

# Functional Joint Models for Longitudinal and Time-to-Event Data

Kan Li

Department of Biostatistics and Data Science  
Univ. of Texas Health Science Center - Houston

Committee members:

Dr. Hulin Wu (Advisor/Committee Chair)  
Dr. Sheng Luo (Dissertation Supervisor)  
Dr. David Lairson (Minor Advisor)  
Dr. Momiao Xiong (Breath Advisor)

February 26, 2018

# Outline

## 1 Introduction

- FDA and JM
- Motivation
- Data Source
- Research Aims

## 2 Dissertation Aim 1

- Methods
- Application

## 3 Dissertation Aim 2

- Methods
- Dynamic Prediction
- Application

## 4 Dissertation Aim 3

- Methods
- Application

## 5 Conclusion and Future Work

## 6 Acknowledgment

# Outline

## 1 Introduction

- FDA and JM
- Motivation
- Data Source
- Research Aims

## 2 Dissertation Aim 1

- Methods
- Application

## 3 Dissertation Aim 2

- Methods
- Dynamic Prediction
- Application

## 4 Dissertation Aim 3

- Methods
- Application

## 5 Conclusion and Future Work

## 6 Acknowledgment

# Functional Data

- Functional data are the type of data whose units of observation are functions defined on certain continuous domains but sampled on discrete grids in practice.
- Current functional data analysis (FDA) methods mainly focus on the statistical inference instead of prediction.

# Joint Models for Longitudinal and Time-to-Event Data

- Joint model is a popular framework to analyze datasets including repeated measurements and time-to-event outcomes appropriately.
- A novel use of joint model is making **personalized dynamic prediction**.
- Current state-of-the-art joint models do not incorporate functional data.

# Outline

## 1 Introduction

- FDA and JM
- Motivation
- Data Source
- Research Aims

## 2 Dissertation Aim 1

- Methods
- Application

## 3 Dissertation Aim 2

- Methods
- Dynamic Prediction
- Application

## 4 Dissertation Aim 3

- Methods
- Application

## 5 Conclusion and Future Work

## 6 Acknowledgment

## Alzheimer's Disease as Motivating Example

- Alzheimer's disease (AD) is a neurodegenerative disorder of the brain and is the most common form of dementia.
- The number of Americans with AD will reach 7.7 million by 2030 and the corresponding total cost of care for AD will increase to \$1.08 trillion each year.
- There is no effective disease-modifying treatments for AD currently.
- Urgency to discover and assess markers for early prognosis and prediction of the disease.

## Alzheimer's Disease as Motivating Example (cont'd)

- In the studies of AD, researchers collect repeated measurements of neurocognitive assessments (scalar data), neuroimaging (functional data), event histories (time-to-event data), and other clinical and genetic information to better understand the disease.
- Mild cognitive impairment (MCI) is a transitional stage between normal cognition and Alzheimer's disease (AD).
- Real-world problem: investigating the effects of combining various type of longitudinal markers for predicting the AD conversion within MCI patients.

# Outline

## 1 Introduction

- FDA and JM
- Motivation
- Data Source**
- Research Aims

## 2 Dissertation Aim 1

- Methods
- Application

## 3 Dissertation Aim 2

- Methods
- Dynamic Prediction
- Application

## 4 Dissertation Aim 3

- Methods
- Application

## 5 Conclusion and Future Work

## 6 Acknowledgment

## Data Source: ADNI Study

- Alzheimer's Disease Neuroimaging Initiative (ADNI) study has large samples, long follow-up period, breadth of cognitive markers and imaging (e.g., MRI), and prospective nature.
- Focus on 355 MCI patients who started from ADNI-1 and were reassessed at 6, 12, 18, 24, 36 months, and then followed annually as part of ADNI-2.
- 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.

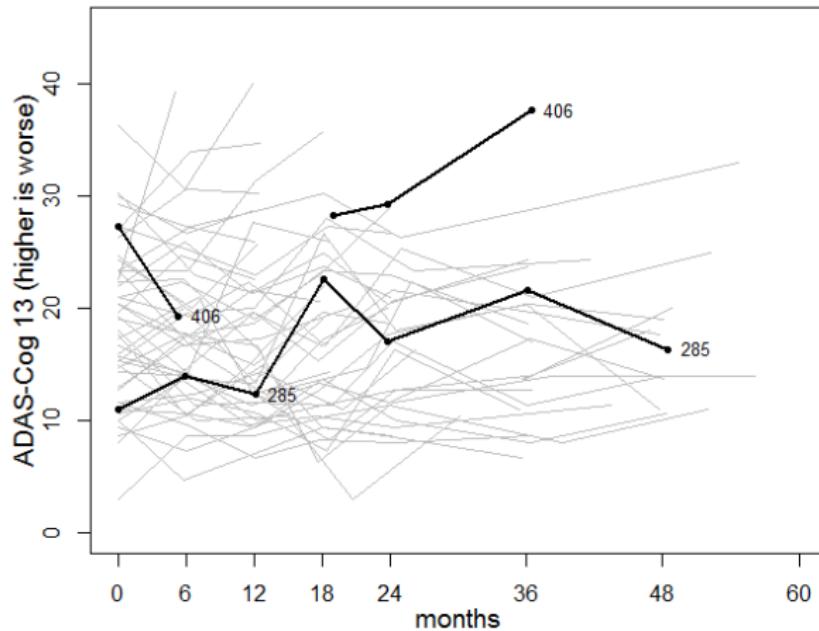
## Data Source: Longitudinal Markers

- Longitudinal **AD Assessment Scale-Cognitive** (ADAS-Cog) score and **Hippocampal volume** (HV) are the strongest predictors of AD conversion from MCI in neurocognitive and neuroimaging domain.<sup>1</sup>
- Enormous information lost occurs when the high dimensional image data are reduced to a single volume.
- Surface-based morphology analysis retains more information about Hippocampus atrophy.
  - ▶ **Hippocampal radial distance** (HRD): the distance from the medial core of the hippocampus to points on the surface and quantifies the thickness of hippocampus relative to its center line.

---

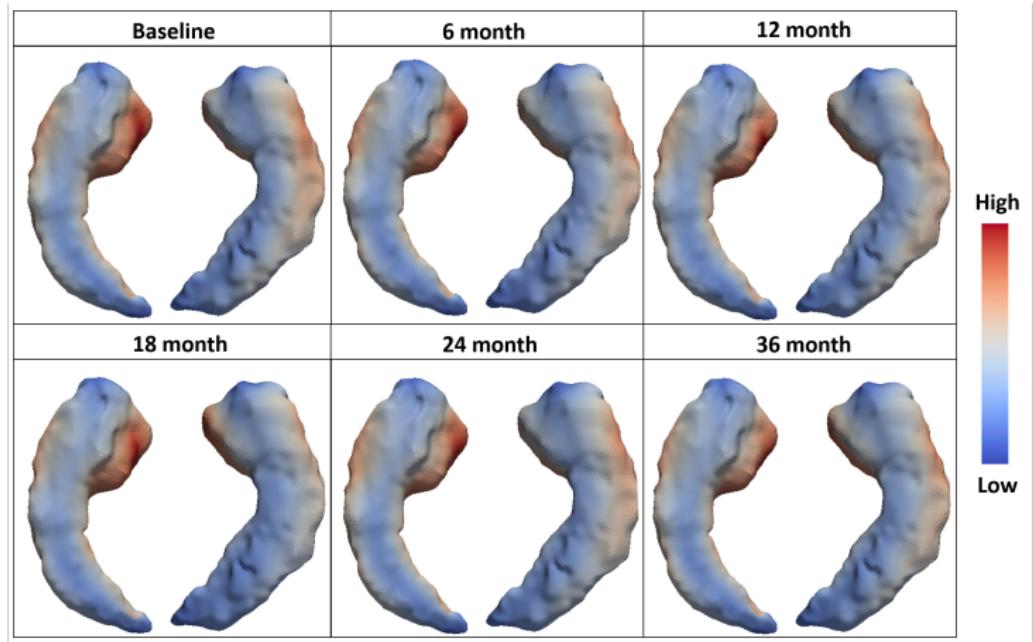
<sup>1</sup> Li K, Chan W, Doody RS, Quinn J, Luo S. (2017). Prediction of conversion to Alzheimer's disease with longitudinal measures and time-to-event data, *Journal of Alzheimer's Disease*, 58(2):361-71.

# Longitudinal ADAS-Cog



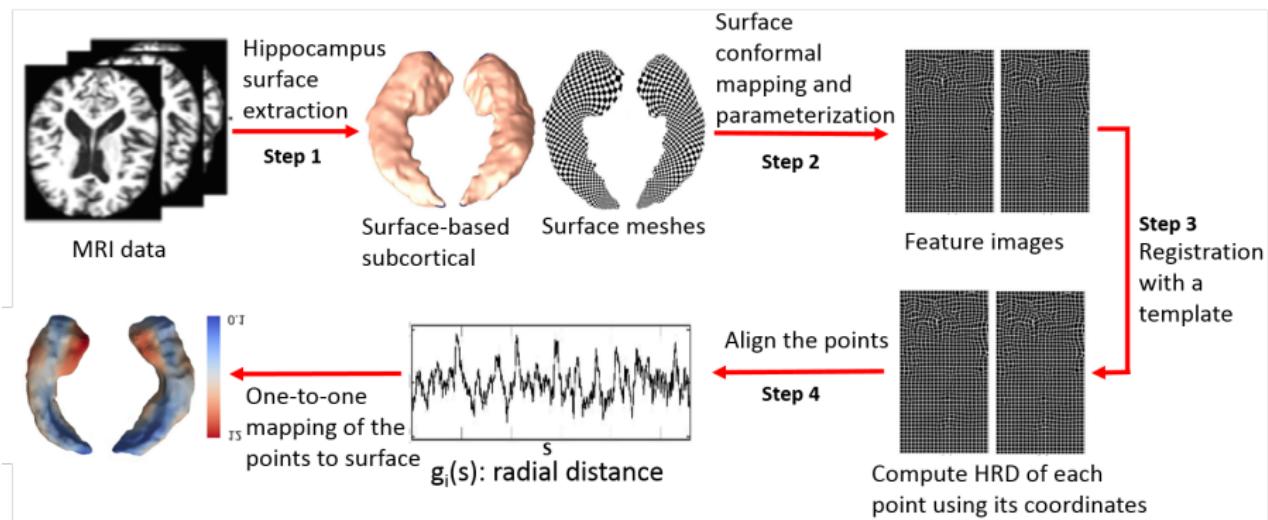
**Figure:** Longitudinal trajectories of ADAS-Cog 13: 50 randomly selected MCI patients (with two patients highlighted) from the ADNI study.

# Longitudinal HRD



**Figure:** The longitudinal profile of surface-based hippocampal images of one MCI patient: hippocampal radial distances are denoted by colors.

# Hippocampus Image Processing



# Outline

## 1 Introduction

- FDA and JM
- Motivation
- Data Source
- Research Aims

## 2 Dissertation Aim 1

- Methods
- Application

## 3 Dissertation Aim 2

- Methods
- Dynamic Prediction
- Application

## 4 Dissertation Aim 3

- Methods
- Application

## 5 Conclusion and Future Work

## 6 Acknowledgment

## Research Aims

Propose a series of novel functional joint models (FJM) to incorporate both scalar outcomes and functional outcomes in the framework of joint modeling of longitudinal and survival data.

- Aim 1: Develop a FJM that accounts for time-invariant functional outcomes as predictors in the joint model.
- Aim 2: Extend the FJM to account for a longitudinal functional outcome in a multivariate joint modeling framework.
- Aim 3: Develop a novel framework for the use of multiple longitudinal scalar outcomes and longitudinal high-dimensional functional outcome in dynamic prediction.

# Outline

## 1 Introduction

- FDA and JM
- Motivation
- Data Source
- Research Aims

## 2 Dissertation Aim 1

- Methods
- Application

## 3 Dissertation Aim 2

- Methods
- Dynamic Prediction
- Application

## 4 Dissertation Aim 3

- Methods
- Application

## 5 Conclusion and Future Work

## 6 Acknowledgment

# FJM with Time-Invariant Functional Data

- A functional joint model (FJM) that integrates time-invariant functional predictors, longitudinal scalar outcome, and time to event.<sup>2, 3</sup>
- For each subject  $i$  ( $i = 1, \dots, I$ ) at visit  $j$  ( $j = 1, \dots, J_i$ ) and on a 1D domain  $s \in [0, S_{max}] = \mathbf{S}$ , we observe data  $\{Y_i(t_{ij}), X_i(s), \mathbf{w}_i, \mathbf{g}_i\}$ ,

$$Y_i(t_{ij}) = m_i(t_{ij}) + \varepsilon_{ij}$$

$$m_i(t_{ij}) = \mathbf{w}_i^\top \boldsymbol{\beta} + \int_{\mathbf{S}} X_i(s) B^{(x)}(s) ds + \mathbf{z}_i^\top \mathbf{u}_i$$

$$h_i(t) = h_0(t) \exp\{\mathbf{g}_i^\top \boldsymbol{\gamma} + \int_{\mathbf{S}} X_i(s) B^{(z)}(s) ds + \alpha m_i(t)\}.$$

- Coefficient functions  $B(s)$  quantify the point-wise association between  $X_i(s)$  and  $Y_i$  or the hazard.

---

<sup>2</sup>Li K, Luo S. (2017). Functional joint model for longitudinal and time-to-event data: an application to Alzheimer's disease, *Statistics in Medicine*, 36(22), 3560-72.

<sup>3</sup>Li K, Luo S. (2017). Dynamic predictions in Bayesian functional joint models for longitudinal and time-to-event data: an application to Alzheimer's disease, *Statistical Methods in Medical Research*, published online on July 28, 2017.

# Methods

- Perform functional principle components analysis (FPCA) on  $X_i(s)$  such that  $X_i(s) \approx \mu(s) + \sum_{l=1}^{K_x} \xi_{il} \phi_l(s) = \mu(s) + \boldsymbol{\xi}_i^\top \boldsymbol{\phi}(s)$ .
  - ▶ Covariance function  $\Sigma(s, s') = \sum_{l=1}^{\infty} \lambda_l \phi_l(s) \phi_l(s')$ , where  $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$  are non-increasing eigenvalues and  $\phi_l(s)$ 's are the corresponding orthonormal eigenfunctions.
  - ▶  $\xi_{il} = \int_S \{X_i(s) - \mu(s)\} \phi_l(s) ds$ , functional principal component (FPC) score;
  - ▶  $K_x$  is truncation number, can be determined by the proportion of variance explained.
- Express coefficient function  $B(s)$  in term of cubic B-spline basis functions (known)  $\psi(s) = [\psi_1(s), \dots, \psi_{K_B}(s)]^\top$ , so that  $B(s) = \sum_{l=1}^{K_B} B_l \psi_l(s) = \boldsymbol{B}^\top \boldsymbol{\psi}(s)$ .

## Methods (cont'd)

- The FJM is rewritten as

$$Y_i(t_{ij}) = m_i(t_{ij}) + \varepsilon_{ij}$$

$$m_i(t_{ij}) = \mathbf{w}_i^\top \boldsymbol{\beta}' + \boldsymbol{\xi}_i^\top \mathbf{J}_{\phi,\psi} \mathbf{B}^{(x)} + \mathbf{z}_i^\top \mathbf{u}_i$$

$$h_i(t) = h_0'(t) \exp\{\mathbf{g}_i^\top \boldsymbol{\gamma} + \boldsymbol{\xi}_i^\top \mathbf{J}_{\phi,\psi} \mathbf{B}^{(z)} + \alpha m_i(t)\},$$

where  $\mathbf{J}_{\phi,\psi}$  is a  $K_x \times K_B$  matrix with the  $(k,l)$ th entry equal to  $\int_{\mathcal{S}} \phi_k(s) \psi_l(s)^\top ds$ .

- Estimate the spline coefficients  $\mathbf{B}^{(x)}$  and  $\mathbf{B}^{(z)}$  as well as other parameters via Bayesian approach.

## Estimation and inference

- Use vague prior distributions on all unknown parameters in the model.
- We choose  $K_B$  sufficient large (e.g., 10), and impose smoothness on coefficient function estimates through the prior specification,

$$\boldsymbol{B} \sim MVN(\mathbf{0}, \sigma_B^2 \boldsymbol{P}^{-1})$$

where  $\boldsymbol{P}$  is a pre-specified  $K_B \times K_B$  combined zeroth- and second-order derivative penalty matrices (Ruppert *et al*, 2003; Goldsmith *et al*, 2015).

- The estimated coefficient function is constructed as  $\hat{B}(s) = \hat{\boldsymbol{B}}^\top \boldsymbol{\psi}(s)$ .

# Outline

## 1 Introduction

- FDA and JM
- Motivation
- Data Source
- Research Aims

## 2 Dissertation Aim 1

- Methods
- Application

## 3 Dissertation Aim 2

- Methods
- Dynamic Prediction
- Application

## 4 Dissertation Aim 3

- Methods
- Application

## 5 Conclusion and Future Work

## 6 Acknowledgment

# Application to the ADNI Study

- Model structure

- ▶ Time from first visit to AD diagnosis as survival outcome;
- ▶ ADAS-Cog 11 as the longitudinal outcome;
- ▶ The baseline hippocampal radial distance ( $HRD$ ) as the functional predictor.
- ▶ Baseline hippocampal volume, age, gender, years of education and presence of the apolipoprotein E ( $APOE$ )  $\varepsilon 4$  allele as scalar covariates.

$$ADAS-Cog_i(t_{ij}) = m_i(t_{ij}) + \varepsilon_{ij}$$

$$m_i(t_{ij}) = \beta_0 + \beta_1 t_{ij} + \beta_2 bAge_i + \beta_3 bHV_i + \int_s HRD_i(s)B^{(x)}(s)ds + u_{i1}$$

$$\begin{aligned} h_i(t) &= h_0(t) \exp\{\gamma_1 gender_i + \gamma_2 bAge_i + \gamma_3 Edu_i + \gamma_4 APOE-\varepsilon 4 \\ &\quad + \gamma_5 bHV_i + \int_s HRD_i(s)B^{(w)}(s)ds + \alpha m_i(t)\}. \end{aligned}$$

# Parameter Estimation

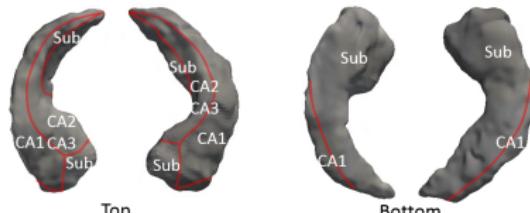
- Parameter estimates from model  $FJM$  with HRD in both longitudinal and survival submodels.

	Parameters	Mean	SE	2.5%	97.5%
For longitudinal outcome					
ADAS-Cog 11	Time (Years)	0.428	0.045	0.338	0.521
	$bAge$	-0.364	0.260	-0.885	0.156
	$bHV (mm^3)$	-1.617	0.295	-2.201	-1.051
For survival process					
MCI to AD	Female	-0.088	0.173	-0.397	0.270
	$bAge$	-0.283	0.042	-0.423	-0.109
	$Edu$ (years)	0.028	0.016	-0.002	0.062
	$APOE-\varepsilon 4$	0.533	0.125	0.239	0.728
	$bHV (mm^3)$	0.056	0.114	-0.185	0.276
	$\alpha$	0.134	0.022	0.079	0.177

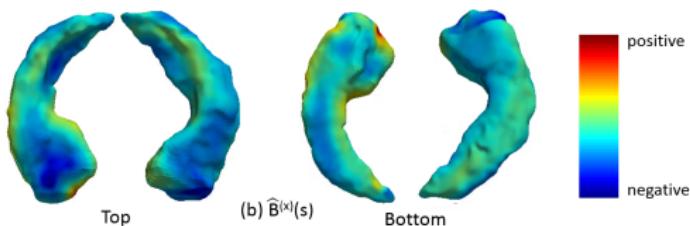
Table: ADNI data analysis results from model  $FJM$ .

# Parameter Estimation (cont'd)

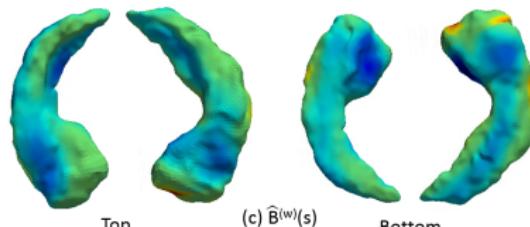
- Estimated coefficient functions for *HRD* in the submodels are mapped back to the hippocampal surfaces.



(a) Hippocampal subfields



(b)  $\hat{B}^{(x)}(s)$



(c)  $\hat{B}^{(w)}(s)$

# Outline

## 1 Introduction

- FDA and JM
- Motivation
- Data Source
- Research Aims

## 2 Dissertation Aim 1

- Methods
- Application

## 3 Dissertation Aim 2

- Methods
- Dynamic Prediction
- Application

## 4 Dissertation Aim 3

- Methods
- Application

## 5 Conclusion and Future Work

## 6 Acknowledgment

## FJM with longitudinal functional data (MFJM)

- A multivariate functional joint model that integrates longitudinal functional outcome, longitudinal scalar outcome, and survival outcome.
- For each subject  $i$  ( $i = 1, \dots, I$ ) at visit  $j$  ( $j = 1, \dots, J_i$ ) and on a 1D domain  $s \in [0, S_{max}] = \mathcal{S}$ , we observe data  $\{y_i^*(t_{ij}), y_i(s, t_{ij}), \mathbf{x}_{ij}\}$

$$\begin{aligned} y_i^*(t_{ij}) &= m_i^*(t_{ij}) + \epsilon_{ij}^*, \\ m_i^*(t_{ij}) &= \beta_0 + t_{ij}\beta_t + \sum_k^p x_{ijk}\beta_k + \sum_{r=1}^R \zeta_r(t - \kappa_r)_+ + b_i^*, \\ y_i(s, t_{ij}) &= m_i(s, t_{ij}) + \epsilon_{ij}(s), \\ \mathbf{m}_i(\mathbf{s}, \mathbf{t}_{ij}) &= B_0(s) + t_{ij}B_t(s) + \sum_k^p x_{ijk}B_k(s) + \mathbf{b}_i(\mathbf{s}), \\ h_i(t) &= h_0(t) \exp\{\mathbf{w}_i^\top \boldsymbol{\gamma} + \alpha^* m_i^*(t) + \int_s h_i(s) \mathbf{m}_i(\mathbf{s}, \mathbf{t}) ds\}, \end{aligned}$$

# Methods

- Similar to the FPCA approach, a truncated approximation for  $b_i(s)$  is given by  $b_i(s) \approx \sum_{k=1}^{K^{(b)}} \xi_{ik}^{(b)} \phi_k^{(b)}(s)$ . Thus,

$$y_i(s, t_{ij}) \approx B_0(s) + t_{ij} B_t(s) + \sum_k^p x_{ijk} B_k(s) + \sum_{k=1}^{K^{(b)}} \xi_{ik}^{(b)} \phi_k^{(b)}(s) + \epsilon_{ij}(s)$$

- Assume the correlation between  $y_i^*(t_{ij})$  and  $y_i(s, t_{ij})$  is manifested by the correlation between  $b_i^*$  and  $\xi_i^{(b)} = [\xi_{i1}^{(b)}, \dots, \xi_{iK^{(b)}}^{(b)}]$ , and  $\mathbf{b}_i = [b_i^*, \xi_i^{(b)}] \sim MVN(\mathbf{0}, \boldsymbol{\Sigma})$ , where

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{b^*}^2, & \rho \sigma_{b^*} \sqrt{\lambda_1^{(b)}}, & \dots, & 0 \\ \rho \sigma_{b^*} \sqrt{\lambda_1^{(b)}}, & \lambda_1^{(b)}, & \dots, & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0, & 0, & \dots, & \lambda_{K^{(b)}}^{(b)} \end{bmatrix}.$$

## Methods (cont'd)

- Express the fixed coefficient functions  $B_0(s)$ ,  $B_t(s)$ ,  $B_k(s)$ 's, and FPC eigenfunctions  $\phi^{(b)}(s)$ 's in term of cubic B-spline basis functions (known)  
 $\psi(s) = [\psi_1(s), \dots, \psi_{K_B}(s)]^\top$ ,
- The functional longitudinal sub-model is rewritten as

$$y_i(s, t_{ij}) \approx m_i(s, t_{ij}) + \epsilon_{ij}(s), \text{ where}$$

$$m_i(s, t_{ij}) \approx \mathbf{B}_0^\top + t_{ij} \mathbf{B}_t^\top \psi(s)^\top + \mathbf{x}_{ij}^\top \mathbf{B}_x^\top \psi(s)^\top + \boldsymbol{\xi}_i^{(b)} \mathbf{B}_{\phi^{(b)}}^\top \psi(s)^\top.$$

- Along with scalar longitudinal submodel and proportional hazard model as:

$$y_i^*(t_{ij}) = m_i^*(t_{ij}) + \epsilon_{ij}^*,$$

$$m_i^*(t_{ij}) = \beta_0 + t_{ij} \beta_t + \sum_k^p x_{ijk} \beta_k + \sum_{r=1}^R \zeta_r (t - \kappa_r)_+ + b_i^*,$$

$$h_i(t) = h_0(t) \exp\{ \mathbf{w}_i^\top \gamma + \alpha^* m_i^*(t) + \int_s \alpha(s) m_i(s, t) ds \}$$

# Outline

## 1 Introduction

- FDA and JM
- Motivation
- Data Source
- Research Aims

## 2 Dissertation Aim 1

- Methods
- Application

## 3 Dissertation Aim 2

- Methods
- **Dynamic Prediction**
- Application

## 4 Dissertation Aim 3

- Methods
- Application

## 5 Conclusion and Future Work

## 6 Acknowledgment

# Dynamic risk prediction

- Given a new subject  $N$ 's outcome histories  
 $\mathbf{y}_N^{\{t\}} = \{y_N^*(t_{Nj}), y_N(s, t_{Nj}); 0 \leq t_{Nj} \leq t\}$  and covariates  
 $\mathbf{X}_N^{\{t\}} = \{\mathbf{x}_N(t_{Nj}), \mathbf{w}_N; 0 \leq t_{Nj} \leq t\}$  up to time  $t$ , and  $\delta_N = 0$  (no event).
- Predict the conditional probability of event-free at time  $t' > t$  (e.g.,  $t' = t + \Delta t$ ) , denoted by  $\pi_N(t'|t) = P(T_N^* \geq t' | T_N^* > t, \mathbf{y}_N^{\{t\}}, \mathbf{X}_N^{\{t\}})$ .
- Obtain samples for subject  $N$ 's random effect vector  $\mathbf{b}_N = [b_N^*, \xi_N^{(b)}]$  from its posterior distribution

$$p(\mathbf{b}_N | T_N^* > t, \mathbf{y}_N, \boldsymbol{\theta}) \propto p(\mathbf{y}_N^{\{t\}}, T_N^* > t, \mathbf{b}_N | \boldsymbol{\theta}^{(d)}) \\ = p(\mathbf{y}_N^{\{t\}} | \mathbf{b}_N, \boldsymbol{\theta}^{(d)}) p(T_N^* > t | \mathbf{b}_N, \boldsymbol{\theta}^{(d)}) p(\mathbf{b}_N | \boldsymbol{\theta}^{(d)}).$$

## Dynamic risk prediction (cont'd)

- The conditional probability of event-free at time  $t'$  is
$$\begin{aligned}\hat{\pi}_N(t'|t) &= \int P(T_N^* \geq t' | T_N^* > t, \mathbf{y}_N^{\{t\}}, \mathbf{X}_N^{\{t\}}, \mathbf{b}_N) p(\mathbf{b}_N | T_N^* > t, \mathbf{y}_N^{\{t\}}, \mathbf{X}_N^{\{t\}}) d\mathbf{b}_N \\ &\approx \frac{1}{D} \sum_{d=1}^D P(T_N^* \geq t' | T_N^* > t, \mathbf{y}_N^{\{t\}}, \mathbf{X}_N^{\{t\}}, \mathbf{b}_N^{(d)}) \\ &= \frac{1}{D} \sum_{d=1}^D \frac{P(T_N^* \geq t' | T_N^* > t, \mathbf{y}_N^{\{t\}}, \mathbf{X}_N^{\{t\}}, \mathbf{b}_N^{(d)})}{P(T_N^* \geq t | T_N^* > t, \mathbf{y}_N^{\{t\}}, \mathbf{X}_N^{\{t\}}, \mathbf{b}_N^{(d)})}\end{aligned}$$
- The performance of the prediction is measured by the time-dependent receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC).

# Outline

## 1 Introduction

- FDA and JM
- Motivation
- Data Source
- Research Aims

## 2 Dissertation Aim 1

- Methods
- Application

## 3 Dissertation Aim 2

- Methods
- Dynamic Prediction
- Application

## 4 Dissertation Aim 3

- Methods
- Application

## 5 Conclusion and Future Work

## 6 Acknowledgment

# Application to the ADNI Study

- Model structure
  - ▶ Time from first visit to AD diagnosis as survival outcome;
  - ▶ ADAS-Cog 11 as one longitudinal predictor for survival;
  - ▶ Brain imaging information, hippocampal volume ( $bHV$ ) as a scalar predictor, the hippocampal radial distance ( $HRD$ ) as a functional predictor.
  - ▶ Age, gender, years of education and presence of the apolipoprotein E ( $APOE$ )  $\varepsilon 4$  allele as scalar covariates.
- *JM*: regular joint model incorporates **baseline  $bHV$**  as a scalar predictor and longitudinal ADAS-Cog in survival sub-model.
- *FJM*: incorporates **baseline  $HRD$**  as a functional predictor and longitudinal ADAS-Cog in survival sub-model.
- *MFJM*: incorporates **longitudinal  $HRD$**  as a functional predictor and longitudinal ADAS-Cog in survival sub-model.

# HRD Processing



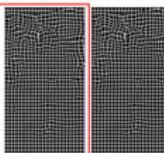
Surface-based  
subcortical



Surface meshes

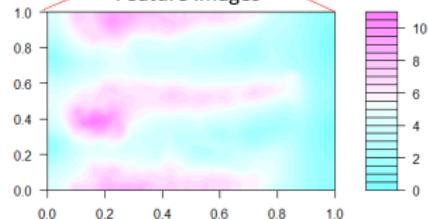
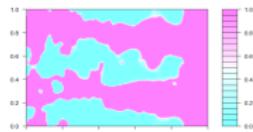
Registration  
with a  
template

Compute HRD of each  
point using its  
coordinates



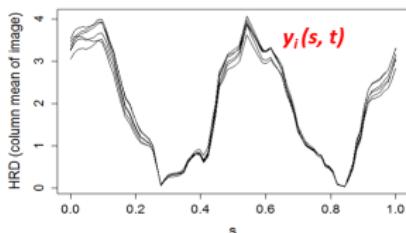
Feature images

Retain the points that have  
significant changes over time



Step 4  
Column mean of  
HRD image matrix  
with selected points

longitudinal HRD of one MCI patient



## Predictive Performance

- Compare the three candidate models by assessing their predictive performance, manifested by the time-dependent AUCs, at different time points over the follow-up period.

$\Delta t$	$t$	JM	FJM	MFJM
6m	12m	0.715	0.754	0.821
	18m	0.691	0.738	0.734
	24m	0.781	0.809	0.812
12m	12m	0.696	0.747	0.792
	18m	0.735	0.776	0.777
	24m	0.749	0.769	0.766

Table: Areas under the ROC curve (AUC) by three candidate models in the ADNI study (based on 10-fold cross validation)

- Including hippocampal imaging information as functional predictor *HRD* in the dynamic prediction framework could improve the capability of risk prediction.

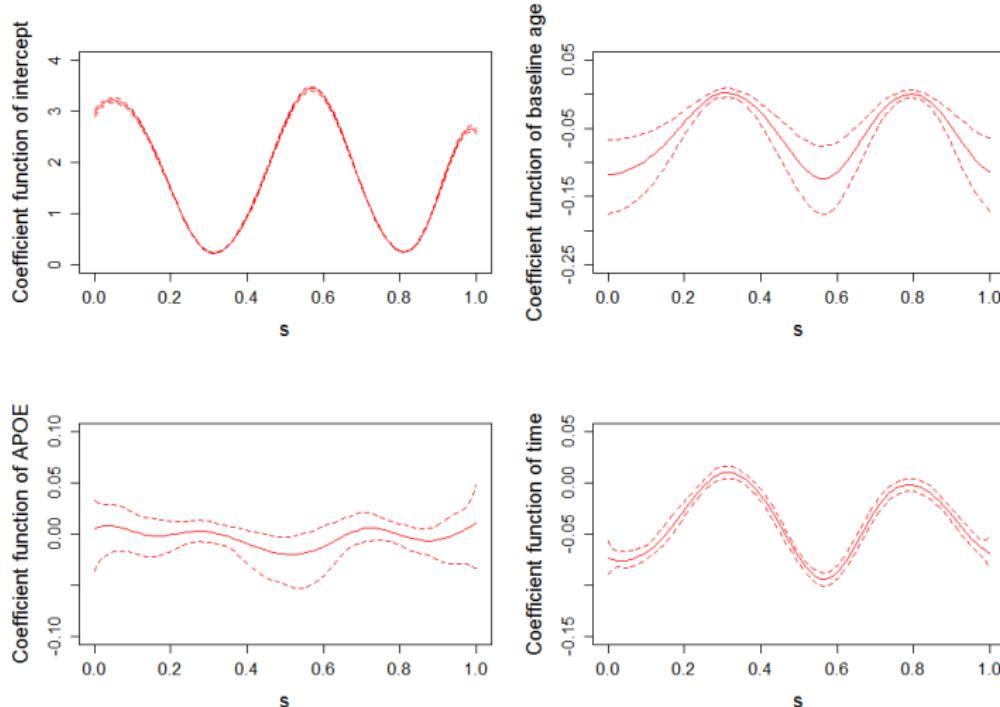
# Parameter Estimation

- Parameter estimates from model *MFJM* with longitudinal HRD.

	Parameters	Mean	SE	2.5%	97.5%
For longitudinal outcome					
ADAS-Cog 11	<i>bAge</i>	0.172	0.258	-0.346	0.668
	<i>APOE-ε</i>	2.207	0.397	1.464	3.006
	Time (Years)	1.121	0.240	0.648	1.584
For survival process					
MCI to AD	Female	0.094	0.172	-0.241	0.442
	<i>bAge</i>	-0.113	0.086	-0.278	0.059
	<i>Edu</i> (years)	0.029	0.026	-0.022	0.080
	<i>APOE-ε</i>	0.409	0.164	0.081	0.736
	$\alpha^*$	0.173	0.021	0.135	0.215
	$\alpha(s) \equiv \alpha$	-1.105	0.437	-1.969	-0.259

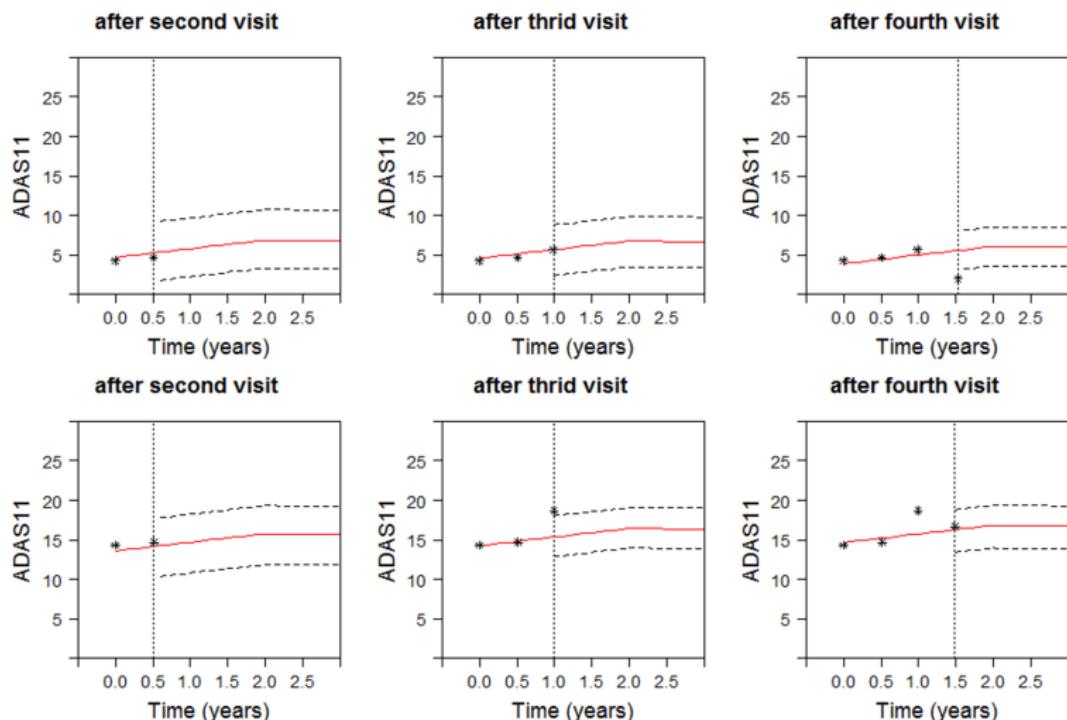
Table: ADNI data analysis results from model *MFJM*.

## Parameter Estimation (cont'd)



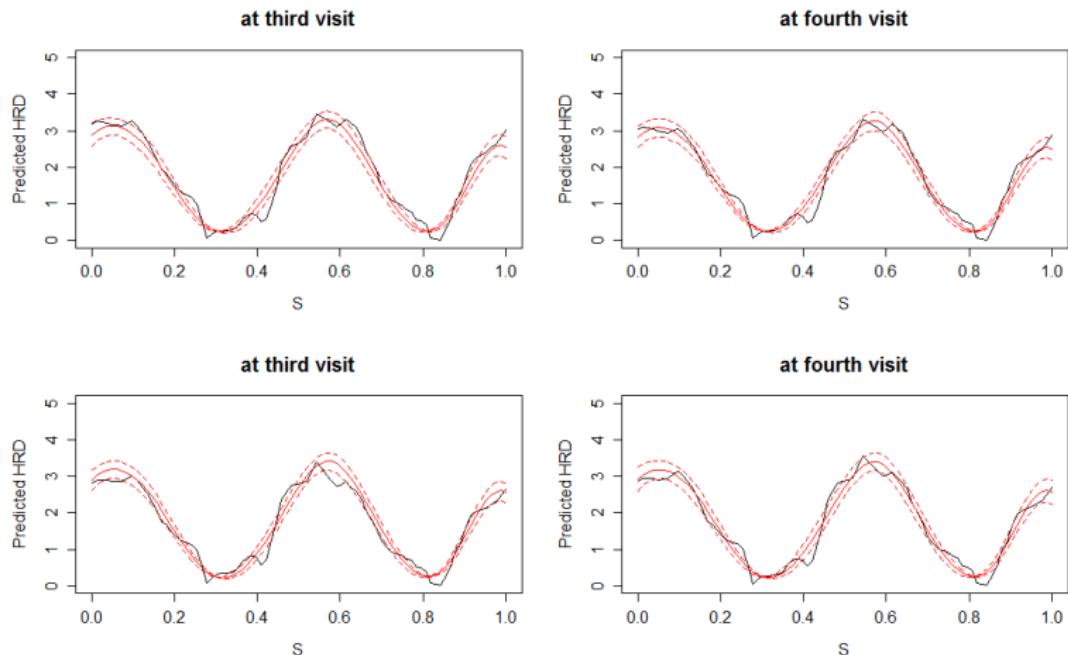
**Figure:** Estimated coefficient functions in the functional longitudinal submodel for HRD.

# Dynamic prediction for new patients using MFJM



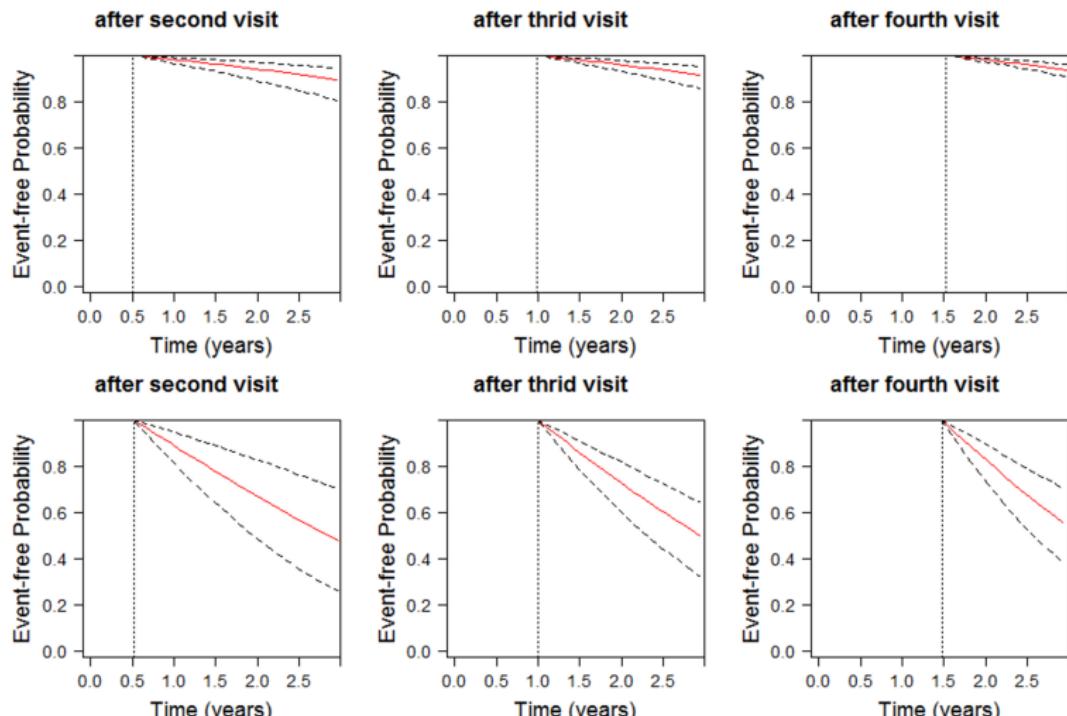
**Figure:** Predicted ADAS-Cog 11 for Patient A (upper panels) and Patient B (lower panels). Solid line is predicted longitudinal trajectories. Dashed lines construct a 95% pointwise uncertainty band. The dotted vertical line represents the time of prediction  $t$ .

# Dynamic prediction for new patients using MFJM (cont'd)



**Figure:** Predicted HRD trace with 95% pointwise uncertainty band for Patient A (upper panels) and Patient B (lower panels).

# Dynamic prediction for new patients using MFJM (cont'd)

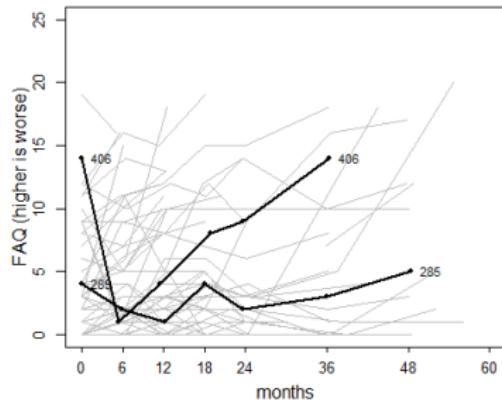
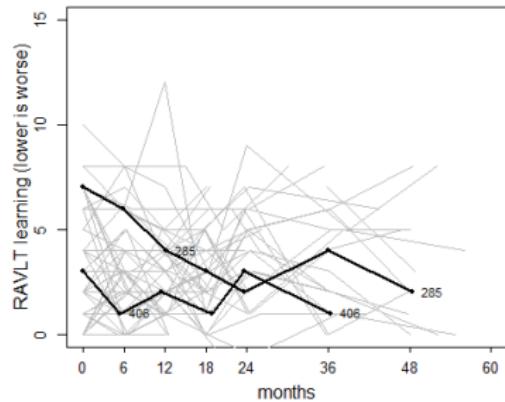
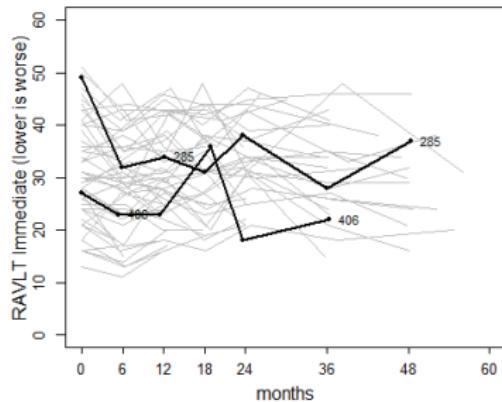
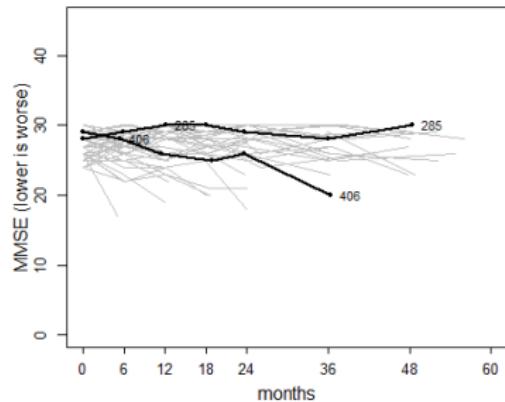


**Figure:** Predicted event-free probability with 95% pointwise uncertainty band for Patient A (upper panels) and Patient B (lower panels).

## Limitations

- The approaches are computationally intensive for high dimensional functional data.
- Hard to extend the model to incorporate a large number ( $>10$ ) of longitudinal biomarkers.
- Difficulty to identify a satisfactory parametric family to model the longitudinal trajectories in all settings.
- Increasing model complexity and computation burden make the predictive tools less attractive for clinical use.

# Multivariate longitudinal markers



# Outline

## 1 Introduction

- FDA and JM
- Motivation
- Data Source
- Research Aims

## 2 Dissertation Aim 1

- Methods
- Application

## 3 Dissertation Aim 2

- Methods
- Dynamic Prediction
- Application

## 4 Dissertation Aim 3

- Methods
- Application

## 5 Conclusion and Future Work

## 6 Acknowledgment

## Joint model based on MFPCA and Cox Model<sup>4</sup>

- A joint model framework that is built on multivariate functional principal component analysis (MFPCA; Happ *et al.*, 2016).
- Consider  $\{Y_{iq}(t_{ij})\}_{1=1,\dots,Q}$ ,  $Y_i^*(s, t_{ij})$  as stochastic functions over the longitudinal visit time or both time and space, and use MFPCA to extract the changing patterns (features) of multiple health outcome trajectories.
- Use the features to extrapolate the health outcome trajectories via FPCA and use score on these features as predictors in a Cox proportional hazards model.

---

<sup>4</sup>Li K, O'Brien R, Lutz M, Luo S, (2018). A Prognostic Model of Alzheimer's Disease Relying on Multiple Longitudinal Measures and Time-to-Event Data, *Alzheimer's & Dementia*.

## Methods: MFPCA

- Conduct FPCA via principal analysis by conditional estimation (PACE, Yao *et al.*, 2005),  $Y_{iq}(t_{ij}) \approx \mu_q(t) + \sum_{l=1}^{L_q} \xi_{iql} \phi_{ql}(t) + \varepsilon_{ijq}$ , where the estimated FPC scores  $\hat{\boldsymbol{\xi}}_{iq} = \{\hat{\xi}_{iql}\}_{l=1,\dots,L_q}$  are

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

- The subject  $i$ 's vector of estimated FPC scores across all outcomes is denoted as  $\hat{\boldsymbol{\xi}}_i = \{\hat{\xi}_{iq}\}_{q=1,\dots,Q}$ , which is of length  $L_+ = \sum_{q=1}^Q L_q$ . Let  $\Theta$  be an  $I \times L_+$  matrix, whose  $i$ -th row is  $\hat{\boldsymbol{\xi}}_i$ .
- 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 MFPC scores  $\hat{\rho}_i = \{\hat{\rho}_{ik}\}_{k=1,\dots,D}$  of subject  $i$  can be calculated via

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

## Methods: high-dimensional MFPCA

- Let  $\mathbf{Y}_{ij}^* = \{Y_{ijm}^*\}_{m=1,\dots,M}$ , where  $Y_{ijm}^*$  is the observed  $Y_i^*(s, t_{ij})$  at the sampling points (vertices or voxels)  $\{s_1, \dots, s_M\}$  and time  $t_{ij}$ . We consider each point with a longitudinal process  $Y_{im}^*(t)$  and the MFPCA approach is applied.
- Apply PACE algorithm independently on the trajectory of each point, which could be run in parallel computing.
- Construct FPC scores matrix  $\Theta^*$  as a  $L \times L_+^*$  centered data matrix and  $H^* = (n - 1)^{-1}\Theta^{*\top}\Theta^*$  is analogous to a  $L_+^* \times L_+^*$  covariance matrix.
- Modify the fast covariance estimation (FACE; Xiao *et al.*, 2016) algorithm on  $H^*$  to get an approximated estimation of the eigenvectors  $\hat{\mathbf{c}}_k^*$  and MFPC scores  $\hat{\boldsymbol{\rho}}_i^*$  of subject  $i$ .

## Methods: Survival analysis

- The estimated scores on the features  $\hat{\rho}_i$  and  $\hat{\rho}_i^*$  can be used as predictors in modeling the relations between the survival time and the patterns of longitudinal scalar and functional/image outcomes.

$$h_i(t) = h_0(t) \exp\{\mathbf{Z}_i^\top \boldsymbol{\gamma} + \hat{\boldsymbol{\rho}}_i^\top \boldsymbol{\beta} + (\hat{\boldsymbol{\rho}}_i^*)^\top \boldsymbol{\eta}\}.$$

## Methods: Dynamic Prediction

- Given a new subject  $N$  who is event-free till time  $t$ , we calculate the subject-specific scores  $\hat{\rho}_N$  and  $\hat{\rho}_N^*$  based on the longitudinal observation till time  $t$ .
- Future outcome trajectories at time  $t'$  is

$$E(Y_{Nq}(t)) = \hat{\mu}_q(t) + \sum_{k=1}^D \hat{\rho}_{Nk} \hat{\psi}_{qk}(t).$$

- The risk of an event occurring within a prediction window is

$$\begin{aligned}\hat{\pi}_N(t'|t) &= p(T_N^* \geq t' | T_N^* > t, \mathbf{Z}_N, \hat{\rho}_N, \hat{\rho}_N^*) \\ &= \left\{ \frac{\hat{S}_0(t')}{\hat{S}_0(t)} \right\}^{\exp\{\mathbf{Z}_N^\top \gamma + \hat{\rho}_N^\top \beta + (\hat{\rho}_N^*)^\top \eta\}}\end{aligned}$$

# Outline

## 1 Introduction

- FDA and JM
- Motivation
- Data Source
- Research Aims

## 2 Dissertation Aim 1

- Methods
- Application

## 3 Dissertation Aim 2

- Methods
- Dynamic Prediction
- Application

## 4 Dissertation Aim 3

- Methods
- **Application**

## 5 Conclusion and Future Work

## 6 Acknowledgment

# Application to the ADNI Study

- Model structure
  - ▶ Time from first visit to AD diagnosis as survival outcome;
  - ▶ Five neurocognitive markers that have strong predictive value: Alzheimer Disease ADAS-Cog 13; RRAVLT immediate; FAQ; MMSE.
  - ▶ Longitudinal hippocampal surface information based on hippocampal radial distance (HRD).
  - ▶ Age, gender, years of education and presence of the apolipoprotein E (*APOE*) ε4 allele as scalar covariates in survival model.
- *Model 1*: predictors are MFPCA scores based on five neurocognitive markers;
- *Model 2*: predictors are MFPCA scores based on five neurocognitive markers and MFPCA scores based on HRD.

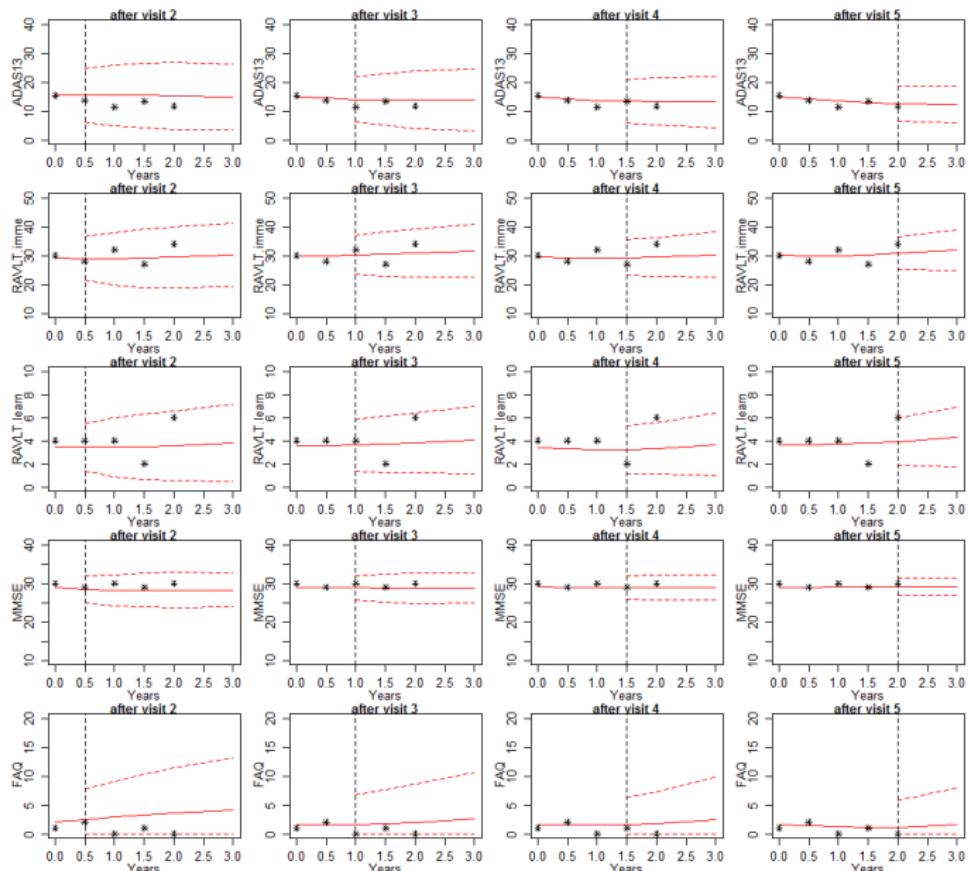
# Predictive Performance

- Comparison of prediction performance by two candidate models in ADNI study. AUC: Areas under the time-dependent ROC curve; BS: dynamic expected Brier scores.

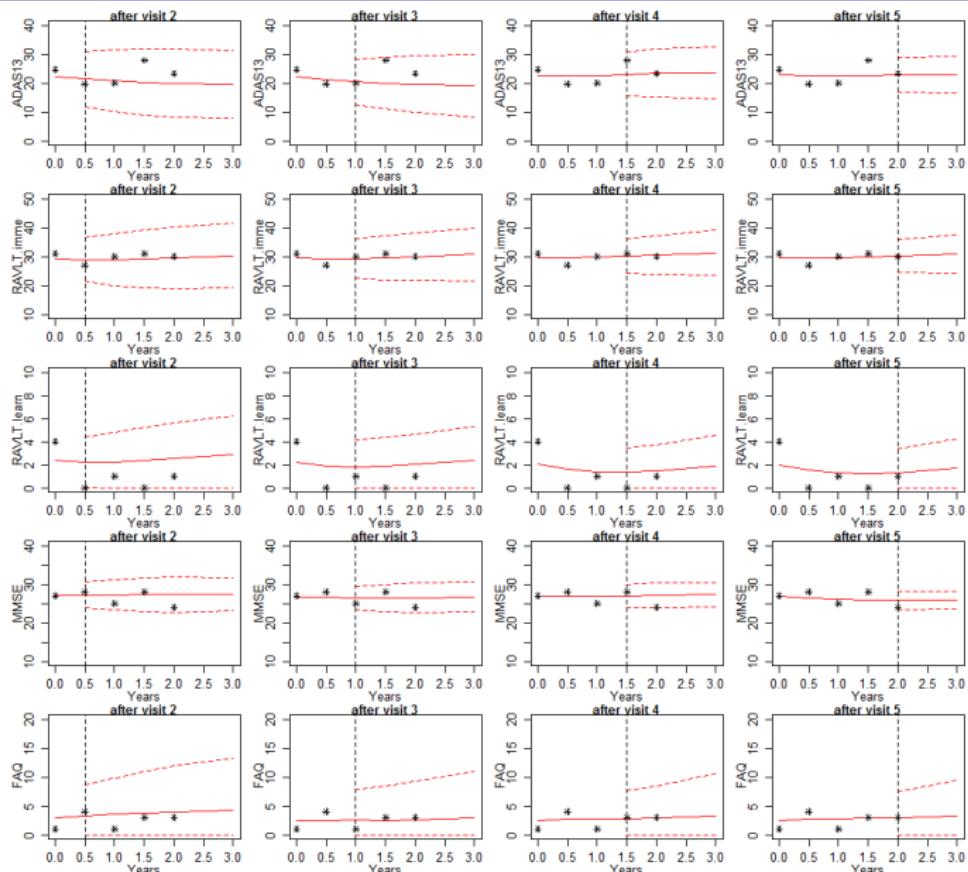
$\Delta t$	$t$	Model 1		Model 2	
		AUC	BS	AUC	BS
6m	6m	0.706	0.078	0.717	0.078
	12m	0.804	0.097	0.824	0.096
	18m	0.754	0.108	0.783	0.107
12m	6m	0.781	0.147	0.800	0.145
	12m	0.788	0.166	0.814	0.163
	18m	0.791	0.159	0.799	0.157

- Including features of longitudinal *HRD* in Cox model may further improve the overall predictive ability of the model.

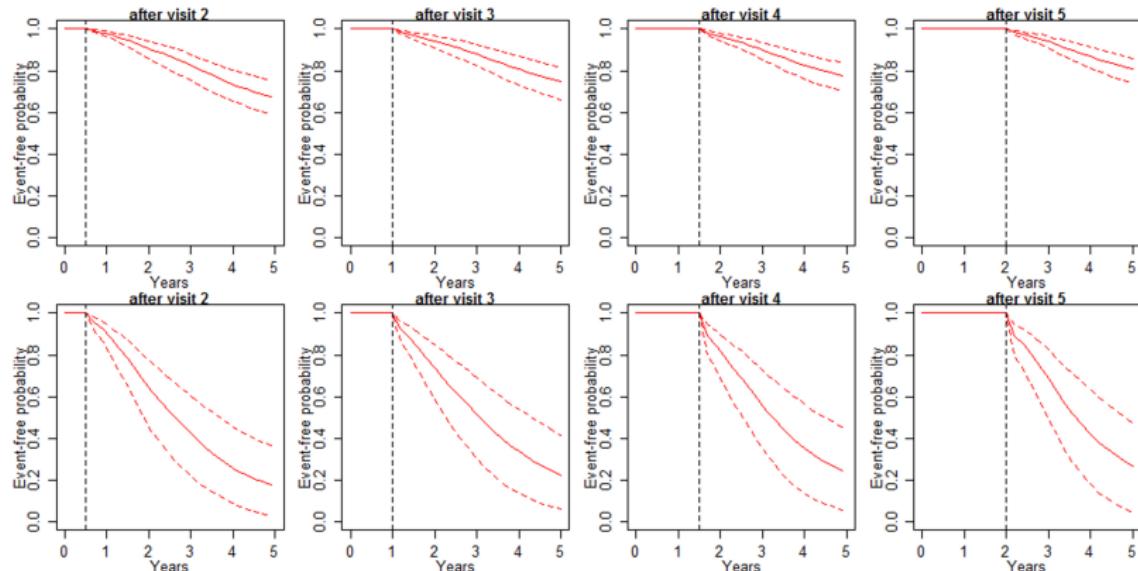
# Dynamic prediction for new patients using MPCA-Cox



# Dynamic prediction for new patients using MPPCA-Cox



# Dynamic prediction for new patients using MPCA-Cox



**Figure:** Predicted event-free probability with 95% bootstrap confidence interval for Patient A (upper panels) and Patient B (lower panels). The dotted vertical line represents the time of prediction  $t$ .

# Conclusion

- The functional joint models (FJM) yield novel insights of the dynamic association between longitudinal functional outcomes and the hazard.
- The personalized dynamic prediction approaches provide valuable guidance for clinical decision making on patient prognosis and targeted treatment.
- The FJM are important developments to both fields of functional data analysis and joint modeling of longitudinal and survival data.
- The practical impact of FJM could be dramatic for the neurodegenerative diseases (e.g., Alzheimer's disease).

## Future work

- Extend the proposed model to a generalized MFPC framework to accommodate other type of longitudinal outcomes, e.g., binary data, ordinal data, or mixture of these.
- Software development and web deployment. To facilitate the use of these approaches by physicians.
- Patient centered care: combine personalized prediction and patient preference information together to facilitate medical decision making.
- Apply the proposed methods to many other studies with a similar data structure.

# Acknowledgment

- Committee members
  - ▶ Dr. Hulin Wu (Advisor/ Committee Chair)
  - ▶ Dr. Sheng Luo (Dissertation Supervisor)
  - ▶ Dr. David Lairson (Minor Advisor)
  - ▶ Dr. Momiao Xiong (Breath Advisor)
  - ▶ Dr. Dejian Lai (External Reviewer)

Thank You