Supplementary Materials to "Prediction of Cognitive Function via Brain Region Volumes with Applications to Alzheimer's Disease Based on Space-Factor-Guided Functional Principal Component Analysis"

Shoudao Wen[1], Yi Li[2], Dehan Kong[3] and Huazhen Lin[1*]
[1]Center of Statistical Research, School of Statistics,
and New Cornerstone Science Laboratory,
Southwestern University of Finance and Economics, Chengdu, China
[2] Department of Biostatistics, University of Michigan, Ann Arbor, USA
[3] Department of Statistical Sciences, University of Toronto, Toronto, Canada

March 8, 2025

S1 Supplementary Material A: Proof of Theorem 1

Noting that $\mathbf{M}$ is given, consider two combinations of parameters $(\mathbf{B}^{(1)}, \boldsymbol{\eta}^{(1)}, \boldsymbol{\Phi}^{(1)}(t), \boldsymbol{\zeta}^{(1)})$ and $(\mathbf{B}^{(2)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\Phi}^{(2)}(t), \boldsymbol{\zeta}^{(2)})$ both satisfying model (5), i.e. $(\mathbf{B}^{(1)} + \mathbf{M}\boldsymbol{\eta}^{(1)})\boldsymbol{\Phi}^{(1)\prime}(t)\boldsymbol{\zeta}^{(1)} = (\mathbf{B}^{(2)} + \mathbf{M}\boldsymbol{\eta}^{(2)})\boldsymbol{\Phi}^{(2)\prime}(t)\boldsymbol{\zeta}^{(2)}$. We consider the similar proofs in Wang et al. (2017) and Bing et al. (2024).

Let $\mathcal{S}$ be the set of indices such that $\mathbf{B}^{(1)}_{\mathcal{S}} = \mathbf{B}^{(2)}_{\mathcal{S}} = \mathbf{0} \in \mathbb{R}^{|\mathcal{S}|\times q}$. By condition (I1), it is easy to show that there exists such $\mathcal{S}$ and $|\mathcal{S}| \geqslant m$ because both $\mathbf{B}^{(1)}$ and $\mathbf{B}^{(2)}$ have at most $\lfloor (p-m)/2 \rfloor$ rows whose elements are not all equal to 0. Let $\mathbf{X}_{\mathcal{S}}(t), \mathbf{u}_{\mathcal{S}}(t)$ be the corresponding

*Corresponding author. Email: linhz@swufe.edu.cn. The research was supported by National Key R&D Program of China (No.2022YFA1003702), National Natural Science Foundation of China (Nos. 11931014), and New Cornerstone Science Foundation.

subvectors of $\mathbf{X}(t), \mathbf{u}(t)$ for fixed t. Denote $\mathbf{\Sigma}_{\mathbf{X}_\mathcal{S}}(t)$ and $\mathbf{\Sigma}_{\mathbf{u}_\mathcal{S}}(t)$ be the covariance matrix of $\mathbf{X}_\mathcal{S}(t)$ and $\mathbf{u}_\mathcal{S}(t)$ respectively for the fixed t. Further, denote $\widetilde{\mathbf{\Sigma}}_{\mathbf{X}_\mathcal{S}} = \int \mathbf{\Sigma}_{\mathbf{X}_\mathcal{S}}(t)dt$ and $\widetilde{\mathbf{\Sigma}}_{\mathbf{u}_\mathcal{S}} = \int \mathbf{\Sigma}_{\mathbf{u}_\mathcal{S}}(t)dt$. Because $\mathbf{B}_\mathcal{S} = \mathbf{0}$, we have

$$\mathbf{X}_\mathcal{S}(t) = (\mathbf{B}_\mathcal{S} + \mathbf{M}_\mathcal{S}\boldsymbol{\eta})\mathbf{\Phi}'(t)\boldsymbol{\zeta} + \mathbf{u}_\mathcal{S}(t) = \mathbf{M}_\mathcal{S}\boldsymbol{\eta}\mathbf{\Phi}'(t)\boldsymbol{\zeta} + \mathbf{u}_\mathcal{S}(t).$$

Thus, $(\boldsymbol{\eta}^{(1)}, \mathbf{\Phi}^{(1)}(t), \boldsymbol{\zeta}^{(1)})$ and $(\boldsymbol{\eta}^{(2)}, \mathbf{\Phi}^{(2)}(t), \boldsymbol{\zeta}^{(2)})$ satisfy $\mathbf{M}_\mathcal{S}\boldsymbol{\eta}^{(1)}\mathbf{\Phi}^{(1)\prime}(t)\boldsymbol{\zeta}^{(1)} = \mathbf{M}_\mathcal{S}\boldsymbol{\eta}^{(2)}\mathbf{\Phi}^{(2)\prime}(t)\boldsymbol{\zeta}^{(2)}$.

By conditions (I3) and (I4),

$$\widetilde{\mathbf{\Sigma}}_{\mathbf{X}_\mathcal{S}} = \mathbf{M}_\mathcal{S}\boldsymbol{\eta}\mathbf{\Lambda}_{\boldsymbol{\zeta}}\boldsymbol{\eta}'\mathbf{M}_\mathcal{S}' + \widetilde{\mathbf{\Sigma}}_{\mathbf{u}_\mathcal{S}}, \tag{S1}$$

where $\mathbf{\Lambda}_{\boldsymbol{\zeta}} = \text{diag}\big\{\sum_{k=1}^K \text{var}(\xi_{1k}), \cdots, \sum_{k=1}^K \text{var}(\xi_{qk})\big\}$ is a diagonal matrix. By condition (I1), we have $\mathbf{M}_\mathcal{S}'\mathbf{M}_\mathcal{S}$ is invertible. Multiplying both sides of (S1) on the left by $(\mathbf{M}_\mathcal{S}'\mathbf{M}_\mathcal{S})^{-1}\mathbf{M}_\mathcal{S}'$ and on the right by $\mathbf{M}_\mathcal{S}(\mathbf{M}_\mathcal{S}'\mathbf{M}_\mathcal{S})^{-1}$, we have

$$(\mathbf{M}_\mathcal{S}'\mathbf{M}_\mathcal{S})^{-1}\mathbf{M}_\mathcal{S}'\widetilde{\mathbf{\Sigma}}_{\mathbf{X}_\mathcal{S}}\mathbf{M}_\mathcal{S}(\mathbf{M}_\mathcal{S}'\mathbf{M}_\mathcal{S})^{-1} - (\mathbf{M}_\mathcal{S}'\mathbf{M}_\mathcal{S})^{-1}\mathbf{M}_\mathcal{S}'\widetilde{\mathbf{\Sigma}}_{\mathbf{u}_\mathcal{S}}\mathbf{M}_\mathcal{S}(\mathbf{M}_\mathcal{S}'\mathbf{M}_\mathcal{S})^{-1} = \boldsymbol{\eta}\mathbf{\Lambda}_{\boldsymbol{\zeta}}\boldsymbol{\eta}',$$

and this imlies $\boldsymbol{\eta}^{(1)}\mathbf{\Lambda}_{\boldsymbol{\zeta}}^{(1)}\boldsymbol{\eta}^{(1)\prime} = \boldsymbol{\eta}^{(2)}\mathbf{\Lambda}_{\boldsymbol{\zeta}}^{(2)}\boldsymbol{\eta}^{(2)\prime}$. By conditions (I2) and (I3), we have $\boldsymbol{\eta}^{(1)\prime}\boldsymbol{\eta}^{(1)} = \boldsymbol{\eta}^{(2)\prime}\boldsymbol{\eta}^{(2)} = \mathbf{I}_q$ and $\mathbf{\Lambda}_{\boldsymbol{\zeta}}^{(1)}$ and $\mathbf{\Lambda}_{\boldsymbol{\zeta}}^{(2)}$ are both diagonal matrix with decreasing element. The first q eigenvectors associated with the first q largest eigenvalues of the matrix $(\mathbf{M}_\mathcal{S}'\mathbf{M}_\mathcal{S})^{-1}\mathbf{M}_\mathcal{S}'\widetilde{\mathbf{\Sigma}}_{\mathbf{X}_\mathcal{S}}$ $\mathbf{M}_\mathcal{S}(\mathbf{M}_\mathcal{S}'\mathbf{M}_\mathcal{S})^{-1} - (\mathbf{M}_\mathcal{S}'\mathbf{M}_\mathcal{S})^{-1}\mathbf{M}_\mathcal{S}'\widetilde{\mathbf{\Sigma}}_{\mathbf{u}_\mathcal{S}}\mathbf{M}_\mathcal{S}(\mathbf{M}_\mathcal{S}'\mathbf{M}_\mathcal{S})^{-1}$ are thus determined by $(\boldsymbol{\eta}^{(1)}, \mathbf{\Lambda}_{\boldsymbol{\zeta}}^{(1)})$ and $(\boldsymbol{\eta}^{(2)}, \mathbf{\Lambda}_{\boldsymbol{\zeta}}^{(2)})$. According to the uniqueness of the matrix eigen decomposition, we have $\boldsymbol{\eta}^{(1)} = \boldsymbol{\eta}^{(2)}$ and $\mathbf{\Lambda}_{\boldsymbol{\zeta}}^{(1)} = \mathbf{\Lambda}_{\boldsymbol{\zeta}}^{(2)}$.

Because $\boldsymbol{\eta}^{(1)} = \boldsymbol{\eta}^{(2)}$, we have $\mathbf{M}_\mathcal{S}\boldsymbol{\eta}\mathbf{\Phi}^{(1)\prime}(t)\boldsymbol{\zeta}^{(1)} = \mathbf{M}_\mathcal{S}\boldsymbol{\eta}\mathbf{\Phi}^{(2)\prime}(t)\boldsymbol{\zeta}^{(2)}$. Along with the rank constraints of $\boldsymbol{\eta}$ and $\mathbf{M}_\mathcal{S}$ in conditions (I1) and (I2), this implies $\mathbf{\Phi}^{(1)\prime}(t)\boldsymbol{\zeta}^{(1)} = \mathbf{\Phi}^{(2)\prime}(t)\boldsymbol{\zeta}^{(2)}$.

Then, we consider the covariance function matrix of $\boldsymbol{\Phi}'(t)\boldsymbol{\zeta}$. By simple calculation, we have

$$\mathrm{cov}\{\boldsymbol{\Phi}'(t)\boldsymbol{\zeta}, \boldsymbol{\Phi}'(s)\boldsymbol{\zeta}\} = \mathrm{diag}\left\{\sum_{k=1}^{K}\mathrm{var}(\xi_{1k})\phi_{1k}(t)\phi_{1k}(s), \cdots, \sum_{k=1}^{K}\mathrm{var}(\xi_{qk})\phi_{qk}(t)\phi_{qk}(s)\right\}.$$

Then, for each $j = 1, \cdots, q$, we have

$$\boldsymbol{\Phi}_j^{(1)\prime}(t)\boldsymbol{\Lambda}_{\boldsymbol{\zeta},j}^{(1)}\boldsymbol{\Phi}_j^{(1)}(s) = \boldsymbol{\Phi}_j^{(2)\prime}(t)\boldsymbol{\Lambda}_{\boldsymbol{\zeta},j}^{(2)}\boldsymbol{\Phi}_j^{(2)}(s), \tag{S2}$$

where $\boldsymbol{\Lambda}_{\boldsymbol{\zeta},j} = \mathrm{diag}\{\mathrm{var}(\xi_{j1}), \cdots, \mathrm{var}(\xi_{jK})\}$ is a diagonal matrix with decreasing elements. Multiplying both sides of equation (S2) on the left by $\boldsymbol{\Phi}_j^{(1)}(t)$ and on the right by $\boldsymbol{\Phi}_j^{(2)\prime}(s)$ and integrating with respect to t and s, then by condition (I4), we have

$$\boldsymbol{\Lambda}_{\boldsymbol{\zeta},j}^{(1)}\int \boldsymbol{\Phi}_j^{(1)}(t)\boldsymbol{\Phi}_j^{(2)\prime}(t)dt = \int \boldsymbol{\Phi}_j^{(1)}(t)\boldsymbol{\Phi}_j^{(2)\prime}(t)dt\boldsymbol{\Lambda}_{\boldsymbol{\zeta},j}^{(2)}. \tag{S3}$$

That is,

$$\begin{aligned}
&\begin{pmatrix}
\mathrm{var}(\xi_{j1}^{(1)})\int \phi_{j1}^{(1)}(t)\phi_{j1}^{(2)}(t)dt & \cdots & \mathrm{var}(\xi_{j1}^{(1)})\int \phi_{j1}^{(1)}(t)\phi_{jK}^{(2)}(t)dt \\
\vdots & \ddots & \vdots \\
\mathrm{var}(\xi_{jK}^{(1)})\int \phi_{jK}^{(1)}(t)\phi_{j1}^{(2)}(t)dt & \cdots & \mathrm{var}(\xi_{jK}^{(1)})\int \phi_{jK}^{(1)}(t)\phi_{jK}^{(2)}(t)dt
\end{pmatrix} \\
= &\begin{pmatrix}
\mathrm{var}(\xi_{j1}^{(2)})\int \phi_{j1}^{(1)}(t)\phi_{j1}^{(2)}(t)dt & \cdots & \mathrm{var}(\xi_{jK}^{(2)})\int \phi_{j1}^{(1)}(t)\phi_{jK}^{(2)}(t)dt \\
\vdots & \ddots & \vdots \\
\mathrm{var}(\xi_{j1}^{(2)})\int \phi_{jK}^{(1)}(t)\phi_{j1}^{(2)}(t)dt & \cdots & \mathrm{var}(\xi_{jK}^{(2)})\int \phi_{jK}^{(1)}(t)\phi_{jK}^{(2)}(t)dt
\end{pmatrix}.
\end{aligned} \tag{S4}$$

Because $\boldsymbol{\Lambda}_{\boldsymbol{\zeta},j}^{(1)}$ and $\boldsymbol{\Lambda}_{\boldsymbol{\zeta},j}^{(2)}$ are not equal to $\mathbf{0}$, it easy to show that the unique solution to (S4) is $\boldsymbol{\Lambda}_{\boldsymbol{\zeta},j}^{(1)} = \boldsymbol{\Lambda}_{\boldsymbol{\zeta},j}^{(2)}$ and $\int \boldsymbol{\Phi}_j^{(1)}(t)\boldsymbol{\Phi}_j^{(2)\prime}(t)dt$ is diagnoal but not equal to $\mathbf{0}$. Then, by (S3), we

have

$$\mathbf{\Lambda}_{\boldsymbol{\zeta},j} = \left\{\int \mathbf{\Phi}_j^{(1)}(t)\mathbf{\Phi}_j^{(2)\prime}(t)dt\right\}\mathbf{\Lambda}_{\boldsymbol{\zeta},j}\left\{\int \mathbf{\Phi}_j^{(2)}(s)\mathbf{\Phi}_j^{(1)\prime}(s)ds\right\} = \mathbf{\Lambda}_{\boldsymbol{\zeta},j}\left\{\int \mathbf{\Phi}_j^{(2)}(t)\mathbf{\Phi}_j^{(1)\prime}(t)dt\right\}^2,$$

which indicates the elements of the diagnoal matrix $\int \mathbf{\Phi}_j^{(1)}(t)\mathbf{\Phi}_j^{(2)\prime}(t)dt$ are only 1 or -1 (Without loss of generality, we assume they are both equal to 1). The last equation is because $\mathbf{\Lambda}_{\boldsymbol{\zeta},j}$ and $\int \mathbf{\Phi}_j^{(1)}(t)\mathbf{\Phi}_j^{(2)\prime}(t)dt$ are both diagnoal matrices. Then, multiplying both sides of equation (S2) on the left by $\mathbf{\Phi}_j^{(1)}(t)$ and integrating with respect to t, we have

$$\mathbf{\Lambda}_{\boldsymbol{\zeta},j}\mathbf{\Phi}_j^{(1)}(s) = \left\{\mathbf{\Phi}_j^{(1)}(t)\mathbf{\Phi}_j^{(1)\prime}(t)dt\right\}\mathbf{\Lambda}_{\boldsymbol{\zeta},j}\mathbf{\Phi}_j^{(1)}(s) = \left\{\int \mathbf{\Phi}_j^{(1)}(t)\mathbf{\Phi}_j^{(2)\prime}(t)dt\right\}\mathbf{\Lambda}_{\boldsymbol{\zeta},j}\mathbf{\Phi}_j^{(2)}(s) = \mathbf{\Lambda}_{\boldsymbol{\zeta},j}\mathbf{\Phi}_j^{(2)}(s).$$

Because $\mathbf{\Lambda}_{\boldsymbol{\zeta},j}$ is invertible, then $\mathbf{\Phi}_j^{(1)}(s) = \mathbf{\Phi}_j^{(2)}(s)$ for each $j = 1, \cdots, q$ and $\mathbf{\Phi}^{(1)}(s) = \mathbf{\Phi}^{(2)}(s)$. This implies

$$(\mathbf{B}^{(1)} + \mathbf{M}\boldsymbol{\eta})\left\{\int \mathbf{\Phi}'(t)\mathbf{\Phi}(t)dt\right\}\mathbb{E}(\boldsymbol{\zeta}\boldsymbol{\zeta}') = (\mathbf{B}^{(2)} + \mathbf{M}\boldsymbol{\eta})\left\{\int \mathbf{\Phi}'(t)\mathbf{\Phi}(t)dt\right\}\mathbb{E}(\boldsymbol{\zeta}\boldsymbol{\zeta}').$$

By the conditions (I3) and (I4), we have $\int \mathbf{\Phi}'(t)\mathbf{\Phi}(t)dt = K\mathbf{I}_q$ and $\mathbb{E}(\boldsymbol{\zeta}\boldsymbol{\zeta}')$ is invertible. Thus, we have $\mathbf{B}^{(1)} = \mathbf{B}^{(2)}$.

S2 Supplementary Material B: Calculation of (11)

Our goal is to obtain $\mathbf{B}^{(r+1)}$ based on (11). For simplicity of notation, we omit the superscript $(r+1)$ of $\mathbf{f}^{(r+1)}(\mathbf{s})$ and $\mathbf{h}_i^{(r+1)}(t)$. Denote $\mathbf{\Gamma} = (\mathbf{\Gamma}_{jj'})_{j'<j}$ and Lagrangian multipliers

$\boldsymbol{\nu} = (\boldsymbol{\nu}_{jj'})_{j'<j}$. Based on (11), we update $(\mathbf{B}, \boldsymbol{\Gamma}, \boldsymbol{\nu})$ by the following optimization problem:

$$\min_{\mathbf{B},\boldsymbol{\Gamma},\boldsymbol{\nu}} \sum_{j=1}^{p}\sum_{i=1}^{n}\sum_{l=1}^{n_i}\{X_{ij}(t_{il}) - \mathbf{f}'(\mathbf{s}_j)\mathbf{h}_i(t_{il}) - \mathbf{b}_j'\mathbf{h}_i(t_{il})\}^2 + \lambda_1\sum_{j=1}^{p}\|\mathbf{b}_j\|_2 + \lambda_2\sum_{j'<j=1}^{p} w(\|\mathbf{s}_j - \mathbf{s}_{j'}\|_2)\cdot\|\boldsymbol{\Gamma}_{jj'}\|_2 + \sum_{j'<j=1}^{p}\boldsymbol{\nu}_{jj'}'(\mathbf{b}_j - \mathbf{b}_{j'} - \boldsymbol{\Gamma}_{jj'}) + \sum_{j'<j=1}^{p}\frac{\mu}{2}\|\mathbf{b}_j - \mathbf{b}_{j'} - \boldsymbol{\Gamma}_{jj'}\|_2^2,$$

where μ is a pre-given hyperparameter.

Let $\mathbf{A}^{(t+1)}$ be the updation of $\mathbf{A}$ in the $(t+1)$-th step, $(x)_+ = \max(x, 0)$, and $\mathbf{A}_{a:b}$ represent the subvector of vector $\mathbf{A}$ corresponding to the positions a through b. We update $(\mathbf{b}_j^{(t+1)}, \boldsymbol{\Gamma}_{jj'}^{(t+1)}, \boldsymbol{\nu}_{jj'}^{(t+1)})$ $(j' < j = 1, \cdots, p)$ as follows:

$$\begin{aligned}
\mathbf{b}_j^{(t+1)} &= (\|\widetilde{\mathbf{b}}_j^{(t+1)}\|_2 - \lambda_1)_+ \cdot \frac{\widetilde{\mathbf{b}}_j^{(t+1)}}{\|\widetilde{\mathbf{b}}_j^{(t+1)}\|_2},\\
\boldsymbol{\Gamma}_{jj'}^{(t+1)} &= \left\{\|\widetilde{\boldsymbol{\Gamma}}_{jj'}^{(t+1)}\|_2 - \lambda_2 w(\|\mathbf{s}_j - \mathbf{s}_{j'}\|_2)\right\}_+ \cdot \frac{\widetilde{\boldsymbol{\Gamma}}_{jj'}^{(t+1)}}{\|\widetilde{\boldsymbol{\Gamma}}_{jj'}^{(t+1)}\|_2},\\
\boldsymbol{\nu}_{jj'}^{(t+1)} &= \boldsymbol{\nu}_{jj'}^{(t)} + \mu(\mathbf{b}_j^{(t+1)} - \mathbf{b}_{j'}^{(t+1)} - \boldsymbol{\Gamma}_{jj'}^{(t+1)}),
\end{aligned} \tag{S5}$$

where

$$\widetilde{\mathbf{b}}_j^{(t+1)} = \left(\left\{\mu\mathbf{D} + 2\sum_{i=1}^{n}\sum_{l=1}^{n}\widetilde{\mathbf{h}}_i(t_{il})\widetilde{\mathbf{h}}_i'(t_{il})\right\}^{-1}\left[2\sum_{i=1}^{n}\sum_{l=1}^{n}\widetilde{\mathbf{h}}_i(t_{il})\{\mathbf{X}_i(t_{il}) - \mathbf{f}\mathbf{h}_i(t_{il})\} - \mathbf{E}^{(t)}\right]\right)_{(j-1)q+1\,:\,jq},$$

$$\widetilde{\boldsymbol{\Gamma}}_{jj'}^{(t+1)} = \frac{\boldsymbol{\nu}_{jj'}^{(t)} + \mu(\mathbf{b}_j^{(t+1)} - \mathbf{b}_{j'}^{(t+1)})}{\mu},$$

and

$\mathbf{D} = (p\mathbf{I}_p - 1_{p\times p})\otimes\mathbf{I}_q$, $1_{p\times p}$ is a matrix with all elements being 1 and $\otimes$ denotes Kronecker product,

$$\mathbf{E}^{(t)} = \left\{\sum_{j'<j}\left(\boldsymbol{\nu}_{1j}^{(t)} - \boldsymbol{\nu}_{j1}^{(t)} - \boldsymbol{\Gamma}_{1j}^{(t)} + \boldsymbol{\Gamma}_{j1}^{(t)}\right), \cdots, \sum_{j'<j}\left(\boldsymbol{\nu}_{pj}^{(t)} - \boldsymbol{\nu}_{jp}^{(t)} - \boldsymbol{\Gamma}_{pj}^{(t)} + \boldsymbol{\Gamma}_{jp}^{(t)}\right)\right\} \in \mathbb{R}^{pq},$$

$\widetilde{\mathbf{h}}_i(t) = \mathrm{diag}\{\mathbf{h}_i(t), \cdots, \mathbf{h}_i(t)\} \in \mathbb{R}^{pq\times p}$.

We update (S5) until convergence and obtain the final $\mathbf{B}$ as the solution of (11).

Among the cross-validation procedure, the components $\mathbf{B}, \mathbf{\Phi}(\cdot)$ and $\mathbf{f}(\cdot)$ are shared among all the individuals and are estimated by the training set. In the test set, we only update $\boldsymbol{\zeta}$ because it is specific to individuals. Particularly, for $\boldsymbol{\zeta}_i$ in the test data, by model (4), we estimate $\boldsymbol{\zeta}_i$ by

$$\boldsymbol{\zeta}_i^{test} = \Big\{ \sum_{l=1}^{n_i} \mathbf{\Phi}^{trian}(t_{il}) \mathbf{\Phi}^{trian\prime}(t_{il}) \Big\}^{-1} \Big[\sum_{l=1}^{n_i} \mathbf{\Phi}^{trian}(t_{il}) \{ (\mathbf{B}^{trian} + \mathbf{f}^{trian})' \times (\mathbf{B}^{trian} + \mathbf{f}^{trian}) \}^{-1} (\mathbf{B}^{trian} + \mathbf{f}^{trian})' \mathbf{X}_i^{test}(t_{il}) \Big],$$

where the superscripts "*train*" and "*test*" represent the components obtained from the training set and the test set, respectively.

S3 Supplementary Material C: Data Description and Explanatory Analysis

ADNI is a longitudinal multicenter study involving scientists from 59 research centers across the United States and Canada. It collects imaging, genetic, clinical, and cognitive data from subjects diagnosed as cognitively normal (CN), with mild cognitive impairment (MCI), and with AD. The study aims to investigate the association between AD and biomarkers, neuropsychological scores, medical imaging data, and genetic variants. It spans four stages: ADNI-1, ADNI-GO, ADNI-2, and ADNI-3. The data considered in this paper are sourced from the ADNI-GO and ADNI-3 studies, which commenced in 2009 and 2016, respectively. In addition to MRI data and MMSE scores, we extract age and gender information from the dataset. The dataset comprises 803 participants, including 413 males and 390 females, with an average age of 81.6 years at the last assessments, ranging from 55 to 103 years.

Among the participants, 292 individuals were diagnosed with AD, 213 with MCI, and 298 were CN.

The transition from CN to MCI and ultimately to AD is often accompanied by progressive atrophy of specific brain structures (Fotenos et al., 2005). Therefore, studying brain volume data can contribute to a better understanding and detection of the pathological processes associated with AD. MRI, utilized in ADNI, is a brain imaging technique primarily used for evaluating and analyzing the structure and function of various brain regions. It allows for the examination of anatomical structures, neuronal density, and metabolic conditions within specific regions of the brain. MRI contains a massive amount of information, posing challenges in accurately assessing the impact of volume changes of ROIs on AD. Brain volume data are more effective in detecting subtle changes in brain structure, thereby enabling a more accurate assessment of the impact of ROI volume changes on AD. Numerous studies (Ferrarini et al., 2008; Henneman et al., 2009; Evans et al., 2010) have used volumetric measurements obtained from MRI to investigate the impact of brain atrophy on cognitive function. Hence, we perform preprocessing on the original MRI data and extract relevant information pertaining to the volumes of the ROIs.

The original MRI data are preprocessed using advanced normalization tools (Avants et al., 2011). The preprocessing procedure consists of several steps: N4 bias correction, registration-based brain extraction, and a prior-based N4-Atropos 6 tissue segmentation using the oasis template. By performing multi-atlas cortical parcellation, we obtain the brain local volumetric measures of 101 ROIs defined by the manually edited labels of the publicly available MindBoggle-101 dataset (Klein and Tourville, 2012). After excluding subjects whose imaging data do not pass standard imaging quality controls and removing six ROIs with large missing values, the log-Jacobian transformation map for each individual in the standard space is divided into 95 ROIs. The 3D coordinates of the center for each ROI are determined by minimizing the distance between all measured loci and the center.

To address the limitations of density functions not residing in a linear space, which actually is required by model (2), we employ the log quantile density transformation (Petersen and Müller, 2016; Li et al., 2023) and take the log quantile density function of the density curves along the 95 ROIs as functional variables. This transformation allows us to effectively represent the density curves as functional variables suitable for analysis. Concretely, let $f, F, Q = F^{-1}$ represent the density function, the cumulative distribution function and the quantile functions of the volume. The log quantile density transformation is given by $\log\{dQ(t)/dt\} = \log\{dF^{-1}(t)/dt\} = -\log[f\{Q(t)\}]$. We randomly sample curves from female participants aged 80 to 90 years old, belonging to the CN and non-CN groups, in the left cerebellum exterior region. Our analysis using the proposed SF-FPCA in Section 4 reveals that atrophy of the left cerebellum exterior exacerbates functional decline. Figure S1(a) illustrates the differences in local volumes between individuals with cognitive impairment and normal subjects, showing evidence of brain volume change among those with cognitive impairment. The brain volume curves provide an intuitive and comprehensive representation of brain volume values for each ROI, retaining essential information such as mode, spread, and shape of these densities.

The MMSE score is commonly used to measure cognitive ability, with lower scores indicating cognitive impairment. In our subsequent regression analysis, we treat the MMSE score as the response variable. The average MMSE score for the participants is 26.5, with a maximum score of 30 and a minimum score of 4. Moreover, age and gender have been validated as important features influencing cognitive function (Gao et al., 1998) and can serve as additional covariates in the subsequent regression analysis. Figure S1(b) displays the density of MMSE scores among individuals with AD, differentiated by gender, revealing that females are more likely to have lower MMSE scores (8-11) and less likely to have higher MMSE scores (18-26) than males. Figure S1(c) illustrates the density of MMSE scores across different age groups in individuals with AD, suggesting a decline in cognitive

function with increasing age.

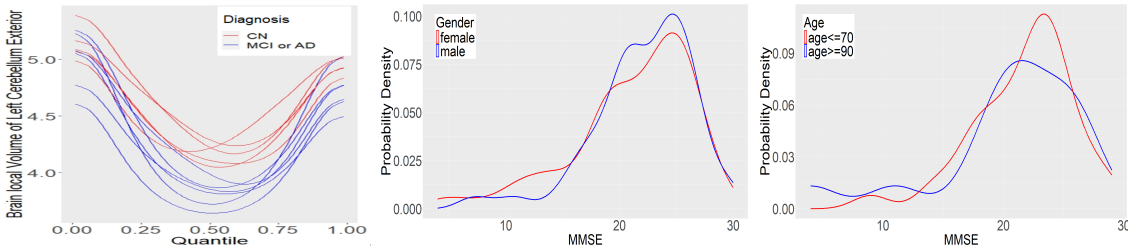


Figure S1: (a): Curves of left cerebellum exterior in selected females with CN (red) or not (blue) between the ages of 80 and 90; (b): The probability density of MMSE scores for females (red) and males (blue) with AD; (c): The probability density of MMSE scores for individuals with AD categorized into two age groups: younger than 70 years old (red) and older than 90 years old (blue).

S4 Supplementary Material D: Simulation results related to the performance of SF-FPCA

We provide supplementary simulation results related to the performance of SF-FPCA, which includes estimation consistency, selection of the number of factors and eigenfunctions, and computational costs.

We generate data from the SF-FPCA model as $\mathbf{X}_i(t) = (\mathbf{B} + \mathbf{f})\mathbf{\Phi}'(t)\boldsymbol{\zeta}_i + \mathbf{u}_i(t)$. To construct $\mathbf{B}$, we set $\mathbf{b}_j = p^{1/2}/2(1, \cdots, 1)'$ for $j = 1, \cdots, 100$ and $\mathbf{b}_j = \mathbf{0}$ for $j = 101, \cdots, p$. To construct $\mathbf{f}(\cdot) = \boldsymbol{\eta}'\mathbf{M}(\cdot)$, we generate $\mathbf{M}(\mathbf{s}) = \widetilde{\mathbf{M}}(s_1) \otimes \widetilde{\mathbf{M}}(s_2) \otimes \widetilde{\mathbf{M}}(s_3)$ as the Kronecker product of three B-spline basis functions, with each basis function corresponds to one dimension of $\mathbf{s} = (s_1, s_2, s_3)'$. Each dimension of the coordinates $\mathbf{s}_j = (s_{j1}, s_{j2}, s_{j3})'$ for location $j = 1, \cdots, p$ is generated from $U(0, 1)$. For $\boldsymbol{\eta} \in \mathbb{R}^\tau$, we firstly generate q random τ-dimensional vectors from $N(\mathbf{0}_\tau, \mathbf{I}_\tau)$ and then apply singular value decomposition (SVD) to obtain the final $\boldsymbol{\eta}$. The components of $\mathbf{\Phi}(t)$ are defined as $\phi_{jk}(t) = \sqrt{2}\sin\{(k+1)\pi t\}$

if k is odd and $\sqrt{2}\cos(k\pi t)$ if k is even $(j = 1, \cdots, q)$. We generate $\boldsymbol{\zeta}_i \sim N(\mathbf{0}_{Kq}, \boldsymbol{\Sigma}_{\boldsymbol{\zeta}})$, where $\boldsymbol{\Sigma}_{\boldsymbol{\zeta}} = \text{diag}(\xi_{ijk})_{j=1,\cdots,q;k=1,\cdots,K}$ and $\text{var}(\xi_{ijk}) = \{3(q+1-j)-1\}(K-k)/(K-1)+1$. We generate $\mathbf{u}_i(t) \sim N(\mathbf{0}_p, 3\mathbf{I}_p)$. For each trajectory $X_{ij}(\cdot)$, we randomly sample 100 observation time points from $U(0,1)$, unless stated otherwise.

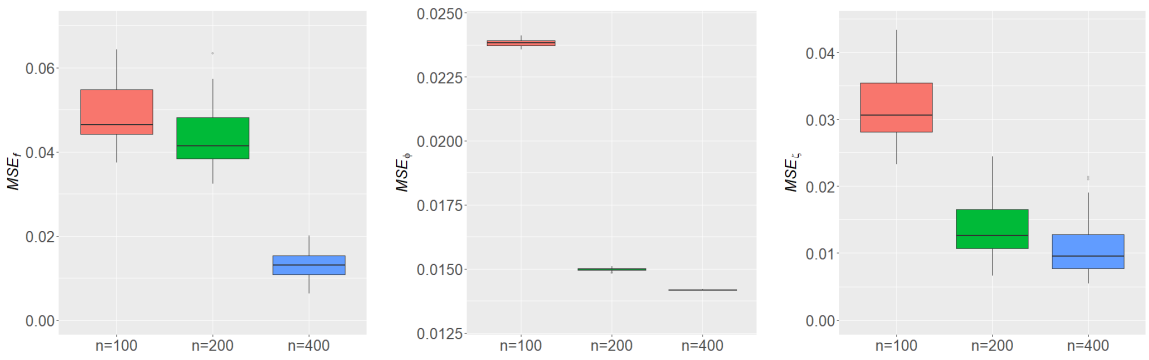
Estimation consistency

We set $(p, q, K) = (500, 2, 2)$ and vary n among $(100, 200, 400)$ to check whether all components of the model can be consistently estimated. To check whether all the pieces specified by $\mathbf{B}$ are consistently identified, we introduce the normalized mutual information (NMI), which is a common measure for similarity between clusterings (Ke et al., 2015). Suppose $\mathcal{D}_1 = \{D_{11}, D_{12}, \cdots\}$ and $\mathcal{D}_2 = \{D_{21}, D_{22}, \cdots\}$ are two sets of disjoint clusters of $\{1, \cdots, p\}$, we define

$$\text{NMI}(\mathcal{D}_1, \mathcal{D}_2) = \frac{2\mathcal{I}(\mathcal{D}_1; \mathcal{D}_2)}{\{H(\mathcal{D}_1) + H(\mathcal{D}_2)\}},$$

where $\mathcal{I}(\mathcal{D}_1; \mathcal{D}_2) = \sum_{k,j}(|D_{1k} \cap D_{2j}|/p)\log\{p|D_{1k} \cap D_{2j}|/(|D_{1k}||D_{2j}|)\}$ is the mutual information between $\mathcal{D}_1$ and $\mathcal{D}_2$, and $H(\mathcal{D}_1) = -\sum_k(|D_{1k}|/p)\log(|D_{1k}|/p)$ is the entropy of $\mathcal{D}_1$. NMI takes values on $[0, 1]$, and large NMI implies that the two grouping structures are close. In our scenario, $\mathcal{D}_1$ denotes the estimated grouping structure by SF-FPCA and $\mathcal{D}_2$ is the true one. The performance of $\widehat{\mathbf{f}}(\cdot), \widehat{\boldsymbol{\Phi}}(\cdot)$ and $\widehat{\boldsymbol{\zeta}}$ are evaluated via the mean square error $\text{MSE}_{\mathbf{f}} = (pq)^{-1}\|\widehat{\mathbf{f}} - \mathbf{f}_0\|_F^2$, $\text{MSE}_{\boldsymbol{\Phi}} = (qK)^{-1}\sum_{j,k}\int\{\widehat{\phi}_{jk}(t) - \phi_{jk0}(t)\}^2\, dt$, and $\text{MSE}_{\boldsymbol{\zeta}} = (nqK)^{-1}\sum_{i=1}^n \|\widehat{\boldsymbol{\zeta}}_i - \boldsymbol{\zeta}_{i0}\|_2^2$.

Based on 200 repetitions, the mean(sd) of NMI values are 0.9243(0.0434), 0.9604(0.0407), and 1(0) for the scenarios with $n = 100, 200, 400$, respectively, which indicates the pieces can be well identified in each setting. The results in Figure S2 show that as n increases, the estimates of $\mathbf{f}(\cdot), \boldsymbol{\Phi}(\cdot)$ and $\boldsymbol{\zeta}$ become more precise, demonstrating that the proposed model can be consistently estimated.



(a) $\text{MSE}_{\mathbf{f}}$ (b) $\text{MSE}_{\mathbf{\Phi}}$ (c) $\text{MSE}_{\boldsymbol{\zeta}}$

Figure S2: The performance of the estimates of $\mathbf{f}(\cdot), \mathbf{\Phi}(\cdot)$ and $\boldsymbol{\zeta}$ over 200 repetitions of SF-FPCA when n varies

Selection of the number of factors and eigenfunctions

We check if the criteria work well by setting $(n, p) = (100, 500)$ and varying q among $(2, 3, 4, 5, 6)$ and K among (2,4,6,8,10). The results in Tables S1 and S2 demonstrate that, under all settings, we accurately identify the true number of factors and eigenfunctions using the criteria (15) and (16).

Table S1: The values of $\widehat{q}$ in different settings with fixed K.

	$q_0 = 2$	$q_0 = 3$	$q_0 = 4$	$q_0 = 5$	$q_0 = 6$
$\widehat{q}$(sd)	2(0)	3(0)	4(0)	5(0)	6(0)

Table S2: The values of $\widehat{K}$ in different settings with fixed q.

	$K_0 = 2$	$K_0 = 4$	$K_0 = 6$	$K_0 = 8$	$K_0 = 10$
$\widehat{K}$(sd)	2(0)	4(0)	6(0)	8(0)	10(0)

Computational costs

The results in Tables S3 and S4 show the computation time required for convergence under each setting with 10 repetitions and are computed in R (version 4.4.0) on a 14-core machine with 32 GB of RAM.

First, we examine the effect of the number of factors q and the number of eigenfunctions

K on computational cost by varying q and K across $(2, 5, 10, 20)$, while keeping (n, p, n_i) fixed at $(100, 100, 50)$. The results in Table S3 show that variations in K have minimal impact on computational cost, as only $\mathbf{\Phi}(\cdot)$ and $\boldsymbol{\zeta}$ depend on K, and their computation involves only eigenvalue decomposition truncated at the first K eigenvalues. In contrast, changes in q significantly increase computation time, mainly due to the high cost of updating the inverses of several $q \times q$ matrices.

Table S3: Average computation time (seconds) over 10 repetitions given different q, K.

	$K = 2$	$K = 5$	$K = 10$	$K = 20$
$q = 2$	1.27	1.28	1.27	0.89
$q = 5$	5.32	4.24	5.08	4.64
$q = 10$	15.70	15.37	15.43	15.75
$q = 20$	72.47	70.64	74.02	74.36

Second, we examine the impact of sample size n, number of variables p, and number of observations n_i on computational cost by varying n, p and n_i across $(100, 200, 500)$, while keeping (q, K) fixed at $(2, 2)$. The results in Table S4 show that increasing n, p and n_i all raise computational costs, with changes in p having a much more significant impact than those in n or n_i. This is mainly due to the ADMM algorithm used in the iteration for $\mathbf{B}$, which involves estimating $p^2/2 \times q$ parameters and thus consumes the largest portion of computation time.

Table S4: Average computation time (seconds) over 10 repetitions given different n, p, n_i.

	$n_i = 100$			$n_i = 200$			$n_i = 500$		
	$p = 100$	$p = 200$	$p = 500$	$p = 100$	$p = 200$	$p = 500$	$p = 100$	$p = 200$	$p = 500$
$n = 100$	1.81	4.21	21.40	2.76	5.03	25.77	3.22	7.45	33.44
$n = 200$	2.55	5.00	24.07	2.67	5.97	28.32	4.90	11.90	40.38
$n = 500$	2.56	6.55	31.19	4.28	10.27	34.92	9.33	20.74	59.02

S5 Supplementary Material E: The transformation from the regression relationships between MMSE and scores to ROI measures.

With the regression coeffcients corresponding to $\boldsymbol{\zeta}_i$ denoted by $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_{Kq})'$, $\boldsymbol{\zeta}_i'\boldsymbol{\beta}$ is viewed as the measurement of the effect of the ROIs on MMSE. By multiplying $\boldsymbol{\Phi}(t)(\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'$ on both sides of (4) with $\mathbf{C} = \mathbf{B} + \mathbf{f}$, we obtain $\boldsymbol{\Phi}(t)(\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'\mathbf{X}_i(t) \approx \boldsymbol{\Phi}(t)\boldsymbol{\Phi}'(t)\boldsymbol{\zeta}_i$. After combining it with the identification condition $\int \boldsymbol{\Phi}(t)\boldsymbol{\Phi}'(t)dt = \mathbf{I}_{Kq}$, we have

$$\int \boldsymbol{\beta}'\boldsymbol{\Phi}(t)(\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}'\mathbf{X}_i(t)dt \approx \boldsymbol{\beta}' \int \boldsymbol{\Phi}(t)\boldsymbol{\Phi}'(t)dt\boldsymbol{\zeta}_i = \boldsymbol{\beta}'\boldsymbol{\zeta}_i.$$

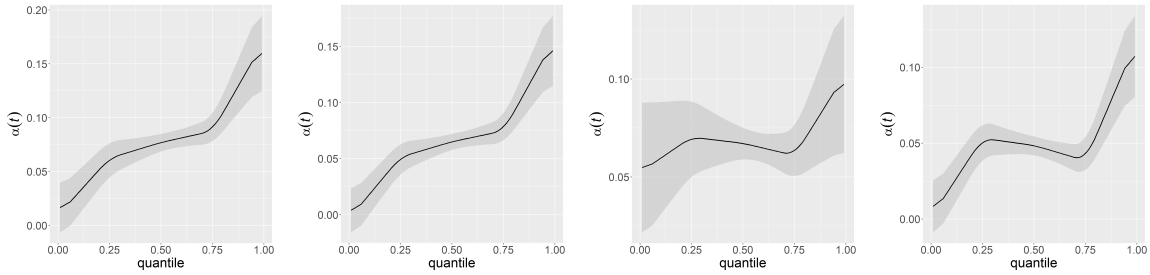
Then, the regression relationship $\boldsymbol{\zeta}_i'\boldsymbol{\beta}$ between response and score $\boldsymbol{\zeta}_i$ can be written as $\int \mathbf{X}_i'(t)\boldsymbol{\alpha}(t)dt$ between response and original functional covariates $\mathbf{X}_i(t)$, where $\boldsymbol{\alpha}(t) = \{\alpha_1(t), \cdots, \alpha_p(t)\}' = \mathbf{C}(\mathbf{C}'\mathbf{C})^{-1}\boldsymbol{\Phi}'(t)\boldsymbol{\beta}$ is the regression coefficient function.

S6 Supplementary Material F: Table and figures in the real data analysis

Table S5 shows the used 95 ROIs in the real data analysis and their pieces based on $\widehat{\mathbf{B}}$. Figures S3, S4 and S5 show the coefficient functions of the selected ROIs in temporal lobe, frontal lobe and cerebellum. Figure S6 shows the coefficient function of other 19 ROIs that are the risk factors for AD.

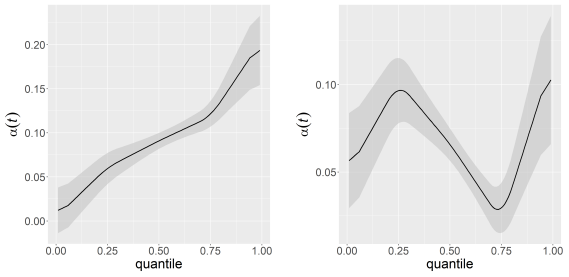
Table S5: The used 95 ROIs in anaylsis and their pieces.

ROI	name	ROI	name	ROI	name
Piece 1					
4	left lateral ventricle	50	right caudate	1009	left inferior temporal
5	left inferior lateral ventricle	91	left basal forebrain	1011	left lateral occipital
10	left thalamus proper	630	cerebellar vermal lobules I-V	2002	right caudal anterior cingulate
11	left caudate	631	cerebellar vermal lobules VI-VII	2011	right lateral occipital
14	3rd ventricle	632	cerebellar vermal lobules VIII-X	2013	right lingual
15	4th ventricle	1002	left caudal anterior cingulate	2021	right pericalcarine
17	left hippocampus	1005	left cuneus	2025	right precuneus
26	left accumbens area	1007	left fusiform	2026	right rostral anterior cingulate
43	right lateral ventricle	1008	left inferior parietal		
Piece 2					
6	left cerebellum exterior	1012	left lateral orbitofrontal	2003	right caudal middle frontal
12	left putamen	1015	left middle temporal	2005	right cuneus
45	right cerebellum exterior	1018	left pars opercularis	2008	right inferior parietal
46	right cerebellum white matter	1020	left pars triangularis	2010	right isthmus cingulate
49	right thalamus proper	1021	left pericalcarine	2012	right lateral orbitofrontal
51	right putamen	1024	left precentral	2018	right pars opercularis
52	right pallidum	1026	left rostral anterior cingulate	2020	right pars triangularis
54	right amygdala	1027	left rostral middle frontal	2024	right precentral
58	right accumbens area	1029	left superior parietal	2027	right rostral middle frontal
60	right ventral DC	1034	left transverse temporal		
Piece 3					
7	left cerebellum white matter	1013	left lingual	2006	right entorhinal
13	left pallidum	1014	left medial orbitofrontal	2007	right fusiform
16	Brain stem	1016	left parahippocampal	2009	right inferior temporal
18	left amygdala	1017	left paracentral	2016	right parahippocampal
24	CSF	1019	left pars orbitalis	2017	right paracentral
28	left ventral DC	1022	left postcentral	2022	right postcentral
44	right inferior lateral ventricle	1023	left posterior cingulate	2028	right superior frontal
53	right hippocampus	1025	left precuneus	2029	right superior parietal
92	right basal forebrain	1028	left superior frontal	2030	right superior temporal
1003	left caudal middle frontal	1030	left superior temporal	2031	right supramarginal
1006	left entorhinal	1031	left supramarginal		
1010	left isthmus cingulate	1035	left insula		
Piece 4					
2014	right medial orbitofrontal	2035	right insula		
Piece 5					
2015	right middle temporal				
Piece 6					
2019	right pars orbitalis				
Piece 7					
2023	right posterior cingulate				
Piece 8					
2034	right transverse temporal				

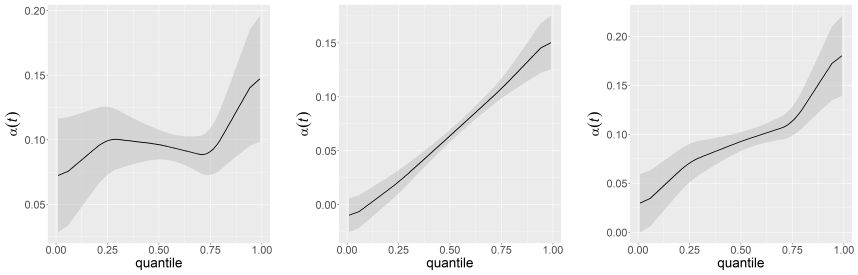


(a) left middle temporal (b) left transverse temporal (c) right inferior temporal (d) right fusiforml

Figure S3: Functional regression coefficient estimates in temporal lobe (a): Left middle temporal; (b): Left transverse temporal; (c): Right inferior temporal; (d): Right fusiforml.

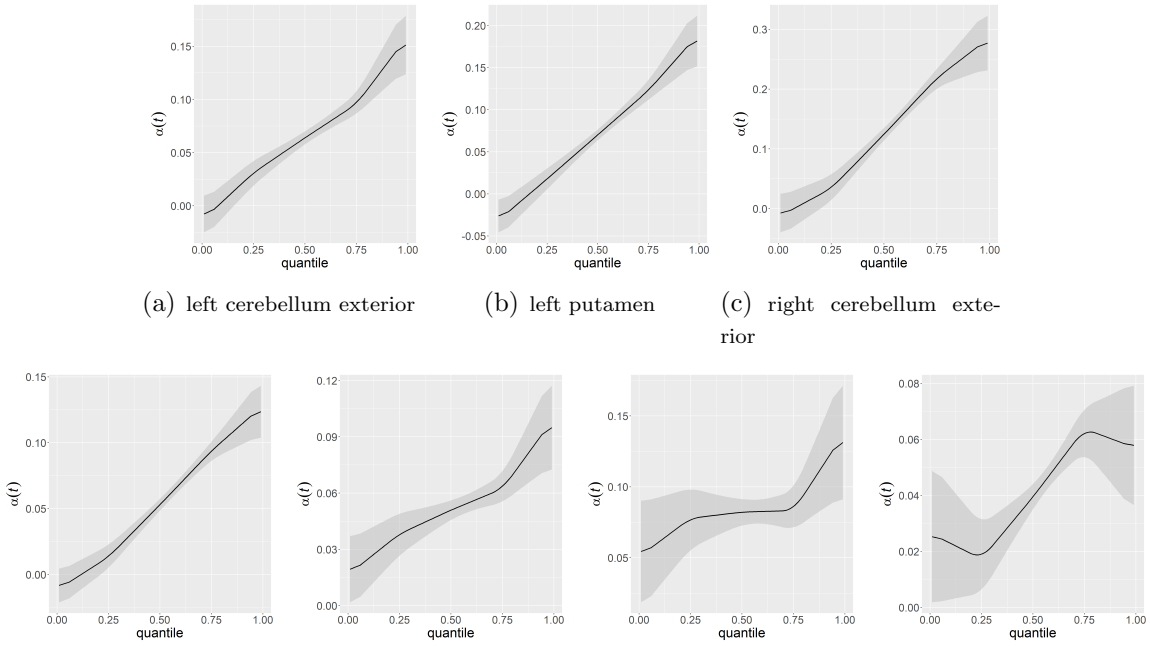


(a) left lateral orbitofrontal (b) left medial orbitofrontal



(c) right superior frontal (d) left pars opercularis (e) left pars triangularis

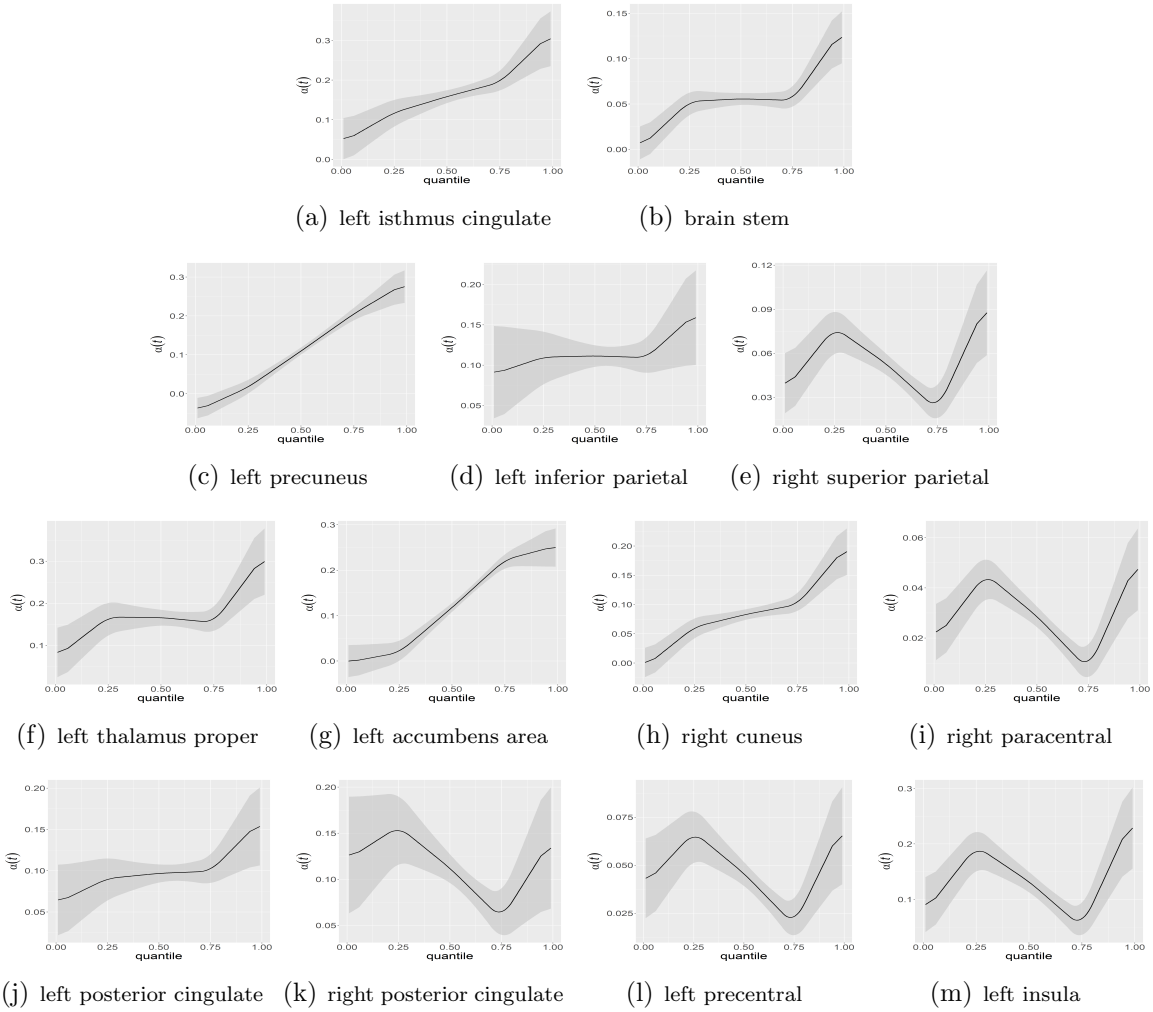
Figure S4: Functional regression coefficient estimates in frontal lobe (a): Left lateral orbitofrontal; (b): Left medial orbitofrontal; (c): Right superior frontal; (d): Left pars opercularis; (e): Left pars triangularis.



(a) left cerebellum exterior (b) left putamen (c) right cerebellum exterior

(d) right putamen (e) left ventral DC (f) cerebellar vermal lobules I-V (g) cerebellar vermal lobules VI-VI

Figure S5: Functional regression coefficient estimates in cerebellum (a): Left cerebellum exterior; (b): Left putamen; (c): Right cerebellum exterior; (d): Right putamen; (e): Left ventral DC; (f): Cerebellar vermal lobules I-V; (g): Cerebellar vermal lobules VI-VII.



(a) left isthmus cingulate (b) brain stem

(c) left precuneus (d) left inferior parietal (e) right superior parietal

(f) left thalamus proper (g) left accumbens area (h) right cuneus (i) right paracentral

(j) left posterior cingulate (k) right posterior cingulate (l) left precentral (m) left insula

Figure S6: Functional regression coefficient estimates and confidence bands of other 13 ROIs.

References

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., and Gee, J. C. (2011). A reproducible evaluation of ants similarity metric performance in brain image registration. *NeuroImage*, 54(3):2033–2044.

Bing, X., Cheng, W., Feng, H., and Ning, Y. (2024). Inference in high-dimensional multivariate response regression with hidden variables. *J. Am. Stat. Assoc.*, 119(547):2066–2077.

Evans, M. C., Barnes, J., Nielsen, C., Kim, L. G., Clegg, S. L., Blair, M., Leung, K. K., Douiri, A., Boyes, R. G., Ourselin, S., and Initiative, A. D. N. (2010). Volume changes in Alzheimer's disease and mild cognitive impairment: cognitive associations. *Eur. Radiol.*, 20(3):674–682.

Ferrarini, L., Palm, W. M., Olofsen, H., van der Landen, R., Blauw, G. J., Westendorp, R. G. J., Bollen, E. L. E. M., Middelkoop, H. A. M., Reiber, J. H. C., van Buchem, M. A., and Admiraal-Behloul, F. (2008). MMSE scores correlate with local ventricular enlargement in the spectrum from cognitively normal to Alzheimer disease. *NeuroImage*, 39(4):1832–1838.

Fotenos, A. F., Snyder, A. Z., Girton, L. E., Morris, J. C., and Buckner, R. L. (2005). Normative estimates of cross-sectional and longitudinal brain volume decline in aging and AD. *Neurology*, 64(6):1032–1039.

Gao, S., Hendrie, H. C., Hall, K. S., and Hui, S. (1998). The relationships between age, sex, and the incidence of dementia and Alzheimer disease: a meta-analysis. *Arch. Gen. Psychiatry.*, 55(9):809–815.

Henneman, W. J., Sluimer, J. D., Barnes, J., van der Flier, W. M., Sluimer, I. C., Fox,

N. C., Scheltens, P., Vrenken, H., and Barkhof, F. (2009). Hippocampal atrophy rates in Alzheimer disease. *Neurology*, 72(11):999–1007.

Ke, T., Fan, J., and Wu, Y. (2015). Homogeneity in regression. *J. Am. Stat. Assoc.*, 110(509):175–194.

Klein, A. and Tourville, J. (2012). 101 labeled brain images and a consistent human cortical labeling protocol. *Front. Neurosci.*, 6(8):171.

Li, T., Zhu, H., Li, T., and Zhu, H. (2023). Asynchronous functional linear regression models for longitudinal data in reproducing kernel Hilbert space. *Biometrics*, 79(3):1880–1895.

Petersen, A. and Müller, H. G. (2016). Functional data analysis for density functions by transformation to a Hilbert space. *Ann. Stat.*, 44(1):183–218.

Wang, J., Zhao, Q., Hastie, T., and Owen, A. B. (2017). Confounder adjustment in multiple hypothesis testing. *Ann. Stat.*, 45(5):1863–1894.