

# Penalized logistic regression using functional connectivity as covariates with an application to mild cognitive impairment

Jae-Hwan Jung<sup>a</sup>, Seong-Jin Ji<sup>a</sup>, Hongtu Zhu<sup>b</sup>, Joseph G. Ibrahim<sup>b</sup>, Yong Fan<sup>c</sup>,  
Eunjee Lee<sup>1,a</sup>, for the Alzheimer's Disease Neuroimaging Initiative<sup>2</sup>.

<sup>a</sup>Department of Information and Statistics, Chungnam National University, Korea;

<sup>b</sup>Department of Biostatistics, University of North Carolina at Chapel Hill, USA;

<sup>c</sup>Department of Radiology, University of Pennsylvania, USA

---

## Abstract

There is an emerging interest in brain functional connectivity (FC) based on functional Magnetic Resonance Imaging in Alzheimer's disease (AD) studies. The complex and high-dimensional structure of FC makes it challenging to explore the association between altered connectivity and AD susceptibility. We develop a pipeline to refine FC as proper covariates in a penalized logistic regression model and classify normal and AD susceptible groups. Three different quantification methods are proposed for FC refinement. One of the methods is dimension reduction based on common component analysis (CCA), which is employed to address the limitations of the other methods. We applied the proposed pipeline to the Alzheimer's Disease Neuroimaging Initiative (ADNI) data and deduced pathogenic FC biomarkers associated with AD susceptibility. The refined FC biomarkers were related to brain regions for cognition, stimuli processing, and sensorimotor skills. We also demonstrated that a model using CCA performed better than others in terms of classification performance and goodness-of-fit.

**Keywords:** resting-state functional magnetic resonance imaging, penalized logistic regression, common component analysis, Alzheimers disease, mild cognitive impairment

---

## 1. Introduction

Alzheimer's disease (AD) is a neurodegenerative disease that affects the elderly's health and places a huge burden on families and society. Its pathophysiological process is thought to begin many years before diagnosis (Morris, 2005). The preclinical phase of AD provides critical opportunities for early diagnosis that could reduce healthcare costs for both patients and governments. If 80–100% of AD patients were diagnosed at an early stage, it would yield a total cumulative savings of \$7 trillion to \$7.9 trillion in medical and long-term care costs (Alzheimer's Association, 2019). Early diagnosed patients could also prepare legal and financial plans while cognitively capable of making those critical decisions. Furthermore, early diagnosis could help the patients to lessen anxieties about their cognitive

---

Jae-Hwan Jung and Seong-Jin Ji equally contributed to this work.

<sup>1</sup> Corresponding author: Department of Information and Statistics, Chungnam National University, 99 Daehak-ro, Yuseong-gu, Daejeon 34134, South Korea. E-mail: eunjee.cnu@gmail.com

<sup>2</sup> Data used in this article's preparation were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As such, the ADNI investigators contributed to the design and implementation of the ADNI and/or provided data but did not participate in this analysis or writing of this report. A complete list of ADNI investigators can be found at [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)

Published 30 November 2020 / journal homepage: <http://csam.or.kr>

© 2020 The Korean Statistical Society, and Korean International Statistical Society. All rights reserved.

and behavioral symptoms by being aware of disease progression. Therefore, there have been many studies to detect AD at mild cognitive impairment (MCI) (Davatzikos *et al.*, 2011; Moradi *et al.*, 2015; Sperling *et al.*, 2011), which refers to an intermediate stage between the expected cognitive decline of normal aging and AD (Ganguli *et al.*, 2004).

This study helps establish promising biomarkers that contribute to the classification of MCI patients and cognitively normal elderly. Patients with MCI have a significantly higher likelihood to progress to probable AD relative to unimpaired individuals (Ganguli *et al.*, 2004), with a conversion rate of 10-15% per year (Petersen *et al.*, 1999) comparing to normal elderly: about 1-2% annually (Bischkopf *et al.*, 2002). Especially if memory loss is the predominant symptom, patients with MCI have an increased risk of developing AD (Dubois and Albert, 2004; Han *et al.*, 2012). Observation of such symptoms means that significant neurodegeneration has already occurred in the brain of the patients, either anatomically or functionally (Wee *et al.*, 2013; Zang *et al.*, 2012). However, there are no pharmacological treatments capable of delaying the long-term progression of MCI to dementia, and it is not recommended the use of medicines for MCI-treatment (Feldman and Jacova, 2005; Chertkow, 2008). MCI patients also show different cognitive decline rates, and even some never convert to AD due to heterogeneity of the etiology for MCI (Grand *et al.*, 2011). Thus, earlier detection of patients who are likely to convert from MCI to probable AD is critical to warn potential patients and guide them through appropriate treatment for delaying or preventing the onset of AD. AD could be diagnosed earlier by monitoring biomarkers associated with the development of MCI (compared to monitoring risk factors for AD) from cognitively normal elderly who have no apparent symptoms, but a higher risk of AD (Shankle *et al.*, 2005). Furthermore, biomarkers could provide insight into the mechanism underlying AD pathogenesis.

Biomarkers from brain imaging methods such as computed tomography (CT), magnetic resonance imaging (MRI), and position emission tomography (PET) have been used in the study of AD over the past decade. For more details about these methods, refer to (Johnson *et al.*, 2012). Among the methods, functional MRI (fMRI), which reflects the spontaneous blood oxygen level-dependent (BOLD) signal fluctuations, is increasingly being used to investigate functional changes between brain regions on AD and MCI patients (Allen *et al.*, 2007; Greicius *et al.*, 2004). Brain regions tend to be temporally correlated at resting state because they are functionally related or work together during a cognitive task (Beckmann *et al.*, 2005). In fMRI, functional connectivity (FC) describes the functional networks of the brain. It indicates temporal connection of brain activity in spatially distinct brain regions (Cordes *et al.*, 2001) by indexing changes in temporal patterns of the neural activity either in rest or task condition (Hutchison *et al.*, 2013).

Resting-state fMRI (rs-fMRI) shows the baseline BOLD variance using resting-state FC. Observation of the resting-state FC has reported all across the spectrum from AD (Wang *et al.*, 2007; Gili *et al.*, 2011) to MCI (Bai *et al.*, 2009; Petrella *et al.*, 2011), to normal control (NC) (Damoiseaux *et al.*, 2008). By the time AD symptoms develop, widespread FC changes are present throughout the brain (Wang *et al.*, 2007). Similarly, resting-state functional brain networks of MCI also show reduced connection strength and efficiency (Wang *et al.*, 2013; Drzezga *et al.*, 2011). For example, the network node attributes in the prefrontal cortex, insula, and white matter connectivity in the parietal cortex are distinct between normal elderly and MCI patients (Wee *et al.*, 2012). MCI patients also have decreased connections between the hippocampus and prefrontal gyrus, temporal gyrus, and parietal gyrus during an episodic memory task (Bai *et al.*, 2009). Hence, the resting-state FCs, as valid biomarkers, could provide unmistakable evidence for disruptive and abnormal brain functional differences between NC and MCI patients.

We focus on the resting-state FC-based biomarkers associated with the classification of MCI and

cognitively normal. Logistic regression is used for the classification, since it provides a the straightforward interpretation of the coefficients, unlike other machine learning methods. The most common way of estimating FC is Pearson's correlation coefficients between two BOLD signals of brain regions. In this study, the brain is segmented into 116 regions of interests (ROIs) by an anatomical parcellation on the brain with the automated anatomical labeling (AAL) template (Tzourio-Mazoyer *et al.*, 2002). The follow-up question is how to quantify the FC as proper covariates for the regression model. One simple way is to half-vectorize each FC matrix, which results in  $116 \times (116 - 1)/2$  unique elements for each subject. Therefore, the number of variables initially available is too many, causing a high-dimensionality issue. One popular way for dimension reduction is to calculate descriptive statistics for FC based on graph theory, which can assess the properties of brain regions and functional connections (Salvador *et al.*, 2005). However, this approach has two major limitations. One is that it does not specify brain regions with different connectivity between patient groups, while it enables exploring FC's overall organization. Another limitation is that it summarizes the FC to a very high degree; it might detect confounders such as systematic group differences in head motion or heart rate instead of genuine differences in the groups (Smith *et al.*, 2013). To address those limitations, we also employ the common component analysis (CCA) for dimension reduction (Wang *et al.*, 2011). It is recently proposed to decompose multiple symmetric matrices such as resting-state FCs. This method can be seen as an extension of the principal component analysis in that it pursues a new coordinate system, which is assumed to be common for all the matrices in the CCA.

Consequently, in order to use FC as covariates in a statistical model, we consider the methods mentioned above: (1) half-vectorization of the Pearson's correlation matrices, (2) the graph-theory based descriptive measures, and (3) half-vectorization of the dimension-reduced correlation matrices by CCA. We acquire three different datasets from the same rs-fMRI data as a result. However, the three FC datasets are still high dimensional, which is not suitable for a classic logistic regression model. That what, penalization using the elastic net penalty is employed to address the high-dimensionality and multicollinearity by compromising between Lasso and Ridge. We used leave-one-out cross-validation (LOOCV) to estimate the area under the curve (AUC) to select proper tuning parameters of the elastic net penalty. The AUC and deviance of a selected optimum model were also employed to assess classification performance and goodness-of-fit of the three datasets.

The goals of this paper are : (1) to establish a pipeline to incorporate functional connectivity as covariates in a logistic regression model, (2) to compare the performance of models using the three datasets, and (3) to investigate the effects of the FC-based biomarkers. It could provide insight into the relationship with the classification of MCI from NC. The rest of this paper is organized as follows. Section 2 provides detailed descriptions of ADNI data regarding rs-fMRI and data-preprocessing procedures. Three quantification methods for FC will be discussed in Section 3. They include two traditional treatments for FC's high-dimensionality and the CCA that have not received adequate attention. In Section 4, we propose a pipeline for modeling a Binary Response with FC Covariates by using the CCA. The section briefly explains penalized logistic regression, a core framework of the pipeline, and provides the modeling pipeline with graphic representation. Section 5 shows the performance comparison among the three models using and findings from biomarkers highlighted from regression models. Finally, Section 6 ties up these together and concludes the paper. Also, since there are many technical terms, we list their abbreviations in Appendix A to improve the readability.

## 2. Materials

### 2.1. Alzheimer's disease neuroimaging initiative (ADNI)

This paper is motivated by the ADNI database. The ADNI was launched in 2003 by the National

Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration, the National Institute on Aging, private pharmaceutical companies and non-profit organizations, as a \$60 million, 5-year public-private partnership (Martínez-Murcia *et al.*, 2013). Michael W. Weiner, MD, VA Medical Center, and the University of California-San Francisco is the Principal Investigator of this initiative. ADNI's primary goal is to test if brain imaging data like serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. All subjects were recruited from over 50 sites across the U.S. and Canada (<http://www.adni-info.org/>). The demographic information and the rs-fMRI data of 120 subjects were obtained from ADNI's publicly available database (<http://adni.loni.usc.edu/ADNI>).

## 2.2. Demographic and clinical variables

Many demographic factors have been known to be associated with the progression of AD, so they have been used as covariates in prior studies to predict conversion from MCI to AD including Age, Education length, ADAS-cog score, Gender, and APOE- $\epsilon 4$ . See the references therein for more details.

Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog) (Mohs, 1983) is a clinical and cognitive assessment score from the ADNI dataset that is potentially useful for predicting MCI-to-AD conversion. A higher ADAS score means a greater degree of cognitive impairment and a higher probability of being MCI. Also, the apolipoprotein E (APOE)  $\epsilon 4$ , especially its homozygosity, is the most potent genetic risk factor for AD for those between 40 and 90 years of age (Farrer *et al.*, 1997). Carriers of APOE- $\epsilon 4$  show worse cognitive performance and more significant cognitive decline over time than non-carriers in elderly patients without dementia (Small *et al.*, 1999).

Table 1 shows a summary of participant demographics. 53 NC (24 males/29 females) and 67 MCI subjects (36 males/31 females) were obtained from the ADNI dataset. In the table, mean and standard error were presented for continuous variables such as Age, Education length, and ADAS-cog. The unit of Age and Education length is a year. For example, the mean of Age in terms of year for MCI and NC groups was 71.531 and 72.955, with standard error 0.872 and 0.825, respectively. The count and the percentage of each category were presented for categorical variables such as Gender and APOE- $\epsilon 4$ . Though we treated APOE- $\epsilon 4$  as a continuous covariate in models, they are often categorized into 0 for homozygous reference allele, 1 for heterozygotes, and 2 for homozygous for alternative allele. Thus, the mean and standard error refer to total APOE- $\epsilon 4$ , and the proportion refers to each category of APOE- $\epsilon 4$ .

## 2.3. Preprocessing of fMRI

The rs-fMRI data, which were acquired using a 3.0 Tesla Philips Medical Systems during the task-free scans, was downloaded in original Directed Components (DICOM) format from the ADNI website. The scanning protocol for the rs-fMRI of all subjects was depicted as follows: flip angle = 80.0 degrees; manufacturing model=Intera; echo time (TE) = 30.001 ms; repetition time (TR) = 3000.0 ms; pixel spacing size =  $3.3125 \times 3.3125$ ; slice thickness = 3.313; slices = 6720.0; matrix size =  $64 \times 64$ ; pulse sequence = GR; the number of anatomical volumes = 140. Detailed acquisition parameters could be referred at the ADNI web site (<http://www.adni-info.org/>).

For the rs-fMRI data preprocessing, we used SPM8 (<https://www.fil.ion.ucl.ac.uk/spm/>). The details are: (1) removing the first 10 image volumes of functional time series manually to ensure magnetization equilibrium. (2) slice acquisition timing was corrected for each volume, followed by head-motion correction (i.e., realignment) with rigid-body transformation. (3) intensity scaling of

Table 1: Demographic information

Demographics		Total	NC	MCI
Age		72.160 ± 0.609	72.955 ± 0.825	71.531 ± 0.872
Education length		16.250 ± 0.234	16.585 ± 0.328	15.985 ± 0.328
ADAS-cog score		7.542 ± 0.369	5.566 ± 0.362	9.104 ± 0.523
Gender	Male	60 (50.00%)	24 (45.28%)	36 (53.73%)
	Female	60 (50.00%)	29 (54.72%)	31 (46.27%)
Total		0.517 ± 0.059	0.377 ± 0.072	0.627 ± 0.087
APOE-ε4	0	68 (56.70%)	34 (64.20%)	34 (50.70%)
	1	42 (35.00%)	18 (34.00%)	24 (35.80%)
	2	10 (8.30%)	1 (1.90%)	9 (13.40%)

each fMRI scan after motion correction to yield a whole-brain mean value of 10000, (4) temporally band-pass filtering with low-frequency range (0.01–0.08Hz) to remove effects of very low-frequency drift and high-frequency noise, (5) regressing out a set of nuisance signals, including the signal averaged over the white matter, signal averaged over the cerebrospinal fluid, global signal averaged over the whole brain, and six motion parameters, and (6) nonlinear normalization to the Montreal Neurological Institute space and spatially smoothing using Gaussian kernel of 6mm full-width.

### 3. Quantification methods for functional connectivity

#### 3.1. Low-order functional network: LON

The preprocessed BOLD time-series signals of all voxels were partitioned into 116 ROIs using the AAL template atlas (Tzourio-Mazoyer *et al.*, 2002).  $\mathbf{r}_i \in \mathbb{R}^T$  denotes the averaged time series of all voxels belonging to the  $i^{\text{th}}$  ROI, where  $T$  is the total number time points of BOLD signals. The AFNI package was used for those calculations (Cox, 1996). A commonly used method for quantifying the FC is the Pearson correlation (Smith *et al.*, 2013). The correlation-based FC matrix for the  $k^{\text{th}}$  subject is defined below :

$$\mathbf{C}^{(k)} = [c_{ij}]_{1 \leq i, j \leq V} = [\text{corr}(\mathbf{r}_i, \mathbf{r}_j)]_{1 \leq i, j \leq V}, \quad (3.1)$$

where  $V$  is the number of ROIs and  $\text{corr}(\mathbf{r}_i, \mathbf{r}_j)$  denotes the Pearson's correlation between the  $i^{\text{th}}$  ROI ( $\mathbf{r}_i$ ) and the  $j^{\text{th}}$  ROI ( $\mathbf{r}_j$ ).  $V = 116$  for the AAL template. Additionally, the resulting correlation coefficients were transformed by Fisher's z-transformation denoted by  $\tilde{\mathbf{C}}^{(k)}$ . This method can be considered a low-order representation of the functional interaction since the relationship investigated is only between two ROIs. Here,  $\mathbf{X}_{(L)}$  is a  $n \times V(V-1)/2$  data matrix, whose  $k^{\text{th}}$  row is the vectorized upper triangular parts of  $\tilde{\mathbf{C}}^{(k)}$ . This whole procedure will be called low-order functional network (LON). The correlation coefficient is the simplest way to quantify the FC; however, this method usually suffers the curse of dimensionality. The number of  $\mathbf{X}_{(L)}$  columns is  $(116 \times 115)/2$  with the AAL template, which is larger than the number of our subjects, 120. Moreover, this approach only considers marginal linear independence and dependence between ROIs without considering complex and time-varying interaction among brain regions.

#### 3.2. Graph-theory-based metrics: METRIC

The FC can be treated as a weighted graph based on the graph theory: a specific ROI (i.e., brain region) corresponds to a node, and an edge is used to characterize the pairwise FC between the ROIs.  $\mathbf{X}_{(M)}$  is a design matrix consisting of descriptive metrics for weighted graphs derived from the

ROI-based functional connectivity matrices,  $\mathbf{C}^{(k)}$ . The correlation coefficients denote edge weights in calculating metrics. Among many measures, we calculated three popular summary measures: diameter (Weissstein, 2003), strength (Barrat *et al.*, 2004), and PageRank (Brin and Page, 1998). The integration of those three measures resulted in the data matrix  $\mathbf{X}_{(M)}$  with  $n \times (2V + 1)$ , where  $2V + 1 = 1(\text{graph diameter}) + V(\text{strength}) + V(\text{PageRank})$  for each subject. This approach will be called METRIC in this paper. However, this approach has two major limitations. One is that it does not specify which brain regions have different connectivity between patient groups, while it enables exploring the overall organization of FC. Another limitation is that it summarizes the FC to a very high degree; therefore, it might detect confounding variables such as systematic group differences due to head motion or heart rate instead of genuine group differences (Smith *et al.*, 2013). Moreover, Tijms *et al.* (2013) reports that graph theory application to rs-fMRI data in AD showed conflicting results.

### 3.3. Common component analysis: DFC

Another recently proposed approach is the *common component analysis* (Wang *et al.*, 2011) that can provide insight into the complex organization of brain networks while conducting a dimension reduction of connectivity. Let  $L_i = [L_{i(g,g')}]_{g,g'=1,\dots,V}$  be the rs-fMRI connectivity matrix of the  $i^{\text{th}}$  subject, where  $g$  is a vertex or an ROI, and  $V$  is the total number of vertices or ROIs. It is supposed that  $L_i$  is symmetric without loss of generality, that is,  $L_{i(g,g')} = L_{i(g',g)}$  holds for all  $g, g' = 1, \dots, V$ . The common component model assumes that

$$L_{i(g,g')} = \sum_{r=1}^R \sum_{s=1}^R \gamma_{r,g} \lambda_{i(r,s)} \gamma_{s,g'} + \epsilon_{i(g,g')}, \quad (3.2)$$

where  $\lambda_{i(r,s)}$  is a subject-specific coefficient that can be non-zero even for  $r \neq s$  and  $\epsilon_{i(g,g')}$  is a measurement error. In addition,  $G = [\gamma_1, \dots, \gamma_R]$  as a  $V \times R$  orthogonal matrix is a common eigenmap across all subjects, where  $\gamma_s = (\gamma_{s,1}, \gamma_{s,2}, \dots, \gamma_{s,V})^T$  for  $s = 1, \dots, R$ . The subject-specific matrix  $\Lambda_i = [\lambda_{i(r,s)}]_{r,s=1,\dots,R}$  maintains an intrinsic network structure in the low-dimensional space spanned by the columns of  $G$ . Equation (3.2) can be written in a matrix form as:

$$L_i = G \Lambda_i G^T + \epsilon_i, \quad \text{where } \epsilon_i = [\epsilon_{i(g,g')}]_{g,g'=1,\dots,V}. \quad (3.3)$$

The unknown parameter matrices  $G$  and  $\Lambda_i$  can be estimated by minimizing the Frobenius norm of  $L_i - G \Lambda_i G^T$  by using iterative optimization steps. This optimization problem is solved by adapting low-rank approximation techniques proposed by Ye (2005). In detail, we consider the following optimization problem

$$\min_{G, \Lambda_i} \sum_{i=1}^n \|L_i - G \Lambda_i G^T\|_F^2 \quad \text{such that } G^T G = I_R, \quad (3.4)$$

where  $\|A\|_F^2 = \text{tr}(AA^T)$  is the Frobenius norm of  $A$  and  $\text{tr}(A)$  is its trace. Then,

$$\hat{\Lambda}_i = \hat{G} L_i \hat{G}^T,$$

where the eigenmap  $G$  can be estimated from the following iterative optimization steps. Under  $\Lambda_i =$

$GL_iG^T$ , the above minimizing problem is equivalent to maximizing

$$\begin{aligned} \sum_{i=1}^n \|G^T L_i G\|_F^2 &= \sum_{i=1}^n \text{tr}(G^T L_i G G^T L_i G) \\ &\approx \sum_{i=1}^n \text{tr}(G^T L_i G_0 \mathbf{G}_0^T L_i G), \end{aligned}$$

where  $G_0$  is the  $G$  matrix from the previous iteration. Then  $G$  can be estimated by the following iterative steps:

1. Let  $G_0$  be the  $G$  matrix from the previous iteration.
2. Calculate  $\mathbf{Q} = \sum_{i=1}^n L_i G_0 G_0^T L_i$ .
3. Compute the  $R$  eigenvectors  $\{\gamma_i\}_{i=1}^R$  of  $\mathbf{Q}$  corresponding to the largest  $R$  eigenvalues.
4. Set  $G = [\gamma_1, \gamma_2, \dots, \gamma_R]$ .
5. Repeat the above iterations until it converges.
6. Calculate  $\hat{\Lambda}_i = \hat{G} L_i \hat{G}^T$  for  $i = 1, \dots, n$ .

We half-vectorize  $\Lambda_i$ , dimension-reduced connectivity, into  $(\lambda_{i(1,1)}, \lambda_{i(2,1)}, \lambda_{i(2,2)}, \dots, \lambda_{i(R,R)})$  to use it as covariates in the regression model. Let this quantification procedure be called Dimension-reduced Functional Connectivity (DFC). The resulting data  $\mathbf{X}_{(D)}$  is a  $n \times R(R+1)/2$  matrix, whose  $i^{\text{th}}$  row is  $(\lambda_{i(1,1)}, \lambda_{i(2,1)}, \lambda_{i(2,2)}, \dots, \lambda_{i(R,R)})$ . One of our goals is to show that a classification model using  $\mathbf{X}_{(D)}$  is more promising in predictive performance and interpretability than the model using  $\mathbf{X}_{(L)}$  or  $\mathbf{X}_{(M)}$ .

#### 4. Penalized logistic regression with functional connectivity covariates

We propose a modeling pipeline for a binary response with FC covariates. Logistic regression is widely used in this case, which is not suitable for the FC data due to its complex matrix structure. The functional connectivity data is refined as appropriate covariates by applying for each of the three quantification methods, LON, METRIC, and DFC. Since the refined covariates are still high-dimensional, penalized logistic regression is employed as a core model in our modeling strategy to address the issue. We summarize the whole procedure as a pipeline for the FC analysis in Subsection 4.2.

##### 4.1. Logistic regression with elastic net penalty

Assume that we have  $n$  observations and  $p$  explanatory variables. Let  $y_i \in \{0, 1\}$  be the response variable value for the  $i^{\text{th}}$  observation. In our study, 1 represents a MCI patient and 0 represents a NC subject.  $\mathbf{x}_i = \left(1, (\mathbf{x}_i^{(1)})^T, (\mathbf{x}_i^{(2)})^T\right)^T$  is the  $i^{\text{th}}$  vector of design matrix  $\mathbf{X}$ , where  $\mathbf{x}_i^{(1)} = (x_{i1}, \dots, x_{iq})^T$  and  $\mathbf{x}_i^{(2)} = (x_{i(q+1)}, \dots, x_{ip})^T$  indicate control variables and the FC biomarkers respectively.  $q$  denotes the number of demographic variables and  $(p - q)$  denotes the number of FC biomarkers.

The design matrix  $\mathbf{X}$  can be varied by the FC quantification method such as  $[\mathbf{X}_{(1)} \ \mathbf{X}_{(L)}]$ ,  $[\mathbf{X}_{(1)} \ \mathbf{X}_{(M)}]$ , or  $[\mathbf{X}_{(1)} \ \mathbf{X}_{(D)}]$ , where  $\mathbf{X}_{(1)} = \left[\mathbf{x}_1^{(1)} \ \mathbf{x}_2^{(1)} \ \dots \ \mathbf{x}_n^{(1)}\right]^T$ . Accordingly,  $\boldsymbol{\beta}_{(1)} = (\beta_1, \dots, \beta_q)^T$  and  $\boldsymbol{\beta}_{(2)} =$

$(\beta_{q+1}, \dots, \beta_p)^T$  are coefficient vectors for control variables and the FC biomarkers respectively. Then, the response variable is related to the explanatory variables by

$$\text{logit}[\pi(\mathbf{x}_i)] = \beta_0 + (\mathbf{x}_i^{(1)})^T \boldsymbol{\beta}_{(1)} + (\mathbf{x}_i^{(2)})^T \boldsymbol{\beta}_{(2)} = \mathbf{x}_i^T \boldsymbol{\beta}, \quad i = 1, \dots, n, \quad (4.1)$$

where  $\pi(\mathbf{x}_i) = P(y_i = 1 | \mathbf{x}_i)$  and  $\boldsymbol{\beta} = (\beta_0, \boldsymbol{\beta}_{(1)}^T, \boldsymbol{\beta}_{(2)}^T)^T$  is a  $(p + 1) \times 1$  coefficient vector. Then, the log-likelihood function is defined as

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^n [y_i \log(\pi(\mathbf{x}_i)) + (1 - y_i) \log(1 - \pi(\mathbf{x}_i))]. \quad (4.2)$$

The penalized logistic regression adds a non-negative penalty term to the log-likelihood function to solve constrained maximization for  $\ell(\boldsymbol{\beta})$ . Lasso with a  $L_1$ -penalty (Tibshirani, 1996) and Ridge with a  $L_2$ -penalty (Hoerl and Kennard, 1970) are widely used. Ridge regression addresses collinearity within explanatory variables, and Lasso efficiently reduces the dimensionality by shrinking some of the regression coefficients to zero. An elastic net penalty (Zou and Hastie, 2005) is constructed to deal with the drawbacks of Lasso and Ridge. In the logistic elastic net regression,  $\boldsymbol{\beta}$  can be estimated by maximizing the penalized likelihood,

$$\hat{\boldsymbol{\beta}}_{\text{Elastic}} = \underset{\boldsymbol{\beta} \in \mathbb{R}^{p+1}}{\text{argmax}} [\ell(\boldsymbol{\beta}) - \lambda P_\alpha(\boldsymbol{\beta}_{(2)})] = (\hat{\beta}_0, \hat{\boldsymbol{\beta}}_{(1)}^T, \hat{\boldsymbol{\beta}}_{(2)}^T)^T, \quad (4.3)$$

where

$$P_\alpha(\boldsymbol{\beta}_{(2)}) = \sum_{j=q+1}^p \left( \frac{1}{2} (1 - \alpha) \beta_j^2 + \alpha |\beta_j| \right), \quad (0 \leq \alpha \leq 1). \quad (4.4)$$

For the model estimation, we used the R *glmnet* package that applies the coordinate descent algorithm for the maximization (Friedman *et al.*, 2010).

Equation (4.3) shows that an elastic net estimator depends on non-negative tuning parameters,  $\lambda$  and  $\alpha$ , which leads to a penalized logistic regression solution. One of our goals is to investigate and control the effects of the demographic and clinical variables mentioned in Section 2.2; therefore, we impose the penalty term only on the FC biomarkers. We estimated AUC by leave-one-out cross-validation (LOOCV) and used it to adjust the tuning parameters to proper values. That is, we choose an  $(\alpha, \lambda)$  pair that shows the highest AUC value. The AUC and a deviance test are used to evaluate the model.

## 4.2. Modeling pipeline

Figure 1 shows the pipeline of the modeling framework. It describes the following steps:

- (1) Quantification of the FC data: This quantification includes simple vectorization (LON), feature extraction based on graph theory (METRIC), and dimension reduction (DFC) for the FC data. Hence, three different biomarkers are refined and denoted by  $\mathbf{X}_{(L)}$ ,  $\mathbf{X}_{(M)}$ , and  $\mathbf{X}_{(D)}$ , respectively.
- (2) Penalized logistic regression: After the first step, there are possibly more variables than the sample size. Therefore, we conduct penalized logistic regression using the elastic net penalty, where

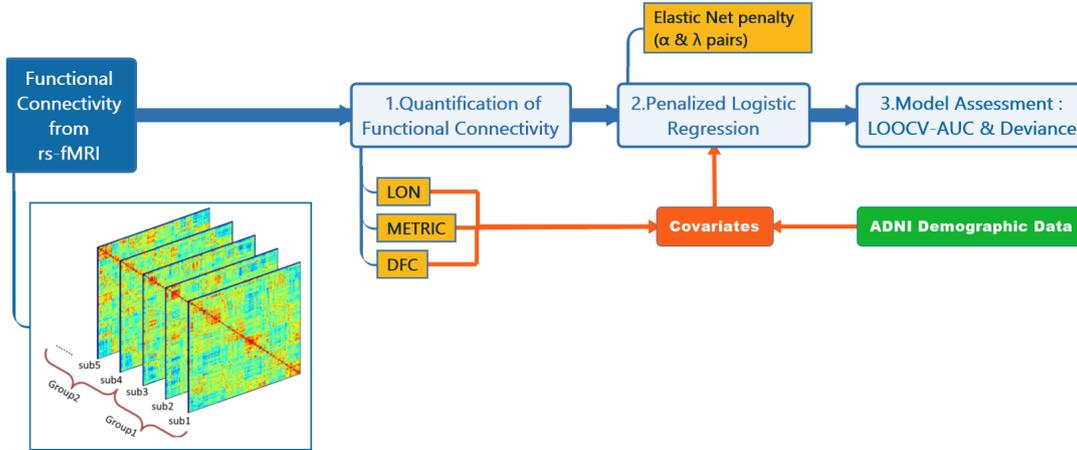


Figure 1: Pipeline of the penalized logistic regression using FC data for classifying MCI and NC.

the refined biomarkers ( $\mathbf{X}_{(L)}$ ,  $\mathbf{X}_{(M)}$ , or  $\mathbf{X}_{(D)}$ ) are penalized except for demographics. The tuning parameters,  $\alpha$  and  $\lambda$ , for the penalty are chosen by comparing their LOOCV-AUC values.

- (3) Model assessment: Each model is assessed by LOOCV-AUC and the deviance test. One can conclude which quantification method shows better performance than the others in terms of classification performance and goodness-of-fit.

This pipeline can be applied to any classification problem, where explanatory variables are symmetric matrices. It can also be easily adapted to generalized linear models.

## 5. Results

We included Age, Education length, ADAS-cog score, Gender, and APOE- $\epsilon 4$  as control variables in the model ( $q = 5$ ). After executing the modeling pipeline, we obtained three sets of refined FC biomarkers,  $\mathbf{X}_{(L)}$ ,  $\mathbf{X}_{(M)}$ , and  $\mathbf{X}_{(D)}$ . The estimated models for each FC quantification method are called  $M_{LON}$ ,  $M_{METRIC}$ , and  $M_{DFC}$ . Since the tuning by LOOCV might not necessarily give the most predictive model, in particular for logistic regression, we also considered to restrict penalized logistic regression models for LON, METRIC, and DFC to have the same number of covariates by manually tuning  $\lambda$  and  $\alpha$ . This strategy would help compare the predictive performance of three quantification methods in various aspects. We manually selected the  $\alpha$  and  $\lambda$  values so that the final model had the predetermined number of covariates (e.g., 3 or 5) with the highest LOOCV-AUC. The models with 3 and 5 covariates are denoted by  $M_{LON}^{(3)}$ ,  $M_{METRIC}^{(3)}$ ,  $M_{DFC}^{(3)}$ ,  $M_{LON}^{(5)}$ ,  $M_{METRIC}^{(5)}$ , and  $M_{DFC}^{(5)}$  for each quantification method. We used 101  $\lambda$  values from 0 to 10 and 6  $\alpha$  values from 0 to 1 to search for candidate ranges for  $\lambda$  and  $\alpha$ . After the rough search, we considered 101  $\lambda$  values from 0 to 1 and 101  $\alpha$  values from 0 to 1 to find the optimal parameters by LOOCV.

The results of the current paper can be divided into three parts. Section 5.1 presents the results of the classification between NC and MCI. By computing AUC and deviance, we assess the classification performance and the goodness-of-fit for each FC quantification method. In Section 5.2, we explore selected biomarkers by penalized logistic regression. We will describe the selected biomarkers from  $M_{DFC}$  in detail since the implementation of DFC is a primary purpose. In Section 5.3, we investigate the estimated effects for demographic and clinical variables.

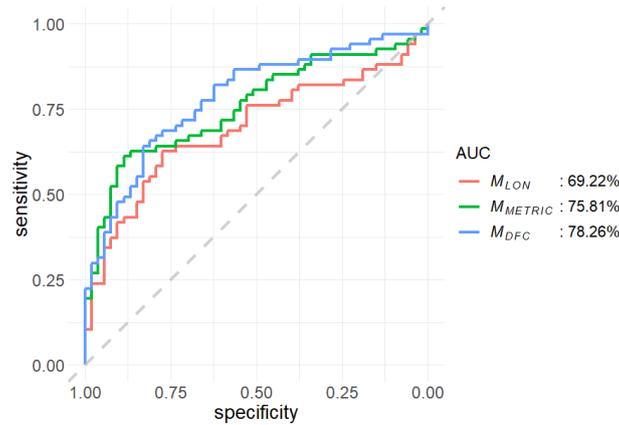


Figure 2: The receiver operating characteristic (ROC) curves of the three classification models.

Table 2: Model assessment

Tuning method	Model	AUC(%)	Deviance	df	$p$ -value
LOOCV	$M_{LON}$	69.22	140.36	113	0.0414*
	$M_{METRIC}$	75.81	129.15	113	0.1421
	$M_{DFC}$	78.26	120.15	111	0.2603
3 covariates	$M_{LON}^{(3)}$	68.68	138.64	111	0.0389*
	$M_{METRIC}^{(3)}$	75.47	127.27	111	0.1385
	$M_{DFC}^{(3)}$	78.26	120.15	111	0.2603
5 covariates	$M_{LON}^{(5)}$	67.33	137.95	109	0.0319*
	$M_{METRIC}^{(5)}$	75.08	121.94	109	0.1871
	$M_{DFC}^{(5)}$	77.76	119.90	109	0.2238

### 5.1. Model assessment

Figure 2 shows the receiver operating characteristic (ROC) curve for each model and Table 2 summarizes the assessment results. The first three rows of the table represent the results when the tuning has been done by LOOCV. The next three rows show performance of the manually tuned model with three covariates. Results for five covariates are summarized in the last three rows. The AUC of  $M_{LON}$  was 69.22%, where  $\alpha = 0.700$  and  $\lambda = 0.180$ . The AUC of  $M_{METRIC}$  was 75.81%, where  $\alpha = 0.850$  and  $\lambda = 0.091$ . Lastly, the AUC of  $M_{DFC}$  was 78.26%, where  $\alpha = 1.000$  and  $\lambda = 0.067$ , showing better performance than the others. The ROC curve of  $M_{DFC}$  was above that of  $M_{LON}$ , while curves of  $M_{DFC}$  and  $M_{METRIC}$  crossed at specificity = 0.83.  $M_{DFC}$  performed better than  $M_{METRIC}$ , when specificity was achieved within a moderate level ( $\leq 0.83$ ). However, its performance was slightly worse when the specificity was higher than 0.83. In this study, sensitivity is more important than specificity, considering that the elderly susceptible to AD are needed to be monitored instead of being missed (Wollman and Prohovnik, 2003). Therefore, those ROC curves showed that the  $M_{DFC}$  performed better than  $M_{LON}$  and  $M_{METRIC}$ . We also conducted the deviance test and obtained the results that  $M_{METRIC}(p = 0.14)$  and  $M_{DFC}(p = 0.26)$  fitted the data well, while  $M_{LON}(p = 0.04)$  did not. Manually tuned models (e.g.,  $M_{LON}^{(3)}$ ,  $M_{METRIC}^{(3)}$ ,  $M_{DFC}^{(3)}$ ) gave similar results in terms of the model performance. The AUC values tended to slightly decrease as the number of covariates increased, while the AUC values stayed the same for  $M_{DFC}$  and  $M_{DFC}^{(3)}$ . The reason was that there were already three covariates in  $M_{DFC}$  as in  $M_{DFC}^{(3)}$ . The deviances and test results were also almost identical. As a result,

Table 3: Coefficients of the three penalized logistic regression models

Coefficients		$M_{LON}$		$M_{METRIC}$		$M_{DFC}$	
		$\beta$	$\exp(\beta)$	$\beta$	$\exp(\beta)$	$\beta$	$\exp(\beta)$
Demographics	Age	-0.055	0.947	-0.071	0.932	-0.079	0.924
	Education	-0.169	0.845	-0.149	0.862	-0.141	0.869
	ADAS	0.070	1.073	0.162	1.176	0.199	1.220
	Gender(male)	0.583	1.791	0.490	1.632	0.411	1.508
	APOE- $\epsilon$ 4	0.679	1.972	0.647	1.910	0.606	1.833
FC Biomarkers <sup>a</sup>	$L_{58,97}$	-0.107	0.898	-	-	-	-
	$M211$	-	-	-0.011	0.989	-	-
	$\Lambda_{2,12}$	-	-	-	-	0.070	1.073
	$\Lambda_{4,12}$	-	-	-	-	-0.095	0.910
	$\Lambda_{11,14}$	-	-	-	-	0.193	1.213

<sup>a</sup> Selected FC biomarkers for each model.

the  $M_{DFC}$  was better than other models in terms of classification performance and goodness-of-fit.

## 5.2. Selected FC biomarkers

### 5.2.1. LON

The estimated coefficients are presented in Table 3. With  $\mathbf{X}_L$ , the model  $M_{LON}$  selected only one variable,  $L_{58,97}$ . It is the FC between the left calcarine fissure and the surrounding cortex (CAL.L) and left anterior cingulate and paracingulate gyri (ACG.L). CAL.L is located on the inside of the occipital lobe between the hemispheres, where the primary visual cortex is concentrated. The primary visual cortex, found in the occipital lobe in both cerebral hemispheres, contains a complete map of the visual field covered by the eyes. ACG.L is related to allocating cognitive resources to concurrent auditory and visual information. The ACG and its neighboring areas are also known to be recruited during conflict monitoring and attentional control (Botvinick *et al.*, 2001; Shenhav *et al.*, 2013). In previous network analysis, (Wang *et al.*, 2016) reported significantly altered betweenness centrality located in CAL.R and ACG.L of MCI subjects compared to NC. Compared with the NC and MCI groups, the nodal centrality in the AD population showed significant increase CAL.L (Yao *et al.*, 2010). In our results, the estimated odds of MCI multiply by  $\exp(\hat{\beta}_{L_{58,97}}) = 0.898$  for each 1 unit increase in the FC between CAL.L and ACG.L. The decrease in the FC of CAL.L and ACG.L affects the increase in the transition to MCI and is consistent with the previous studies.

### 5.2.2. METRIC

The model  $M_{METRIC}$ , which was fitted by graph-theoretical metrics of 116 ROIs, selected only one variable,  $M211$ .  $M211$  was the PageRank of the right superior parietal gyrus (SPG.R), representing the rank of SPG.R in terms of importance in the graph. The superior parietal lobule is one of the three subdivisions of the parietal lobe, which is critical in manipulating information in working memory (Koenigs *et al.*, 2009). In the previous study, compared to the NC groups, a graph theory measure of MCI showed a significant decrease in SPG.R (Liu *et al.*, 2012). The estimated odds of being MCI multiply by  $\exp(\hat{\beta}_{M211}) = 0.989$  for each 1 unit increase in  $M211$  (1.1% decrease). Therefore, the decrease in the PageRank of SPG.R affects the increase in the odds of having MCI.

### 5.2.3. DFC

For the model  $M_{DFC}$ , three covariates were selected by the elastic net penalty. The selected covariates involved 5 eigenvectors: the 2<sup>nd</sup>, 4<sup>th</sup>, 11<sup>th</sup>, 12<sup>th</sup>, and 14<sup>th</sup> eigenvectors. Hence, to interpret how the

Table 4: ROIs mainly associated with the 5 eigenvectors

	2nd eigenvector	4th eigenvector	11th eigenvector
(1)	right inferior occipital gyrus	dorsolateral area of left superior frontal gyrus	orbital part of right middle frontal gyrus
(2)	right middle occipital gyrus	left middle frontal gyrus	orbital part of right inferior frontal gyrus
(3)	flocculonodular lobe of left cerebellum	right supramarginal gyrus	right inferior parietal
(4)	left calcarine fissure and surrounding cortex	triangular part of left inferior frontal gyrus	dorsolateral part of right superior frontal gyrus
(5)	left gyrus rectus		right insula
(6)	right olfactory cortex		right Heschl gyrus
(7)	orbital part of left inferior frontal gyrus		right rolandic operculum
(8)			bilateral anterior cingulate and paracingulate gyri
(9)			right Inferior parietal, but supramarginal and angular gyri
	12th eigenvector	14th eigenvector	
(1)	left precentral gyrus	orbital part of left inferior frontal gyrus	
(2)	left rolandic operculum	temporal pole of left superior temporal gyrus	
(3)	right superior occipital gyrus	left vermis 3	
(4)	dorsolateral part of left superior frontal gyrus	temporal pole of the left middle temporal gyrus	
(5)	left vermis 9	left cuneus	
(6)	left Heschl gyrus		
(7)	right angular gyrus		
(8)	right inferior parietal, but supramarginal and angular gyri		
(9)	right Lenticular nucleus and putamen		

selected covariates were associated with disease status, the 5 eigenvectors should be examined. Table 4 list the ROIs associated with each eigenvector.

The 2<sup>nd</sup> eigenvector had high weights on the 7 ROIs as in Table 4. They play central roles in visual processing from the basic level to the higher level, such as face recognition (Mechelli *et al.*, 2000; Renier *et al.*, 2010), and olfaction (Menini, 2009). The function of (5) is unclear but it may be involved in higher cognitive function (Orrison, 2008). Therefore, the 2<sup>nd</sup> eigenvector represents brain regions related to *odor and visual information processing*. The weights of the 4<sup>th</sup> eigenvector had high values on the 4 ROIs, which were associated with the default mode network (DMN) and cognitive execution network (CEN), memory and attention, and semantic tasks (Li *et al.*, 2013; Ben-Shabat *et al.*, 2015). This eigenvector, therefore, would represent *cognition* related brain regions. The 11<sup>th</sup> eigenvector was mainly associated with the 9 ROIs related to integrative work of audio and visual processing, language production, attention related tasks, DMN and CEN. This eigenvector could be considered to play a role in *processing and responding to external stimuli*. The 12<sup>th</sup> eigenvector had high weights on the 9 ROIs. The ROIs were related to motor skills, auditory processing, and language functions from simple to complex levels (Blefari *et al.*, 2017; Li *et al.*, 2013; Warrier *et al.*, 2009; Hall, 2010). Therefore, this eigenvector would represent *auditory-motor integration*. The weights of the 14<sup>th</sup> eigenvector were high on the 5 ROIs, where those ROIs were charge of reorienting to unexpected stimuli, language, semantic processing, visual processing (Ardila *et al.*, 2014, 2017; Coffman *et al.*, 2011). It implied that the eigenvector represented *stimuli processing from unexpected to expected*.

The estimated odds of MCI multiply by  $\exp(\hat{\beta}_{\Lambda_{11,14}}) = 1.213$  for each 1 unit increase in  $\Lambda_{11,14}$ ; that is, about 20% increase. It would imply that MCI had stronger connections among stimuli related brain regions compared to NC. The estimated odds of MCI decrease by a factor of  $\exp(\hat{\beta}_{\Lambda_{2,12}}) = 1.073$  as  $\Lambda_{2,12}$  increases by 1 unit. It would suggest that MCI had a slightly stronger connection among brain regions related to sensorimotor skills. The estimated odds of being MCI multiply by  $\exp(\hat{\beta}_{\Lambda_{4,12}}) = 0.910$  for each 1 unit increase in  $\Lambda_{4,12}$ . A weaker connection between *cognition* function and *auditory-*

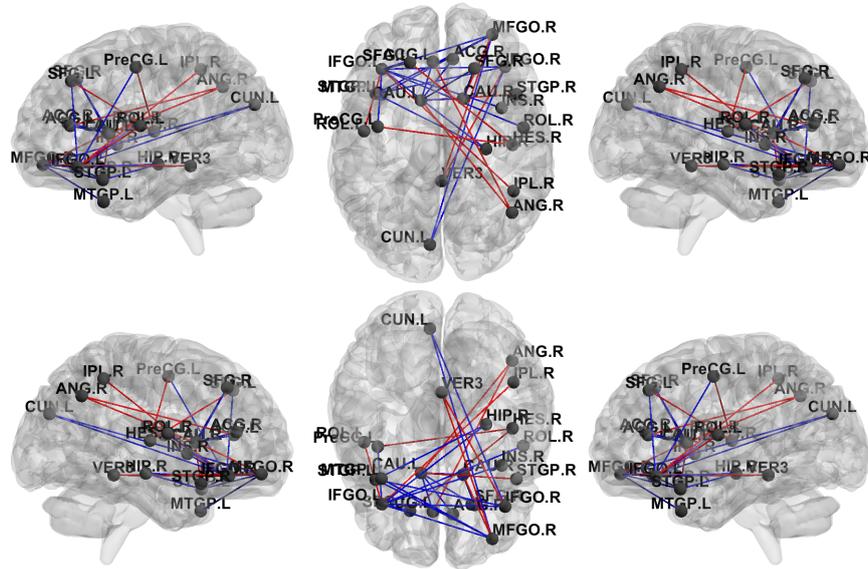


Figure 3: The functional connections of the shown pairs of ROIs were selected as important factors for classifying MCI and NC groups.

motor integration was expected in MCI. In order to examine which connections were altered for MCI compared to NC, we mapped the estimated coefficients from the eigenvector space to the original ROI space. Figure 3 depicts the largest 1% regression coefficients in the ROI space, where the blue line implies a weaker connection for MCI and the red line means a stronger connection for MCI. While we could observe that various brain regions are involved in the altered connections, CAL.L and ACG.L that were detected by  $M_{LON}$  were also implicated. It suggested us that  $M_{DFC}$  gave more encompassing results. In conclusion, the highlighted biomarkers from  $M_{DFC}$  were related to altered FC among brain regions that were in charge of external stimuli processing, language functions, and sensorimotor skills.

### 5.3. Coefficients of demographic and clinical variables

In AD study, demographic and clinical variables have allegedly been considered to be critical factors. Since AGE, APOE-4, Education length, Gender, and ADAS-cog score could be causing-factors of AD, we added them as control variables to each model not to be penalized. Table 3 shows that the signs of estimated coefficients were the same in the three models, indicating the same directions of the effects on the probability of being MCI. Hence, we only focused on the variables from  $M_{DFC}$ . The coefficients showed the change in log odds due to incremental-unit changes in the predictors (DeMaris, 1992). Also, logistic curves were plotted to visualize the effects of particular variables, as shown in Figure 4. It showed that male subjects have lower predicted probabilities of MCI than female subjects, overall. As the ADAS-cog score increases, regardless of gender, the predicted probability of being MCI for subjects with 2 APOE- $\epsilon$ 4 alleles tends to increase more sharply than for APOE- $\epsilon$ 4 non-carriers. The rest of the logistic curves were in Appendix B.

The regression coefficients of Gender(Male = 1, Female = 0), APOE- $\epsilon$ 4, and ADAS-cog demonstrated the increments of these variables means the higher probability of becoming MCI. The regres-

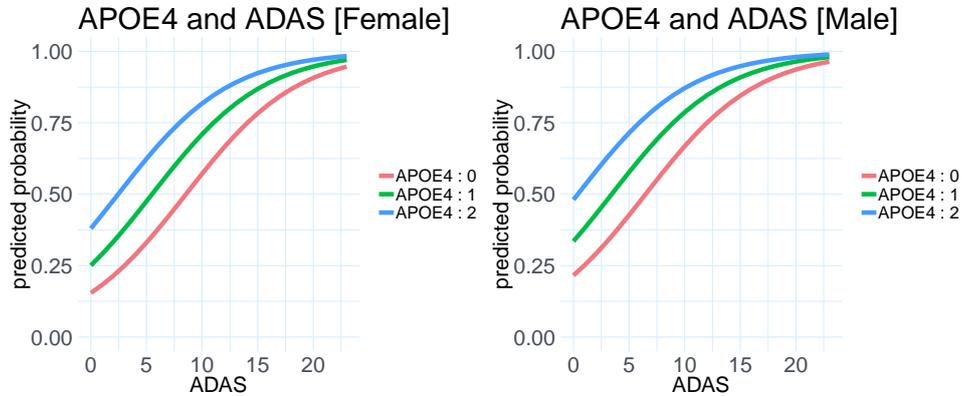


Figure 4: The logistic curves of ADAS-cog score stratified by APOE- $\epsilon$ 4. The left and the right panels show the estimated curves for males and females, respectively. The other covariates are fixed at their mean values.

sion coefficients of Age and Education demonstrated that decrements of these variables implied the higher probability of becoming MCI. In our results, the estimated odds of being MCI multiply by  $\exp(\hat{\beta}_{\text{ADAS}}) = \exp(0.199) \approx 1.220$  for each 1 score increase in ADAS (22% increase). The estimated odds of being MCI multiply by  $\exp(\hat{\beta}_{\text{APOE-}\epsilon 4}) = \exp(0.606) \approx 1.833$  for each score also increase in APOE- $\epsilon$ 4 (83% increase). Many studies have shown that the prevalence of MCI increases with age (Ganguli *et al.*, 2013; Kryscio *et al.*, 2006) However, our results were the opposite. The estimated odds of being MCI multiply by  $\exp(\hat{\beta}_{\text{Age}}) = \exp(-0.079) \approx 0.924$  for each score increase in Age (7.6% decrease).

## 6. Conclusions

In this work, we developed the pipeline for classification using FC. This pipeline can be applied to any classification problem, where explanatory variables are symmetric matrices. Three different methods were used to quantify FC and construct classification models using FC biomarkers as covariates. Penalized logistic regression with the elastic net penalty was applied in the MCI classification to simultaneously tackle model estimation and FC biomarker selection. We compared the performance of three methods for the ADNI data. The AUCs of 69.22%, 75.81%, and 78.26%, for LON, METRIC, and DFC, respectively. Therefore, in comparison with other alternatives, the  $M_{\text{DFC}}$  showed higher AUC value. From the ROC analysis,  $M_{\text{DFC}}$  showed a better performance because its ROC curve was the highest located one within a moderate level of specificity. We also found that the selected FC biomarkers of DFC were related to cognition, stimuli processing, and sensorimotor skills.

The ADNI dataset could include heterogeneous MCI subjects. Diagnosis for MCI has followed the criteria formulated by the Mayo Alzheimer's Disease Research Center (Hänninen *et al.*, 2002). However, the criteria usually depend on observations by clinicians, experience, and individuals subjective reports. Accordingly, MCI patients might be categorized to be heterogeneous and show different brain degeneration in disease progress that makes the early diagnosis challenging. Furthermore, there are few studies in the identification of MCI converter and MCI non-converter. Therefore, our work may have a possibility that heterogeneous MCI subjects are contained, so our results might not be reliable. Consequently, more detailed criteria for segmenting MCI patients could improve the performance results.

Longitudinal fMRI studies in patients with dementia have multiple challenges. Changes in the

brain of NC and MCI have very weak signals to diagnose if a subject is cognitively normal or MCI (Johnson *et al.*, 2006; Li *et al.*, 2011). BOLD signals are known to be variable across subjects because fMRI techniques are quite sensitive to head motion. Due to those reasons, the classification between NC and MCI is a more challenging problem in brain-imaging data. A multi-modal approach may improve these weak points since the biomarkers based on brain-imaging modalities have complementary information on fMRI (Rathore *et al.*, 2017). However, brain-imaging is an expensive tool and thus has limited accessibility as frontline screening and diagnostic tools for AD. It could be the reason for subjects without fMRI information and accordingly cause a limited sample size.

### **Acknowledgements**

This work was supported by 2018 research fund of Chungnam National University.

### **Appendix A: Abbreviations**

- AAL** Anatomical automatic labeling
- ACG.L** Left Anterior cingulate and paracingulate gyri
- AD** Alzheimer's disease
- ADAS-Cog** Alzheimer's Disease Assessment Scale-Cognitive Subscale
- ADNI** Alzheimer's Disease Neuroimaging Initiative
- APOE** Apolipoprotein E
- AUC** Area under the ROC curve
- BOLD** Blood oxygen level-dependent
- CAL.L** Left calcarine fissure and surrounding cortex
- CCA** Common component analysis
- DFC** Dimension-reduced Functional Connectivity
- DICOM** Directed Components
- FC** Functional connectivity
- fMRI** Functional Magnetic Resonance Imaging
- LON** Low-Order functional Network
- LOOCV** Leave-One-Out Cross-Validation
- MCI** Mild cognitive impairment
- MRI** Magnetic Resonance Imaging
- NC** Normal control
- PET** Positron emission tomography
- ROC** Receiver operating characteristic
- ROIs** Region of interests
- rs-fMRI** Resting-state fMRI
- SPG.R** Right superior parietal gyrus

### Appendix B: Adjusted logistic curves of a continuous covariate stratified by a categorical covariate

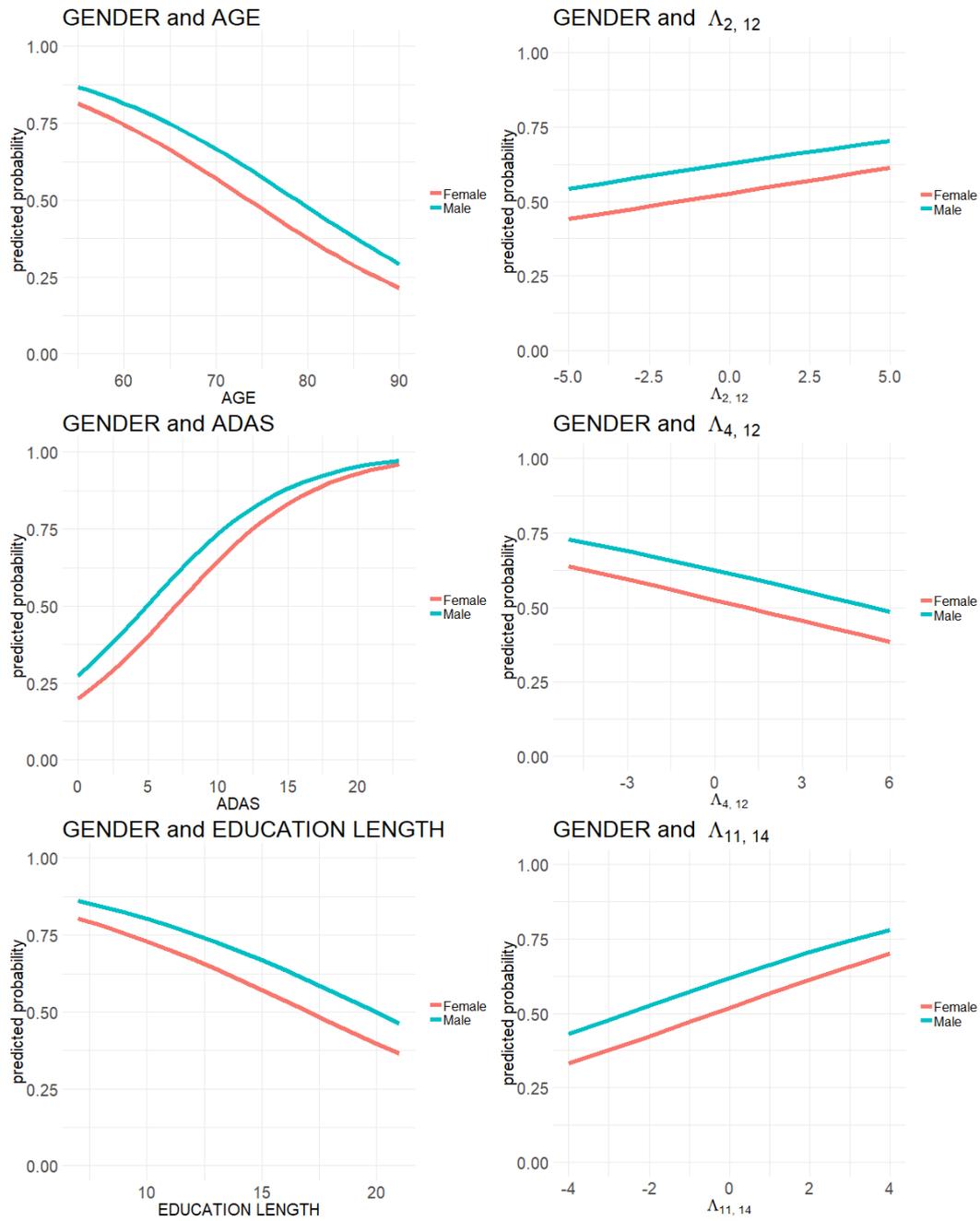


Figure B.1: Logistic curves of gender variable.

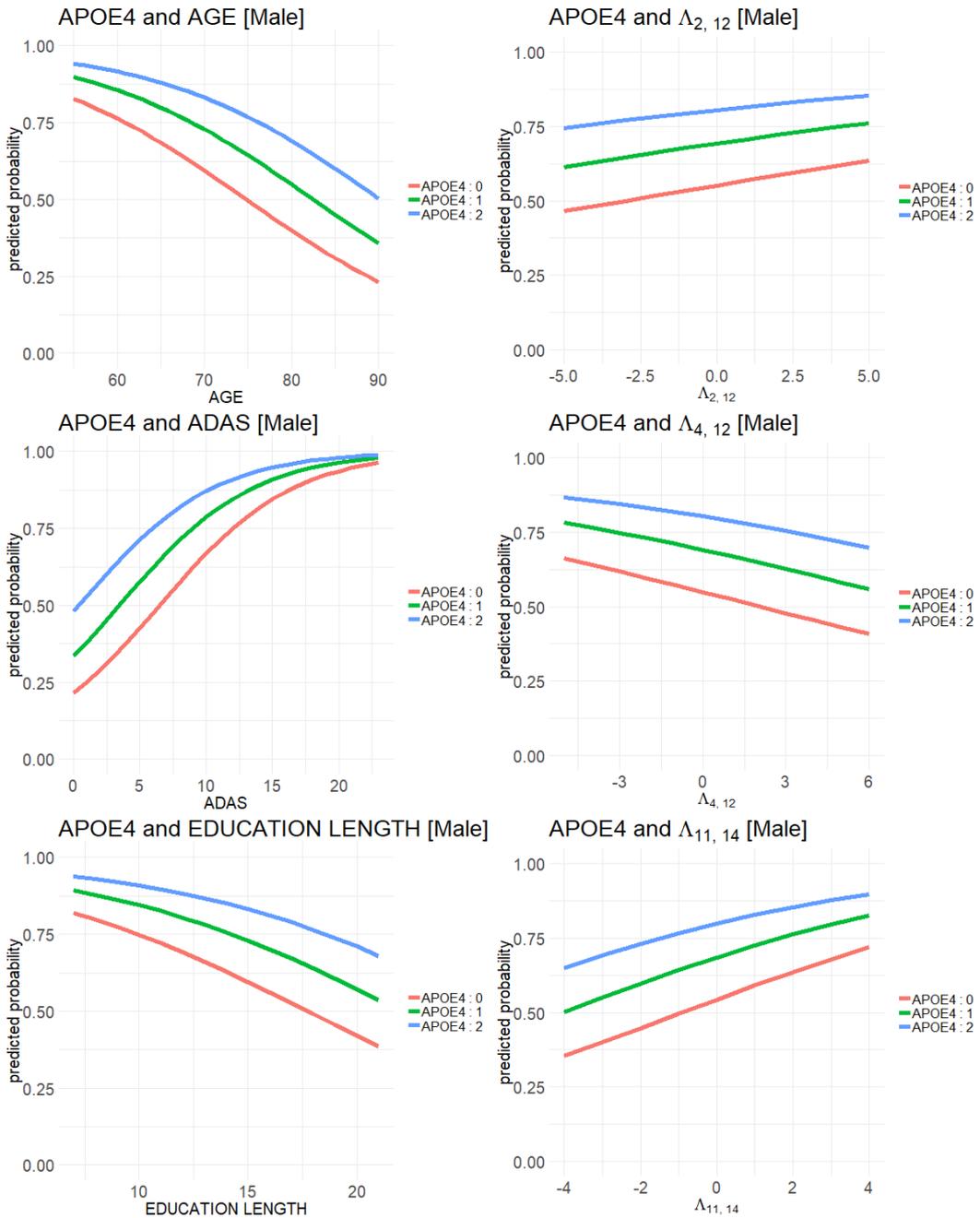


Figure B.2: Logistic curves using APOE- $\epsilon 4$  as a categorical variable (Male).

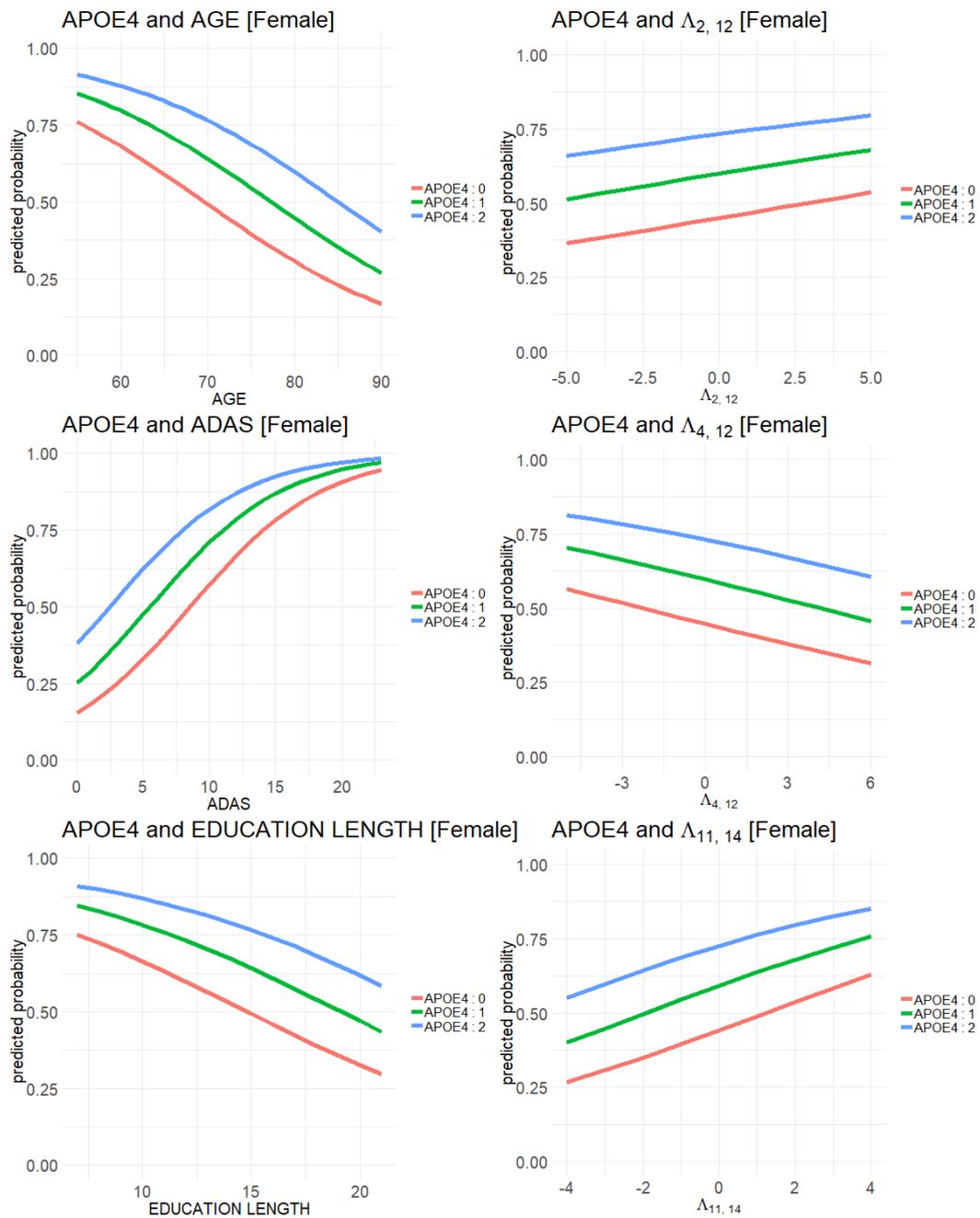


Figure B.3: Logistic curves using APOE- $\epsilon$ 4 as a categorical variable (Female).

## References

- Allen G, Barnard H, and McColl R, *et al.* (2007). Reduced hippocampal functional connectivity in Alzheimer disease, *Archives of Neurology*, **64**, 1482–1487.
- Alzheimer's Association (2019). 2019 Alzheimer's disease facts and figures, *Alzheimer's & Dementia*, **15**, 321–387.
- Ardila A, Bernal B, and Rosselli M (2014). The elusive role of the left temporal pole (BA38) in language: a preliminary meta-analytic connectivity study, *International Journal of Brain Science*, **2014**, 946039.
- Ardila A, Bernal B, and Rosselli M (2017). Should Broca's area include Brodmann area 47?, *Psicothema*, **29**, 73–77.
- Bai F, Zhang Z, Watson DR, Yu H, Shi Y, Yuan Y, Zang Y, Zhu C, and Qian Y (2009). Abnormal functional connectivity of hippocampus during episodic memory retrieval processing network in amnesic mild cognitive impairment, *Biological Psychiatry*, **65**, 951–958.
- Barrat A, Barthelemy M, Pastor-Satorras R, and Vespignani A (2004). The architecture of complex weighted networks. In *Proceedings of the National Academy of Sciences*, **101**, 3747–3752.
- Beckmann CF, DeLuca M, Devlin JT, and Smith SM (2005). Investigations into resting-state connectivity using independent component analysis, *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**, 1001–1013.
- Ben-Shabat E, Matyas TA, Pell GS, Brodtmann A, and Carey LM (2015). The right supramarginal gyrus is important for proprioception in healthy and stroke-affected participants: a functional MRI study, *Frontiers in Neurology*, **6**, 248.
- Bischkopf J, Busse A, and Angermeyer MC (2002). Mild cognitive impairment 1—a review of prevalence, incidence and outcome according to current approaches, *Acta Psychiatrica Scandinavica*, **106**, 403–414.
- Blefari ML, Martuzzi R, Salomon R, Bello-Ruiz J, Herbelin B, Serino A, and Blanke O (2017). Bilateral Rolandic operculum processing underlying heartbeat awareness reflects changes in bodily self-consciousness, *European Journal of Neuroscience*, **45**, 1300–1312.
- Botvinick MM, Braver TS, Barch DM, Carter CS, and Cohen JD (2001). Conflict monitoring and cognitive control, *Psychological Review*, **108**, 624–652.
- Brin S and Page L (1998). The anatomy of a large-scale hypertextual Web search engine, *Computer Networks and ISDN Systems*, **30**, 107–117.
- Coffman KA, Dum RP, and Strick PL (2011). Cerebellar vermis is a target of projections from the motor areas in the cerebral cortex. In *Proceedings of the National Academy of Sciences*, **108**, 16068–16073.
- Cordes D, Haughton VM, Arfanakis K, Carew JD, Turski PA, Moritz CH, Quigley MA, and Meyerand ME (2001). Frequencies contributing to functional connectivity in the cerebral cortex in resting-state data, *American Journal of Neuroradiology*, **22**, 1326–1333.
- Chertkow H (2008). Diagnosis and treatment of dementia: introduction. Introducing a series based on the Third Canadian Consensus Conference on the Diagnosis and Treatment of Dementia, *CMAJ*, **178**, 316–321.
- Cox RW (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages, *Computers and Biomedical Research*, **29**, 162–173.
- Damoiseaux JS, Beckmann CF, Arigita ES, Barkhof F, Scheltens P, Stam CJ, Smith SM, and Rombouts S (2008). Reduced resting-state brain activity in the default network in normal aging, *Cerebral Cortex*, **18**, 1856–1864.

- Davatzikos C, Priyanka B, Shaw LM, Batmanghelich KN, and Trojanowski JQ (2011). Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification, *Neurobiology of Aging*, **32**, 2322–e19.
- DeMaris A (1992). *Logit Modeling: Practical Applications*, Sage, **86**.
- Drzezga A, Becker JA, and Van Dijk KR, (2011). Neuronal dysfunction and disconnection of cortical hubs in non-demented subjects with elevated amyloid burden, *Brain*, **134**, 1635–1646.
- Dubois B and Albert ML (2004). Amnestic MCI or prodromal Alzheimer's disease?, *The Lancet Neurology*, **3**, 246–248.
- Farrer LA, Cupples LA, Haines JL, *et al.* (1997). Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease: a meta-analysis, *JAMA*, **278**, 1349–1356.
- Feldman HH and Jacova C (2005). Mild cognitive impairment, *The American Journal of Geriatric Psychiatry*, **13**, 645–655.
- Friedman J, Hastie T, and Tibshirani R (2010). Regularization paths for generalized linear models via coordinate descent, *Journal of Statistical Software*, **33**, 1–22.
- Ganguli M, Dodge HH, Shen C, and DeKosky ST (2004). Mild cognitive impairment, amnestic type: an epidemiologic study, *Neurology*, **63**, 115–121.
- Ganguli M, Fu B, Snitz BE, Hughes TF, and Chang CCH (2013). Mild cognitive impairment: incidence and vascular risk factors in a population-based cohort, *Neurology*, **80**, 2112–2120.
- Gili T, Cercignani M, Serra L, Perri R, Giove F, Maraviglia B, Caltagirone C, and Bozzali M (2011). Regional brain atrophy and functional disconnection across Alzheimer's disease evolution, *Journal of Neurology, Neurosurgery & Psychiatry*, **82**, 58–66.
- Grand JH, Caspar S, and MacDonald SW (2011). Clinical features and multidisciplinary approaches to dementia care, *Journal of Multidisciplinary Healthcare*, **4**, 125–147.
- Greicius MD, Srivastava G, Reiss AL, and Menon V (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI, *Proceedings of the National Academy of Sciences*, **101**, 4637–4642.
- Hall JE (2010). *Guyton and Hall Textbook of Medical Physiology e-Book*, Elsevier Health Sciences.
- Hänninen T, Hallikainen M, Tuomainen S, Vanhanen M, and Soininen H (2002). Prevalence of mild cognitive impairment: a population-based study in elderly subjects, *Acta Neurologica Scandinavica*, **106**, 148–154.
- Han Y, Lui S, Kuang W, Lang Q, Zou L, and Jia J (2012). Anatomical and functional deficits in patients with amnestic mild cognitive impairment, *PLoS One*, **7**, e28664.
- Hoerl AE and Kennard RW (1970). Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics*, **12**, 55–67.
- Hutchison RM, Womelsdorf T, Allen EA, *et al.* (2013). Dynamic functional connectivity: promise, issues, and interpretations, *Neuroimage*, **80**, 360–378.
- Johnson KA, Fox NC, Sperling RA, and Klunk WE (2012). Brain imaging in Alzheimer disease, *Cold Spring Harbor Perspectives in Medicine*, **2**, a006213.
- Johnson SC, Schmitz TW, Moritz CH, *et al.* (2006). Activation of brain regions vulnerable to Alzheimer's disease: the effect of mild cognitive impairment, *Neurobiology of Aging*, **27**, 1604–1612.
- Koenigs M, Barbey AK, Postle BR, and Grafman J (2009). Superior parietal cortex is critical for the manipulation of information in working memory, *Journal of Neuroscience*, **29**, 14980–14986.
- Kryscio RJ, Schmitt FA, Salazar JC, Mendiondo MS, and Markesbery WR (2006). Risk factors for transitions from normal to mild cognitive impairment and dementia, *Neurology*, **66**, 828–832.

- Li C, Wang J, Gui L, Zheng J, Liu C, and Du H (2011). Alterations of whole-brain cortical area and thickness in mild cognitive impairment and Alzheimer's disease, *Journal of Alzheimer's Disease*, **27**, 281–290.
- Li W, Qin W, Liu H, Fan L, Wang J, Jiang T, and Yu C (2013). Subregions of the human superior frontal gyrus and their connections, *Neuroimage*, **78**, 46–58.
- Liu Z, Zhang Y, Yan H, *et al.* (2012). Altered topological patterns of brain networks in mild cognitive impairment and Alzheimer's disease: a resting-state fMRI study, *Psychiatry Research: Neuroimaging*, **202**, 118–125.
- Martínez-Murcia FJ, Górriz JM, Ramírez J, Puntonet CG, Illán IA, and Alzheimer's Disease Neuroimaging Initiative (2013). Functional activity maps based on significance measures and independent component analysis, *Computer Methods and Programs in Biomedicine*, **111**, 255–268.
- Mechelli A, Humphreys GW, Mayall K, Olson A, and Price CJ (2000). Differential effects of word length and visual contrast in the fusiform and lingual gyri during. In *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **267**, 1909–1913.
- Menini A (2009). *The Neurobiology of Olfaction*, CRC Press.
- Mohs RC (1983). The Alzheimer's disease assessment scale: an instrument for assessing treatment efficacy, *Psychopharmacol Bull*, **19**, 448–450.
- Moradi E, Pepe A, Gaser C, Huttunen H, Tohka J, and Alzheimer's Disease Neuroimaging Initiative (2015). Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects, *Neuroimage*, **104**, 398–412.
- Morris JC (2005). Early-stage and preclinical Alzheimer disease, *Alzheimer Disease and Associated Disorders*, **19**, 163–165.
- Orrison WW (2008). *Atlas of Brain Function* (2nd ed), Thieme, New York.
- Petrella JR, Sheldon FC, Prince SE, Calhoun VD, and Doraiswamy PM (2011). Default mode network connectivity in stable vs progressive mild cognitive impairment, *Neurology*, **76**, 511–517.
- Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, and Kokmen E (1999). Mild cognitive impairment: clinical characterization and outcome, *Archives of Neurology*, **56**, 303–308.
- Rathore S, Habes M, Iftikhar MA, Shacklett A, and Davatzikos C (2017). A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages, *NeuroImage*, **155**, 530–548.
- Renier LA, Anurova I, De Volder AG, Carlson S, VanMeter J, and Rauschecker JP (2010). Preserved functional specialization for spatial processing in the middle occipital gyrus of the early blind, *Neuron*, **68**, 138–148.
- Salvador R, Suckling J, Coleman MR, Pickard JD, Menon D, and Bullmore ED (2005). Neurophysiological architecture of functional magnetic resonance images of human brain, *Cerebral Cortex*, **15**, 1332–1342.
- Shankle WR, Romney AK, Hara J, Fortier D, Dick MB, Chen JM, Chan T, and Sun X (2005). Methods to improve the detection of mild cognitive impairment. In *Proceedings of the National Academy of Sciences*, **102**, 4919–4924.
- Shenhav A, Botvinick MM, and Cohen JD (2013). The expected value of control: an integrative theory of anterior cingulate cortex function, *Neuron*, **79**, 217–240.
- Small GW, Chen ST, Komo S, *et al.* (1999). Memory self-appraisal in middle-aged and older adults with the apolipoprotein E-4 allele, *American Journal of Psychiatry*, **156**, 1035–1038.
- Smith SM, Vidaurre D, Beckmann CF, *et al.* (2013). Functional connectomics from resting-state fMRI, *Trends in Cognitive Sciences*, **17**, 666–682.
- Sperling RA, Aisen PS, Beckett LA, *et al.* (2011). Toward defining the preclinical stages of Alzheimer's

- disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease, *Alzheimer's & Dementia*, **7**, 280–292.
- Tibshirani R (1996). Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society: Series B (Methodological)*, **58**, 267–288.
- Tijms BM, Wink AM, de Haan W, van der Flier WM, Stam CJ, Scheltens P, and Barkhof F (2013). Alzheimer's disease: connecting findings from graph theoretical studies of brain networks, *Neurobiology of Aging*, **34**, 2023–2036.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, and Joliot M (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain, *Neuroimage*, **15**, 273–289.
- Wang H, Banerjee A, and Boley D (2011). Common component analysis for multiple covariance matrices. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 956–964.
- Wang J, Zuo X, Dai Z, *et al.* (2013). Disrupted functional brain connectome in individuals at risk for Alzheimer's disease, *Biological Psychiatry*, **73**, 472–481.
- Wang K, Liang M, Wang L, Tian L, Zhang X, Li K, and Jiang T (2007). Altered functional connectivity in early Alzheimer's disease: A resting-state fMRI study, *Human Brain Mapping*, **28**, 967–978.
- Wang XN, Zeng Y, Chen GQ, *et al.* (2016). Abnormal organization of white matter networks in patients with subjective cognitive decline and mild cognitive impairment, *Oncotarget*, **7**, 48953–48962.
- Warrier C, Wong P, Penhune V, Zatorre R, Parrish T, Abrams D, and Kraus N (2009). Relating structure to function: Heschl's gyrus and acoustic processing, *Journal of Neuroscience*, **29**, 61–69.
- Wee CY, Yap PT, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang L, and Shen D (2012). Resting-state multi-spectrum functional connectivity networks for identification of MCI patients, *PLoS One*, **7**, e37828.
- Wee CY, Yap PT, Shen D, and Alzheimer's Disease Neuroimaging Initiative (2013). Prediction of Alzheimer's disease and mild cognitive impairment using cortical morphological patterns, *Human Brain Mapping*, **34**, 3411–3425.
- Weisstein EW (2003). *Graph Diameter*, Wolfram Research, Inc., <https://mathworld.wolfram.com/>
- Wollman DE and Prohovnik I (2003). Sensitivity and specificity of neuroimaging for the diagnosis of Alzheimer's disease, *Dialogues in Clinical Neuroscience*, **5**, 89–99.
- Yao Z, Zhang Y, Lin L, Zhou Y, Xu C, Jiang T, and Alzheimer's Disease Neuroimaging Initiative (2010). Abnormal cortical networks in mild cognitive impairment and Alzheimer's disease, *PLoS Computational Biology*, **6**, e1001006.
- Ye J (2005). Generalized low rank approximations of matrices, *Machine Learning*, **61**, 167–191.
- Zhang D, Shen D, and Alzheimer's Disease Neuroimaging Initiative (2012). Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers, *PLoS One*, **7**, e33182.
- Zou H and Hastie T (2005). Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **67**, 301–320.