of an event not occurring within a prediction window $(s, s']$, we use the conditional probability of event-free at any time s' after the landmark time s given by

$$\hat{\pi}_N(s'|s) = p(T_N^* \geq s' | T_N^* > s, \mathbf{Z}_N, \hat{\rho}_N) = \left\{ \frac{\hat{S}_0(s')}{\hat{S}_0(s)} \right\}^{\exp \left\{ \mathbf{Z}_N^\top \hat{\gamma} + \hat{\rho}_N^\top \hat{\beta} \right\}},$$

where $\hat{S}_0(s) = \exp \{- \int_0^s \hat{h}_0(t) dt\}$ is the baseline survival function. These risk predictions can be updated when further measurements are taken, leading to dynamic risk predictions that change over time as more data accumulate. For example, suppose that subject N has not experienced the event of interest by time s' and has a new set of Q observations $\{Y_{Nq}(t_{N(j'+1)})\}_{q=1,\dots,Q}$ at time point $t_{N(j'+1)}$. Following the procedures in Section 2.2, we first update the estimated MFPC scores $\hat{\rho}_N$ by taking into account the new observations. Since MFPC scores quantify how strongly the longitudinal outcomes of subject N follow patterns $\hat{\psi}_k$, the scores are updated to reflect the trends that are updated by the new observations. With the updated MFPC scores, the new predictions of longitudinal process and event-free probability can be calculated.

The predictive performance of the model is assessed in terms of discrimination (how well the models discriminate between subjects who had the event from those who did not) and calibration (the agreement between the predicted and true risks). We use time-dependent area under the curve (AUC)²⁴ as a discrimination measure, which has been shown to be a more proper scoring rule, compared with C-index, to evaluate t -year predicted risks.²⁵ In addition, we use dynamic expected Brier score^{26,27} as calibration measure to assess the accuracy of individual risk prediction. The comparison of dynamic predictive accuracy of multiple models is conducted using the method proposed by Blanche et al.²⁷ This method derives the confidence regions of AUCs and Brier scores. It also enables the significance tests of the differences of

AUCs or Brier scores obtained from two candidate models. In general, higher AUC and lower Brier score indicate better discrimination and calibration, respectively.

Deriving the analytical expression of the sampling variability of the predictions $\hat{X}_{Nq}(s')$ and $\hat{\pi}_N(s'|s)$ in the current context is challenging. Instead, we implement a bootstrap procedure to examine the uncertainty in both the model and FPC decomposition objects. Specifically, if B bootstrap iterations are to be performed, for $b = 1, \dots, B$, we re-sample subjects from the index set $I^* = \{1, \dots, I\}$ with replacement. Let D_b be the b th bootstrap dataset that includes re-sampled subjects, we obtain the model inference using D_b and calculate the b th predictions for subject N . The process is repeated B times and B samples of model parameters and predictions are obtained. The pointwise confidence intervals for the estimated trajectories $\hat{X}_{Nq}(s)$ are constructed using a corrected confidence bands methods for sparsely observed curves proposed by Goldsmith et al.²⁸ The confidence interval for $\hat{\pi}_N(s'|s)$ is constructed from the quantile of the bootstrap samples. To facilitate easy reading and implementation of the proposed MFPC Cox model, a sample R code with simulated data is provided on https://github.com/kan-li/MFPC_Cox.

3 | APPLICATION TO THE ADNI STUDY

The primary goal of the ADNI study is to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), cerebrospinal fluid (CSF) markers, and neuropsychological assessments can be combined to measure the progression of AD. Because MCI is commonly considered as a transitional stage between normal cognition and AD, numerous recent studies assess various clinical markers and neuroimaging techniques to predict AD diagnosis among MCI patients.^{29,30} To this end, our analysis focuses on 355 MCI patients in the ADNI-1 study, and we consider AD diagnosis among MCI patients to be the survival event of interest. In the ADNI-1 study, the 355 MCI patients were reassessed at 6, 12, 18, 24, and 36 months, and additional follow-ups were conducted annually as part of ADNI-2. The average follow-up period is 3.2 years (SD 2.6; range 0.4–9.3) before AD diagnosis or censoring. The mean number of visits is 5.4 times (SD, 2.5; range 2–11) and the overall nonmonotone missing rate of longitudinal neurocognitive outcomes is 11.6%. Among them, 180 patients were diagnosed with AD (survival event) and 175 had stable MCI over a mean follow-up period of 2.3 years and 4.2 years, respectively.

We apply the proposed MFPC Cox method to the motivating ADNI-1 study to investigate whether the combination of multivariate longitudinal outcomes improves the predictive performance of the risk of AD conversion. We focus on the five neurocognitive markers that have strong predictive value and are commonly measured in AD studies⁶: Alzheimer Disease Assessment Scale-Cognitive 13 items (ADAS-Cog 13); Rey Auditory Verbal Learning Tests (RAVLT immediate score and RAVLT learning score); Functional Assessment Questionnaire (FAQ); Mini Mental State Examination (MMSE). We also include other relevant demographic and genetic variables, ie, baseline age ($bAge$), gender, years of education (Edu), and presence of at least one apolipoprotein E allele ($APOE-\epsilon 4$), given their potential effects on AD progression.^{31–33}

Two candidate Cox regression models are specified. Model 1 incorporates the baseline information of the five selected neurocognitive markers as predictors, and Model 2 incorporates the longitudinal information of the five neurocognitive markers by including the MFPC scores derived from the longitudinal observations. Along with other covariates, Model 1 is

$$\begin{aligned} h_i(t) = & h_0(t) \exp\{\gamma_1 \text{gender}_i + \gamma_2 bAge_i + \gamma_3 Edu_i + \gamma_4 APOE - \epsilon 4_i \\ & + \gamma_5 ADAS - \text{Cog}_i + \gamma_6 RAVLT_immed_i + \gamma_7 RAVLT_learn_i \\ & + \gamma_8 FAQ_i + \gamma_9 MMSE_i\}, \end{aligned}$$

and Model 2 is

$$h_i(t) = h_0(t) \exp \left\{ \gamma_1 \text{gender}_i + \gamma_2 bAge_i + \gamma_3 Edu_i + \gamma_4 APOE - \epsilon 4_i + \hat{\rho}_i^\top \beta \right\},$$

where $\hat{\rho}_i$ are MFPC scores of five neurocognitive tests. We select seven scores that account for 98% total variation of the longitudinal outcomes in Model 2. Web Figure 1 displays the first three multivariate eigenfunctions $\hat{\psi}_{q1}(t)$, $\hat{\psi}_{q2}(t)$, and $\hat{\psi}_{q3}(t)$ for the five selected neurocognitive markers, respectively. Almost 70% of the variability in the data are explained by the first functional principal component, which can be considered as the degree of cognitive impairment. The first functional principal components show the general trends relative to means during follow-up, while the second and the third functional principal components may display the trends of changing acceleration. The proportional hazards assumption

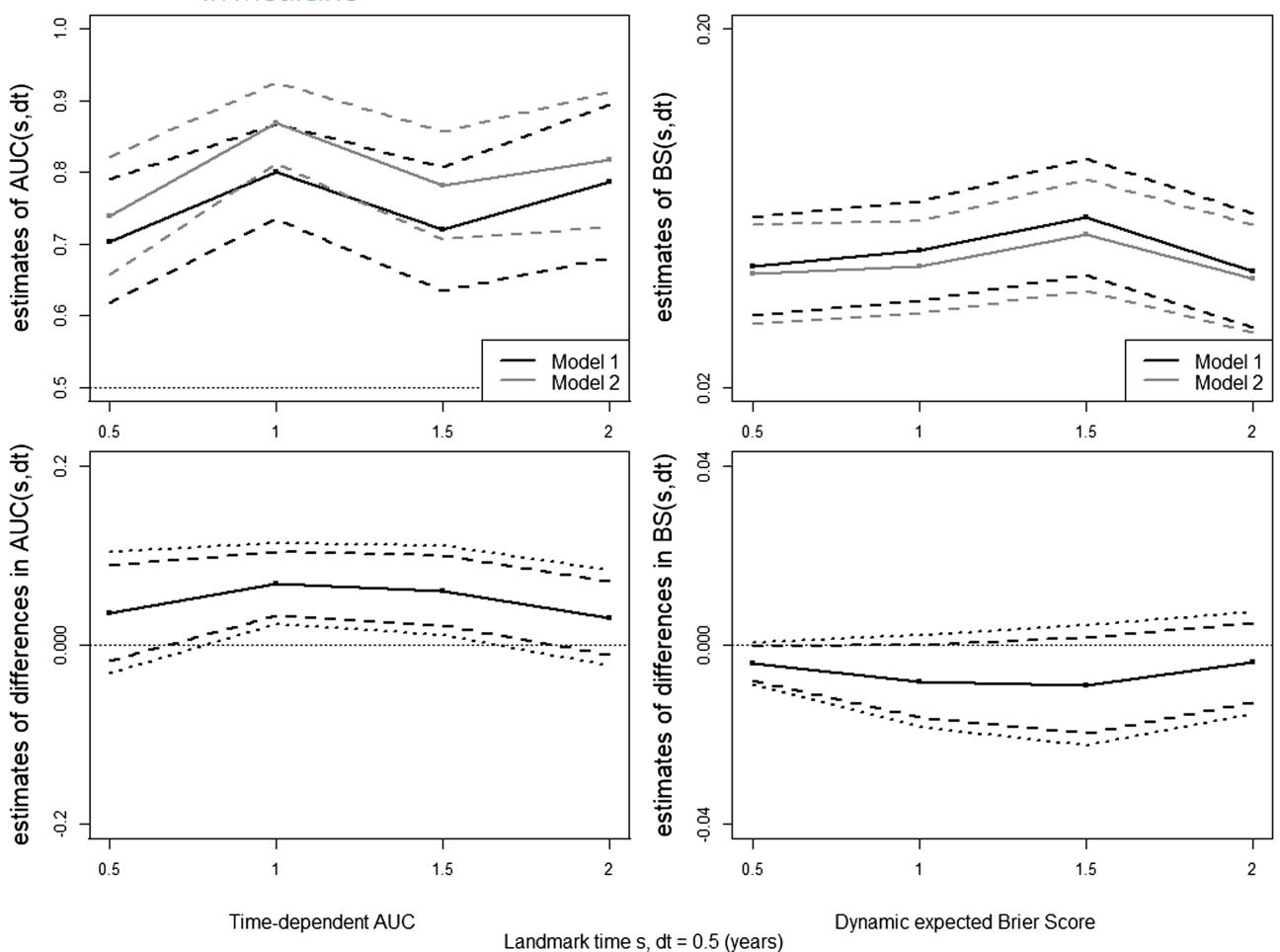


FIGURE 2 Comparison of discrimination and calibration performance of Model 1 and Model 2 within time window $(s, s + \Delta t]$, where $s = \{0.5, 1, 1.5, 2\}$ and $\Delta t = 0.5$ year. Dashed lines construct 95% pointwise confidence intervals. AUC, area under the curve; BS, Brier score

is evaluated by examining the plots of Schoenfeld residuals for each baseline covariate in each candidate model. The residual plots do not show extreme nonlinearity or other irregularities, suggesting that the proportional hazards assumption is reasonable.

We compare the two candidate models by assessing their predictive performance, in terms of the time-dependent AUC and dynamic expected Brier score (BS), at different time points over the follow-up period. To avoid overestimation of the prediction, we conduct a 10-fold cross validation. Parameters of the models are estimated from the training dataset and applied to the validation dataset. The conditional event-free probability corresponding to the time frame $(s, s + \Delta t]$ is predicted for each patient in the validation datasets as described in Section 2.4. Because the ADNI patients were reassessed approximately every half a year, we select landmark time s at 0.5, 1, 1.5, and 2 years, and let predicted window Δt be 0.5 and 1 year. The time-dependent AUC, $AUC(s, \Delta t)$, and dynamic expected Brier scores, $BS(s, \Delta t)$, are calculated based on the predicted probabilities of all patients.

Figures 2 and 3 display the curves of the estimated $AUC(s, \Delta t)$, $BS(s, \Delta t)$, difference in AUCs, $\Delta AUC(s, \Delta t)$, and difference in Brier scores, $\Delta BS(s, \Delta t)$, between Model 1 and Model 2 for different predicted windows Δt . For a $\Delta t = 0.5$ year prediction (Figure 2), Model 2 has relatively larger $AUC(s, \Delta t)$ than Model 1 for all the combinations of s and Δt . The 95% pointwise confidence intervals of the differences in $AUC(s, \Delta t)$ do not cover the zero line at landmark time $s = \{1, 1.5\}$, indicating that the null hypotheses $\Delta AUC(s, \Delta t) = 0$ are rejected with a significance level of $\alpha = 5\%$. Model 2 has significantly improved predictive performance in terms of discrimination in this interval. Although not significant, Brier score for predictions from the longitudinal health outcomes is also estimated to be lower than those from the baseline information. For a $\Delta t = 1$ year prediction (Figure 3), the 95% pointwise confidence intervals of the differences in $AUC(s, \Delta t)$, as well as the confidence intervals of the differences in $BS(s, \Delta t)$, do not cover the zero lines at landmark time $s = \{0.5, 1, 1.5\}$. It suggests that Model 2 has significantly improved predictive performance in terms of both discrimination and calibration

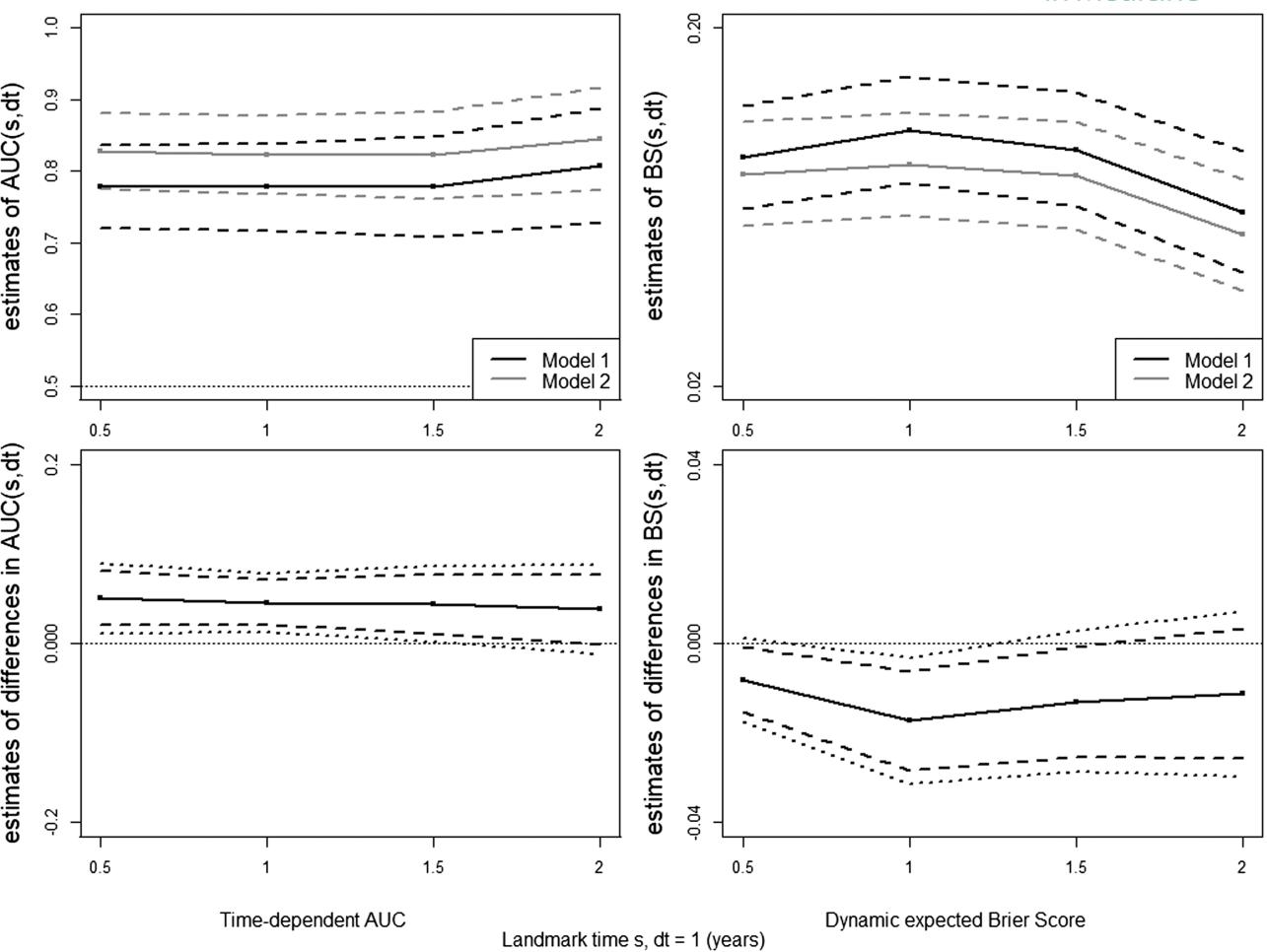


FIGURE 3 Comparison of discrimination and calibration performance of Model 1 and Model 2 within time window $(s, s + \Delta t]$, where $s = \{0.5, 1, 1.5, 2\}$ and $\Delta t = 1$ year. Dashed lines construct 95% pointwise confidence intervals. AUC, area under the curve; BS, Brier score

in these intervals. Including longitudinal information of health outcomes in the Cox model improves the capability of the model in predicting risk of AD diagnosis. The wide confidence regions may be caused by the small sample size of the study ($I = 355$).²⁷ We have also investigated another model (Model 3) to incorporate the neuroimaging information, ie, longitudinal HV, in MFPCA. The predictive performance of Model 3 has no significant improvement compared to Model 2, and it needs additional efforts to collect neuroimaging data in clinical practice. Please refer to Web Supplement for details.

Web Table 1 in the Supplementary Material displays the values of time-dependent AUC and dynamic expected Brier score from the two candidate models for each landmark time s and predicted window Δt . Both models have acceptable AUCs (range from 0.71 to 0.88) and small Brier scores (less than 0.15) indicating a good prediction performance in terms of discrimination and calibration. We select Model 2 as the final model because it has better predictive performance than Model 1. Note that Model 2 does not necessarily display an increase in $AUC(s, \Delta t)$ curve for an increasing of s . Although the models utilize more information as s increases and thus provide more accurate prediction, the decrease in discrimination is realistic and reasonable. It is probably the consequence of a selection process that makes the at-risk population more and more homogeneous as s increases.

To illustrate the personalized dynamic predictions, we set aside two target patients as validation data, and predict their future health outcomes and event-free probability based on Model 2 estimates using the remaining data as training set. Patient A has a baseline age of 78, male, no *APOE-ε4*, as compared with a more severe Patient B, 61 years old at baseline, female, and with *APOE-ε4*. The cubic spline smoothed trajectory functions of the five neurocognitive makers are plotted in Figures 4 and 5, which demonstrate how the predictions of the longitudinal outcomes are updated over time for these two patients. From the left to the right in these two figures, by using more follow-up data, predictions are closer to the true observed values and the smoothed 95% pointwise confidence intervals are narrower. It also suggests that Patient A has more stable and better predicted neurocognitive scores (better cognitive function) than Patient B. Figure 6 displays

TABLE 1 Simulation results:

comparison of prediction performance of MJM and MFPCox approaches in each scenario: (1) linear longitudinal outcomes without missing observations; (2) nonlinear longitudinal outcomes without missing observations; (3) linear longitudinal outcomes with random missing observations; (4) nonlinear longitudinal outcomes with random missing observations

	<i>t</i>	<i>Δt</i>	MJM				MFPCox		
			True AUC	AUC	BS	MSE	AUC	BS	MSE
Scenario 1	6	3	0.820	0.815	0.085	0.270	0.815	0.089	0.316
		6	0.862	0.857	0.150	0.270	0.858	0.152	0.316
		9	0.849	0.845	0.134	0.211	0.845	0.146	0.247
	12	6	0.874	0.869	0.175	0.211	0.869	0.178	0.247
		3	0.834	0.829	0.169	0.175	0.827	0.182	0.213
		6	0.817	0.812	0.274	0.175	0.813	0.284	0.213
Scenario 2	6	3	0.825	0.818	0.104	1.339	0.823	0.112	0.325
		6	0.858	0.854	0.172	1.339	0.855	0.175	0.325
		9	0.840	0.838	0.131	1.145	0.836	0.148	0.286
	12	6	0.871	0.868	0.168	1.145	0.867	0.172	0.286
		3	0.833	0.831	0.142	0.932	0.832	0.162	0.255
		6	0.832	0.829	0.191	0.932	0.832	0.198	0.255
Scenario 3	6	3	0.820	0.817	0.085	0.279	0.814	0.090	0.329
		6	0.862	0.856	0.150	0.279	0.858	0.153	0.329
		9	0.849	0.844	0.136	0.224	0.846	0.147	0.267
	12	6	0.874	0.867	0.175	0.224	0.870	0.179	0.267
		3	0.834	0.829	0.170	0.188	0.826	0.181	0.232
		6	0.817	0.814	0.275	0.188	0.809	0.285	0.232
Scenario 4	6	3	0.825	0.819	0.105	1.271	0.819	0.113	0.422
		6	0.858	0.854	0.172	1.271	0.853	0.178	0.422
		9	0.840	0.835	0.133	1.120	0.834	0.149	0.339
	12	6	0.871	0.869	0.168	1.120	0.868	0.174	0.339
		3	0.833	0.831	0.143	0.929	0.830	0.163	0.288
		6	0.832	0.828	0.191	0.929	0.830	0.200	0.288

Abbreviations: AUC, area under the curve; BS, Brier score; MFPCox, multivariate functional principal component analysis and Cox model; MJM, multivariate joint modeling; MSE, mean squared error.

the predicted probability of being free of AD diagnosis, and survival curves are estimated using cubic splines to provide smooth curves over the follow-up period. For Patient A (upper panels), the event-free probability curve does not show large changes because Patient A's predicted neurocognitive scores indicate a relatively better and stable AD status. In comparison, Patient B (lower panels) has a worse cognitive function, and thus has considerably drop in the event-free probability, suggesting that Patient B has a higher risk of AD diagnosis and should be monitored frequently.

4 | SIMULATION STUDY

We evaluate the performance of our method by using simulations studies in which the simulated data mimic the motivating application. In the simulation, we compare our proposed MFPCox to multivariate JM (MJM) approach with parametric models for longitudinal outcomes. The simulation study is designed to evaluate predictive performance of the MJM and the proposed MFPCox in four scenarios: (1) linear longitudinal outcomes without missing observations; (2) nonlinear longitudinal outcomes without missing observations; (3) linear longitudinal outcomes with random missing observations (missing completely at random); (4) nonlinear longitudinal outcomes with random missing observations. Each scenario is based on 100 datasets with sample size $I = 300$ subjects and each subject has three longitudinal outcomes, with observations at time points 0, 3, 6, 9, 12, 15, and 18 ($J_i=7$).

In Scenarios 1 and 3, the linear longitudinal measurements $\{Y_{iq}(t_{ij})\}_{q=1,2,3}$ are simulated from the longitudinal submodels

$$\begin{aligned} Y_{iq}(t_{ij}) &= X_{iq}(t_{ij}) + \epsilon_{ijq}, \\ X_{iq}(t_{ij}) &= \beta_{0q} + \beta_{tq} t_{ij} + \beta_{1q} w_i + b_{iq}. \end{aligned}$$

We let $[\beta_{01}, \beta_{02}, \beta_{03}] = [1.5, 2, 0.5]$, $[\beta_{11}, \beta_{12}, \beta_{13}] = [1.5, -1, 0.6]$, and $[\beta_{11}, \beta_{12}, \beta_{13}] = [2, -1, 1]$. We generate scalar predictors using $w_{i1} \sim N(3, 1)$, with 1 being the variance. The subject-specific random effects $[b_{i1}, b_{i2}, b_{i3}]$ are generated from

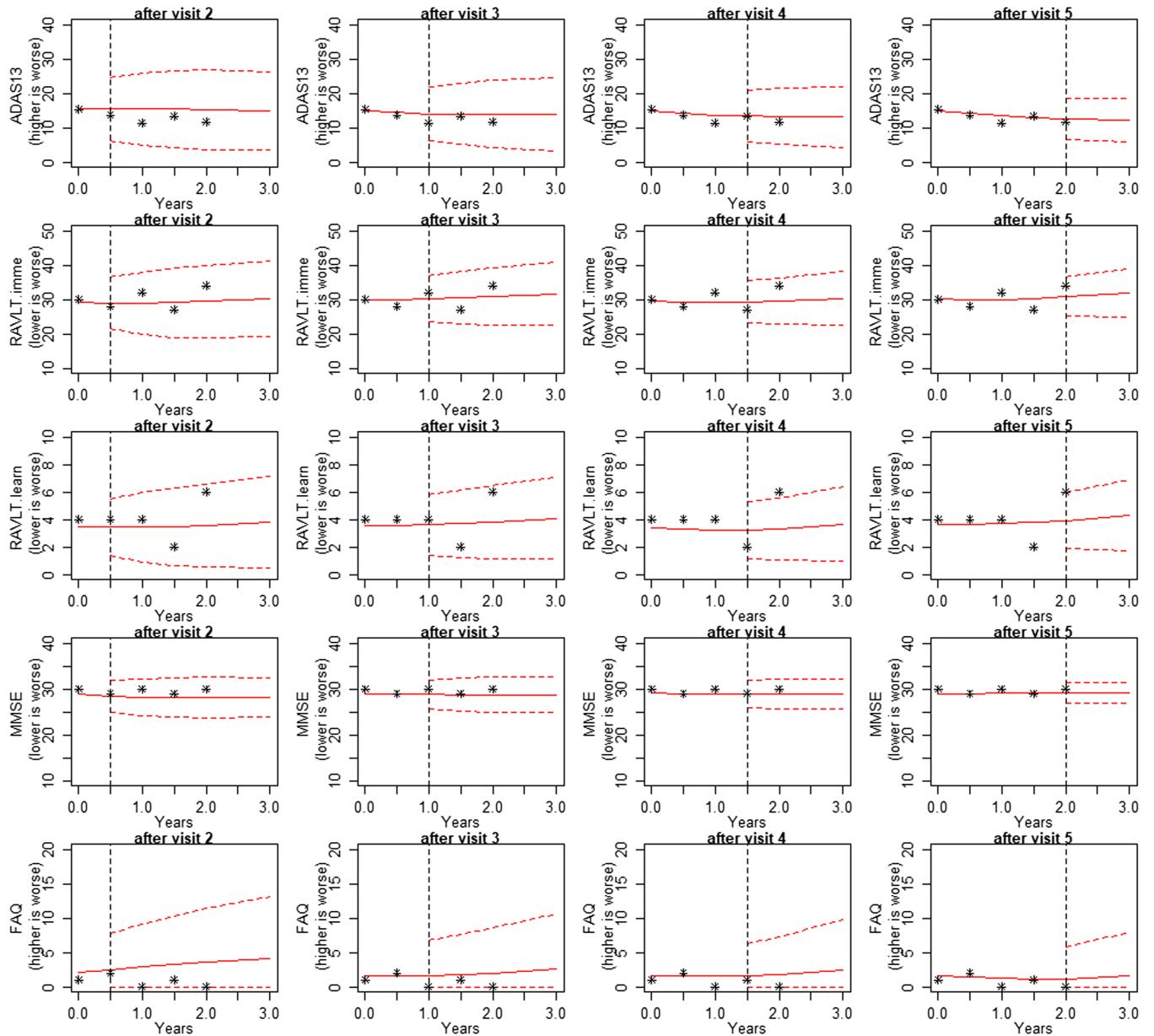


FIGURE 4 Predicted five longitudinal neurocognitive markers for Patient A. In each plot, solid line is smoothed predicted longitudinal trajectory. Dashed lines are smoothed 95% pointwise confidence intervals. The dashed vertical lines represent the landmark time s . ADAS-Cog 13, Alzheimer's Disease Assessment Scale-Cognitive 13 items; FAQ, Functional Assessment Questionnaire; MMSE, Mini Mental State Examination; RAVLT, Rey Auditory Verbal Learning Tests [Colour figure can be viewed at wileyonlinelibrary.com]

multivariate normal distribution with zero-mean and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2, & \eta_{12}\sigma_1\sigma_2, & \eta_{13}\sigma_1\sigma_3 \\ & \sigma_2^2, & \eta_{23}\sigma_2\sigma_3 \\ & & \sigma_3^2 \end{bmatrix},$$

where $[\sigma_1^2, \sigma_2^2, \sigma_3^2] = [1, 1.5, 2]$, $\eta_{12} = -0.2$, $\eta_{13} = 0.1$, and $\eta_{23} = -0.3$, respectively. The measurement error ϵ_{ijq} is simulated from $N(0, 1)$ for $q = 1, 2, 3$. In Scenarios 2 and 4 with nonlinear longitudinal outcomes, the longitudinal submodels are

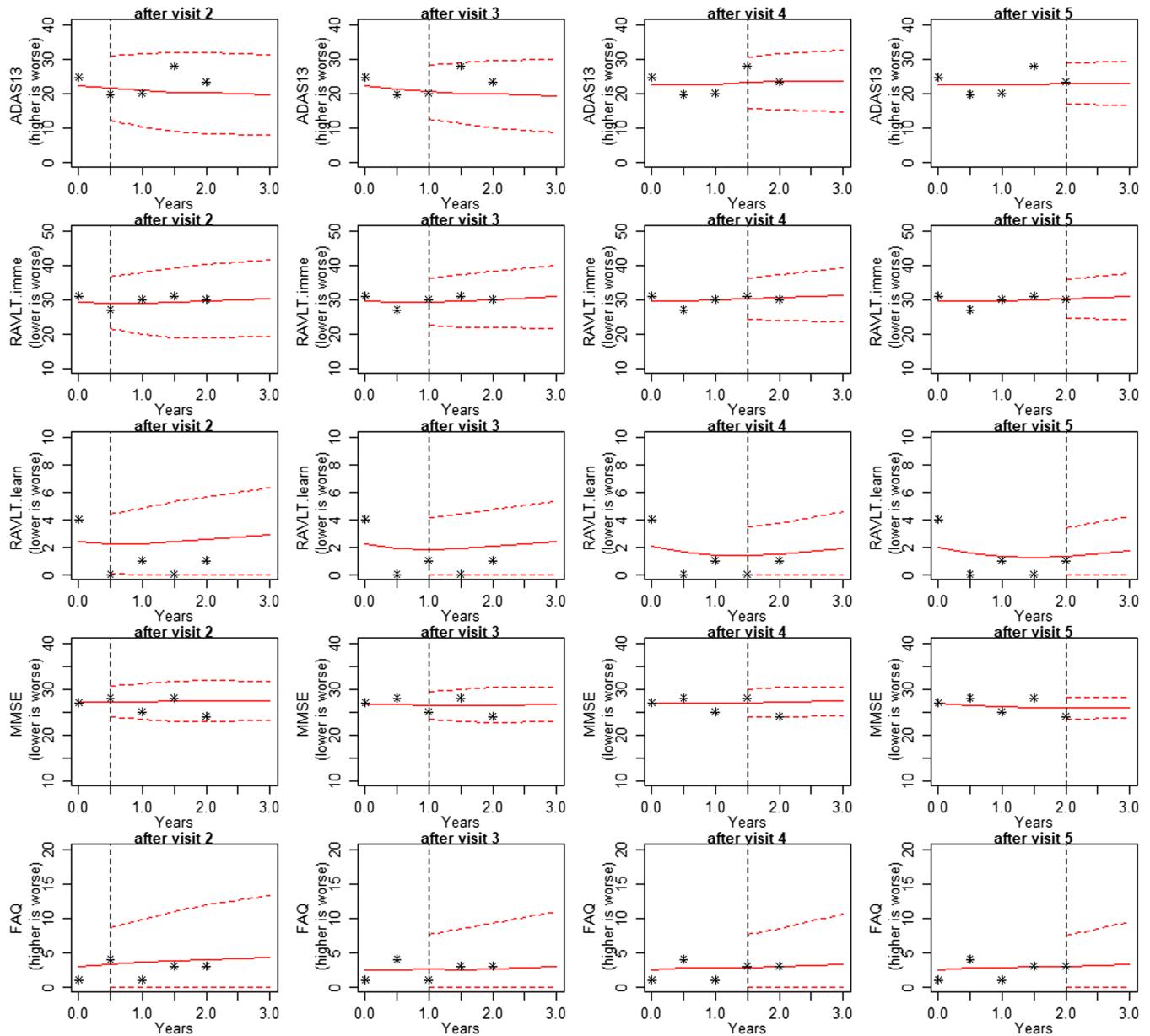


FIGURE 5 Predicted five longitudinal neurocognitive markers for Patient B. In each plot, solid line is smoothed predicted longitudinal trajectory. Dashed lines are smoothed 95% pointwise confidence intervals. The dashed vertical lines represent the landmark time s .
 ADAS-Cog 13, Alzheimer's Disease Assessment Scale-Cognitive 13 items; FAQ, Functional Assessment Questionnaire; MMSE, Mini Mental State Examination; RAVLT, Rey Auditory Verbal Learning Tests [Colour figure can be viewed at wileyonlinelibrary.com]

defined as

$$Y_{iq}(t_{ij}) = X_{iq}(t_{ij}) + \epsilon_{ijq},$$

$$X_{iq}(t_{ij}) = \beta_{0q} + \beta_{1q} \sum_{r=1}^3 c_r(t_{ij} - \kappa_r)_+ + \beta_{1q} x_i + b_{iq},$$

where the spline coefficients $[c_1, c_2, c_3] = [1.2, 0.7, 0.5]$, knots $[\kappa_1, \kappa_2, \kappa_3] = [0, 6, 13]$, and the truncated basis functions of time are $(t - \kappa_r)_+ = t - \kappa_r$ if $t > \kappa_r$ and 0 otherwise.

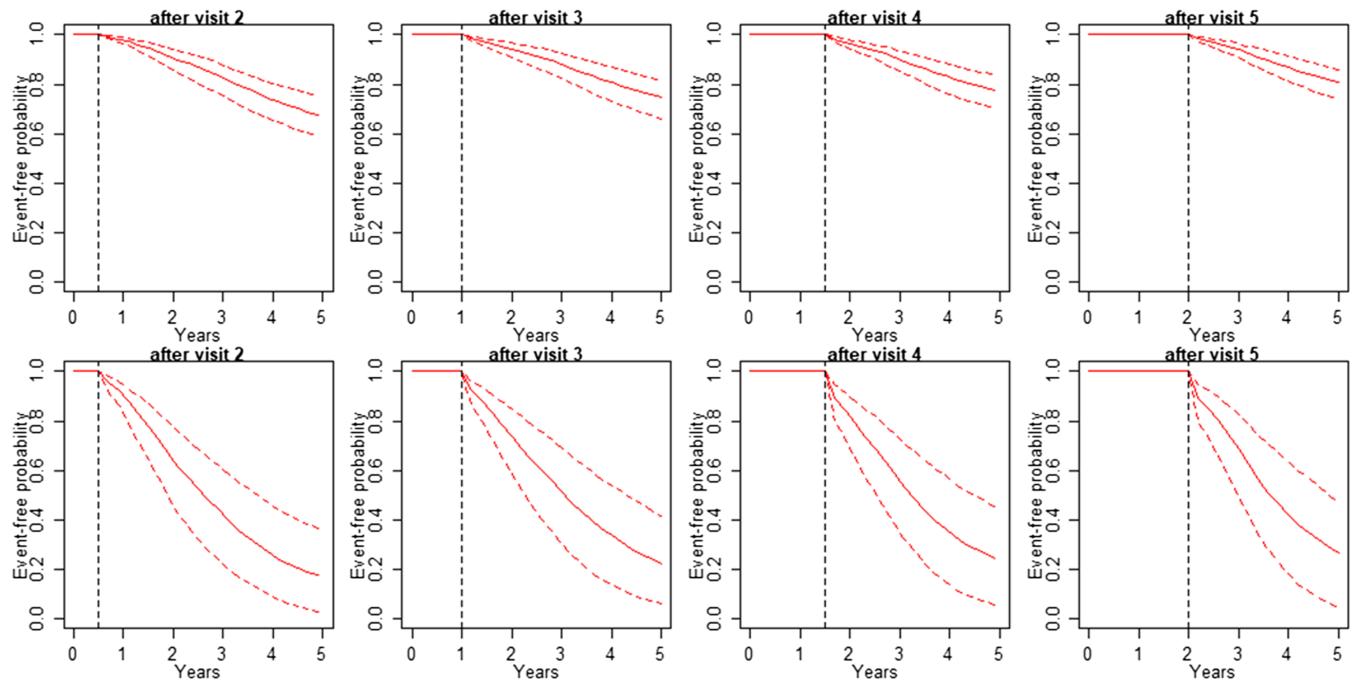


FIGURE 6 Predicted event-free probability with 95% bootstrap confidence intervals for Patient A (upper panels) and Patient B (lower panels). The dashed vertical lines represent the landmark time s [Colour figure can be viewed at wileyonlinelibrary.com]

In all four scenarios, we choose a constant baseline hazard function $h_0(t) = \exp(-7)$ and the survival submodel is

$$h_i(t) = h_0(t) \exp \left\{ z_i \gamma_1 + \sum_{q=1}^3 \alpha_q X_{iq}(t) \right\},$$

where z_i is sampled from Bernoulli distribution with event probability being 0.5, $\gamma_1 = -2.5$, and $[\alpha_1, \alpha_2, \alpha_3] = [0.1, -0.1, 0.2]$. We randomly generate survival probability $S_i(T_i)$ from standard uniform distribution and then solve for event times T_i from the survival function $S_i(T_i) = \exp\{-\int_0^{T_i} h_i(t) dt\}$ using the `uniroot()` and `integrate()` functions in R. Censoring time is independently simulated from a uniform distribution to achieve a censoring rate about 30%. Due to censoring, each subject has four longitudinal measurements, on average. To introduce missing data, we randomly select 10% of the measurements and delete them. This level of missingness is similar to the ADNI-1 dataset.

For each of the 100 simulated datasets of size 300, we randomly select 200 subjects as the training dataset and set aside the remaining 100 subjects for prediction. In each scenario, the model inference is first conducted based on all observations in a training dataset using two frameworks: MJM framework with linear longitudinal submodels and our proposed MFPC Cox framework combining MFPCA and Cox model. After fitting the model, the dynamic predictions are made for subjects $n, n = 1, \dots, 100$, in the validation datasets using methods corresponding to each framework. We use the longitudinal data until the landmark time s to reconstruct the longitudinal process $\hat{X}_{nq}(t)$ and predict the subject-specific survival probability $\hat{\pi}_n(s'|s)$ at a future time point s' . The predicted AUCs are calculated based on the predicted risks of all subjects in the validation dataset. The true conditional survival probabilities $\hat{\pi}_n(s'|s)$ and the true time-dependent AUCs are computed using the prespecified parameter values and the generated random effects. We use the dynamic expected Brier score to assess the bias between the predicted and true risks. The ability to reconstruct the true longitudinal process is assessed by the mean squared error (MSE) as $MSE = \sum_{q=1}^3 \int_0^{T_i} [\hat{X}_{nq}(t) - X_{nq}(t)]^2 du$, where the integral is approximated by summation. To fit the multivariate joint model and make dynamic prediction, we use the latest R package `JMbayes` that is available on `github` through the package's author.¹³ The `JMbayes` adopts the Bayesian approach and speeds up the code via calling functions written in C++. The proposed MFPC Cox method is implemented using available R package `MFPCA`³⁴ for the MFPC analysis and package `survival`³⁵ for survival model.

Table 1 presents the time-dependent AUCs, dynamic expected BS, and MSE by averaging the results from 100 simulated validation datasets. In all scenarios, the predicted AUCs from both MJM and MFPC Cox are only slightly smaller than the true AUCs. It suggests a good prediction performance of the MFPC Cox, compared to the MJM, in terms of discrimination. The small values of BS also indicate a comparable calibration performance of the MFPC Cox in all scenarios. In Scenarios 1 and 3 with linear longitudinal outcomes, the MJM model provides slightly larger AUCs, smaller BSs, and smaller MSEs

as it is the true model. However, in Scenarios 2 and 4 with nonlinear outcomes, although both MJM and MFPCCox have comparable performance in predicting the event probability, MFPCCox outperforms MJM in robustness of reconstructing the longitudinal processes with much smaller MSEs when the MJM misspecifies the parametric form. The simulation results in Scenarios 3 and 4 with missing values suggest that a small percentage of missing in observed data has a limited effect on the prediction performance of both methods. Overall, our proposed method has a robust predictive performance in terms of both calibration and discrimination. Although the proposed MFPCCox, whose model inference is conducted using a two-stage approach, ignores the informative censoring on the longitudinal outcomes due to drop-out process,³⁶ the simulation results show limited effect of informative censoring issue on the prediction performance of the model. The two methods are further compared under scenario with more complex correlation structure of longitudinal outcomes, ie, including the random slopes in the longitudinal submodels. The MJM and the MFPCCox still have comparable predict performance (Web Table 2 in the Supplementary Material). However, with the number of parameters in the covariance matrix Σ increases, the complexity and the computational burden for the model inference of MJM approach also increase. Please refer to the Web Supplement for details.

The proposed MFPCCox method can be implemented using available software with much less computational burden. Using a personal computer (RAM 8G, CPU 3.30 GHz) for the first simulation scenario, it takes about 6.25 hours for fitting a Bayesian MJM using JMBayes, while it takes about 80 seconds for a frequentist-based MFPCCox. As the number of longitudinal outcomes increases, the computational time for MJM would increase exponentially. Additional simulation studies are conducted to assess the capabilities of MJM and MFPCCox to incorporate a large number of longitudinal outcomes relying on existing software. The simulation study is designed to extend the Scenario 1 in this section with total number of longitudinal outcomes Q increases from 4 to 10. Although MJM is theoretically implementable, the JMBayes package fails to fit a MJM model on a simulated dataset including six or more longitudinal outcomes in a reasonable time frame and has an error message reported. However, the proposed MFPCCox method can be easily implemented on the same dataset and provide accurate predictions on a validation dataset (Web Table 3 in the Supplementary Material). The attempt to fit a multivariate joint model (with five longitudinal neurocognitive markers as Section 3) on the ADNI dataset also failed when using JMBayes package. It suggests that MFPCCox is more robust and reliable in terms of implementation. To further demonstrate the scalability of MFPCCox approach, additional simulation studies on the comparison of the computing times for varying sample sizes and number of outcomes are available in the Web Supplement.

5 | DISCUSSION

We have developed a MFPCCox framework to use the trajectory patterns of multiple longitudinal health outcomes to conduct real-time dynamic prediction of future outcome trajectories and risks of target events. The multivariate functional components analysis (MFPCA) effectively deals with measurement errors and captures the temporal correlation in the outcomes as well as correlations among outcomes. It does not require a prespecified parametric functional form for the longitudinal trajectory and it allows missingness and irregularities of the data. In the real data analysis, our results suggest that including the history of multivariate longitudinal neurocognitive markers, in addition to baseline variables, greatly improves predictive performance for AD progression. In the simulation studies, our proposed method has a good dynamic predictive accuracy and discrimination. It is more robust against misspecification of the models for longitudinal outcomes, and thus well-suited for characterizing the features of multiple longitudinal trajectories than the parametric MJM. Moreover, the proposed MFPCCox model is capable to incorporate a large number of longitudinal outcomes. In contrast, the MJM would quickly become intractable as the number of outcomes and the size of their covariance matrix increase. The MFPCCox models can be implemented in available software and are computationally attractive. Such a framework is applicable to many medical conditions with similar multimodal data structure (eg, Parkinson's Progression Markers Initiative Study³⁷ and TRACK-HD study of Huntington's disease³⁸).

To facilitate the personalized dynamic predictions in clinical setting, we develop a web-based calculator available at https://kanli.shinyapps.io/AD_prediction/. A screenshot of the user interface is presented in Web Figure 4. The calculator requires as input the patient's baseline characteristics and his/her longitudinal outcome values up to the present time. The online calculator will then produce time-dependent predictions of future health outcome trajectories and the risk of AD progression, in addition to the 95% confidence intervals. Moreover, additional data generated from more follow-up visits can be input to obtain updated predictions. The calculator is a user friendly and easily accessible tool to provide clinicians with dynamically-updated patient-specific predictions. The practical impact of these dynamic prediction tools can be profound for the neurodegenerative diseases because they provide unique insight and valuable guidance for clinical decision making on patient prognosis, treatment initiation, and counseling to facilitate a targeted treatment.

One limitation of the proposed two-stage approach could be the unignorable informative censoring on the longitudinal model due to drop-out process. Many studies have compared the two-stage modeling and JM approaches in the context of a single longitudinal outcome and suggested that unignorable censoring leads to parameter bias.³⁶ However, the informative censoring has limited effect on prediction performance as shown in our simulation study that mimics the real data. Moreover, there are less than 10% of the MCI population in our analyses that discontinued because of death before progression to AD. Death is treated as a noninformative right-censoring event in our analyses. Cox proportional hazard model may produce biased parameter estimates of the predictor effects when failing to account for the competing risk of death in elder population. The competing risk approaches^{12,39} could be adopted to improve our method. Further improvement of the model could be made by considering AD conversion as an interval-censoring problem, as AD diagnosis is available only at visit times. In addition, the proposed model only applies to continuous multivariate longitudinal outcomes. However, the types of outcomes could also be binary and ordinal. In the future work, we will extend the proposed model to a generalized MFPC and Cox modeling framework to accommodate other types of longitudinal outcomes. The effect of correlation strength of the longitudinal outcomes on MFPCA will also be investigated in the future work.

ACKNOWLEDGEMENTS

The research of Sheng Luo was supported by National Institute of Neurological disorders and Stroke (grant number: R01NS091307) and National Institute of Aging (grant number: R56AG062302). Authors Kan Li and Sheng Luo (for the Alzheimer's Disease Neuroimaging Initiative*) acknowledge the Texas Advanced Computing Center (TACC) for providing high-performing computing resources. Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.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 article. A complete listing of ADNI investigators can be found at: https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

DATA AVAILABILITY STATEMENT

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

ORCID

Kan Li  <https://orcid.org/0000-0002-5762-7440>

Sheng Luo  <https://orcid.org/0000-0003-4214-5809>

REFERENCES

1. Weiner MW, Veitch DP, Aisen PS, et al. The Alzheimer's disease neuroimaging initiative: a review of papers published since its inception. *Alzheimer's Dement.* 2013;9(5):e111-e194.
2. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol.* 1999;56(3):303-308.
3. Robinson L, Tang E, Taylor JP. Dementia: timely diagnosis and early intervention. *BMJ.* 2015;350.
4. Fauchet CL, Thomas DC. Simultaneously modelling censored survival data and repeatedly measured covariates: a Gibbs sampling approach. *Statist Med.* 1996;15(15):1663-1685.
5. Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. *Biometrics.* 1997;53(1):330-339.
6. Li K, Chan W, Doody RS, Quinn J, Luo S. Prediction of conversion to Alzheimer's disease with longitudinal measures and time-to-event data. *J Alzheimer's Dis.* 2017;58(2):361-371.
7. Li K, Luo S. Functional joint model for longitudinal and time-to-event data: an application to Alzheimer's disease. *Statist Med.* 2017;36(22):3560-3572.
8. Li K, Luo S. Dynamic predictions in Bayesian functional joint models for longitudinal and time-to-event data: an application to Alzheimer's disease. *Stat Methods Med Res.* 2019;28(2):327-342.
9. Rizopoulos D, Ghosh P. A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. *Statist Med.* 2011;30(12):1366-1380.
10. Tang NS, Tang AM, Pan DD. Semiparametric Bayesian joint models of multivariate longitudinal and survival data. *Comput Stat Data Anal.* 2014;77(Supplement C):113-129.

11. He B, Luo S. Joint modeling of multivariate longitudinal measurements and survival data with applications to Parkinson's disease. *Stat Methods Med Res.* 2016;25(4):1346-1358.
12. Proust-Lima C, Dartigues JF, Jacqmin-Gadda H. Joint modeling of repeated multivariate cognitive measures and competing risks of dementia and death: a latent process and latent class approach. *Statist Med.* 2016;35(3):382-398.
13. Rizopoulos D. The R package JMbayes for fitting joint models for longitudinal and time-to-event data using MCMC. *J Stat Softw.* 2016;72(7):1-46.
14. Müller H-G. Functional modelling and classification of longitudinal data. *Scand J Stat.* 2005;32(2):223-240.
15. Yao F. Functional principal component analysis for longitudinal and survival data. *Statistica Sinica.* 2007;17(3):965-983.
16. Ding J, Wang JL. Modeling longitudinal data with nonparametric multiplicative random effects jointly with survival data. *Biometrics.* 2008;64(2):546-556.
17. Holte SE, Randolph TW, Ding J, et al. Efficient use of longitudinal CD4 counts and viral load measures in survival analysis. *Statist Med.* 2012;31(19):2086-2097.
18. Ye J, Li Y, Guan Y. Joint modeling of longitudinal drug using pattern and time to first relapse in cocaine dependence treatment data. *Ann Appl Stat.* 2015;9(3):1621-1642.
19. Fang H-B, Wu TT, Rapoport AP, Tan M. Survival analysis with functional covariates for partial follow-up studies. *Stat Methods Med Res.* 2016;25(6):2405-2419.
20. Yan F, Lin X, Huang X. Dynamic prediction of disease progression for leukemia patients by functional principal component analysis of longitudinal expression levels of an oncogene. *Ann Appl Stat.* 2017;11(3):1649-1670.
21. Yan F, Lin X, Li R, Huang X. Functional principal components analysis on moving time windows of longitudinal data: dynamic prediction of times to event. *J R Stat Soc Ser C Appl Stat.* 2018;67(4):961-978.
22. Happ C, Greven S. Multivariate functional principal component analysis for data observed on different (dimensional) domains. *J Am Stat Assoc.* 2018;113(522):649-659.
23. Yao F, Müller H-G, Wang J-L. Functional data analysis for sparse longitudinal data. *J Am Stat Assoc.* 2005;100(470):577-590.
24. Hung H, Chiang C. Estimation methods for time-dependent AUC models with survival data. *Can J Stat.* 2009;38(1):8-26.
25. Blanche P, Kattan MW, Gerdts TA. The c-index is not proper for the evaluation of t-year predicted risks. *Biostatistics.* 2018;20(2):347-357.
26. Sène M, Taylor JMG, Dignam JJ, Jacqmin-Gadda H, Proust-Lima C. Individualized dynamic prediction of prostate cancer recurrence with and without the initiation of a second treatment: development and validation. *Stat Methods Med Res.* 2016;25(6):2972-2991.
27. Blanche P, Proust-Lima C, Loubère L, Berr C, Dartigues JF, Jacqmin-Gadda H. Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and competing risks. *Biometrics.* 2015;71(1):102-113.
28. Goldsmith J, Greven S, Crainiceanu C. Corrected confidence bands for functional data using principal components. *Biometrics.* 2013;69(1):41-51.
29. Weiner MW, Veitch DP, Aisen PS, et al. 2014 update of the Alzheimer's disease neuroimaging initiative: a review of papers published since its inception. *Alzheimer's Dement.* 2015;11(6):e1-e120.
30. Weiner MW, Veitch DP, Aisen PS, et al. Recent publications from the Alzheimer's disease neuroimaging initiative: reviewing progress toward improved AD clinical trials. *Alzheimer's Dement.* 2017;13(4):e1-e85.
31. Risacher SL, Saykin AJ, Wes JD, Shen L, Firpi HA, McDonald BC. Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. *Curr Alzheimer Res.* 2009;6(4):347-361.
32. Cui Y, Liu B, Luo S, et al. Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. *PLOS ONE.* 2011;6(7):e21896.
33. Li S, Okonkwo O, Albert M, Wang MC. Variation in variables that predict progression from MCI to AD dementia over duration of follow-up. *Am J Alzheimer's Dis.* 2013;2(1):12-28.
34. Happ C. MFPCA: multivariate functional principal component analysis for data observed on different dimensional domains. R package version 1.3-1. 2018. <https://github.com/ClaraHapp/MFPCA>
35. Therneau TM. A package for survival analysis in S. R package version 2.37-7. 2014.
36. Wu L, Liu W, Yi GY, Huang Y. Analysis of longitudinal and survival data: joint modeling, inference methods, and issues. *J Probab Stat.* 2012;2012.
37. Marek K, Jennings D, Lasch S, et al. The Parkinson progression marker initiative (PPMI). *Prog Neurobiol.* 2011;95(4):629-635.
38. Tabrizi SJ, Reilmann R, Roos RAC, et al. Potential endpoints for clinical trials in premanifest and early Huntington's disease in the TRACK-HD study: analysis of 24 month observational data. *Lancet Neurol.* 2012;11(1):42-53.
39. Elashoff RM, Li G, Li N. A joint model for longitudinal measurements and survival data in the presence of multiple failure types. *Biometrics.* 2008;64(3):762-771.

SUPPORTING INFORMATION

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

How to cite this article: Li K, Luo S. Dynamic prediction of Alzheimer's disease progression using features of multiple longitudinal outcomes and time-to-event data. *Statistics in Medicine.* 2019;38:4804-4818.
<https://doi.org/10.1002/sim.8334>