Functional Joint Model for Longitudinal and Time-to-Event Data: An Application to Alzheimer's Disease

Kan Li
Dept. of Biostat, Univ of Texas at Houston

Aug. 2, 2016

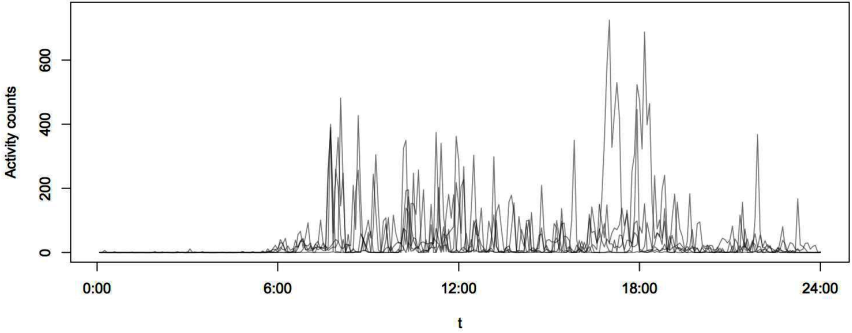
Outline

Functional Data

- Functional Data: data for which units of observation are functions defined on certain continuous domains and recorded on discrete grids.
  - These functions can be curves (1D), images (2D or 3D), or higher dimension object data (e.g. functional MRI).

Examples of Functional Data

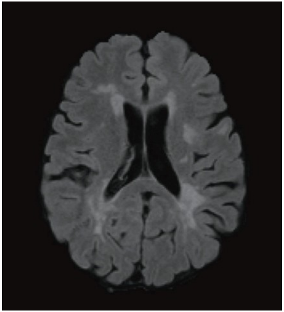
- Physical activity information



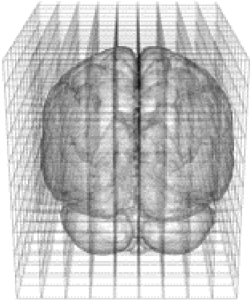
Examples of Functional Data

- Brain imaging

A slice of Magnetic Resonance Imaging (MRI)



Voxel-based whole-brain image



- The analysis of functional data is termed "Functional Data Analysis" (FDA)

Functional Regression

- Functional Regression: regression analysis involving functional data.
  - Functional predictor regression (scalar-on-function)
    $y_i = \beta_0 + \int x_i(s)\beta(s)ds + \varepsilon_i$
  - Functional response regression (function-on-scalar)
    $y_i(s) = \beta_0(s) + x_i\beta(s) + \varepsilon_i(s)$
  - Function-on-function regression (function-on-function)
    $y_i(s) = \beta_0(s) + \int x_i(s)\beta(s)ds + \varepsilon_i(s)$

Functional Regression

- Functional predictor regression: $y_i = \beta_0 + \int x_i(s)\beta(s)ds + \varepsilon_i$
  - Most existing work deal only with cross-sectional functional data;
  - Goldsmith *et al.*(2012), longitudinal functional regression;
  - Gellar *et al.*(2015), Cox model with cross-sectional functional covariate;
  - No previous functional regression modeling attempted under joint models framework for longituindal and time-to-event data.

Joint Models for Longitudinal and Time-to-Event data

- Why use Joint Models?

  The evolution of a biomarker is directly informative about the time to the event.

- Intuitive idea behind Joint Models:
  - mixed effects submodel to describe the evolution of the biomarker;
  - Cox proportional hazard submodel for survival outcome;
  - link the two submodels using a common latent structure.

- Current joint models in the literature only include scalar variables as responses and do not account for functional covaraites.

Objectives

- Develop a joint model for data where outcomes are longitudinal scalar measures and time to event, and the exposure involve both functional/image and scalar covariates.

Alzheimer's Disease as Motivation Example

- Neurodegenerative disorder and is the most common form of dementia.
- 5.4 million American have AD and the number will reach 7.7 million by 2030.
- $172 billion for the total cost of care for Americans with AD in 2010, and will increase to $1.08 trillion by 2050 each year.
- Thus, many resources are invested to accelerate the search for cures while improving diagnosis of Alzheimer's Disease.

Alzheimer's Disease Neuroimaging Initiative (ADNI) study

- Ongoing multisite longitudinal study.
- Collects serial clinical, imaging (MRI, PET, fMRI), genetic, biospecimen, neuropsychological assessments data.
- Phase I of ADNI study (ADNI1)
  - 229 normal cognition (NC) patients
  - 397 mild cognitive impairment (MCI) patients
  - 193 Alzheimer's disease (AD) patients
  - Patients were reassessed at 6, 12, 18, 24 and 36 months, and followed annually as part of ADNI GO and ADNI2

Alzheimer's Disease Neuroimaging Initiative (ADNI) study

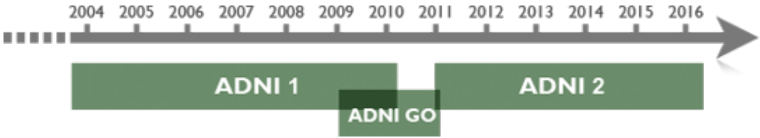
- ADNI GO Study

  Enrolled 128 new patients, all of which were MCI patients.

- ADNI2 Study

  Enrolled 925 new patients, 311 NC patients, 451 MCI patients, and 163 AD patients.

- All data can be downloaded from `http://www.adni-info.org`.

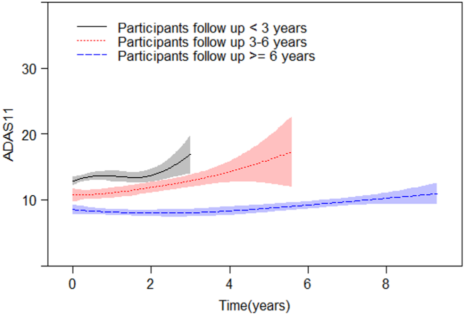


Alzheimer's Disease Neuroimaging Initiative (ADNI) study

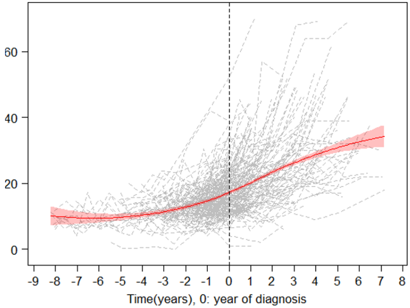
- Mild cognitive impairment (MCI)
  - An intermediate stage between NC and AD;
  - Target population for evaluating prognosis and early treatment.
- Predict the conversion from MCI to AD
- In literature:
  - Cox regression models: predicting time to AD conversion.
  - Linear mixed model: exploiting association between longitudinal markers and cognitive decline

Alzheimer's Disease Neuroimaging Initiative (ADNI) study

- Alzheimer Disease Assessment ScaleCognitive (ADAS-Cog)
  - Assesses written and verbal responses of subjects that are related to fundamental cognitive functions.



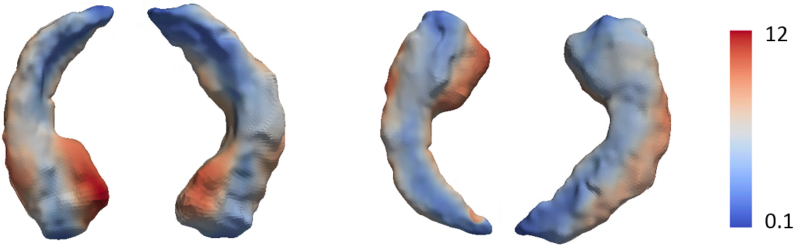
(a)



(b)

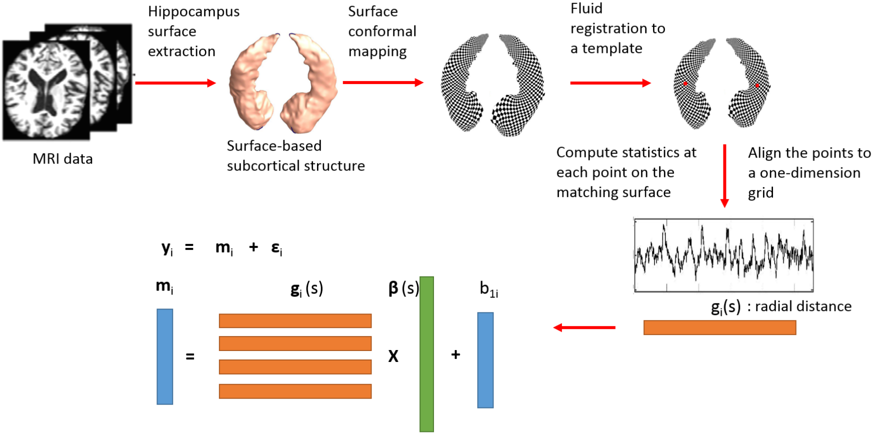
Alzheimer's Disease Neuroimaging Initiative (ADNI) study

- Hippocampus surface morphology data
  - Surface-based subcortical structure analysis.
  - Hippocampal radial distance, the distance from the medial core to each surface point and measures hippocampal thickness.



(a) Top (b) Bottom

Hippocampus Image Processing



$$h_i(t) = h_0(t)\, exp\, \{ w_i \gamma + g_i(s)\, \gamma(s) + \alpha\, m_i(t) \}$$

Joint modeling framework

- Integrate time-independent functional covariates in both longitudinal part and survival part of joint model.
- Model Specification

$$
\begin{aligned}
y_i(t) &= m_i(t) + \varepsilon_{ij}, \\
m_i(t) &= \beta_0 + \boldsymbol{x}_{ij}\boldsymbol{\beta} + \int_S g_i^{(x)}(s)\beta(s)ds + \boldsymbol{z}_{ij}\boldsymbol{b}_i \\
h(t) &= h_0(t)\exp\{\boldsymbol{w}_i\boldsymbol{\gamma} + \int_S g_i^{(w)}(s)\gamma(s)ds + \alpha m_i(t)\}.
\end{aligned}
$$

  - $y_i(t)$, observed longitudinal outcome;
  - $m_i(t)$, true unobserved patient specific longitudinal trajectory;
  - $\boldsymbol{x}_{ij}$, $\boldsymbol{w}_i$, scalar covariates; $\boldsymbol{z}_{ij}$, random effects;
  - $g_i^{(x)}(s)$, $g_i^{(w)}(s)$, time-independent functional covariates defined over the domain $S \in [0, S_{max}]$.

Functional Principal Component

- Let $\mu_X(s)$ be the mean of the functional predictor $g_i^{(x)}(s)$ taken over all subjects.
- Let $\Sigma^{(x)}(s, s') = cov\{g_i^{(x)}(s), g_i^{(x)}(s')\}$ be the covariance functions providing the covariance between two locations of $g_i^{(x)}(s)$.
- Let $\sum_{l=1}^{\infty} \lambda_l^{(x)} \phi_l^{(x)}(s) \phi_l^{(x)}(s')$ be the spectral decomposition of $\Sigma^{(x)}(s, s')$.
  - $\lambda_1^{(x)} \geq \lambda_2^{(x)} \geq \cdots \geq 0$ are the non-increasing eigenvalues;
  - $\boldsymbol{\phi}^{(x)}(s) = [\phi_1^{(x)}(s), \cdots, \phi_{K_x}^{(x)}(s)]^T$ are the corresponding orthonormal eigenfunctions.

Functional Principal Component

- Truncated version of Karhunen-Loève approximation for $g_i^{(x)}(s)$

  $g_i^{(x)}(s) \approx \mu^{(x)}(s) + \sum_{l=1}^{K_x} \xi_{il}^{(x)} \phi_l^{(x)}(s) = \mu^{(x)}(s) + \boldsymbol{\xi}_i^{(x)} \boldsymbol{\phi}^{(x)}(s).$

  - $\xi_{il}^{(x)} = \int_S \{g_i^{(x)}(s) - \mu^{(x)}(s)\} \phi_l(s) ds$, functional principal component (FPC) score;
  - $\xi_{il}^{(x)} \sim N(0, \lambda_l)$
  - K_x is truncation number, can be determined by the proportion of variance explained.

Functional Principal Component

- Expand $\beta(s)$ in the truncated principal component basis $\beta(s) = \sum_{l=1}^{K_x} \phi_l^{(x)}(s)\beta_l^{(x)} = [\boldsymbol{\phi}^{(x)}(s)]^T \boldsymbol{\beta}^{(x)}$.
- Functional term is converted to a scalar term, $\int_S g_i^{(x)}(s)\beta(s)ds = \int_S \mu^{(x)}(s)\beta(s)ds + \int_S \boldsymbol{\xi}_i^{(x)} J_{\phi,\phi}\boldsymbol{\beta}^{(x)} ds$, where $J_{\phi,\phi} = \int \boldsymbol{\phi}^{(x)}(s)[\boldsymbol{\phi}^{(x)}(s)]^T ds = I$ because of the orthonormal of basis functions.
- Similar notation holds for the functional predictor $\int_S g_i^{(w)}(s)\gamma(s)ds$ in survival submodel.

Functional Regression

- Joint model using FPC scores as scalar covariates

$$
\begin{aligned}
y_i(t) &= m_i(t) + \varepsilon_{ij}, \text{ where} \\
m_i(t) &\approx \beta_0' + \boldsymbol{x}_{ij}\boldsymbol{\beta} + \boldsymbol{\xi}_i^{(x)}\boldsymbol{\beta}^{(x)} + \boldsymbol{z}_{ij}\boldsymbol{b}_i, \\
\text{and } h(t) &\approx h_0^*(t)\exp\{\boldsymbol{w}_i\boldsymbol{\gamma} + \boldsymbol{\xi}_i^{(w)}\boldsymbol{\gamma}^{(w)} + \alpha m_i(t)\}.
\end{aligned}
$$

  - $h_0^*(t) = h_0(t)\exp\{\int_S \mu^{(w)}(s)\gamma(s)ds\}$;
  - $\beta_0' = \beta_0 + \int_S \mu^{(x)}(s)\beta(s)ds$.

Estimation and inference

- Estimated FPC scores $\boldsymbol{\xi}_i^{(x)}$ and $\boldsymbol{\xi}_i^{(w)}$ of all subjects.
- Full likelihood of joint model

$$
\begin{aligned}
L(\boldsymbol{\theta}) &= \prod_{i=1}^{n} p(T_i, \delta_i, \boldsymbol{y}_i | \boldsymbol{\theta}) \\
&= \prod_{i=1} \int p(T_i, \delta_i, | \boldsymbol{\theta}, \boldsymbol{b}_i) \prod_{j=1}^{n_i} p(y_{ij}, \boldsymbol{\theta}, \boldsymbol{b}_i) p(\boldsymbol{b}_i | \boldsymbol{\theta}) d\boldsymbol{b}_i
\end{aligned}
$$

- Maximization of the log-likelihood function $\ell(\boldsymbol{\theta}) = \sum_i \log p(T_i, \delta_i, \boldsymbol{y}_i | \boldsymbol{\theta})$ using Expectation Maximization (EM) algorithm.
- Estimated coefficient function is given by $\hat{\beta}(s) = [\hat{\boldsymbol{\phi}}^{(x)}(s)]^T \hat{\boldsymbol{\beta}}^{(x)}$.

Dynamic risk prediction

- The probability of survival at time u conditional on survival up to time t (e.g. $u = t + \Delta t > t$),

$$\begin{aligned}\pi_i(u|t) &= P(T_i^* \geq u | T_i^* > t, Y_i(t); \boldsymbol{\theta}) \\ &= \int P(T_i^* \geq u | T_i^* > t, Y_i(t), \boldsymbol{b}_i; \boldsymbol{\theta}) p(\boldsymbol{b}_i | T_i^* > t, Y_i(t); \boldsymbol{\theta}) d\boldsymbol{b}_i \\ &= \int \frac{S_i\{u | M_i(u, \boldsymbol{b}_i, \boldsymbol{\theta}); \boldsymbol{\theta}\}}{S_i\{t | M_i(t, \boldsymbol{b}_i, \boldsymbol{\theta}); \boldsymbol{\theta}\}} p(\boldsymbol{b}_i | T_i^* > t, Y_i(t); \boldsymbol{\theta}) d\boldsymbol{b}_i,\end{aligned}$$

- A first-order estimate for $\pi_i(u|t)$ is $\pi_i(u|t) = \frac{S_i\{u|M_i(u,\hat{\boldsymbol{b}}_i,\hat{\boldsymbol{\theta}});\hat{\boldsymbol{\theta}}\}}{S_i\{t|M_i(t,\hat{\boldsymbol{b}}_i,\hat{\boldsymbol{\theta}});\hat{\boldsymbol{\theta}}\}} + O(n_i^{-1})$.
- A Monte Carlo estimate of $\pi_i(u|t)$ can be obtain by following sample scheme. For $l = 1, \cdots, L$ repetitions:
  - Draw $\boldsymbol{\theta}^{(l)} \sim N(\hat{\boldsymbol{\theta}}, \hat{var}(\hat{\boldsymbol{\theta}}))$
  - Draw $\boldsymbol{b}_i^{(l)} \sim \{\boldsymbol{b} | T_i^* > t, Y_i(t); \boldsymbol{\theta}^{(l)}\}$
  - Calculate $\pi_i^{(l)}(u|t) = \frac{S_i\{u|M_i(u,\boldsymbol{b}_i^{(l)},\boldsymbol{\theta}^{(l)});\boldsymbol{\theta}^{(l)}\}}{S_i\{t|M_i(t,\boldsymbol{b}_i^{(l)},\boldsymbol{\theta}^{(l)});\boldsymbol{\theta}^{(l)}\}}$

Application to the ADNI Study

- Baseline characteristics of ADNI-1 participants with mild cognitive impairment (MCI)

	Progressed to AD during the study (n = 200)	Did not progress to AD during the study (n = 184)	Combined (n = 384)
Women	75 (37.50%)	62 (33.50%)	137 (35.7%)
Age (years)	74.44 (7.09)	75.03 (7.55)	74.71 (7.31)
APOE4 present	127 (63.50%)	81 (44.00%)	208 (54.16%)
Education (years)	15.82 (2.86)	15.33 (3.19)	15.58 (3.03)
Time in study (years)	2.25 (1.74)	4.24 (2.91)	3.20 (2.57)
Data are meam (SD) or n (%)			

Joint model without functional covariate

- *JM*
  - ADAS-Cog 11 as longitudinal outcome;
  - Time from first visit to AD conversion as survival outcome;
  - Age, gender, years of education and presence of the apolipoprotein E (*APOE*) $\varepsilon 4$ allele as scalar covariates.
  - Including baseline Hippocampal volume as a covariate in both longitudinal and survival submodel provides the best model fitting.
- Specifically, *JM* is

$$
\begin{aligned}
ADAS\text{–}Cog_i(t_{ij}) &= m_i(t_{ij}) + \varepsilon_{ij} \\
m_i(t_{ij}) &= \beta_0 + \beta_1 t_{ij} + \beta_2 bage_i + \beta_3 bHV_i + b_{0i} \\
h(t) &= h_0(t)\exp\{\gamma_1 gender_i + \gamma_2 bage_i + \gamma_3 Edu_i + \\
&\quad \gamma_4 APOE\text{–}\varepsilon 4 + \gamma_5 bHV_i + \alpha m_i(t)\}.
\end{aligned}
$$

Joint model with functional covariate

- Baseline hippocampal surface data based on radial distance as functional covariate.
  - Perform FPCA to the hippocampal radial distance (HRD) and choose the first 20 FPC which explain 82.6% of the total variance in the hippocampus surface data.
- $FJM1$: only include HRD as a functional covariate in the longitudinal submodel.
- $FJM2$: only include HRD as a functional covariate in the survival submodel.
- $FJM3$: include HRD as functional covariates in both the longitudinal and the survival submodel.

Joint model with functional covariate

- The corresponding models with functional covariate are specified accordingly as

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

$$\begin{aligned} m_i(t_{ij}) &= \beta_0 + \beta_1 t_{ij} + \beta_2 bage_i + \beta_3 bHV_i + \int_S HRD_i(s)\beta(s)ds + b_{0i} \\ h(t) &= h_0(t)\exp\{\gamma_1 gender_i + \gamma_2 bage_i + \gamma_3 Edu_i + \gamma_4 APOE\text{–}\varepsilon 4 + \\ &\quad \gamma_5 bHV_i + \int_S HRD_i(s)\gamma(s)ds + \alpha m_i(t)\}. \end{aligned}$$

Models comparison

- Including hippocampal surface data as a functional covariate improve the model fitting.

	JM	*FJM*1	*FJM*2	*FJM*3
AIC	10454	10440	10452	10446

Table: ADNI-1 data analysis results under the four models: AICs

Parameter Estimation

- Parameter estimates based on $FJM1$ with HRD as functional covariates in longitudinal submodel.

	Parameters	Estimated	SE	p value
For longitudinal outcome				
ADAS-Cog 11	Time (Years)	0.425	0.047	$<$0.001
	Baseline Age	-0.389	0.254	0.125
	Hippocampal Volume	-1.878	0.221	$<$0.001
For survival process				
MCI to AD	Gender (Female)	-0.161	0.167	0.331
	Baseline Age	-0.181	0.0867	0.037
	Education Years	-0.004	0.026	0.866
	APOE-ε	0.397	0.167	0.018
	Hippocampal Volume	-3.559	0.928	$<$0.001
	α	0.108	0.019	$<$0.001

Table: ADNI-1 data analysis results for proposed functional joint model $FJM1$

Parameter Estimation

- Estimated coefficient function associated with hippocampal surfaces.

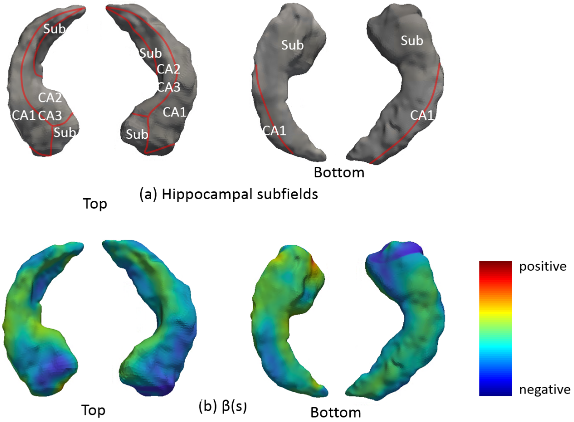


Figure: ADNI-1 data analysis (from *FJM*1) of estimated coefficient function associated with hippocampal surfaces. Each side of the left and right hippocampal surfaces.

Dynamic risk prediction

- Predictive performance was evaluated via a 10-fold cross validation.
  AUC: time-dependent areas under the ROC curves.
  DDI: dynamic discrimination index, which summarizes the discrimination power of the measure over the whole follow-up period.

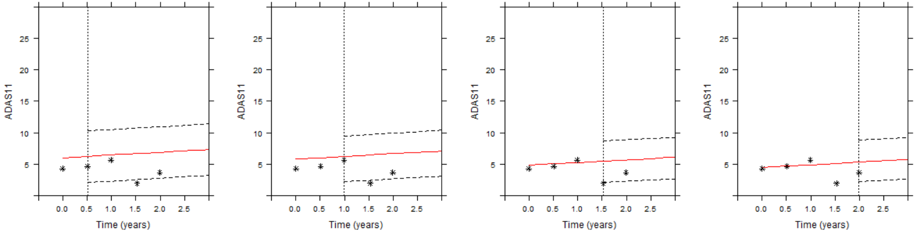
Δt	t	*JM* AUC	*JM* DDI	*FJM*1 AUC	*FJM*1 DDI
0.5	1	0.830	0.758	0.834	0.772
	1.5	0.705		0.762	
	2	0.861		0.910	
1	1	0.781	0.774	0.820	0.795
	1.5	0.769		0.837	
	2	0.789		0.817	

Table: Areas under the ROC curve and estimated dynamic discrimination index for joint model with/without functional covariate.

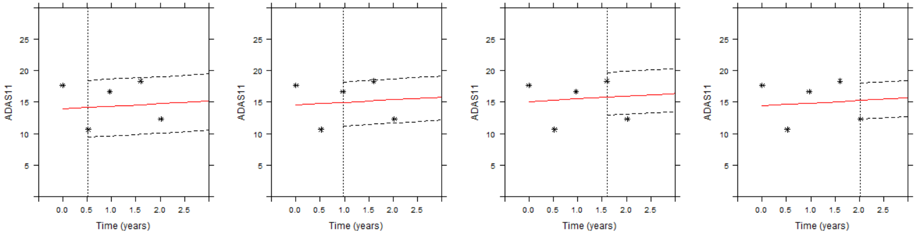
Dynamic prediction for new patients using *FJM*1

- Predict future health outcome ADAS11 trajectories.

MCI patient with low risk of the disease.



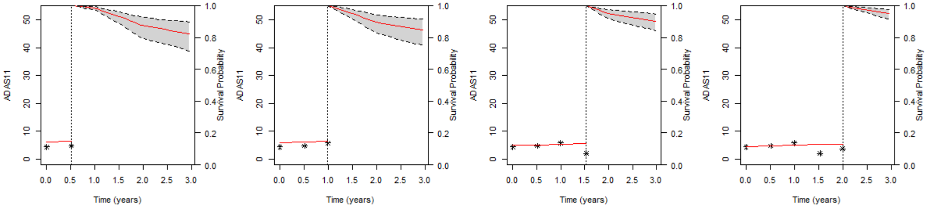
MCI patient with high risk of the disease.



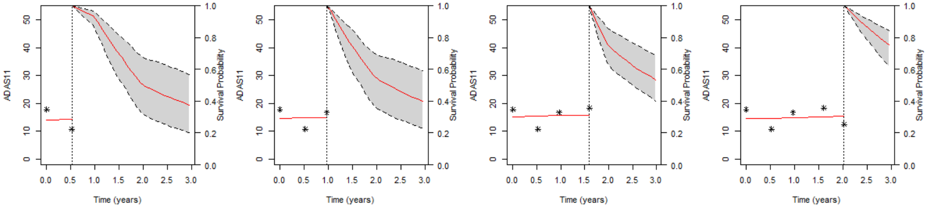
Dynamic prediction for new patients using $FJM1$

- Predict future risk of AD conversion.

MCI patient with low risk of the disease.



MCI patient with high risk of the disease.



Simulation Setting

- Longitudinal Model:

$$
\begin{aligned}
y_i(t_{ij}) &= m_i(t_{ij}) + \varepsilon_{ij}, \\
m_i(t_{ij}) &= \beta_0 + \beta_1 \times t_{ij} + \int_0^{10} g_i^{(x)}(s)\beta(s)ds + b_i, \\
g_i^{(x)}(s) &= u_{i1} + u_{i2} \times s + \sum_{k=1}^{10} \{\nu_{is1} \times \sin(\frac{\pi k}{5}s) + \nu_{is2} \times \cos(\frac{\pi k}{5}s)\}.
\end{aligned}
$$

- Survival Model:
  $h(t) = h_0(t)\exp\{\gamma_1 \times w_1 + \int_0^{10} g_i^{(w)}(s)\gamma(s)ds + \alpha m_i(t)\}.$

- Coefficient functions:
  $\beta(s) = 2\sin(\pi s/5)$ and $\gamma(s) = 1.2\sin(\pi s/4)$.

Simulation results

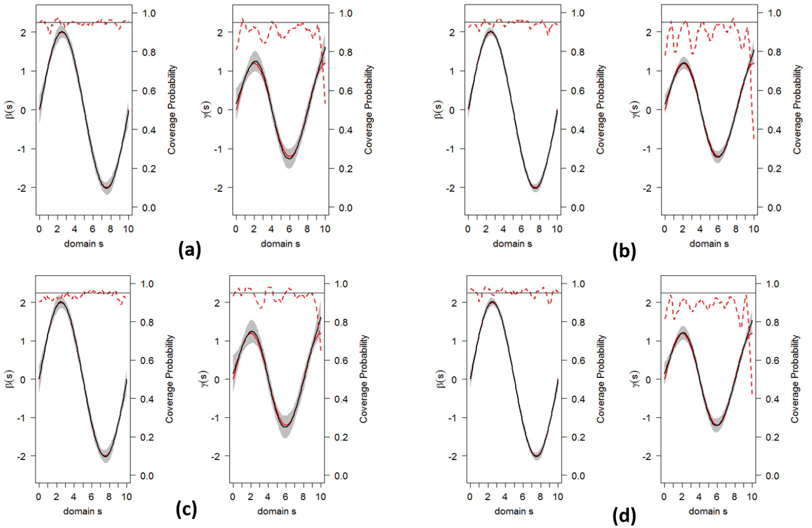
- Simulation results based on different sample size and censoring rate.

	n=200, c=0.3					n=500, c=0.3				
	Bias	AMSE	SE	SD	CP	Bias	AMSE	SE	SD	CP
For longitudinal outcomes										
$\beta_1 = 0.78$	<0.001	<0.001	0.002	0.002	0.945	<0.001	<0.001	0.002	0.002	0.970
$\beta(s) = 2\sin(\pi s/5)$		0.008					0.003			
$\sigma_Y^2 = 1$	<0.001	0.001	0.035	0.038	0.93	<0.001	<0.001	0.022	0.024	0.920
$\sigma_b^2 = 1$	0.036	0.017	0.141	0.127	0.985	0.023	0.007	0.088	0.085	0.955
For survival										
$\gamma_1 = -1.75$	0.119	0.077	0.222	0.251	0.910	0.012	0.027	0.152	0.163	0.940
$\alpha = 0.29$	0.014	0.001	0.022	0.020	0.895	0.003	<0.001	0.013	0.015	0.910
$\gamma(s) = 1.2\sin(\pi s/4)$		0.023					0.012			

	n=200, c=0.5					n=500, c=0.5				
	Bias	AMSE	SE	SD	CP	Bias	AMSE	SE	SD	CP
For longitudinal outcomes										
$\beta_1 = 0.78$	<0.001	<0.001	0.002	0.002	0.954	<0.001	<0.001	0.001	0.001	0.950
$\beta(s) = 2\sin(\pi s/5)$		0.009					0.003			
$\sigma_Y^2 = 1$	0.004	0.002	0.038	0.040	0.965	0.004	<0.001	0.024	0.025	0.955
$\sigma_b^2 = 1$	0.047	0.021	0.146	0.138	0.960	0.025	0.009	0.091	0.097	0.930
For survival										
$\gamma_1 = -1.75$	0.091	0.073	0.254	0.255	0.940	0.042	0.021	0.134	0.139	0.925
$\alpha = 0.29$	0.013	0.001	0.023	0.024	0.91	0.004	<0.001	0.012	0.015	0.905
$\gamma(s) = 1.2\sin(\pi s/4)$		0.025					0.023			

Simulation results

- Estimation of coefficient functions



Future work

- Include multiple brain regions as functional covariances in our application.
- Compare the performance of other basis functions, e.g., splines, Fourier, wavelet, or some combination, to represent the predictor function and/or coefficient function.
- Incorporate both repeated observations of outcome and functional covariate in longitudinal submodel.
- Use longitudinal function-on-scalar model for longitudinal submodel.
- Software development.

Acknowledgement

Supported by grants from:

- NINDS-R01NS091307.
- NCATS-5KL2TR000370.
- Parkinsons Disease Foundation.