Mining Associations between MRI Morphometry Measurements and Beta-Amyloid/tau Burden

by

Jianfeng Wu

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2022 by the
Graduate Supervisory Committee:

Yalin Wang, Chair
Baoxin Li
Jianming Liang
Junwen Wang
Teresa Wu

ARIZONA STATE UNIVERSITY

August 2022

ABSTRACT

Beta-Amyloid(Aβ) plaques and tau protein tangles in the brain are now widely recognized as the defining hallmarks of Alzheimer's disease (AD), followed by structural atrophy detectable on brain magnetic resonance imaging (MRI) scans. However, current methods to detect Aβ/tau pathology are either invasive (lumbar puncture) or quite costly and not widely available (positron emission tomography (PET)). And one of the particular neurodegenerative regions is the hippocampus to which the influence of Aβ/tau on has been one of the research projects focuses in the AD pathophysiological progress.

In this dissertation, I proposed three novel machine learning and statistical models to examine subtle aspects of the hippocampal morphometry from MRI that are associated with Aβ /tau burden in the brain, measured using PET images. The first model is a novel unsupervised feature reduction model to generate a low-dimensional representation of hippocampal morphometry for each individual subject, which has superior performance in predicting Aβ/tau burden in the brain. The second one is an efficient federated group lasso model to identify the hippocampal subregions where atrophy is strongly associated with abnormal Aβ/Tau. The last one is a federated model for imaging genetics, which can identify genetic and transcriptomic influences on hippocampal morphometry. Finally, I stated the results of these three models that have been published or submitted to peer-reviewed conferences and journals.

TABLE OF CONTENTS

Page

CHAPTER Page

CHAPTER Page

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

Alzheimer's disease (AD) is now viewed as a gradual process that begins many years before the onset of detectable clinical symptoms. Measuring brain biomarkers and intervening at preclinical AD stages are believed to improve the probability of therapeutic success (Brookmeyer et al., 2007; Jack et al., 2016). Amyloid-β (Aβ) plaques and tau tangles are two specific protein pathological hallmarks of AD and are believed to induce neurodegeneration and structural brain atrophy consequentially observable from volumetric magnetic resonance imaging (MRI) scans. Brain Aβ and tau pathology can be measured using positron emission tomography (PET) with amyloid/tau-sensitive radiotracers or by using lumbar puncture to measure these proteins in samples cerebrospinal fluid (CSF). Even so, these invasive and expensive measurements are less attractive to subjects in the preclinical stage, and PET scanning is also not as widely available as MRI.

In the A/T/N system - a recently proposed research framework for understanding the biology of AD - the presence of abnormal levels of Aβ (A in A/T/N) in the brain or cerebrospinal fluid (CSF) is used to define the presence of biological Alzheimer's disease (Jack et al., 2016). An imbalance between production and clearance of Aβ occurs early in AD and is typically followed by the accumulation of tau (T in A/T/N) protein tangles (another key pathological hallmark of AD) and neurodegeneration (N in A/T/N) detectable on brain MRI scans (Hardy & Selkoe, 2002; Jack et al., 2016; Sperling et al., 2011). Therefore, there has been great interest in developing techniques to associate Aβ

and tau deposition with MRI measures (Ansart et al., 2020; Dahl et al., 2021; Ezzati et al., 2020; Petrone et al., 2019; W. Sun et al., 2020; Ten Kate et al., 2018; Tosun et al., 2013, 2014, 2016, 2021). In the structural MRI, the hippocampus is a primary target region across the spectrum of dementia research from clinically normal to late stages of AD (Cullen et al., 2020; Dong et al., 2019; B. Li et al., 2016; Shi et al., 2011). Cognitively unimpaired individuals with abnormally high Aβ burden have faster progression of hippocampal volume atrophy (Insel et al., 2017; L. Zhang et al., 2020). Additionally, tau burden in the brain, assessed using PET tracers, also strongly correlates with subsequent hippocampal volume atrophy (La Joie et al., 2020). However, the influence of Aβ/tau pathology on regional hippocampal atrophy in AD is still not fully understood.

Many hippocampal surface morphometry measures can be extract from MRI scans, such as radial distances (RD, distance between each surface points to its medial center) (Thompson et al., 2004), local area differences (Woods, 2003), and spherical harmonic analysis (Styner et al., 2005). Surface tensor-based morphometry (TBM) (Chung et al., 2008a) is an intrinsic surfacer statistic that examines spatial derivatives of the deformation maps that register brains to a common template and construct morphological tensor maps. And we recently proposed a multivariate TBM (mTBM) (Y. Wang et al., 2010) and later further combined RD and mTBM into surface multivariate morphometry statistics (MMS) (Y. Wang et al., 2011), which also show excellent performance in AD diagnosis prediction (Y. Fu et al., 2021; J. Zhang et al., 2021; J. Zhang, Li, et al., 2017).

Additionally, mounting evidence suggests that germline mutations (e.g., DNA SNPs) play an important role in AD etiology and progression (Jack et al., 2016). People with a parent or sibling with AD are more likely to develop the disease. The gene with the strongest association to AD is apolipoprotein E (*APOE*) and the e4 allele is associated with increased risk, whereas the e2 allele is associated with decreased risk. The odds ratios (ORs) for developing AD are 3.2 for individuals with one e4 allele and 14.9 with two e4 alleles, but only 0.6 with one e2 allele, when compared with individuals with two common e3 alleles in Caucasians (Bertram et al., 2007). Besides, progressive tau accumulation is more prominent in e4 carriers (Baek et al., 2020).

Therefore, I proposed three open problems. Firstly, whether the hippocampal morphometry features can be used to predict Aβ/tau deposition. Secondly, what is the association between Aβ/tau Measurements and hippocampal morphometry. Finally, what is the associations between Aβ/tau Measurements and hippocampal morphometry are different in the cohorts with different genotype.

In this dissertation, I will introduce three novel models I developed to address three problems. The first model is an unsupervised feature reduction model, named Patch Analysis-based Surface Correntropy-induced Sparse coding and max-pooling (PASCS-MP). This model can effectively generate a low-dimensional representation of hippocampal morphometry for each individual subject. Then, I trained machine learning models with these individual representations to predict brain Aβ positivity and measures for tau deposition of each person. These results compare favorably relative to measures derived from traditional algorithms, including hippocampal volume and surface area,

shape measures based on spherical harmonics (SPHARM) and our prior Patch Analysis-based Surface Sparse-coding and Max-Pooling (PASS-MP) methods. However, there are two main limitations for this work. Firstly, it cannot figure out the disease-related region of interest (ROI). Secondly, it can only work on a local machine.

Integrating data from multi-sites is common practice for large sample sizes and increased statistical power. An important direction of interest in multi-site neuroimaging research is federated learning – which offers an approach to learn from data spread across multiple sites without having to share the raw data directly or to centralize in any one location. In many cases, different institutions may not be readily able to share biomedical research data due to patient privacy concerns, data restrictions based on patient consent or IRB regulations, and legal complexities; this can present a major obstacle to pooling large scale datasets to discover robust and reproducible signatures of specific brain disorders. To remedy this distributed problem, a large-scale collaborative network, ENIGMA consortium, was built (Thompson et al., 2020). Accordingly, I also built the next two proposed models on a federated framework.

The second model is Federated Morphometry Feature Selection (FMFS) model to examine subtle aspects of hippocampal morphometry that are associated with Aβ/tau burden in the brain. Experimental results indicate that FMFS achieves an 89x speedup compared to other published state-of-the-art methods under five independent hypothetical institutions. In addition, the subiculum and cornu ammonis 1 (CA1 subfield) were identified as hippocampal subregions where atrophy is strongly associated with abnormal Aβ/Tau. As potential biomarkers for Aβ/tau pathology, the features from the identified

ROIs have greater power for predicting cognitive assessment and for survival analysis than five other imaging biomarkers. All the results indicated that FMFS is an efficient and effective tool to reveal associations between Aβ/tau burden and hippocampal morphometry. However, genetics is also a significant factor for AD and this model can only analyze the associations between Aβ/tau measurements and hippocampal morphometry.

To study the trios among Genotypes, Aβ/tau Measurements and Hippocampal Morphometry, I proposed a novel federated model for imaging genetics, Genotype-Expression-Imaging Data Integration (GEIDI), to identify genetic and transcriptomic influences on brain sMRI measures. I applied this model on two different projects. The first one is to discover and infer multimodal relationships among sMRI, GWAS, and transcriptomics. Experimental results demonstrated our proposed method outperformed state-of-the-art expression quantitative trait loci (eQTL) methods for detecting genetic and transcriptomic factors related to AD and has stable performance when data are integrated from multiple sites. Our GEIDI approach may offer novel insights into the relationship among image biomarkers, genotypes, and gene expression and help discover novel genetic targets for potential AD drug treatments. The second project is exploring how the genetic and transcriptomic factor influence the morphometry of the hippocampal surface. Experimental results demonstrate the strong association among APOE, tau deposition and the morphometry of hippocampus. And the identified hippocampal ROIs are quite consistent with the second project.

CHAPTER 2

RELATED WORK

2.1.Hippocampal Morphometry Measures

There are many surface morphometry measures, such as radial distances (RD, distance between each surface points to its medial center) (Thompson et al., 2004), local area differences (Woods, 2003), and spherical harmonic analysis (Styner et al., 2005). Surface tensor-based morphometry (TBM) (Chung et al., 2008a) is an intrinsic surfacer statistic that examines spatial derivatives of the deformation maps that register brains to a common template and construct morphological tensor maps. We recently proposed a multivariate TBM (mTBM) (Y. Wang et al., 2010) and later further combined RD and mTBM into surface multivariate morphometry statistics (MMS) (Y. Wang et al., 2011).

2.2.Unsupervised Feature Reduction

These vertex-wise surface morphometry features are high-fidelity measures to describe the local deformation of the surface and can provide detailed localization and visualization of regional atrophy or expansion (Yao et al., 2018) and development (Thompson et al., 2000a). However, the high dimensionality of such features is likely to cause problems for classification. Feature reduction methods proposed by (Davatzikos et al., 2008; D. Sun et al., 2009) may ignore the intrinsic properties of a structure's regional morphometry. I use the technique of sparse coding and dictionary learning (Mairal et al., 2009; Q. Zhang & Li, 2015) to learn meaningful features. Dictionary learning has been successful in many image processing tasks as it can concisely model natural image patches. In this work, I propose a novel sparse coding and dictionary learning method

with an l_1-regularized correntropy loss function named Correntropy-induced Sparse-coding (CS), which is expected to improve the computational efficiency compared to Stochastic Coordinate Coding (SCC) (Lin et al., 2014). The model is named **Patch Analysis-based Surface Correntropy-induced Sparse coding and max-pooling (PASCS-MP).**

2.3.Federated Morphometry Feature Selection

Federated learning models, such as decentralized independent component analysis (B. T. Baker et al., 2015), sparse regression (Plis et al., 2016), and distributed deep learning (Kaissis et al., 2021; Stripelis et al., 2021; Warnat-Herresthal et al., 2021), have made solid progress with leveraging multivariate image features for statistical inferences, allowing iterative computation on remote datasets. Some other recent studies focus on multivariate linear modeling (Silva et al., 2020), federated gradient averaging (Remedios et al., 2020), and unbalanced data for multi-site (Yeganeh et al., 2020). To our knowledge, these methods have not yet been applied to detect multimodal associations in Alzheimer's disease research, such as finding anatomically abnormal regions on MRI that are associated with Aβ/tau pathology defined using PET. Here I propose a novel framework, **Federated Morphometry Feature Selection (FMFS)**, to detect the association between hippocampal morphometry markers and Aβ/tau burden.

2.4.Imaging Genetics

Various imaging genetics methods have been developed to integrate imaging and genetic data. However, most studies have focused on imaging, imaging combined with GWAS data (Chauhan et al., 2015; Grasby et al., 2020; J. Q. Li et al., 2017), imaging

with transcriptomics (Ritchie et al., 2018), or GWAS with transcriptomics (Albert & Kruglyak, 2015). For example, imaging genetics methods have been used to link SNPs with image features (Stein et al., 2010), and expression quantitative trait loci (eQTL) have been used to discover *APOE*-related genes (A. Zhang et al., 2018). However, relatively few methods have been developed to integrate GWAS/WES/WGS, imaging, and transcriptomic data to infer multimodal relationships. For instance, [49] (Liu et al., 2017) use a brain-wide gene expression profile available in the Allen Human Brain Atlas (AHBA) as a 2-D prior to guide the brain imaging genetics association analysis. Their transcriptome-guided SCCA (TG-SCCA) framework incorporates the gene expression information into the traditional SCCA model. Such a multimodal approach may give us a more holistic view of the evidence from multiple sources to provide novel insights on the molecular mechanisms of AD pathogenesis and prognosis. Besides, both gene expression and imaging features are dynamic and change with time and throughout the disease, whereas germline SNPs are unchanged over an individual's lifetime. Therefore, I proposed a novel federated model for imaging genetics, **Genotype-Expression-Imaging Data Integration (GEIDI)**, to study SNP-image-gene expression relationships to consider both the dynamic changes in imaging and gene expression features and understand how they are affected by an individual's SNPs. Such knowledge will provide novel insights into the relationship among image biomarkers, genotypes and gene expression, and may help discover novel genetic targets for pharmaceutical interventions.

CHAPTER 3

METHODS

3.1.Image Preprocessing

Firstly, we use FIRST (FMRIB's Integrated Registration and Segmentation Tool) (Patenaude et al., 2011) to segment the original MRI data and map the hippocampus substructure. After obtaining a binary segmentation of the hippocampus, we use a topology-preserving level set method (Han et al., 2003) to build surface models. Based on that, the marching cubes algorithm (Lorensen & Cline, 1987) is applied to construct triangular surface meshes. Then, to reduce the noise from MR image scanning and to overcome partial volume effects, surface smoothing is applied consistently to all surfaces. Our surface smoothing process consists of mesh simplification using *progressive meshes* (Hoppe, 1996) and mesh refinement by the Loop subdivision surface method (Loop, 1987). Similar procedures adopted in a number of our prior studies (Colom et al., 2013; Luders et al., 2013; Monje et al., 2013; Shi et al., 2015; Shi, Thompson, et al., 2013; Shi, Wang, et al., 2013; Y. Wang et al., 2010, 2012) have shown the smoothed meshes are accurate approximations to the original surfaces, with a higher signal-to-noise ratio (SNR).

To facilitate hippocampal shape analysis, we generate a conformal grid (150×100) on each surface, which is used as a canonical space for surface registration. On each hippocampal surface, we compute its conformal grid with a holomorphic 1-form basis (Wang et al., 2010, 2007). We adopt surface conformal representation (Shi et al., 2015; Shi, Thompson, et al., 2013) to obtain surface geometric features for automatic surface

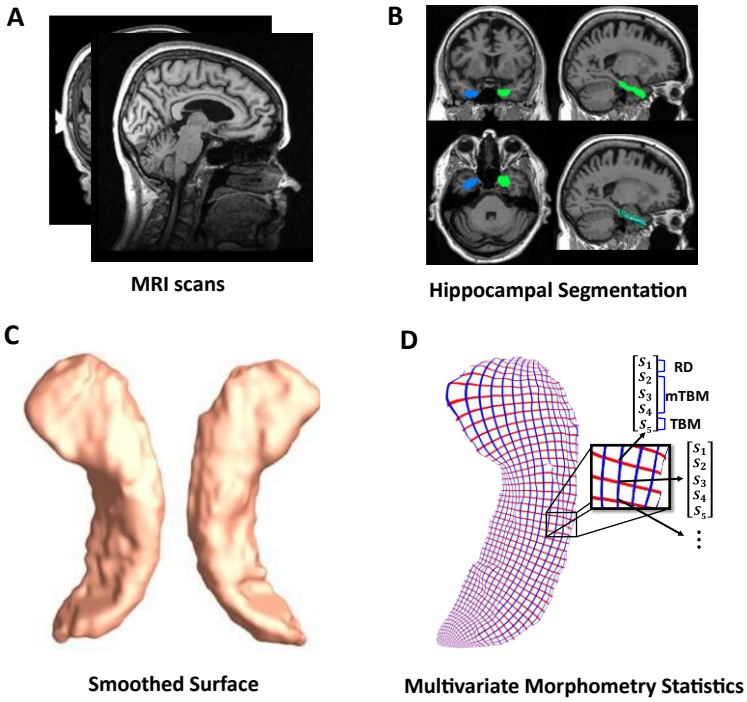

Figure 3.1.1. Image Preprocessing

registration. This consists of the conformal factor and mean curvature, encoding both intrinsic surface structure and information on its 3D embedding. After we compute these two local features at each surface point, we compute their summation and then linearly scale the dynamic range of the summation into the range 0-255, to obtain a feature image for the surface. We further register each hippocampal surface to a common template surface. With surface conformal parameterization and conformal representation, we generalize the well-studied image fluid registration algorithm (Bro-Nielsen and Gramkow, 1996; D'Agostino et al., 2003) to general surfaces. Furthermore, most of the image registration algorithms in the literature are not symmetric, i.e., the correspondences between the two images depending on which image is assigned as the deforming image and which is the non-deforming target image. An asymmetric algorithm can be problematic as it tends to penalize the expansion of image regions more than shrinkage

(Rey et al., 2002). Thus, in our system, we further extend the surface fluid registration method to an inverse-consistent framework (Leow et al., 2005). The obtained surface registration is diffeomorphic. For details of our inverse-consistent surface fluid registration method, we refer to (Shi, Thompson, et al., 2013).

After parameterization and registration, we establish a one-to-one correspondence map between hippocampal surfaces. This makes it effective for us to compare and analyze surface data. Besides, each surface has the same number of vertices (150×100) as shown in **Figure. 3.1.1**. The intersection of the red curve and the blue curve is a surface vertex, and at each vertex, I calculate features, the radial distance (RD) and the surface metric tensor used in tensor-based morphometry (TBM) (Chung et al., 2008a; Davatzikos, 1996; Thompson et al., 2000b; Woods, 2003) and multivariate tensor-based morphometry (mTBM). The RD (a scalar at each vertex) represents the thickness of the shape at each vertex to the medical axis (Pizer et al., 1999; Paul M. Thompson et al., 2004), this reflects the surface differences along the surface normal directions. The medial axis is determined by the geometric center of the isoparametric curve on the computed conformal grid (Y. Wang et al., 2011). The axis is perpendicular to the isoparametric curve, so the thickness can be easily calculated as the Euclidean distance between the core and the vertex on the curve. The mTBM statistics (a 3×1 vector at each vertex) have been frequently studied in our prior work (Shi et al., 2015; Shi, Wang, et al., 2013; Y. Wang et al., 2009, 2010). They measure local surface deformation along the surface tangent plane and show improved signal detection sensitivity relative to more standard tensor-based morphometry (TBM) measures computed as the determinant of the

Jacobian matrix (Y. Wang et al., 2013). RD and mTBM jointly form a new feature, known as the surface multivariate morphometry statistics (MMS). Therefore, MMS is a 4×1 vector at each vertex. The hippocampal surface in each hemisphere has $15{,}000$ vertices, so the feature dimensionality for each hippocampus in each subject is 60,000.

3.2.Surface Feature Dimensionality Reduction

This work develops the **Patch Analysis-based Surface Correntropy-induced Sparse coding and max-pooling (PASCS-MP)** to predict individual Aβ burden (see **Figure 3.2.1** for the processing pipeline). In panel (1), hippocampal structures are segmented from registered brain MR images with FIRST from the FMRIB Software Library (FSL) (Paquette et al., 2017; Patenaude et al., 2011). Hippocampal surface meshes are constructed with the marching cubes algorithm (Lorensen & Cline, 1987). In panel (2), hippocampal surfaces are parameterized with the holomorphic flow segmentation method (Y. Wang et al., 2007). After the surface fluid registration algorithm, the hippocampal MMS features are calculated at each surface point. We propose a PASCS-MP and classification system to refine and classify MMS patches in individuals with different Aβ status. We randomly select patches on each hippocampal surface and generate a sparse code for each patch with our novel PASCS. Next, we adopt a max-pooling operation on the learned sparse codes of these patches to generate a new representation (a vector) for each subject. Finally, we train binary random forest classifiers on individual sparse codes in people with different Aβ status; we validate them with 10-fold cross-validation. The whole system is publicly available, http://gsl.lab.asu.edu/software/pass-mp/ .

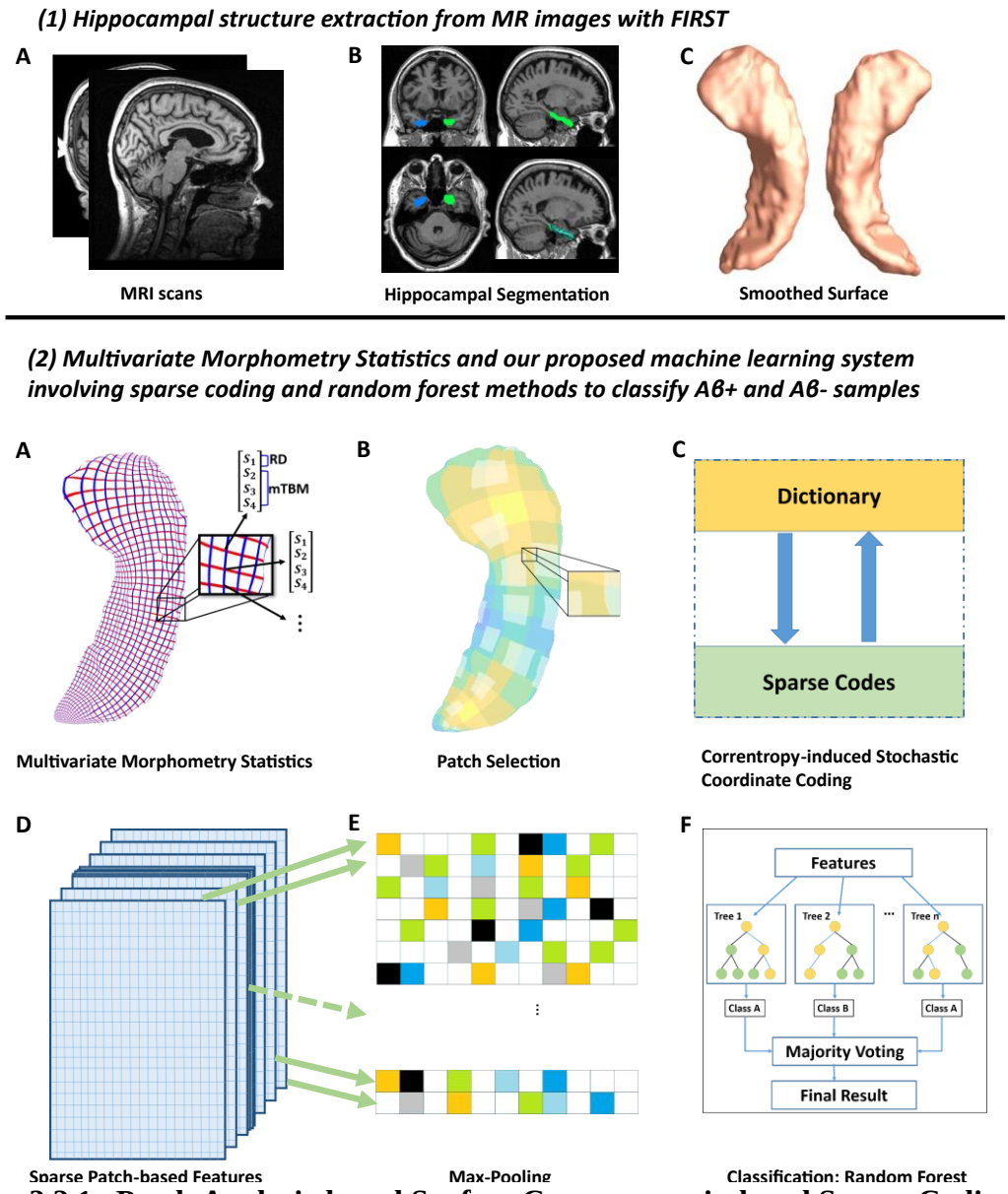

Figure 3.2.1. Patch Analysis-based Surface Correntropy-induced Sparse Coding and Max-pooling (PASCS-MP) Pipeline. Panel (1) shows hippocampal surfaces generated from brain MRI scans. In subfigure A of panel (2), surface-based Multivariate Morphometry Statistics (MMS) are calculated after fluid registration of surface coordinates across subjects. MMS is a 4×1 vector on each vertex, including radial distance (scalar) and multivariate tensor-based morphometry (3×1 vector). In subfigure B and C, we randomly select patches on each hippocampal surface and generate a sparse code for each patch with our novel Patch Analysis-based Surface Correntropy-induced Sparse-coding (PASCS) method. In subfigure D and E, we apply the max pooling operation to the learned sparse codes to generate a new representation (a vector) for each subject. In subfigure F, we train binary random forest classifiers on these representations and validate them with 10-fold cross-validation.

3.2.1. Patch Analysis-based Surface Correntropy-induced Sparse coding

The above-mentioned vertex-wise surface morphometry feature, MMS, is a high-fidelity measure to describe the local deformation of the surface and can provide detailed localization and visualization of regional atrophy or expansion (Yao et al., 2018) and development (Thompson et al., 2000a). However, the high dimensionality of such features is likely to cause problems for classification. Feature reduction methods proposed by (Davatzikos et al., 2008; D. Sun et al., 2009) may ignore the intrinsic properties of a structure's regional morphometry. Therefore, we introduce the following feature reduction method for the vertex-wise surface morphometry features.

The surface MMS feature dimension is typically much larger than the number of subjects, i.e., the so-called *high dimension-small sample problem*. To extract useful surface features and reduce the dimension before making predictions, this work first randomly generates square windows on each surface to obtain a collection of small image patches with different amounts of overlap. In our prior AD studies (Wu et al., 2018; J. Zhang, Shi, et al., 2016; J. Zhang, Stonnington, et al., 2016), we discuss the most suitable patch size and number. Therefore, in this work, we adopt the same optimal experimental settings, as $1{,}008$ patches (patch size=10×10 vertices) for each subject (504 patches for each side of the hippocampal surface). As these patches are allowed to overlap, a vertex may be contained in several patches. The zoomed-in window in subfigure (b) of panel (2) in **Figure 3.2.1** shows overlapping areas on selected patches. After that, we use the technique of sparse coding and dictionary learning (Mairal et al., 2009) to learn meaningful features. Dictionary learning has been successful in many image processing

tasks as it can concisely model natural image patches. In this work, we propose a novel sparse coding and dictionary learning method with an l_1-regularized correntropy loss function named *Correntropy-induced Sparse-coding (CS)*, which is expected to improve the computational efficiency compared to Stochastic Coordinate Coding (SCC) (Lin et al., 2014). Formally speaking, correntropy is a generalized similarity measure between two scalar random variables U and V, which is defined by $\mathcal{V}_\sigma(U,V) = \mathbb{E}\mathcal{K}_\sigma(U,V)$. Here, $\mathcal{K}_\sigma$ is a Gaussian kernel given by $\mathcal{K}_\sigma(U,V) = \exp\{-(\mathrm{u} - \mathrm{v})^2/\sigma^2\}$ with the scale parameter $\sigma > 0$, (u-v) being a realization of (U, V) (Feng et al., 2015; Gui et al., 2017). Utilizing the correntropy measure as a loss function will reduce the negative influence of non-Gaussian noise in the data.

Classical dictionary learning techniques (H. Lee et al., 2007; Olshausen & Field, 1997) consider a finite training set of feature maps, $X = (x_1, x_2, \dots, x_n)$ in $R^{p\times n}$. In our study, X is the set of MMS features from *n* surface patches of all the samples. All the MMS features on each surface patch, x_i, is reshaped to a p-dimensional vector. And we desire to generate a new set of sparse codes, $Z = (z_1, z_2, \dots, z_n)$ in $R^{m\times n}$ for these features. Therefore, we aim to optimize the empirical cost function as **Eq. (3.2.1)**.

$$f(D, z_i) \triangleq \sum_{i=1}^{n} l(x_i, D, z_i) \tag{3.2.1}$$

where $D \in R^{p\times m}$ is the dictionary and $z_i \in R^m$ is the sparse code of each feature vector. $l(x_i, D, z_i)$ is the loss function that measures how well the dictionary D and the sparse code z_i can represent the feature vector x_i. Then, x_i can be approximated by $x_i = Dz_i$. In this way, we convert the p-dimensional feature vector, x_i, to a m-dimensional sparse code, z_i, where m is the dimensionality of the sparse code and the dimensionality could

be arbitrary. In this work, we introduce the correntropy measure (Gui et al., 2017) to the loss function and define the l_1-sparse coding optimization problem as **Eq. (3.2.2)**

$$\min_{D,z_i} \frac{1}{2}\sum_{i=1}^{n} exp\left(-\frac{\|Dz_i - x_i\|_2^2}{\sigma^2}\right) + \lambda \sum_{i=1}^{n} \|z_i\|_1 \tag{3.2.2}$$

where λ is the regularization parameter, σ is the kernel size that controls all properties of correntropy. $\|\cdot\|_2$ and $\|\cdot\|_1$are the l_2-norm and l_1-norm and *exp()* represents the exponential function. The first part of the loss function measures the degree of the image patches' goodness and the correntropy may help remove outliers. Meanwhile, the second part is well known as the l_1 penalty (W. J. Fu, 1998) that can yield a sparse solution for z_i and select robust and informative features. Specifically, there are *m* columns (atoms) in the dictionary D and each atom is $d_j \in R^p, j = 1, 2, \dots, m$. To avoid D from being arbitrarily large and leading to arbitrary scaling of the sparse codes, we constrain each l_2-norm of each atom in the dictionary no larger than one. We will let *C* become the convex set of matrices verifying the constraint as **Eq. (3.3)**.

$$\mathrm{C} \triangleq \left\{D \in R^{p\times m} s.t. \forall j = 1,2, \dots, m, d_j^T d_j \leq 1\right\} \tag{3.2.3}$$

Note that, the empirical problem cost $f(D, z_i)$ is not convex when we jointly consider the dictionary D and the coefficients Z. But the function is convex concerning each of the two variables, D, and Z, when the other one is fixed. Since it takes much time to solve D and Z when dealing with large-scale data sets and a large-size dictionary, we adopt the framework in the stochastic coordinate coding (SCC) algorithm (Lin et al., 2014), which can dramatically reduce the computational cost of the sparse coding, while keeping a comparable performance.

To solve this optimization problem, we reformulate the first part of the equation by the half-quadratic technique (Nikolova & Ng, 2006) and then the objective can be solved as the minimization problem **Eq. (3.2.4)**:

$$\min_{D,z_i} \frac{1}{2}\sum_{i=1}^{n} h_i \|Dz_i - x_i\|_2^2 + \lambda \sum_{i=1}^{n} \|z_i\|_1 , h_i = \exp\left(-\frac{\|Dz_i - x_i\|_2^2}{\sigma^2}\right). \quad (3.2.4)$$

Here the auxiliary variable, h_i, will be updated in each update iteration. At each iteration, we update D and Z alternately, which means we firstly fix D and update the sparse code Z with coordinated descent (CD) and then fix Z to update the dictionary D via stochastic gradient descent (SGD). As our optimization method is stochastic, we only update the sparse code and dictionary with only one signal for each iteration. In the following paragraphs, we will discuss the optimization in one iteration with only one signal.

If a signal, $x = (x_1, x_2, \dots, x_p)^T \in R^p$, is given, we first update its corresponding sparse code, $z = (z_1, z_2, \dots, z_m)$, via CD. Let z_l denote the l-th entry of z and d_{kl} represents the k-th item of d_l. d_l is the l-th atom/column of the dictionary D. Then, we can calculate the partial derivative of z_l in the first part of the function, $f(D, z_i)$, as

$$\begin{aligned}
\frac{\partial}{\partial z_l} c(D, \mathrm{z}) &= \frac{\partial}{\partial z_l}\frac{1}{2} h\|Dz - x\|_2^2 = -h\sum_{k=1}^{p} d_{kl}\left(x_k - \sum_{r=1}^{m} d_{kr} z_r\right) \\
&= -h\sum_{k=1}^{p} d_{kl}\left(x_k - \sum_{r\neq l}^{m} d_{kr} z_r - d_{kl} z_l\right) \\
&= -h\sum_{k=1}^{p} d_{kl}\left(x_k - \sum_{r\neq l}^{m} d_{kr} z_r\right) + h z_l \sum_{k=1}^{p} (d_{kl})^2 \\
&= -\rho_l + h z_l v_l \quad (3.2.5)
\end{aligned}$$

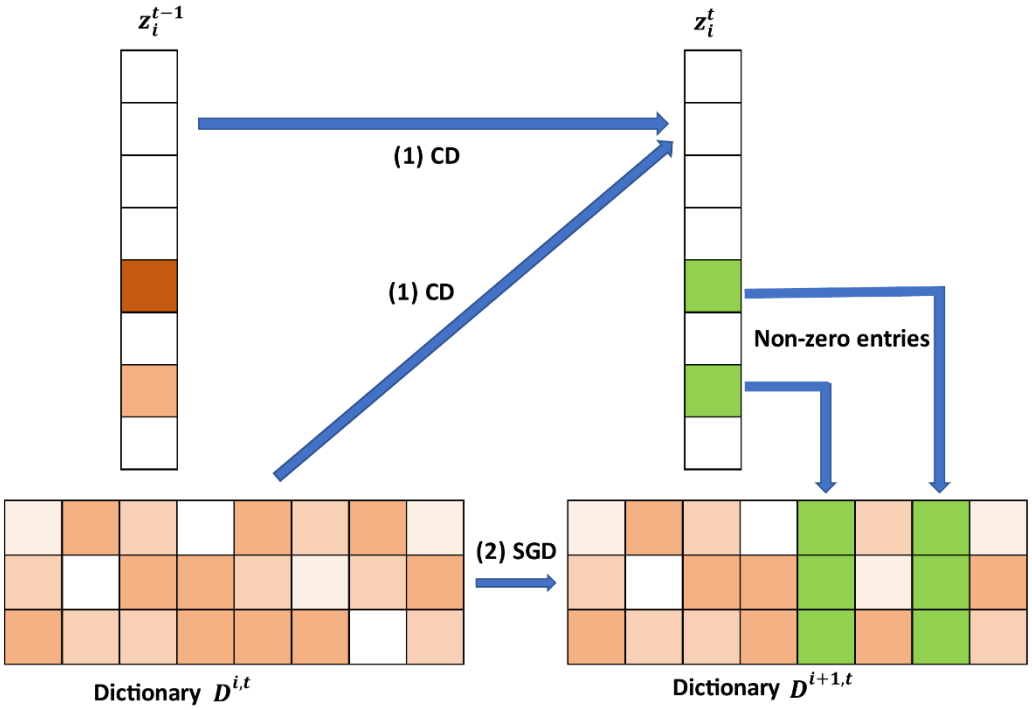

Figure 3.2.2. Illustration of One Iteration of the Proposed Patch Analysis-based Surface Correntropy-induced Sparse-coding (PASCS) Algorithm. The input is many 10 × 10 patches on each surface based on our multivariate morphometry statistics (MMS). With an image patch x_i, PASCS performs one step of coordinate descent (CD) to find the support and the sparse code. Meanwhile, PASCS performs a few steps of CD on supports (non-zero entries) to obtain a new sparse code z_i^k. Then, PASCS updates the supports (green boxes in the figure) of the dictionary by stochastic gradient descent (SGD) to obtain a new dictionary $D^{i+1,t}$. Here, t represents the t-th epoch; i represents the i-th patch.

where $\rho_l = h\sum_{k=1}^{p} d_{kl}(x_k - \sum_{r\neq l}^{m} d_{kr}z_r)$, $v_l = \sum_{k=1}^{p}(d_{kl})^2$ and h is the auxiliary variable for the signal. Since we normalize the atom, d_l, in each iteration, v_l can be ignored. Then, we compute the subdifferential of the lasso loss function and equate it to zero to find the optimal solution as follows:

$$\frac{\partial}{\partial z_l} f(D, \mathrm{z}) = \frac{\partial}{\partial z_l} c(D, \mathrm{z}) + \frac{\partial}{\partial z_l}\lambda\|z\|_1 = -\rho_l + hz_l v_l + \frac{\partial}{\partial z_l}\lambda\|z\|_1 = 0 \qquad (3.2.6)$$

Then, according to the derivative of the l_1-norm, we can have the following equations.

$$\begin{cases} -\rho_l + hz_l v_l - \lambda = 0 & if\ z_l < 0 \\ -\rho_l - \lambda \le 0 \le -\rho_l + \lambda\ if\ z_l = 0 \\ -\rho_l + hz_l v_l + \lambda = 0 & if\ z_l > 0 \end{cases} \tag{3.2.7}$$

Finally, we can get the soft thresholding function as:

$$z_l = \begin{cases} \frac{\rho_l + \lambda}{hv_l} & for\ \rho_l < -\lambda \\ 0 \quad for\ -\lambda \le \rho_l \le \lambda \\ \frac{\rho_l - \lambda}{hv_l} & for\ \rho_l > \lambda \end{cases} \tag{3.2.8}$$

After we update the sparse code, we propose the following strategy to accelerate the convergence for updating the dictionary D. The atom, d_l will stay unchanged if z_l is zero since $\nabla_{d_l} = h(Dz - x)z_l = 0$. Otherwise, as shown in **Figure 3.3.2**, we can update the l-th atom of the dictionary D as $d_l \leftarrow d_l - \gamma_l h(Dz - x)z_l$. γ_l is the learning rate provided by an approximation of the Hessian: $R \leftarrow R + zz^T$ and γ_l is given by $1/r_{ll}$, where r_{ll} is the item at the l-th row and l-th column of the Hessian matrix R. The pseudo-code of the model was shown in **Alg.3.2.1**, dubbed as PASCS.

3.2.2. Pooling and Predicting

After we get the sparse code (the dimension is *m*) for each patch, the dimensionality of sparse codes for each subject is still too large for classification, which is $m \times 1008$. Therefore, we apply Max-pooling to reduce the feature dimensionality for each subject. Max-pooling (Boureau et al., 2010) is a way of taking the most responsive node of a given region of interest and serves as an important layer in the convolutional neural network architecture. In this work, we compute the maximum value of a particular feature over all sparse codes of a subject and generate a new representation for each subject, which is an *m*-dimensional vector. These summary representations are much

lower in dimension, compared to using all the extracted surface patch features; this can improve results generalizability via less over-fitting. Finally, I choose the random forest algorithm (Liaw and Wiener, 2002) for the binary classification and ridge regression for predicting measurements of Aβ and tau.

Alg. 3.2.1 Patch Analysis-based Surface Correntropy-induced Sparse-coding

Require: Data set $\boldsymbol{X} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_n})$ in $\boldsymbol{R^{p\times n}}$

Ensure: Dictionary $\boldsymbol{D} \in \mathbb{R}^{\boldsymbol{p\times m}}$ and sparse codes $\boldsymbol{Z} = (\boldsymbol{z_1}, \boldsymbol{z_2}, \dots, \boldsymbol{z_n}) \in \mathbb{R}^{\boldsymbol{m\times n}}$

Initialize: $\boldsymbol{D^{1,1}}, \boldsymbol{R} = \mathbf{0}, \boldsymbol{z_i^0} = \mathbf{0}, \boldsymbol{h_i^0} = \mathbf{1}, \boldsymbol{i} = \mathbf{1}, \dots, \mathbf{n}$

1: for $\boldsymbol{t} = \mathbf{1}$ to $\boldsymbol{\tau}$ do

2: for $\boldsymbol{i} = \mathbf{1}$ to $\boldsymbol{n}$ do

3: Get an image patch $\boldsymbol{x_i}$ from $\boldsymbol{X}$.

4: Update $\boldsymbol{z_i^t}$ via coordinate descent:

$$\boldsymbol{z_i^t} \leftarrow \boldsymbol{CD}\left(\boldsymbol{x_i}, \boldsymbol{D^{i,t}}, \boldsymbol{z_i^{t-1}}\right).$$

5: Update Hessian matrix and the learning rate:

$$\boldsymbol{R} \leftarrow \boldsymbol{R} + \boldsymbol{z_i^t}(\boldsymbol{z_i^t})^{\boldsymbol{T}}, \boldsymbol{\gamma_{i,l}} = \mathbf{1}/\boldsymbol{r_{ll}}.$$

6: Update the support of the dictionary via SGD for non-zero entry $\boldsymbol{z_{i,l}^t}$

(and normalize it):

$$\boldsymbol{d_l^{i+1,t}} \leftarrow \boldsymbol{d_l^{i,t}} - \boldsymbol{\gamma_{i,l}}\boldsymbol{h_i}(\boldsymbol{D^{i,t}}\boldsymbol{z_i^t} - \boldsymbol{x_i})\boldsymbol{z_{i,l}^t}.$$

7: Update auxiliary variable $\boldsymbol{h_i}$:

$$\boldsymbol{h_i} = \mathbf{exp}\left(-\left\|\boldsymbol{D^{i,t}}\boldsymbol{z_i^t} - \boldsymbol{x_i}\right\|_2^2/\boldsymbol{\sigma^2}\right).$$

8: If $\boldsymbol{i} = \boldsymbol{n}$, Then $\boldsymbol{D^{1,t+1}} = \boldsymbol{D^{n,t}}$.

9: end for

10: end for

Output: $\boldsymbol{D} = \boldsymbol{D^{n,\tau}}$ and $\boldsymbol{z_i} = \boldsymbol{z_i^\tau}$ for $\boldsymbol{i} = \mathbf{1}, \dots, \mathbf{n}$

3.3.Federated Morphometry Feature Selection

In this work, we develop a **Federated Morphometry Feature Selection (FMFS)** model to detect the influence of Aβ and tau deposition on hippocampal shape deformity and to better support the future prediction of AD pathology as shown in **Figure 3.3.1**. In panel (1), each institution first extracts the morphometric features locally. The hippocampal structures are segmented from registered brain MR images and smoothed hippocampal surfaces are further generated. After the surface parameterization and fluid registration, the hippocampal radial distance (RD) and tensor-based morphometry (TBM) features are calculated at each surface point. Each institution selects patches on each hippocampal surface and reshapes the grouped features (RD or TBM on each patch are one group) of each subject to a vector. Next, in panel (2), taking each Aβ/tau measurement as the dependent variable, the institutions perform the federated feature selection model on these patches of features to generate local hippocampal regions of interest (ROIs) for each Aβ/tau measurement.

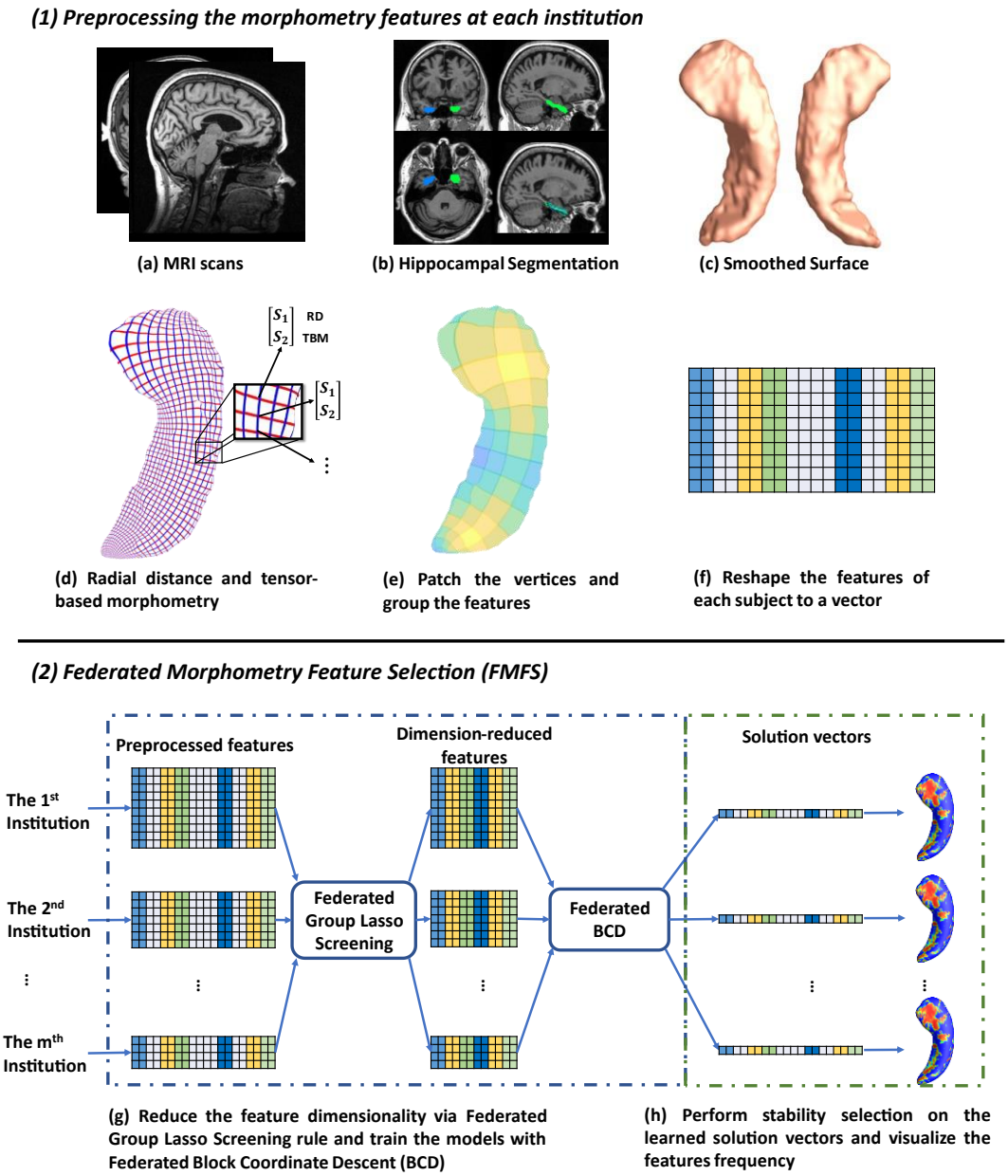

Figure 3.3.1. Federated Morphometry Feature Selection (FMFS) Pipeline. Panel (1) shows the steps for each institution to extract morphometric features locally. The hippocampal structures are segmented from registered brain MR images and smoothed hippocampal surfaces are then generated. After the surface parameterization and fluid registration, the hippocampal radial distance (RD) and tensor-based morphometry (TBM) features are calculated at each surface point. Each institution selects patches on each hippocampal surface and reshapes the grouped features of each subject into a vector. Next, in panel (2), taking Aβ/tau measurements as dependent variables, the institutions perform the federated feature selection model on these patches of features to generate hippocampal local regions of interest (ROIs) for each Aβ/tau measurement.

3.3.1. Federated Group Lasso Regression

Group LASSO (Yuan & Lin, 2006) is a widely-used technique for group-wise feature selection in high dimensional data. A group-LASSO linear regression has the following optimizing problem:

$$\min_{\beta \in \mathrm{R^p}} F(\beta) \ = \frac{1}{2}\left\|y - \sum_{g=1}^{G} X_g \beta_g\right\|_2^2 + \lambda \sum_{g=1}^{G} w_g \left\|\beta_g\right\|_2, \qquad (3.3.1)$$

where $X_g \in \mathrm{R}^{N \times p_g}$ is the feature matrix, and y denotes the N dimensional response vector. Group LASSO divides the original feature matrix $X \in \mathrm{R}^{N \times p}$ into G different groups, where X_g represents the features in gth group and w_g is the weight for this group. After solving the group LASSO problem, we get the G solution vectors, $\beta_1, \beta_2, \ldots, \beta_G$. The dimensionality of each group, p_g, can be arbitrary and the whole solution vector β is $[\beta_1, \beta_2, \ldots, \beta_G] \in \mathrm{R}^p$. Additionally, λ is a positive regularization parameter to control the sparsity of the solution vector, and w_g is the weight for gth group of features.

There are many optimization methods to solve the group LASSO problem; block coordinate descent (BCD) (Qin et al., 2013) is one of the most efficient. Instead of updating all the variables at the same time, BCD only updates one or several blocks of variables at each epoch. Therefore, for the group LASSO problem, it can optimize one group of variables while keeping the other ones fixed. Based on this idea, we proposed a federated block coordinate descent (FBCD) to solve our problem.

(Q. Li et al., 2016) proposed an optimization model, the local query model (LQM), which preserves the data privacy at each institution. We assume that there are *I*

institutions, and each of them owns a private data set (X^i, y^i). We can reformulate the problem (3.3.1) as

$$min_{\beta} \sum_{i=1}^{I} f^i\left(X^i, y^i; \beta\right) + \lambda \sum_{g=1}^{G} w_g \left\|\beta_g\right\|_2, \tag{3.3.2}$$

where $f^i\left(X^i, y^i; \beta\right) = \frac{1}{2}\left\|y^i - \sum_{g=1}^{G} X_g^i \beta_g\right\|_2^2$ is the least square loss of the ith institution. We then have the global gradient,

$$\nabla f(X, y; \beta) = \sum_{i=1}^{I} \nabla f^i\left(X^i, y^i; \beta\right) = \sum_{i=1}^{I} \left(X^i\right)^T \left(X^i \beta - y^i\right). \tag{3.3.3}$$

Each of the local institutions calculates its own gradient locally and uploads it to the master server. The master server will compute the global gradient, $\nabla f(X, y; \beta)$, by adding all $\nabla f^i\left(X^i, y^i; \beta\right)$. It then assigns the global update gradient $\nabla f(X, y; \beta)$ back to all the local institutions to compute β. Then, β is updated locally with the shrinkage function at the 6th line of Algorithm 1. Our proposed Federated Block Coordinate Descent (FBCD) method is outlined in **Alg. 3.3.1**.

Algorithm 3.3.1. Federated Block Coordinate Descent (FBCD)

Input: Data pairs from the i institutions $(X^1, y^1), \dots, (X^i, y^i), \dots, (X^I, y^I)$ with group information and the regularization parameter λ

Output: The learned solution β

Initialize: $\beta = \mathbf{0} \in R^p$

1: **while** *convergence and maximum number of iterations are not reached* **do**

2: Randomly select gth group to optimize

3: Compute the local gradient of gth group: $\nabla f^i(X_g^i) = [X_g^i]^T (X_g^i \beta - y^i)$

4: Compute the global gradient by LQM: $\nabla f(X_g) = \sum_{i=1}^{I} \nabla f^i(X_g^i)$

5: $\beta_g = \beta_g - \nabla f(X_g) / \|X_g\|_2^2$

6: $$\beta_g = \begin{cases} \beta_g - \frac{\lambda w_g}{\|\beta_g\|_2} \beta_g, & if \ \|\beta_g\|_2 > \frac{\lambda w_g}{\|X_g\|_2^2} \\ 0 \in \mathbb{R}^{p_g}, & if \ \|\beta_g\|_2 \leq \frac{\lambda w_g}{\|X_g\|_2^2} \end{cases}$$

3.3.2. Federated Screening for Group Lasso

Finding the optimal value for the regularization parameter λ is a common problem in LASSO techniques. The most frequently used methods, such as cross-validation and stability selection, solve it by trying a sequence of regularization parameters, $\lambda_1 > \cdots > \lambda_\kappa$; this can be very time-consuming. Instead, the enhanced dual polytope projection rule (EDPP) (J. Wang et al., 2015) achieved a 200x speedup on the cross-validation in real-world applications, by using information derived from the solution of the previously tried regularization parameter. For the group LASSO problem, the gth group of features X_g

can be discarded if it satisfies the screening rule, $\|J_g\|_2 \leq w_g(2\lambda - \lambda_{max})$ where $\lambda_{max} = \max_g \frac{\|L_g\|_2}{w_g}$ and J_g and L_g are the elements of J and L defined in **Alg. 3.3.2**. The screening rule is based on the uniqueness and non-expansiveness of the optimal dual solution, because the feasible set in the dual space is a convex and closed polytope. More information on EDPP may be found at the following GitHub: http://dpc-screening.github.io/glasso.html

Following the screening rule, we further propose a federated screening rule for group LASSO, named federated dual polytope projection for group LASSO (FDPP-GL), to rapidly locate the inactive features in a distributed manner while preserving data privacy at each institution (**Figure 3.3.1** (g)). We summarize the method in **Alg. 3.3.2**. In the algorithm, we estimate the maximum regularization parameter, λ_{max}. The input sequence of parameters, $\lambda_1, \lambda_{2,\dots}, \lambda_\kappa$, should be no greater than λ_{max}. Based on the solutions with the sequence of regularization parameters, we can then perform stability selection (Meinshausen & Bühlmann, 2010) to select significant features that are most related to the corresponding y (**Figure 3.3.1** (h)).

Algorithm 3.3.2. Federated Dual Polytope Projection for Group Lasso (FDPP-GL)

Input: Data pairs of the i institutions $(X^1, y^1), \dots, (X^i, y^i), \dots, (X^I, y^I)$ with information of G groups and regularization parameters $\lambda_1 > \lambda_2 > \cdots > \lambda_\kappa$

Output: The optimal solution $\beta^*(\lambda_1), \beta^*(\lambda_2), \dots, \beta^*(\lambda_\kappa)$

1: $L^i = (X^i)^T y^i$, then computes $L = \sum_{i=1}^{I} L^i$ by LQM

2: $\lambda_0 = \lambda_{max} = \max_g \frac{\|L_g\|_2}{w_g}$, L_g represents all the elements in gth group

3: $Q^i = argmax_{x_g^i} \frac{\|L_g\|_2}{w_g}$, compute $S = \sum_{i=1}^{I} (Q^i)^T y^i$ by LQM

4: **for** $k \leftarrow 1$ to κ **do**

5: $$\theta^i(\lambda_{k-1}) = \begin{cases} \frac{y^i - \sum_{g=1}^{G} x_g^i \beta_g^*(\lambda_{k-1})}{\lambda_{k-1}}, & if\ \lambda_{k-1} \in (0,\ \lambda_0) \\ \frac{y^i}{\lambda_0}, & if\ \lambda_{k-1} = \lambda_0 \end{cases}$$

6: $$v_1^i(\lambda_{k-1}) = \begin{cases} \frac{y^i}{\lambda_{k-1}} - \theta^i(\lambda_{k-1}), & if\ \lambda_{k-1} \in (0,\ \lambda_0) \\ Q^i S, & if\ \lambda_{k-1} = \lambda_0 \end{cases}$$

7: $v_2^i(\lambda_k,\ \lambda_{k-1}) = \frac{y^i}{\lambda_k} - \theta^i(\lambda_{k-1})$

8: $M^i = \left\| v_1^i(\lambda_{k-1}) \right\|_2^2$, , then compute $M = \sum_{i=1}^{I} M^i$ by LQM

9: $\hat{v}_2^i(\lambda_k,\ \lambda_{k-1}) = v_2^i(\lambda_k,\ \lambda_{k-1}) - \frac{\langle v_1^i(\lambda_{k-1}), v_2^i(\lambda_k, \lambda_{k-1}) \rangle}{M} v_1^i(\lambda_{k-1})$

10: **for** $g \leftarrow 1$ to G **do**

11: $J^i = \left[X_g^i\right]^T \left[\theta^i(\lambda_{k-1}) + \frac{1}{2} \hat{v}_2^i(\lambda_k,\ \lambda_{k-1})\right]$

12: Compute $J = \sum_{i=1}^{I} J^i$ by LQM

13: **if** $J < 1 - \frac{1}{2} \left\| \hat{v}_2^i(\lambda_k,\ \lambda_{k-1}) \right\|_2 \left\| X_g \right\|_2$ **then**

14: All elements in $\beta_g(\lambda_k)$ are zero

15: Discard features from data if the corresponding coefficients in $\beta_g(\lambda_k)$ are 0

16: Solve the optimal solution, $\beta^*(\lambda_k)$, by Alg. 3.3.1

3.3.3. Morphometry Feature Selection and Visualization

We carry out the proposed federated group LASSO method to measure how significantly the patches of features are related to the response y. Given a decreasing sequence of regularization parameters, $\lambda_1, \dots, \lambda_\kappa$, we learn a set of corresponding models, $\beta(\lambda_1), \dots, \beta(\lambda_\kappa)$. We perform stability selection by counting the frequency of nonzero entries in the learned models and visualize the frequency on the surface. The counted frequency on each vertex is normalized to 1 to 100 and then mapped to a color bar. For better visualization, we smooth the values on each surface with a 2×3 averaging filter. The regions with higher frequency values will be assigned with warmer color as illustrated in the subfigure (h) of **Figure 3.3.1**. In other words, these areas have more significant associations with y.

3.4.Federated Chow Test Framework for Imaging Genetics

In this work, we propose a novel Federated Genotype-Expression-Image Data Integration model (GEIDI) based on the Chow test (Chow, 1960). The intuition behind our multi-omics framework is illustrated in **Figure 3.4.1**. Some important image-expression relationships (correlations) may be diluted when the population is mixed together. Still, when we stratify the population based on their genotypes (a gene like *APOE* or a SNP like *rs942439*), we can observe strong correlations (AA and BB groups) across subgroups. Accordingly, as shown in **Figure 3.4.2**, our model is designed to detect if the relationships between X (imaging biomarker) and Y (gene expression) are different among the subgroups. The *p*-value of the model is then used to prioritize the trios (genotype-expression-image).

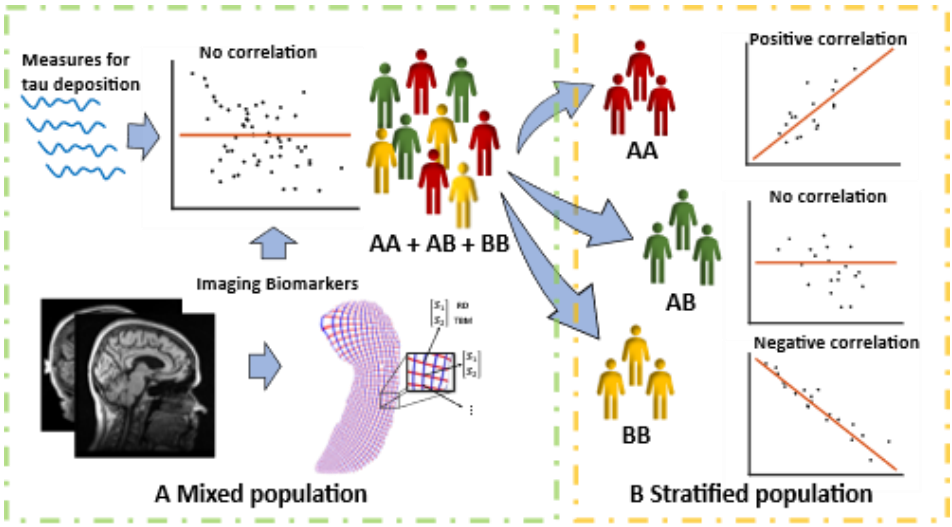

Figure 3.4.1. The Intuition of Our Multi-omics Approach. The image-tau relationship (correlation) is diluted when the population is mixed, but when we stratify the population based on their genotypes, we can observe strong correlations (AA and BB) across subgroups.

Econometrician Gregory Chow first proposed the Chow test in 1960 (Chow, 1960) to determine whether correlation coefficients estimated in two subgroups are significantly different. In econometrics, it is most commonly used in time series analysis to test for the presence of a structural break at a period that can be assumed to be known as *a priori* (for instance, a significant historical event such as a war). For example, we can model the data as $y = wX + \epsilon$. Then, the data can be broken into two groups according to some event and fitted to the regression model as, $y_1 = w_1 x_1 + \epsilon$ and $y_2 = w_2 x_2 + \epsilon$. The null hypothesis of the Chow test asserts that $w_1 = w_2$ and the model errors ϵ are independent and identically distributed from a normal distribution with unknown variance. Let S_C, S_1, and S_2 be the sum of squared residuals for the three regression models respectively, N_1 and N_2 are the number of observations in each group,

and k is the number of parameters. The Chow test statistic is $F = \frac{(S_C-(S_1+S_2))/k}{(S_1+S_2)/(N_1+N_2-2k)}$, which follows the *F*-distribution with k and $N_1 + N_2 - 2k$ degrees of freedom.

Although the Chow test is commonly used in the financial industry, it is seldom used in the biomedical field (B. Lee et al., 2021). In this work, we first generalize the Chow test model to estimate the multi-subgroup condition and further introduce a federated learning technique to the model. We apply the proposed model to the ADNI dataset to detect the significant trios among genotype, gene expression, and imaging biomarkers and discover the dominant genetic and transcriptomics factors for brain structures.

3.4.1. Standardization

We simulate the multi-site condition by separating all the samples into I hypothetical institutions ($I = 5$) on Apache Spark (spark.apache.org), a state-of-the-art distributed computing platform. (Although the ADNI data can be centralized, such a federated analysis would allow the method to be scaled up to much larger datasets, including genomic data that is difficult to centralize for logistic or regulatory reasons). As illustrated in **Figure 3.4.2**, the samples in each institution can be further partitioned into at most three subgroups ($g = 1,2,3$) according to the subject's genotype at certain SNP loci (e.g. GG, GA, AA) or a gene (e.g., stratified by the three *APOE* genotypes considered in this study). Accordingly, X_i^g and y_i^g respectively represent the image biomarkers and gene expression values in the gth group of the ith institution. The data

from the gth group in all I institutions will be fitted into a regression model in a federated strategy.

3.4.2. Federated Chow Test Analysis

Using federated linear regression, we can calculate four linear models for all the I institutions, including three models for three subgroups and one for all samples in the three subgroups. $\overline{w^{(1)}}$, $\overline{w^{(2)}}$, $\overline{w^{(3)}}$ and $\overline{w^{(C)}}$ are their optimal coefficient vectors. The Chow test assumes that the errors ϵ are independent and identically distributed from a normal distribution by an unknown variance. The null hypothesis of the Chow test asserts that $\overline{w^{(1)}}$, $\overline{w^{(2)}}$, and $\overline{w^{(3)}}$ are equal. The predictive test suggested by Chow is then:

$$F = \frac{\left(S^{(C)} - \left(S^{(1)} + S^{(2)} + S^{(3)}\right)\right)/(2k)}{\left(S^{(1)} + S^{(2)} + S^{(3)}\right)/\left(N^{(1)} + N^{(2)} + N^{(3)} - 3k\right)}, \tag{3.4.1}$$

where $S^{(C)}$ is the sum of squared residuals from the combined data from the three subgroups, $S^{(1)}$ is the sum of squared residuals from the first group, and so on for $S^{(2)}$ and $S^{(3)}$. $N^{(1)}$, $N^{(2)}$, and $N^{(3)}$ are the number of samples in each subgroup, and k is the number of parameters. Under the null hypothesis, the test statistic follows the *F*-distribution with $2k$ and $N^{(1)} + N^{(2)} + N^{(3)} - 3k$ degrees of freedom. The global center will calculate *F* by gathering all the least square losses and the number of subjects for each subgroup and combined data from each institution. For example, for the first subgroup, the global least-square loss is $S^{(1)} = \sum_{i=1}^{I} S_i^{(1)}$ and the global subject number is $N^{(1)} = \sum_{i=1}^{I} N_i^{(1)}$. Eventually, the *p*-value will be calculated at the global coordinating center and assigned to each institution.

3.4.3. Federated Linear Regression

Many regression models may be selected for the Chow test model, such as linear regression (Barbur et al., 1994), polynomial regression (Rawlings et al., 1998), ridge regression (Hoerl & Kennard, 1970), and so on. In this study, we focus on studying the differences in the relationships between imaging biomarkers and gene expression among different groups. Complex regression models, like polynomial regression, may lead to over-fitting and meaningless results. Also, sparse or penalized regression methods, such as ridge regression, require an appropriate regularization parameter. Therefore, in this work, linear regression would be the most rational choice.

Since the federated regression models for each subgroup are the same, we omit the group superscripts here. For the data in one subgroup of all the I institutions, we can calculate the linear regression equation as: $y = Xw + \epsilon$, where $X \in R^{N\times k}$ represents the independent variables, $y \in R^N$ is a vector of the observations on a dependent variable, $w \in R^k$ is a coefficient vector, and $\epsilon \in R^N$ is the disturbance vector. N is the number of observations in the group, and k is the number of parameters. Then, the coefficient vector w can be estimated by minimizing the least squared function, $S(w) = \frac{1}{2}\|Xw - y\|_2^2$.

To avoid centralizing the data, (X_i, y_i), from each institution, we first rewrite the minimization problem as, $\min_w \sum_{i=1}^{I} S_i(w; X_i, y_i) = \frac{1}{2}\sum_{i=1}^{I}\|X_i w - y_i\|_2^2$. Then, the global gradient can be calculated as, $\nabla S(w) = X^T(Xw - y) = \sum_{i=1}^{I} X_i^T(X_i w - y_i) = \sum_{i=1}^{I} \nabla S_i(w)$. Therefore, instead of centralizing the data, the global center only needs to gather the partial gradient, $\nabla S_i(w)$, which is calculated with (X_i, y_i) at each local

institution. After computing the global gradient, $\nabla S(w)$, the global center will send it back to ith local institution. Finally, w will be updated at each institution by gradient descent with the same learning rate, $w \leftarrow w - \eta \nabla S(w)$. The reason for not updating w at the global center is to avoid possible data reconstruction. When w is zero, the local gradient sent to the center is $-X_i^T y_i$. Then, the global center can easily acquire $X_i^T X_i w$ and X_i might be reconstructed if w is known to the center. Consequently, our optimization strategy is able to preserve data privacy for all institutions. The whole framework of our federated Genotype-Expression-Image Integration model is summarized in **Alg. 3.4.1**. **And the code can be downloaded at our website, http://gsl.lab.asu.edu/data/uploads/software/GEIDI/GEIDI.zip .**

Algorithm 3.4.1. Federated Genotype-Expression-Image Data Integration Model.

Input: Data pairs of the I institutions, $(X_1, y_1), \ldots, (X_i, y_i), \ldots, (X_I, y_I)$ and the sample numbers of each group, $\left(N_1^{(1)}, N_1^{(2)}, N_1^{(3)}\right), \ldots, \left(N_i^{(1)}, N_i^{(2)}, N_i^{(3)}\right), \ldots, \left(N_I^{(1)}, N_I^{(2)}, N_I^{(3)}\right)$

Output: *p*-value of the studying Genotype-Expression-Image trio

Initialize: $w^{(1)}, w^{(2)}, w^{(3)}, w^{(C)} = \mathbf{0}$

1: **for** $g = \{1,2,3, C\}$ **do**

2: **while** *convergence and maximum number of iterations are not reached* **do**

3: Each institution computes the gradient:

$$\nabla S_i^{(g)}\left(w^{(g)}\right) = [X_i^{(g)}]^T\left(X_i^{(g)} w^{(g)} - y_i^{g(g)}\right).$$

4: Global center computes and sends global gradient to each institution:

$$\nabla S^{(g)}\left(w^{(g)}\right) = \sum_{i=1}^{I} \nabla S_i^{(g)}\left(w^{(g)}\right).$$

5: Each institution updates the coefficient with the global gradient:

$$w^{(g)} \leftarrow w^{(g)} - \eta \nabla S^{(g)}\left(w^{(g)}\right).$$

6: **end while**

7: Each institution calculates the sum of squared residual: $S_i^{(g)}\left(w^{(g)}; X_i^{(g)}, y_i^{(g)}\right)$.

8: Global center gathers the global sum of squared residual: $S^{(g)} = \sum_{i=1}^{I} S_i^{(g)}$.

9: Global center gathers the global sample numbers: $N^{(g)} = \sum_{i=1}^{I} N_i^{(g)}$.

10: **end for**

11: Global center calculates *F* value with equation (1) and then computes and sends *p*-value to all institutions.

CHAPTER 4

EXPERIMENTAL RESULTS

4.1.Results for PASCS-MP

4.1.1. Data Description

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu) and the Open Access Series of Imaging Studies (OASIS) database (Marcus et al., 2010a). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org.

Table 4.1.1 shows demographic information we analyze from the ADNI and OASIS cohorts. From the ADNI cohort, we analyze 841 age and sex-matched subjects with florbetapir PET data and T1-weighted MR images, including 151 AD patients, 342 MCI and 348 asymptomatic CU individuals. Among them, all the 151 AD patients, 171 people with MCI and 116 CU individuals were Aβ positive. The remaining 171 MCI and 232 CU individuals were Aβ negative. From OASIS database, we analyze age-and-sex-matched 260 subjects with florbetapir PET data and T1-weighted MR images, including 52 Aβ positive CU and 208 Aβ negative CU. To match the age and sex to the control group, we randomly select subjects from the majority group millions of times. For each

selected group and control group, we analyze the age with t-test and sex with chi-square test. We first select the groups, of which the p-value of Chi-squared test is 1. Among these groups, we select the one with the largest p-value of age. In this way, we try to minimize the statistical difference in age and sex. The *p*-values of Chi-squared test for sex and the *p*-values of t-test for age are shown in the columns after sex and age in Table 1.

In addition to each MRI scan, we also analyze centiloid measures (Navitsky et al., 2018) from florbetapir PET data in ADNI and OASIS. Operationally, the *positivity* of Aβ biomarkers is defined using standard cut-offs, with some efforts to reconcile differences among different Aβ radiotracers using a norming approach called the centiloid scale (Klunk et al., 2015a; Rowe et al., 2017). ADNI florbetapir PET data are processed using AVID pipeline (Navitsky et al., 2018), and OASIS florbetapir PET data are processed using PUP (Y. K. Lee et al., 2013; Su et al., 2015). Both are converted to the centiloid scales according to their respective conversion equations (Navitsky et al., 2018; Su et al., 2019). A centiloid cutoff of 37.1 is used to determine Aβ positivity, this threshold corresponds to pathologically determined moderate to frequent plaques (Fleisher et al., 2011).

For flortaucipir tau-PET - in a similar fashion to Aβ - tau data are reprocessed using a single pipeline consistent with (Sanchez et al., 2020), so that the standardized uptake value ratio (SUVR) from different ADNI study sites can be analyzed together. In this work, we examine two regional SUVR for tau deposition, corresponding to Braak12, and Braak34 (S. L. Baker, Lockhart, et al., 2017; S. L. Baker, Maass, et al., 2017; Maass et

al., 2017; Schöll et al., 2016). **Table 4.1.2** shows the demographic information from the two cohorts that we analyzed. For amyloid PET, we utilize centiloid measures (Navitsky et al., 2018). Operationally, there have been widely accepted efforts to reconcile differences among different amyloid radiotracers using a norming approach called the centiloid scale (Klunk et al., 2015b; Rowe et al., 2017). ADNI florbetapir PET data are processed using AVID pipeline (Navitsky et al., 2018), which are converted to the Centiloid scales according to their respective conversion equations (Navitsky et al., 2018; Su et al., 2019).

Table 4.1.1 Demographic Information for the Subjects We Study from the ADNI and OASIS Cohorts.

Database	Group	Sex (M/F)	*p*-val	Age	*p*-val	MMSE	Centiloid
ADNI **(n=841)**	Aβ+ AD (n=151)	79/72		74.6±7.8		22.6±3.1	86.3±27.4
	Aβ+ MCI (n=171)	92/79	1.00	74.1±7.4	0.90	27.7±1.7	76.8±26.4
	Aβ- MCI (n=171)	92/79		74.0±7.4		28.3±1.6	8.9±14.9
	Aβ+ CU (n=116)	45/71	1.00	75.9±6.1	0.78	28.9±1.1	71.1±26.4
	Aβ- CU (n=232)	90/142		75.7±6.3		29.0±1.3	7.5±14.5
OASIS **(n=260)**	Aβ+ CU (n=52)	22/30	1.00	70.5±7.5	0.08	29.0±1.3	71.4±20.9
	Aβ- CU (n=208)	88/120		68.5±6.8		29.0±1.3	8.5±9.5

Values are mean ± standard deviation where applicable.

Table 4.1.2 Demographic Information for the Participants We Studied From the ADNI Cohort.

Cohort	Group	Sex (M/F)	Age	MMSE	Braak12	Braak34
tau	AD (n=115)	62/53	76.0±8.5	22.0±4.5	2.39±0.60	2.51±0.73
(n=925)	MCI (n=278)	158/120	74.6±7.9	27.9±2.1	1.82±0.46	1.92±0.46
	CU (n=532)	210/322	73.4±7.1	29.1±1.1	1.58±0.23	1.73±0.21
Cohort	Group	Sex (M/F)	Age	MMSE	Centiloid	
	AD (n=173)	98/75	75.0±7.8	22.7±2.9	72.0±40.2	
Aβ	MCI (n=516)	291/225	72.6±7.8	28.0±1.7	42.0±40.7	
(n=1,127)	CU (n=438)	200/238	74.5±6.5	29.0±1.2	24.4±33.3	

Values are mean ± standard deviation, where applicable.

4.1.2. Key Parameter Estimations for the PASCS-MP Method

To apply PASCS-MP method on hippocampal MMS, four parameters need to be empirically assigned, namely: the patch size, the dimensionality of the learned sparse coding, the regularization parameter for the l_1-norm (λ) and the kernel size (σ) in the exponential function. Selecting suitable parameters will lead to superior performance in refining lower dimensional MMS representations related to AD pathology. With 10-fold cross-validation, these key parameters are evaluated from PASCS-MP based classification performance on 109 AD patients and 180 CU subjects of ADNI-2 cohort. To avoid data leakage, these subjects are not used in the following study of Aβ burden classification.

We perform grid search on the data set to explore the optimal parameter settings. In **Figure 4.1.1**, we only illustrate part of the classification accuracy for different values of each parameter in grid search since the combinations of four different parameters will lead to 5^4 results. For each parameter setting, we also repeat 10-fold cross-validation five times, and the average and 95% confidence interval of the accuracy are shown in Fig.3. When we evaluate one parameter, we fix the rest parameters. For example, in the first bar chart in **Figure 4.1.1**, we try different patch sizes including 5×5,10×10,15×15,20×20 and 30×30 while we fix the sparse code dimensionality as to 1800, and set λ to 0.22, and σ to 3.6. By testing varied sets of parameters, we find that the optimal patch size is 10×10, the optimal sparse code dimensionality is 1800, the optimal λ is 0.22 and the optimal σ is 3.6 and these optimal parameters will be adopted in the study of Aβ burden classification.

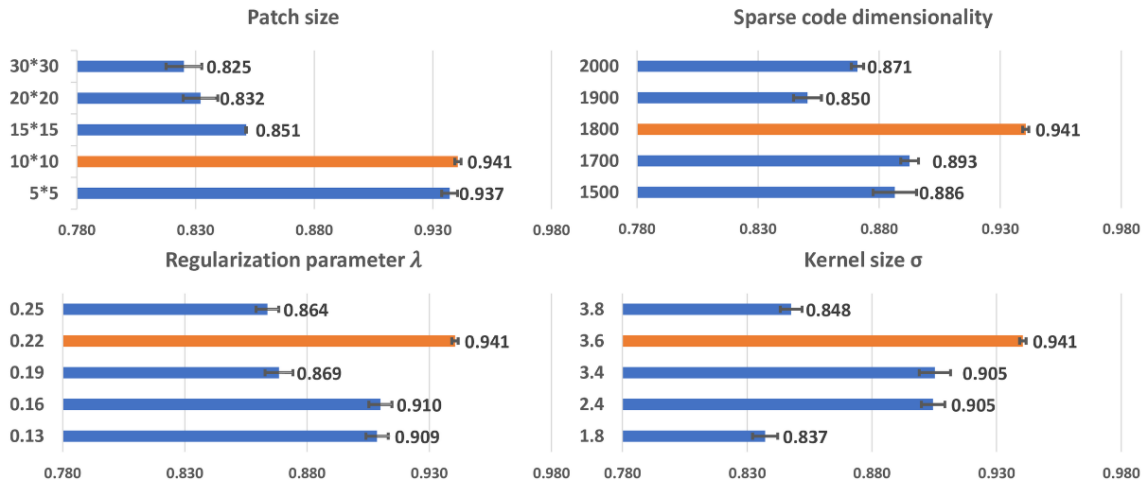

Figure 4.1.1. The Relationship of Each Parameter to Classification Accuracy. The y-axis represents the value for each parameter. The orange bars represent the classification performances using the optimal parameters. Each bar represents the average and 95% confidence interval of classification accuracy.

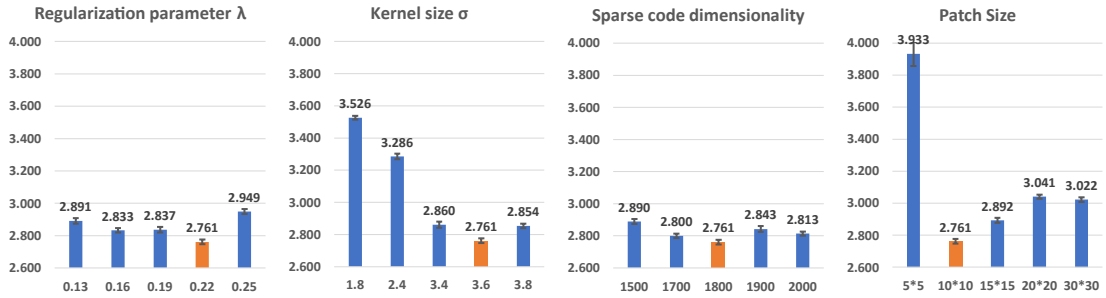

Figure 4.1.2. The Relationship of Each Parameter to RMSE. The x-axis represents the value for each parameter. The orange bars represent the classification performances using the optimal parameters. Each bar represents the average and 95% confidence interval for RMSE.

Instead of predicting Tau measurements, we train ridge regression models to predict MMSE on a separate dataset from ADNI (100 AD patients, 100 MCI, and 100 CN). We perform grid search to explore the optimal parameter settings. After performing 10-fold cross-validation ten times, we compare the average root mean squared errors (RMSE) of MMSE for each parameter setting. In **Figure 4.1.2**, we only illustrate the

average and 95% confidence interval of RMSE for part of the grid search result. In each subfigure, we only compare one parameter and fix the remaining three. Eventually, we find that the optimal patch size is 10×10, the optimal sparse code dimensionality is 1800, the optimal λ is 0.22, and the optimal σ is 3.6; these optimal parameters are subsequently adopted for predicting tau and Aβ measurements.

4.1.3. Classification of Aβ Burden

To explore whether there is a significant gain in classification power with our new system, based on our surface MMS, we generate two different kinds of sparse codes with our previous framework (PASS-MP) (Dong et al., 2020a; Fu et al., 2021; Zhang et al., 2017a, 2016b) and the new framework (PASCS-MP). The parameter settings for the two sparse coding methods are the same. Additionally, we apply the popular SPHARM method (Chung et al., 2008b; Shi, Thompson, et al., 2013) to calculate hippocampal shape features. Based on these three kinds of feature sets, we apply the random forest classifier to detect individuals with different Aβ status. Moreover, we also examine the classification performances using hippocampal MMS, surface area and volume measures. These classification performances are evaluated using ACC, B-ACC, SPE, SEN. For each binary classification of ADNI cohort, we repeat the 10-fold cross-validation 5 times; the mean and 95% confident interval of the evaluation measures are calculated as (Vanwinckelen & Blockeel, 2012) and shown in the middle three columns of **Table 4.1.3**.

We leverage the OASIS dataset as an external validation set to further evaluate the performance of our new framework. We firstly generate new representations with our

proposed PASCS-MP for all the CU subjects from ADNI and OASIS cohorts. Then, we train a binary random forest model on the ADNI dataset and test it with the OASIS dataset. Since there is no cross-validation here, there is no confident interval in the last column of **Table 4.1.3**. We also compute the area-under-the-curve (AUC) of the receiver operating characteristic (ROC). The ROC curve and AUC for these classification tasks are illustrated in **Figure 4.1.3**. This comparison analysis classification performance shows that the combination of PASCS-MP and hippocampal MMS measures have superior performance when detecting individuals with different Aβ status, compared to other similar methods.

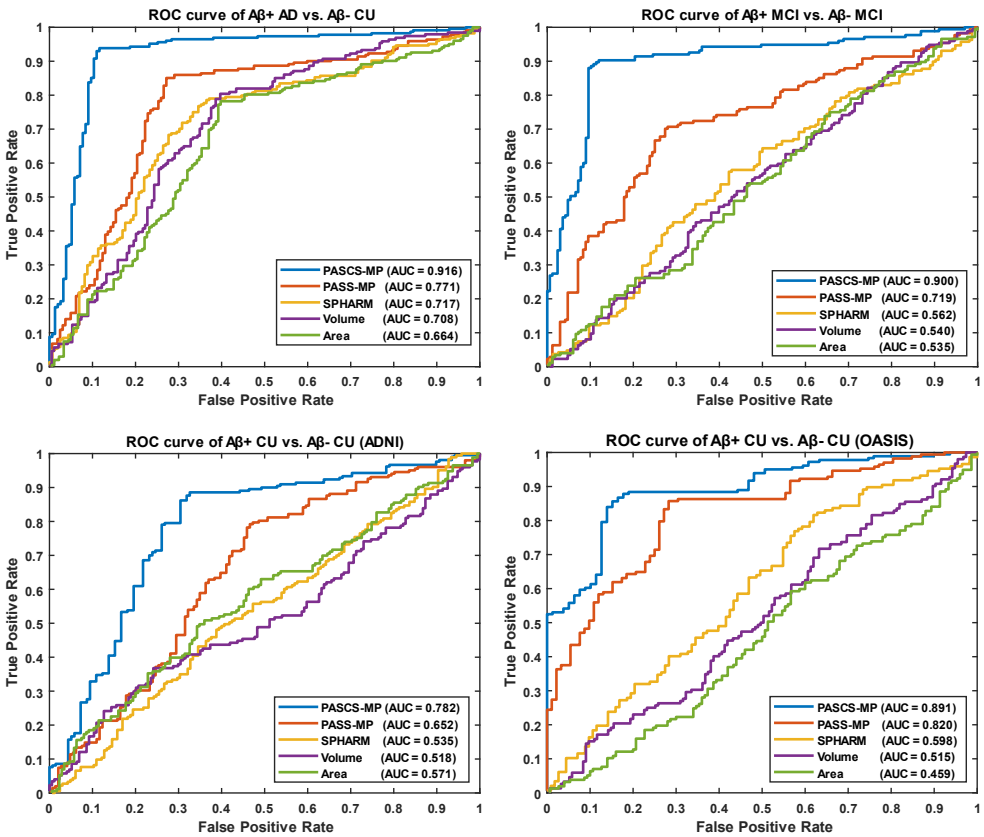

Figure 4.1.3. ROC Curves for the Classification Tasks, Aβ+ AD vs. Aβ- CU, Aβ+ MCI vs. Aβ- MCI, Aβ+ CU vs. Aβ- CU (ADNI), and Aβ+ CU vs. Aβ- CU (OASIS). OASIS is used as an external validation set for the model trained by ADNI CU.

Table 4.1.3. Classification Results for Four Contrasts.

Area	Aβ+ AD vs. Aβ- CU	Aβ+ MCI vs. Aβ- MCI	Aβ+ CU vs. Aβ- CU (ADNI)	Aβ+ CU vs. Aβ- CU (OASIS)
ACC	0.68±0.01	0.55±0.02	0.54±0.01	0.47
B-ACC	0.69±0.02	0.55±0.02	0.54±0.02	0.43
SPE	0.66±0.02	0.54±0.02	0.55±0.02	0.49
SEN	0.71±0.03	0.56±0.03	0.53±0.04	0.37
Volume	Aβ+ AD vs. Aβ- CU	Aβ+ MCI vs. Aβ- MCI	Aβ+ CU vs. Aβ- CU (ADNI)	Aβ+ CU vs. Aβ- CU (OASIS)
ACC	0.71±0.01	0.53±0.02	0.50±0.03	0.51
B-ACC	0.72±0.01	0.53±0.01	0.50±0.03	0.52
SPE	0.68±0.01	0.52±0.01	0.51±0.02	0.54
SEN	0.75±0.01	0.54±0.02	0.49±0.04	0.50
SPHARM	Aβ+ AD vs. Aβ- CU	Aβ+ MCI vs. Aβ- MCI	Aβ+ CU vs. Aβ- CU (ADNI)	Aβ+ CU vs. Aβ- CU (OASIS)
ACC	0.71±0.02	0.56±0.02	0.52±0.02	0.60
B-ACC	0.71±0.02	0.56±0.03	0.51±0.04	0.60
SPE	0.74±0.02	0.61±0.03	0.56±0.03	0.61
SEN	0.68±0.04	0.51±0.03	0.46±0.05	0.60
PASS-MP	Aβ+ AD vs. Aβ- CU	Aβ+ MCI vs. Aβ- MCI	Aβ+ CU vs. Aβ- CU (ADNI)	Aβ+ CU vs. Aβ- CU (OASIS)
ACC	0.79±0.01	0.73±0.02	0.71±0.02	0.74
B-ACC	0.79±0.01	0.73±0.02	0.70±0.03	0.73
SPE	0.78±0.02	0.75±0.02	0.73±0.03	0.74
SEN	0.79±0.01	0.72±0.03	0.67±0.03	0.73
PASCS-MP	Aβ+ AD vs. Aβ- CU	Aβ+ MCI vs. Aβ- MCI	Aβ+ CU vs. Aβ- CU (ADNI)	Aβ+ CU vs. Aβ- CU (OASIS)
ACC	0.91±0.01	0.89±0.01	0.79±0.02	0.81
B-ACC	0.91±0.01	0.89±0.01	0.79±0.03	0.80
SPE	0.91±0.01	0.91±0.01	0.80±0.02	0.82
SEN	0.90±0.01	0.88±0.01	0.79±0.05	0.79

Values are mean ± 95% confident interval where applicable.

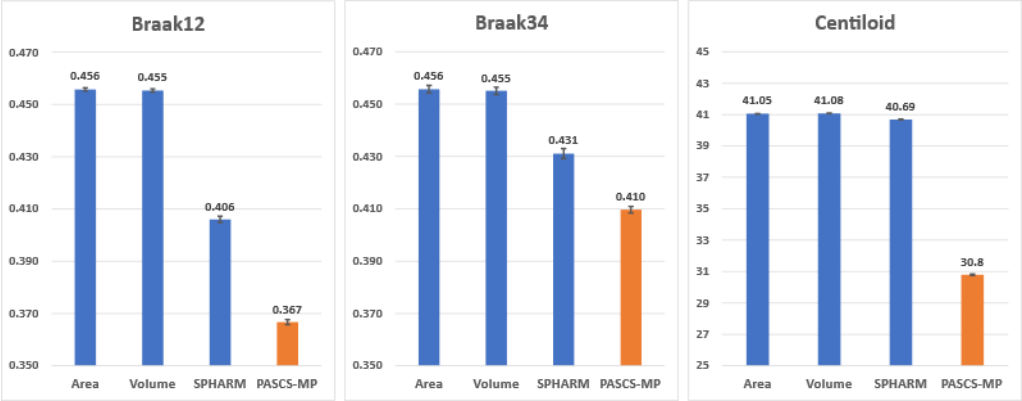

Figure 4.1.4. RMSE for Predicted Braak12 and Braak34 from Four Measurements, Hippocampal Surface Area, Volume, SPHARM and Our MMS-based PASCS-MP Representations. Each bar represents the mean and 95% confidence interval of RMSE for ten 10-fold cross-validations.

4.1.4. Prediction of Tau/Aβ Measurements

After performing PASCS-MP on MMS of 925 subjects from ADNI, we obtain 925 new representations, of which the dimensionality is 1,800. These representations are utilized for training ridge regression models to predict two Tau measurements, Braak12 and Braak34. For each measurement, we also repeat the 10-fold cross-validation ten times. The mean and 95% confidence interval of the RMSE for the two measurements are illustrated in **Figure 4.1.4**. To demonstrate that our representations have stronger predictive power, we train ridge regression models with hippocampal surface area, hippocampal volume, and the hippocampal shape features calculated by the popular SPHARM method (Chung et al., 2008b; Shi, Thompson, et al., 2013). As shown in **Figure 4.1.4**, our PASCS-MP always has the minimum RMSE.

4.1.5. Analysis of the Predicted Tau/Aβ Measurements

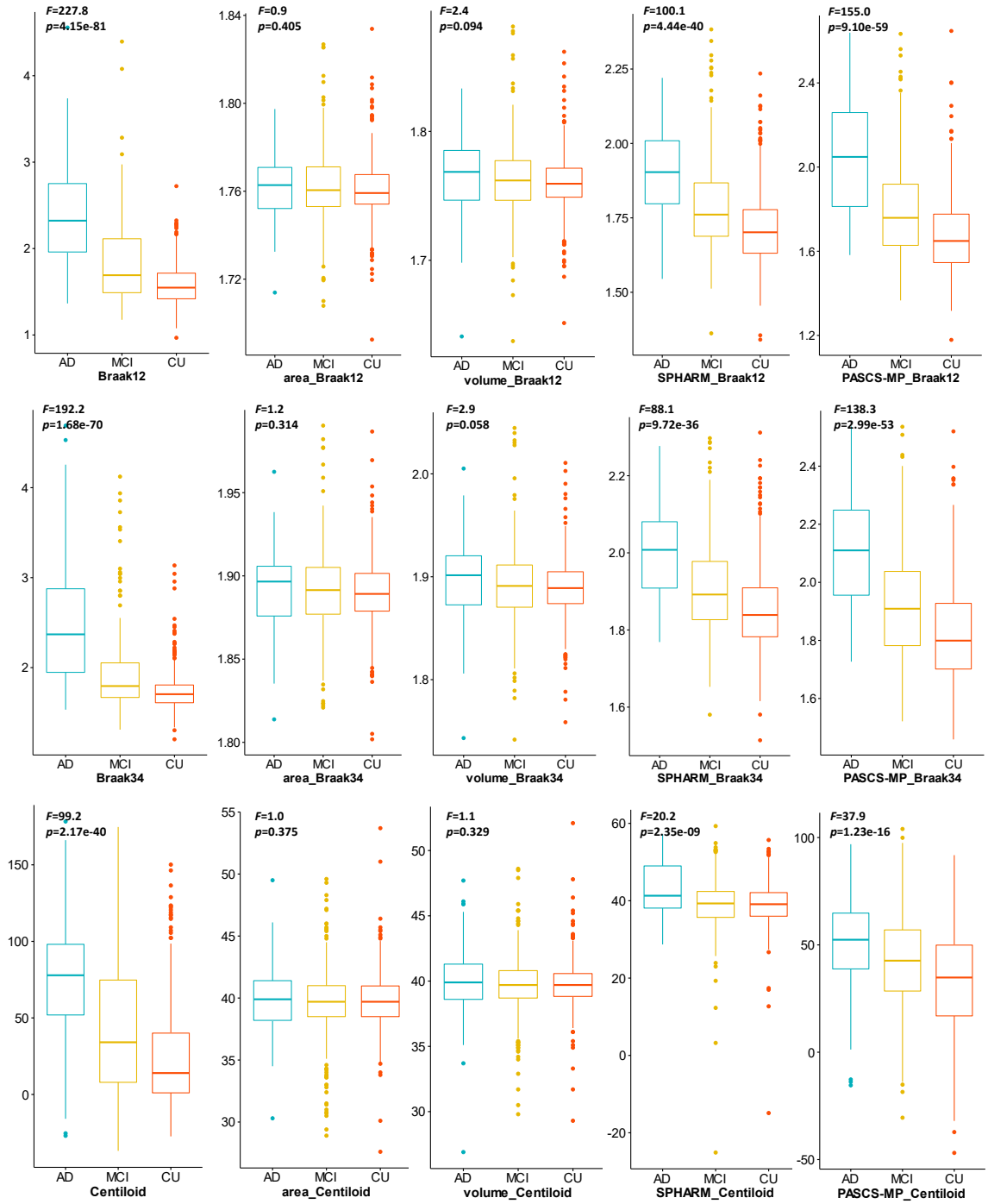

Figure 4.1.5. ANOVA Analysis for Real and Predicted Tau/Aβ Measurements. The first column is the distribution of real Braak12, Braak34 and Centiloid measures. The remaining columns are the predicted measurements from hippocampal surface area, hippocampal volume, SPHARM, and our MMS-based PASCS-MP representations. The F-value and p-value of ANOVA among AD, MCI, and CU are illustrated on the top of each subfigure.

To evaluate the predicted tau/Aβ measurements from different features, we first

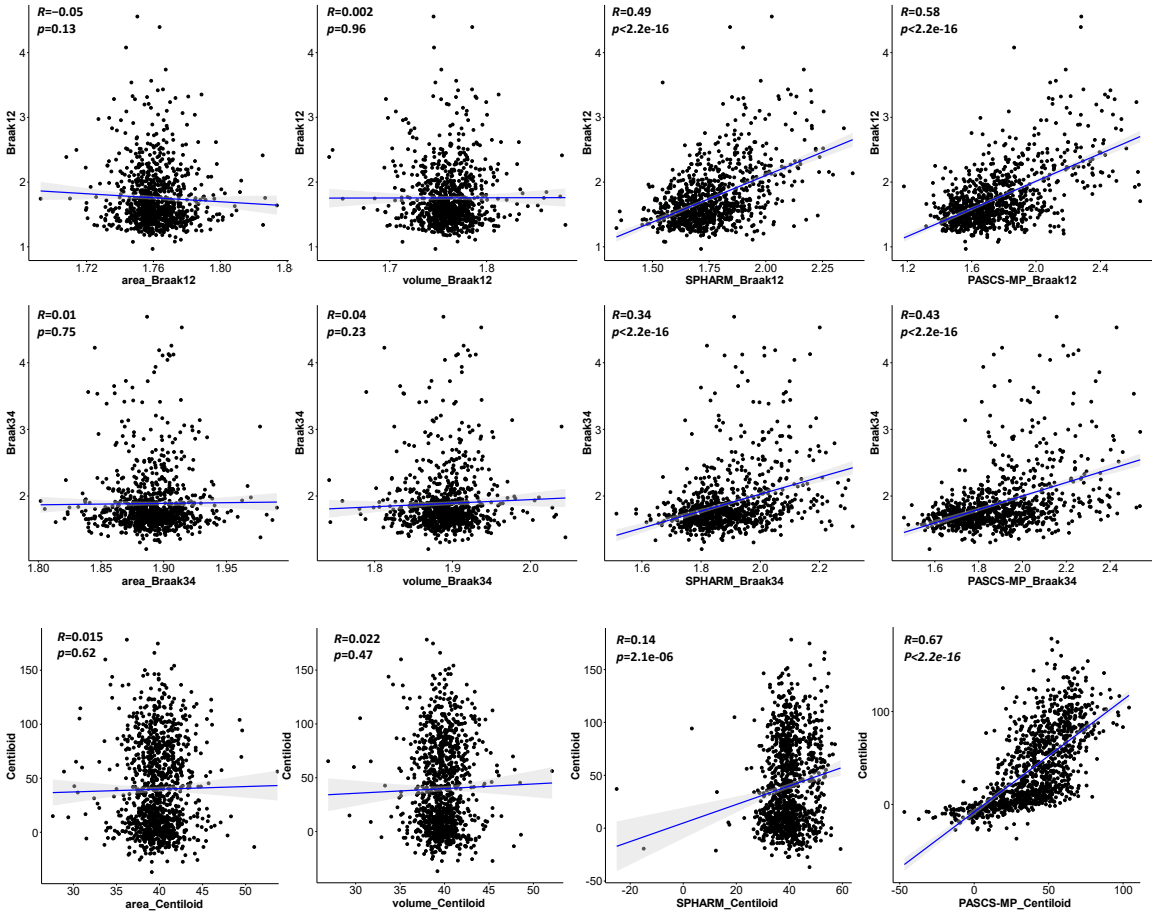

Figure 4.1.6. Pearson Correlation Between Real Tau/Aβ Measurement and Predicted Measurement. The first row shows the Pearson correlation between real Braak12 and predicted Braak12 from hippocampal surface area, hippocampal volume, SPHARM, and our MMS-based PASCS-MP representations. The second row shows the Pearson correlation between real Braak34 and predicted Braak34. The third is for Centiloid. The y-axis is the real measurement and x-axis shows the predicted measurement. The Pearson correlation coefficient, R, and p-values are in the top left corner of each subfigure.

perform analysis of variance (ANOVA) among the three clinical groups, AD, MCI, and CU. The distributions of the predicted measurements are shown in **Figure 4.1.5**. The first column is the distribution of real Braak12, Braak34 and Centiloid. Other columns are the predicted measurements from hippocampal surface area, hippocampal volume, SPHARM, and our PASCS-MP. The *F*-value and *p*-value of ANOVA among the three

clinical groups are illustrated in each subfigure. Our PASCS-MP achieves the most significant group difference among all the predicted measurements.

In addition, we leverage the Pearson correlation to evaluate the relations between real tau/Aβ measurements and each of the predicted tau/β measurements. In **Figure 4.1.6**, we visualize the linear relationships. The vertical axis is the real measurement, and the horizontal axis is the predicted one. The correlation coefficient, *R*, and *p*-value for each analysis are also illustrated in each subfigure. Our PASCS-MP always has the largest correlation coefficients, in these experiments, compared to the traditional measurements, which means the measurements predicted by our MMS-based PASCS-MP representations are close to the real measurements. Both experiments demonstrate that our MMS-based PASCS-MP representations have the best accuracy of the approaches we examined for predicting tau/Aβ measurements.

4.2.Results for FMFS

4.2.1. Data Description

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessments can be combined to measure the progression of MCI and early AD. For up-to-date information, see www.adni-info.org. From the multiple phases of ADNI - ADNI 1, ADNI 2, ADNI GO, and ADNI 3 - we analyzed two sets of scans for the study of Aβ deposition and tau

deposition. For the Aβ deposition study, we analyzed a total of 1,127 pairs of images from 1109 subjects (18 of them have two pairs from different visiting dates), including 1,127 T1-weighted MR images and 1,127 florbetapir PET images. Similarly, we obtained 925 pairs from 688 subjects (191 of them have more than one pair from different visiting dates) of MRI scans and AV1451 PET images for the study of tau deposition.

In addition to each brain MRI scan, we also analyze the corresponding Mini-Mental State Exam (MMSE) scores (Folstein et al., 1975). For amyloid PET, we utilize centiloid measures (Navitsky et al., 2018). Operationally, there have been widely accepted efforts to reconcile differences among different amyloid radiotracers using a norming approach called the centiloid scale (Klunk et al., 2015b; Rowe et al., 2017). ADNI florbetapir PET data are processed using AVID pipeline (Navitsky et al., 2018), which are converted to the Centiloid scales according to their respective conversion equations (Navitsky et al., 2018; Su et al., 2019). For flortaucipir tau-PET - in a similar fashion to Aβ - tau data are reprocessed using a single pipeline consistent with (Sanchez et al., 2020), so that the standardized uptake value ratio (SUVR) from different ADNI study sites can be analyzed together. In this work, we examine two regional SUVR for tau deposition, corresponding to Braak12, and Braak34 (S. L. Baker, Lockhart, et al., 2017; S. L. Baker, Maass, et al., 2017; Maass et al., 2017; Schöll et al., 2016). **Table 4.2.1** shows the demographic information from the two cohorts that we analyzed.

Table 4.2.1 Demographic Information for the Participants We Studied from the ADNI.

Cohort	Group	Sex (M/F)	Age	MMSE	Centiloid	
Aβ (n=1,127)	AD (n=173)	98/75	75.0±7.8	22.7±2.9	72.0±40.2	
	MCI (n=516)	291/225	72.6±7.8	28.0±1.7	42.0±40.7	
	CU (n=438)	200/238	74.5±6.5	29.0±1.2	24.4±33.3	
Cohort	Group	Sex (M/F)	Age	MMSE	Braak12	Braak34
tau (n=925)	AD (n=115)	62/53	76.0±8.5	22.0±4.5	2.39±0.60	2.51±0.73
	MCI (n=278)	158/120	74.6±7.9	27.9±2.1	1.82±0.46	1.92±0.46
	CU (n=532)	210/322	73.4±7.1	29.1±1.1	1.58±0.23	1.73±0.21

Values are mean ± standard deviation, where applicable.

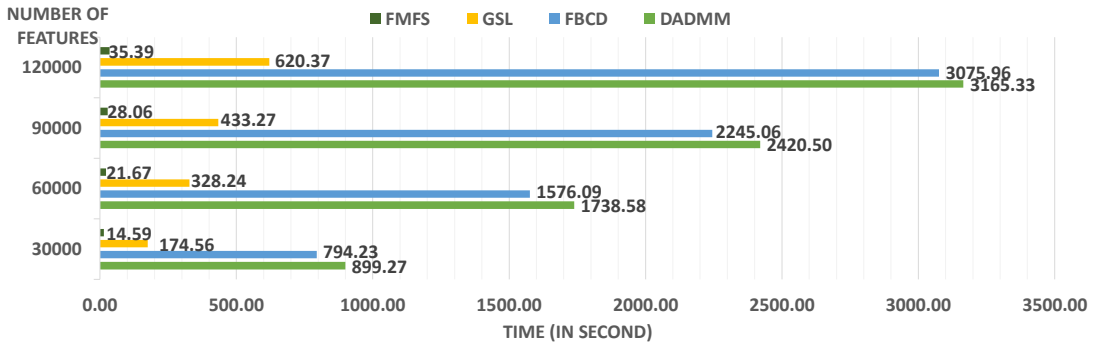

Figure 4.2.1. Comparison Analysis of Computation Efficiency. For the morphometry features with different resolutions, our framework achieves a speedup of 62-fold, 80-fold, 86-fold, and 89-fold, compared to DADMM. For FBCD, our FMFS has a speedup of 54-fold, 72-fold, 80-fold, and 86-fold. For GSL, our model has a speedup of 12-fold, 15-fold, 15-fold, and 17-fold.

4.2.2. Efficiency Evaluation

A significant innovation of FMFS is that we introduce a screening rule during the group LASSO feature selection stage, which highly improves the computation speed compared to the distributed alternating direction method of multipliers algorithm (DADMM) (Boyd et al., 2011). Moreover, we also compare FMFS with the Gauss-Southwell-Lipschitz rules (GSL) for block coordinate descend in (Nutini, 2017). Besides, we also tested the running time of FBCD in our federated framework without the screening rule.

We simulated the distributed condition on a cluster with several conventional x86 nodes, of which each contains two Intel Xeon E5-2680 v4 CPUs running at 2.40 GHz. Each parallel computing node has a full-speed Omni-Path connection to every other node in its partition. 1,127 subjects for the Aβ deposition study were randomly assigned to five simulated institutions, each of which has almost the same number of subjects and one computation node. We uniformly selected 100 regularization parameters from 1.0 to 0.1

and ran all three methods with the same experimental set-up. Under different morphometry feature resolutions (where we randomly down-sampled or up-sampled the dimension of the features), our FMFS method achieved a speedup of 62-fold, 80-fold, 86-fold, and 89-fold, compared to DADMM as shown in **Figure 4.2.1**. For FBCD, our FMFS has a speedup of 54-fold, 72-fold, 80-fold, and 86-fold. For GSL, our model has a speedup of 12-fold, 15-fold, 15-fold, and 17-fold.

4.2.3. Aβ and Tau Associated Hippocampal Morphometry

We employed stability selection with our FMFS model to select the ROIs (subregions of the hippocampal surfaces) related to Aβ and tau. We respectively standardize the two types of input features, RD and TBM, for each subject, using Z-scores, and adopt the centiloid value as the measure of Aβ burden and Braak12 and Braak34 measures for tau deposition. Since the regularization parameters can control the sparsity of the solution vector and further influence the area of the ROIs, we uniformly generated 100 regularization parameters between 0.01 to 0.001, which may lead to a reasonable size for the selected ROIs. After training the model, we got 100 solution vectors, of which the dimensionality is 60,000, since each of the left and right hippocampal surfaces contains 15,000 vertices, and each vertex has two features. Then, we performed stability selection by counting the nonzero entries for RD and TBM on the same vertex. The counted frequency on each vertex was normalized to 0 to 100 and then mapped to a color bar, as shown in **Figure 4.2.2-4.2.4**. For better visualization, we smoothed the values on each surface by a 2×3 averaging filter. The warmer color areas

have a higher frequency value and have stronger associations with the responses, i.e., brain global Aβ and tau burden.

In this experiment, we first ran the proposed model on the cohorts for Aβ and tau deposition. As illustrated in the top left picture of **Figure 4.2.2-4.2.4**, the morphometric abnormalities mainly happen in specific hippocampal subregions, namely the subiculum and CA1. Additionally, we separately studied the ROIs for groups of CU, MCI, and AD subjects. As shown in the rest of the three panels in **Figure 4.2.2-4.2.4**, the morphometric associations are strongest in the subiculum and CA1 at the early stages; but with the progression of AD, the distortions are more focal in subiculum. Specifically, the results for CU subjects are shown in the top right panel of each figure, where the warmer colored regions are widespread in both the subiculum and CA1 areas. However, in the results for the AD group, the warmer colored regions mainly focus on the area of the subiculum.

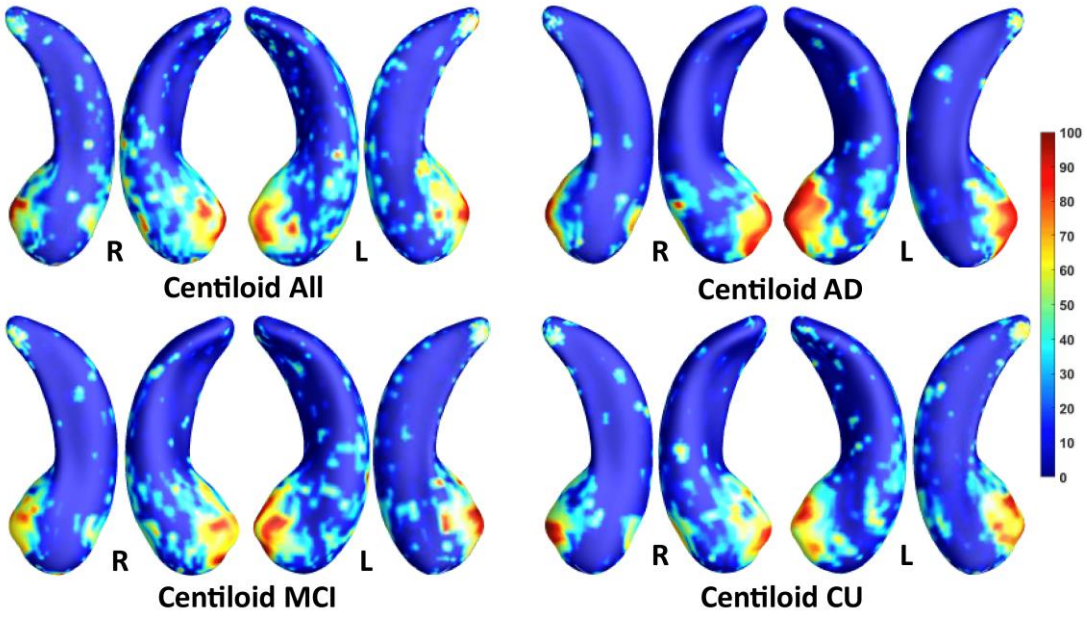

Figure 4.2.2. Visualization of ROIs Associated with Centiloid (Aβ Measurement). The top left figure is the results for all subjects. The top right is for AD patients. The bottom two are for participants with MCI and for the CU group.

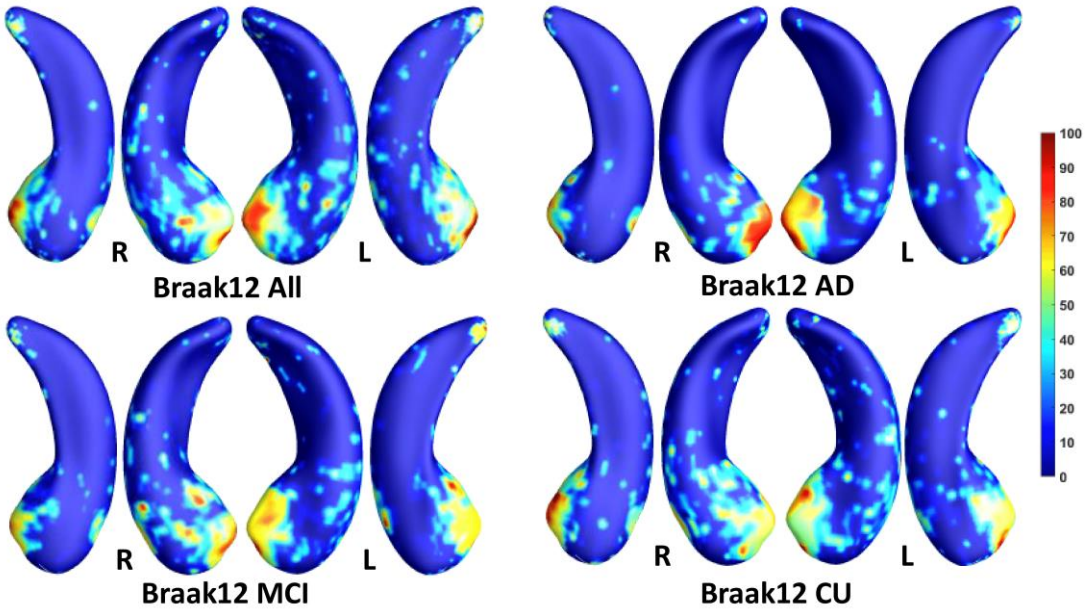

Figure 4.2.3. Visualization of ROIs Associated with Braak12 (Tau Measurement). The top left figure shows the results for all subjects. The top right is for AD patients. The bottom two figures are for participants with MCI and for the CU group.

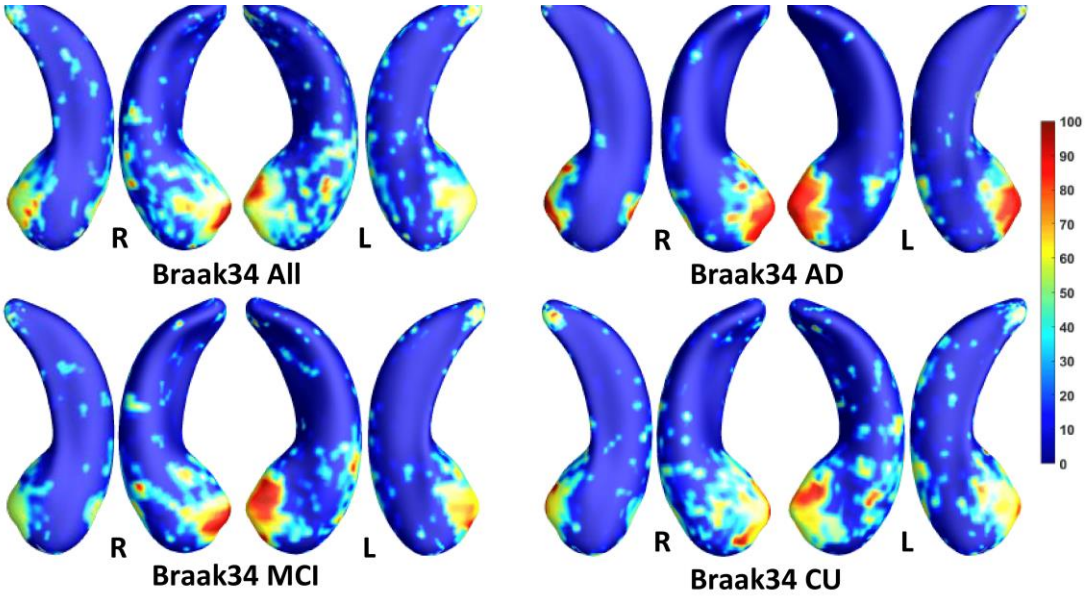

Figure 4.2.4. Visualization of ROIs Associated with Braak34 (Tau Measurement). The top left figure shows the results for all subjects. The top right is for AD patients. The bottom two figures are for participants with MCI and for the CU group.

4.2.4. Association Analysis between Features on ROIs and Measure for Aβ and Tau Deposition

In this experiment, we try to demonstrate the morphometric features of our selected ROIs have stronger correlations with the measures for Aβ and tau deposition than the other hippocampal surface features. After performing stability selection, we were able to rank the vertices related to each measurement of Aβ/tau deposition. We selected the 3,000 features from the 1,500 top-ranked vertices for each subject (1,500 RD and 1,500 TBM). For a fair comparison, we also selected 3,000 features from 1,500 randomly selected vertices for each subject and used them as features representing differences across the entire hippocampus. To fit the Pearson Correlation analysis, we converted the features on ROIs to a single value for each subject. First, as the features on the ROIs should have stronger predictive power, we used the frequency on each vertex as a weight to multiply the RD and TBM on the vertex. And then, we respectively summed up the weighted RD and weighted TBM on the ROIs for each subject. The value for RD and the value for TBM were further reduced to a scalar with principal components analysis (PCA). PCA is an unsupervised model to reduce the dimensionality of the data while minimizing information loss. It creates new uncorrelated features which maximize the variance successively. For the randomly selected features on the whole hippocampal surface, the RD and TBM were directly summed up without multiplying the frequency and reduced to a single value with PCA. In **Figure 4.2.5**, we illustrate the results of Pearson correlation between morphometric features and measures for Aβ and tau deposition. The top three subfigures illustrate the correlation between the values on our

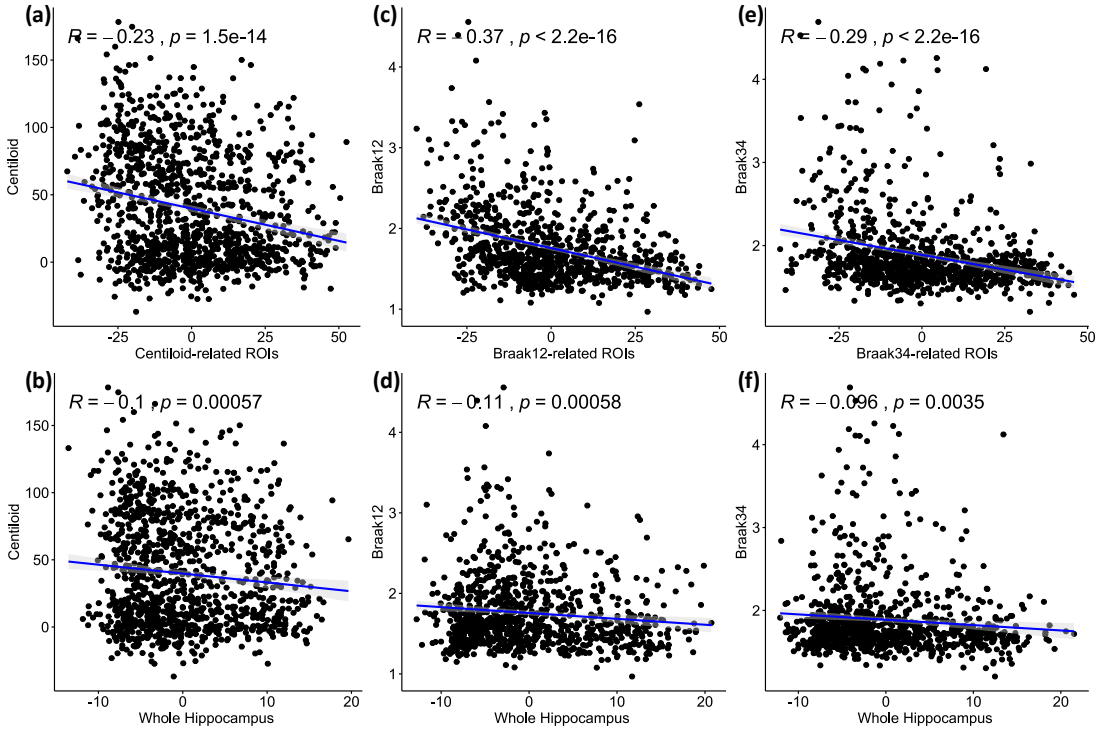

Figure 4.2.5. Pearson Correlation Between Morphometry Features and Measures for Aβ and Tau Deposition. The top three subfigures illustrate the correlation between the values on our selected ROIs and the measure for Aβ or tau deposition. The bottom three are between the value on the whole hippocampal surface and the measure for Aβ or tau deposition. The correlation coefficients and p-values are shown in each subfigure.

selected ROIs and the measure for Aβ or tau deposition. The bottom three are between the value on the whole hippocampal surface and the measure for Aβ or tau deposition. The correlation coefficients and *p*-values are shown in each subfigure. The correlation coefficient of Centiloid-related ROIs is -0.23, and the coefficient for the whole surface is only -0.1. For Braak12 and Braak34, the coefficients of our selected ROIs are -0.37 and -0.29 and the ones for the whole surface are -0.11 and -0.096. Consequently, the features on our selected ROIs have more associations to the measure for Aβ or tau deposition than the other features on the surface.

4.2.5. Predicting MMSE Scores Based on Hippocampal ROIs

In the model of (Jack et al., 2016), an abnormal level of Aβ and tau deposition tends to occur earlier than abnormal cognitive decline can be detected. In this experiment, we further validated the ROIs selected by our proposed model in terms of their prediction accuracy for MMSE score in cohorts where Aβ and tau deposition were measured separately. After performing stability selection, we were able to rank the vertices related to each measurement of Aβ/tau deposition. We selected the 3,000 features from the 1,500 top-ranked vertices for each subject (1,500 RD and 1,500 TBM). Then, we used these features to predict the MMSE score as described in **Section 4.2.3**. For a fair comparison, we also selected 3,000 features from 1,500 randomly selected vertices for each subject and used them as features representing differences across the entire hippocampus. Moreover, we also leveraged the measurements for Aβ or tau deposition to predict MMSE. In addition, we compare our FMFS with recursive feature elimination (RFE) (Guyon et al., 2002). The feature dimensionality of our morphometry feature is 60,000 and RFE may take tens of days to rank features for such a big dataset. For equal comparison, we also selected 1,500 RD and 1,500 TBM for each measurement of Aβ/tau deposition. To accelerate the feature selection speed, we randomly select 300 features from the 30,000 RD and use RFE to select the top 15 RD. We repeated the step 100 times and selected 1,500 RD. With the same strategy, we also select 1,500 TBM. Then, we used these features to train machine learning models, including random forest, multilayer perceptron (MLP), and Lasso regression. 10-fold cross-validation was adopted to evaluate the performance of the models, and root mean squared error (RMSE) was used

for measuring the prediction accuracy. In **Table 4.2.2**, the top five rows indicate the results for Aβ deposition, and the rest of the rows are for different measurements of tau deposition. Hippocampal ROIs represent the features on our selected ROIs and RFE-selected represents the features selected by recursive feature elimination. The RMSEs of our framework are always the smallest. It is worth noting that comparing to the RFE method, our proposed FMFS framework demonstrated significant efficiency improvement. Specifically, the average running time of the RFE method is 49,319 seconds while ours FMFS method 22 seconds, roughly with 2240-fold efficiency improvement. These results demonstrate that the features in the ROIs selected by our model can always have a stronger predictive power and predict the MMSE score better than the measurements of Aβ and tau deposition.

We also perform Pearson correlation between the morphometry features and MMSE and between the measure for Aβ or tau deposition and MMSE. We also utilize the same method as Section 4.2.3 to convert the multivariate features to a scalar. The results are shown in **Figure 4.2.6**. The first column is the correlation between the measures for Aβ and tau deposition and MMSE. And the second column is the correlation between the features on our selected ROIs and MMSE. The last column is between the feature on the whole surface and MMSE. The correlation coefficients and *p*-values are shown in each subfigure. In the study of Aβ deposition, the coefficient for centiloid is -0.36 and the ones for the features on centiloid-related ROIs and the whole surface are 0.3 and 0.11. In the study of tau deposition, the coefficient for Braak12 and Braak34 are -0.58 and -0.59. And the ones for the features on Braak12-related ROIs and

Braak12-related ROIs are 0.29 and 0.28. The coefficient for the features on the whole surface is 0.057. The correlation of the features on our selected ROIs is stronger than the features on the whole surface, which demonstrates the effectiveness of our model. The measures for Aβ or tau deposition have the strongest correlation with MMSE. But our selected features are multivariate, which can have better performance in predicting MMSE. Our work validated the AD progression model and may provide unique insights for accurately estimating clinical disease burden.

Table 4.2.2 RMSEs for Predicting MMSE Based on Various Kinds of Biomarkers and Models.

Aβ associated	Random forest	MLP	LASSO
Hippocampal ROIs	**2.58**	**3.00**	**2.59**
Whole Hippocampal	2.79	3.96	2.79
Centiloid	3.15	4.01	2.85
RFE selected	2.68	3.67	2.68
Braak12 associated	Random forest	MLP	LASSO
Hippocampal ROIs	**2.61**	**3.20**	**2.90**
Whole Hippocampal	3.11	4.24	3.00
Braak12	3.03	4.98	3.03
RFE selected	2.70	3.62	2.98
Braak34 associated	Random forest	MLP	LASSO
Hippocampal ROIs	**2.62**	**3.26**	**2.86**
Whole Hippocampal	3.09	4.16	3.02
Braak34	2.81	4.94	3.02
RFE selected	2.85	3.83	2.98

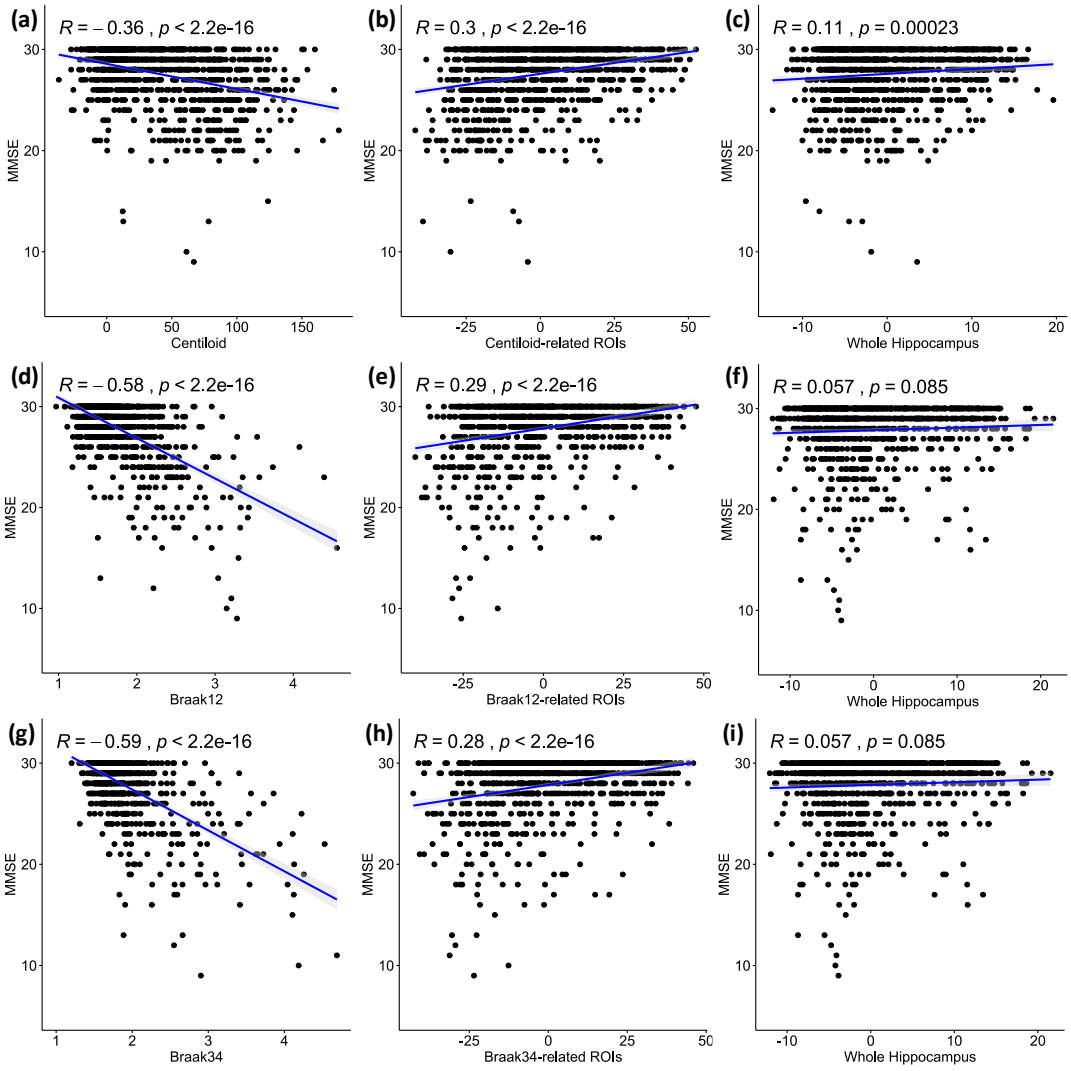

Figure 4.2.6. Pearson Correlation Between the Measures for Aβ and Tau Deposition and MMSE and Between the Morphometry Features and MMSE. The first column is the correlation between the measures for Aβ and tau deposition and MMSE. And the second column is the correlation between the features on our selected ROIs and MMSE. The last column is between the feature on the whole surface and MMSE. The correlation coefficients and p-values are shown in each subfigure.

4.2.6. Predicting Clinical Decline in Participants with MCI

In this experiment, we evaluated the performance of our features on the ROI in survival analysis by using 118 MCI participants' data from a separate dataset (G. Wang et al., 2021) from ADNI (**Table 4.2.3**), including 63 MCI converters, who converted to probable AD in the next six years, and 55 MCI non-converters. Similar to **Section 3.3**, we also chose 1,500 RD and 1500 TBM from the four ROIs (Aβ, Braak12, Braak34) and 3000 features from 1500 random-selected vertices on the whole hippocampal surface to predict the conversion rates from MCI to AD, respectively. For comparison, we also performed the same experiment with the surface area and volume of the hippocampus. The hippocampal volume and surface area were calculated with the smoothed hippocampal structures after linearly registered to the MNI imaging space (Patenaude et al., 2011; Shi, Thompson, et al., 2013), and the sum of the bilateral hippocampal volume and the sum of the bilateral hippocampal surface area for each subject were used for this experiment.

Table 4.2.3 Demographic Information for Participants with MCI.

Group	Sex (M/F)	Age	MMSE
MCI converter (n=63)	42/21	75.2±7.0	26.7±1.7
MCI non-converter (n=55)	38/17	74.7±7.8	27.7±1.4

Values are mean ± standard deviation, where applicable.

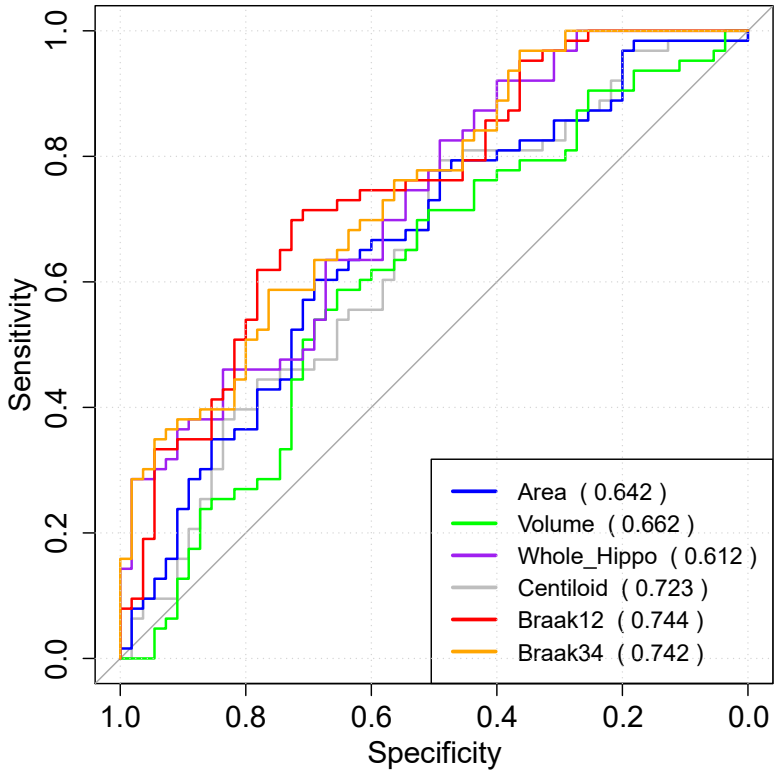

Figure 4.2.7. The ROC Analysis Results for Hippocampal Surface Area, Volume, the Whole Hippocampal Feature, and the Features on ROIs Associated with Aβ, Braak12, and Braak34. The AUC for each measurement is shown in parentheses.

To fit the univariate Cox model, we converted the features on ROIs to a single value for each subject. First, as the features on the ROIs should have stronger predictive power, we used the frequency on each vertex as a weight to multiply the RD and TBM on the vertex. And then, we respectively summed up the weighted RD and weighted TBM on the ROIs for each subject. The value for RD and the value for TBM were further reduced to a scalar with principal components analysis (PCA). PCA is an unsupervised model to reduce the dimensionality of the data while minimizing information loss. It creates new uncorrelated features which maximize the variance successively. For the

randomly selected features on the whole hippocampal surface, the RD and TBM were directly summed up without multiplying the frequency and reduced to a single value with PCA.

Then, the optimal cutoffs for these measurements were determined with the maximum sensitivity and specificity for distinguishing MCI converters and non-converters based on Receiver Operating Characteristic (ROC) analysis (Robin et al., 2011). The ROC curves are illustrated in **Figure 4.2.7,** and the AUC, 95% confidence interval (CI) of AUC, and the optimal cutoffs are shown in **Table 4.2.4**.

Table 4.2.4 AUC for ROC Analysis, Optimal Cutoffs, and Estimated Hazards Ratios (HRs) for Conversion to AD in MCI Patients with High-value and Low-value Biomarkers Based on a Univariate Cox Model.

Measurements	AUC (95% CI)	Cutoff	β	HR (95% CI)	*p*-value
Area	0.64 (0.54, 0.74)	8037.8	0.92	2.5 (2.3, 3.2)	0.001
Volume	0.66 (0.56, 0.76)	7814.9	0.89	2.5 (2.2, 3.1)	4.00E-04
Whole_hippo	0.61 (0.51, 0.72)	1.7	0.69	2.0 (1.7, 2.8)	0.007
Centiloid	0.72 (0.63, 0.81)	13.3	1.56	4.7 (4.6, 5.2)	4.00E-05
Braak12	0.74 (0.66, 0.83)	-3.0	1.35	3.8 (3.6, 4.3)	4.00E-07
Braak34	0.74 (0.65, 0.83)	-7.6	1.24	3.5 (3.3, 4.0)	1.00E-06

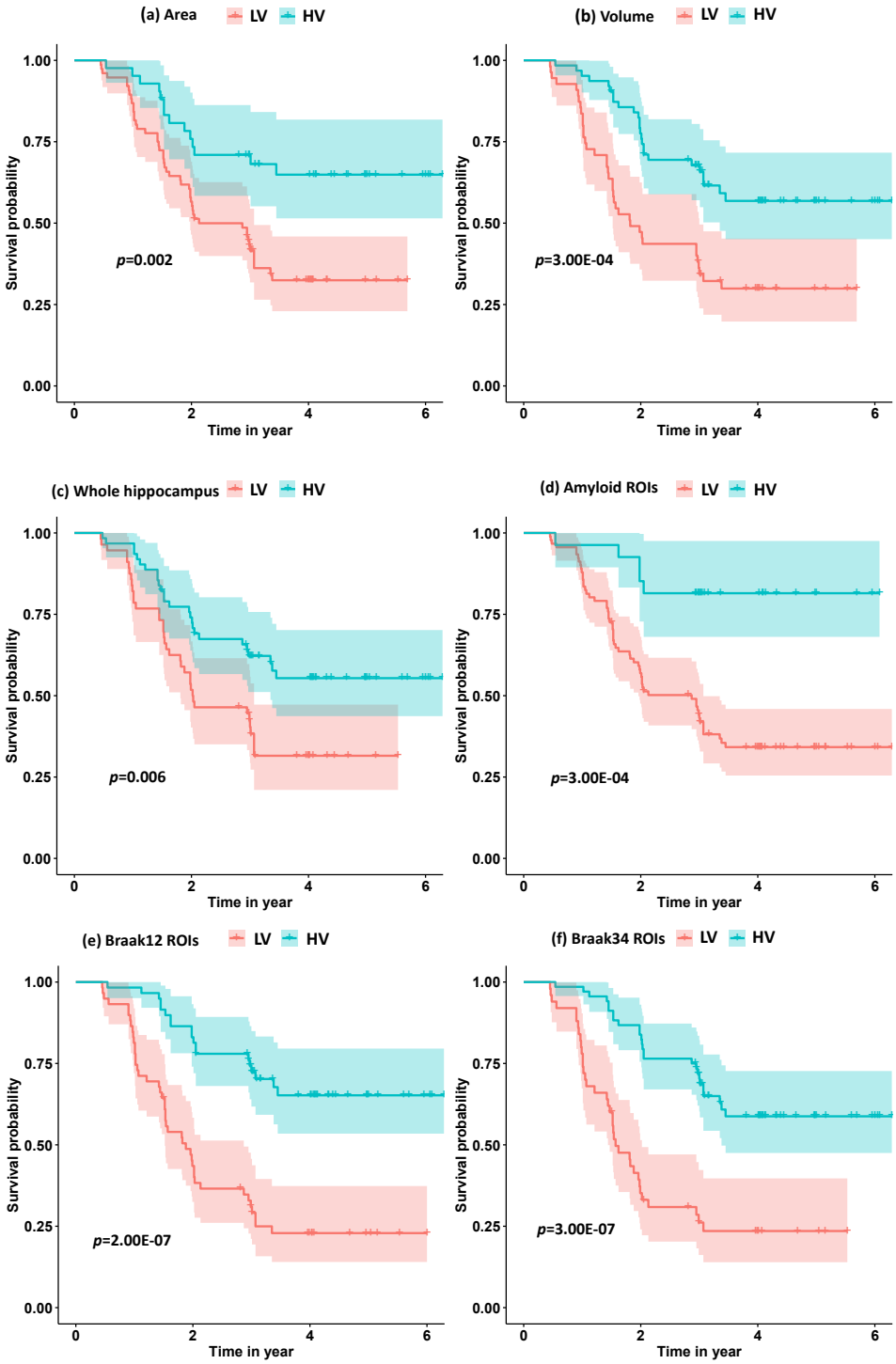

Figure 4.2.8. The Survival Probability Analysis for Progression to AD in MCI Patients Based on Hippocampal Surface Area, Volume, the Whole Hippocampal Features, and the Features on ROIs Related to Aβ, Braak12 and Braak34. The p-values are from the log-rank test. The red curve represents the high-value (HV) group for each measurement. and the blue one represents low-value (LV) group**.**

With the optimal cutoffs, we could divide the whole cohort into two groups with different measurements. For example, the subjects with hippocampal volume higher than 7814.9 mm^3 were assigned to a high-value (HV) group, and the rest were into a low-value (LV) group. As expected, AD may decrease the hippocampal volume as well as the other measurements. Next, we fitted a Cox proportional hazard model (Moore, 2008) with the six measurements separately, and the regression beta coefficients (β), the hazard ratios (HRs), and statistical significance (*p*-values) are shown in **Table 4.2.4**.

Moreover, we calculated the survival probabilities for conversion to AD in the HV group and the LV group by fitting Kaplan-Meier curves. The survival probabilities of the subjects based on hippocampal surface area, volume, the whole hippocampal features, and the features on ROIs related to Aβ, Braak12, and Braak34 are shown in **Figure 4.2.8.** Each color represents the survival curve and 95% CI of one group. Here a log-rank test was used to compare the survival group differences based on a χ^2 test, and the *p*-values are illustrated in each plot. A result with a *p*-value < 0.05 indicates that the two groups are significantly different in terms of survival time. The features from our selected ROIs tended to always yield stronger significant results than the hippocampal surface area, volume, and the whole hippocampal features.

4.2.7. Stability Analysis

To test whether the performance of our FMFS model could be affected by kinds of data distribution across institutions, we performed 5-fold cross-validation on the dataset for the study of Aβ under three conditions, including a data-centralized condition and data distributed across three institutions and five institutions. We simulated the

distributed condition on a cluster with several conventional x86 nodes, of which each contains two Intel Xeon E5-2680 v4 CPUs running at 2.40 GHz. Each institution is assigned one computing node. For each training data set, we randomly assigned the subjects to three institutions, five institutions, or a single institution. Besides, we also compared the performance of FMFS with DADMM and FBCD under the five-institution conditions. We perform cross-validation a total of 10 times with a sequence of regularization parameters, 1, 0.5, and 0.1, and with all the other experimental set-ups being the same as in the previous experiment. The average RMSE for the prediction of MMSE was employed to evaluate the prediction accuracy during training and testing, as shown in **Table 4.2.5**.

Table 4.2.5 Average RMSE for Predicting MMSE with FMFS Across Different Institutional Settings.

	λ	FMFS (1)	FMFS (3)	FMFS (5)	FBCD (5)	DADMM (5)
Train	1.0	2.80	2.80	2.80	2.80	2.80
	0.5	2.70	2.70	2.70	2.70	2.70
	0.1	2.43	2.43	2.43	2.43	2.44
Test	1.0	2.79	2.79	2.79	2.79	2.79
	0.5	2.71	2.71	2.71	2.71	2.71
	0.1	2.60	2.60	2.60	2.60	2.61

Additionally, we tried to collect datasets from different institutions and studies to validate the stability of our federated model in real-world conditions. Besides ADNI, we also collected MRI scans from other institutions, including 307 cognitively unimpaired

subjects from Open Access Series of Imaging Studies (OASIS) (Marcus et al., 2010b) and 38 MCI patients from Arizona APOE cohort study (AZ) (Caselli et al., 2009). The datasets for the study of Aβ and tau are treated as two institutions' data. Therefore, in this experiment, we have four institutions, Aβ for ADNI, tau for ADNI, OASIS and AZ. In the four-institution condition, each institution was assigned one computing node and all the other parameter settings were the same. Then, we fit the features of these data and MMSE in our FMFS as well as FBCD and DADMM. FMFS are validated at data-centralized and the four-institution condition. FBCD and DADMM are only under the four-institution condition. The average RMSE for the prediction of MMSE was used to evaluate the prediction accuracy. The results of the training loss are shown in **Table 4.2.6**. The results indicated that different kinds of institutional distributions did not strongly influence our FMFS model.

Table 4.2.6 Average RMSE for Predicting MMSE with FMFS Across Datasets from Different Institutions.

λ	Centralized	FMFS (4)	FBCD (4)	DADMM (4)
1.0	2.73	2.73	2.73	2.73
0.5	2.56	2.56	2.56	2.56
0.1	2.33	2.33	2.33	2.33

4.3. Results for Federated GEIDI

4.3.1. Data Description

The data in this work are from the Alzheimer's Disease Neuroimaging Initiative (ADNI, adni.loni.usc.edu) and the TADPOLE challenge (tadpole.grand-challenge.org)

(Marinescu et al., 2020). The ADNI was launched in 2003 as a public-private partnership led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessments can be combined to measure the progression of MCI and early AD. The genome-wide association study of ADNI is designed to provide researchers with the opportunity to combine genetics with imaging and clinical data to help investigate the mechanisms of the disease. For up-to-date information, see adni.loni.usc.edu/data-samples/data-types/genetic-data/.

From the ADNI GWAS, we analyzed data from 697 subjects, including AD patients, people with mild cognitive impairment (MCI), and cognitively unimpaired (CU) subjects, for whom the demographic information is shown in **Table 4.3.1**. Each sample has three types of modalities of data: genotypes of known AD risk genes (e.g., *APOE*) and SNPs from genome-wide association studies (GWAS), gene expression measurements (for 20,211 genes) from microarray-based transcriptomic profiling of samples' blood, and imaging biomarkers from structural magnetic resonance imaging (sMRI) data of subjects' brains. We use *plink* to perform a quality check of the genotype data. The SNPs in the normal group that deviate significantly from Hardy-Weinberg equilibrium are removed (Purcell et al., 2007). The LINNORM package (Yip et al., 2017) was adopted to perform data transformation on the expression data for normality and homoscedasticity. Recent evaluations (Huang et al., 2018; Yip et al., 2018) show that LINNORM typically performs better than current DEG analysis methods for both single-

cell and bulk RNA-Seq, such as Seurat (Satija et al., 2015) and DESeq2 (Love et al., 2014).

Eventually, we get 2,059,586 SNPs, *APOE* genotype, and expression data for 20,211 genes for each sample. Besides, from the TADPOLE challenge, we obtained two brain imaging biomarkers for each subject calculated using FreeSurfer (Fischl et al., 1999) with sMRI, including the volume of the hippocampus and middle temporal gyrus (MidTemp). To adjust for individual differences in head size, the volume of each region is adjusted by the intracranial vault volume (ICV) of each subject (volume/ICV). The difference between the dates for gene expression collection and MRI scan is less than five months.

Table 4.3.1 Demographic Information for the Subjects We Study from the ADNI.

Group	**Sex (M/F)**	**Age**	**MMSE**
AD (n=96)	59/37	74.8 ± 7.5	21.8 ± 4.1
MCI (n=366)	209/157	72.0 ± 7.5	28.0 ± 1.7
CU (n=235)	115/120	74.4 ± 5.8	29.1 ± 1.2

Values are mean ± standard deviation, where applicable.

4.3.2. APOE-related Gene Expressions

APOE genotype is a well-known genetic biomarker for predicting subjects' risk for AD. We stratify 697 subjects into three subgroups based on their *APOE* genotype status: non-carriers (e3/e3), heterozygotes (e3/e4), and homozygotes (e4/e4). Federated GEIDI is then adopted to discover genes correlated with hippocampus volume

differentially across the three subgroups. We first run federated GEIDI with the volume of both sides of the hippocampus and the expression measures for 20,211 genes. Next, 1,625 gene expression measures are selected with $p < 0.05$. We evaluate the enrichment of these genes and the 632 AD-related genes annotated on alzgene.org and find 73 overlapping genes, yielding a hypergeometric enrichment $p = 0.00039$. Among the 73 overlapping genes, the top ten gene expressions are those measured for *CAST, CST3, GSTO1, LSS, MS4A4A, NPC1, PMVK, PPM1H, PPP2R2B,* and *SORCS2*. Besides, the top ten genes in the 1,625 gene expressions are *IK, BRPF3, BTN3A2, LOC101929275, TDRG1, PAFAH1B1, SERINC3, ALKBH6, VPS45, and LGALS1*. We also perform the false discovery rate (FDR) (Benjamini & Hochberg, 1995) test on the 20,211 *p*-values but none of the corrected *p*-values are significant.

Additionally, we perform the same experiments on the volume of the middle temporal gyrus (MidTemp); the results are shown in Table 2. 2,415 gene expressions are significant and 92 of them overlap with the 632 AD-related genes - with a hypergeometric enrichment $p = 0.00624$. The top ten gene expressions are those measured for *ABCA2, COL11A1, CST3, GNA11, HMOX1, HSPA1B, MAOA, MS4A4A, PRKAB2,* and *SORCS2*. And the top ten genes in the 2,415 gene expressions are *GLRA3, CAMK2N2, MCOLN2, BPIFA1, KIT, CST3, SLC20A2, LGALS4, TNFSF8,* and *LCOR*. After performing FDR on the 20,211 genes, three gene expressions are significant, including *GLRA3, CAMK2N2,* and *BPIFA1*.

Matrix eQTL (Shabalin, 2012) is a state-of-the-art software to study the association between genotype and gene expression. We also leverage the linear model

and the ANOVA model in Matrix eQTL to evaluate the *APOE* genotype and the measured expression levels of the 20,211 genes. For the linear model, there are 2,657 significant gene expressions and 98 overlapping genes, leading to a hypergeometric enrichment $p = 9.76E-03$. For the ANOVA model, 3,234 gene expressions are selected, and 110 known genes are found, which leads to a *p*-value $= 2.665E-02$. The results show that our federated GEIDI can detect the most gene candidates that are significantly enriched for known AD genes. As the volume of hippocampus has the best performance in detecting AD-related genes, we use it as the imaging biomarker for all the remaining experiments.

Table 4.3.2 Hypergeometric Statistics for APOE.

Structures	**Selected genes**	**Overlapping genes**	***p*-value**
Hippocampus	1,625	73	0.00039
MidTemp	2,415	92	0.00624
Linear Regression	2,657	98	0.00976
ANOVA	3,234	110	0.02665

4.3.3. SNP-related Gene Expressions

In this experiment, we stratify the subjects into three subgroups based on their SNP status. We choose *rs942439*, as this SNP was reported in alzgene.org, and also one of the top hits in our experiment of discovering AD-related SNPs (the details about selecting AD-related SNPs will be introduced in Sec. 4.3.2). And we use the volume of both sides of hippocampus as the imaging biomarker because of its superior performance

in the first experiment. Federated GEIDI is used to detect any known AD gene whose expression is differentially associated with hippocampus volume in the subgroups stratified by the genotype at *rs942439* locus.

As shown in **Table 4.3.3**, 1,587 gene expressions are significant and 60 of them are reported in alzgene.org and IGAP GWAS results, leading to a hypergeometric enrichment $p = 0.017$. Of these 60 gene expression measures, the top ten genes are *ADRB1, ALOX5, ATXN1, CBS, FGF1, FLOT1, HSPA1A, RFTN1, SORL1,* and *XRCC1*. Besides, the top ten genes in the 1,587 gene expressions are *AIF1L, KRT23, CA2, C2ORF88, HSPA1A, LRGUK, LGALS3BP, IFT46, DDX23,* and *FAM166B.* After performing the FDR test on the 20,211 genes, none of the gene expression is significant.

We also perform eQTL analysis on the SNP, *rs942439*. For a linear regression model, 1,794 gene expressions are selected, and, of these, 66 genes are reported in alzgene.org and IGAP GWAS results, yielding a hypergeometric enrichment $p = 0.021$. For the ANOVA model, 1,347 gene expression values are significant, and, of these, 49 genes are reported in alzgene.org and IGAP; in this case, the hypergeometric enrichment was $p = 0.033$.

Table 4.3.3 Hypergeometric Statistics for *rs942439*.

Structures	**Selected genes**	**Overlapping genes**	***p*-value**
Hippocampus	1,587	60	0.017
Linear Regression	1,794	66	0.021
ANOVA	1,347	49	0.033

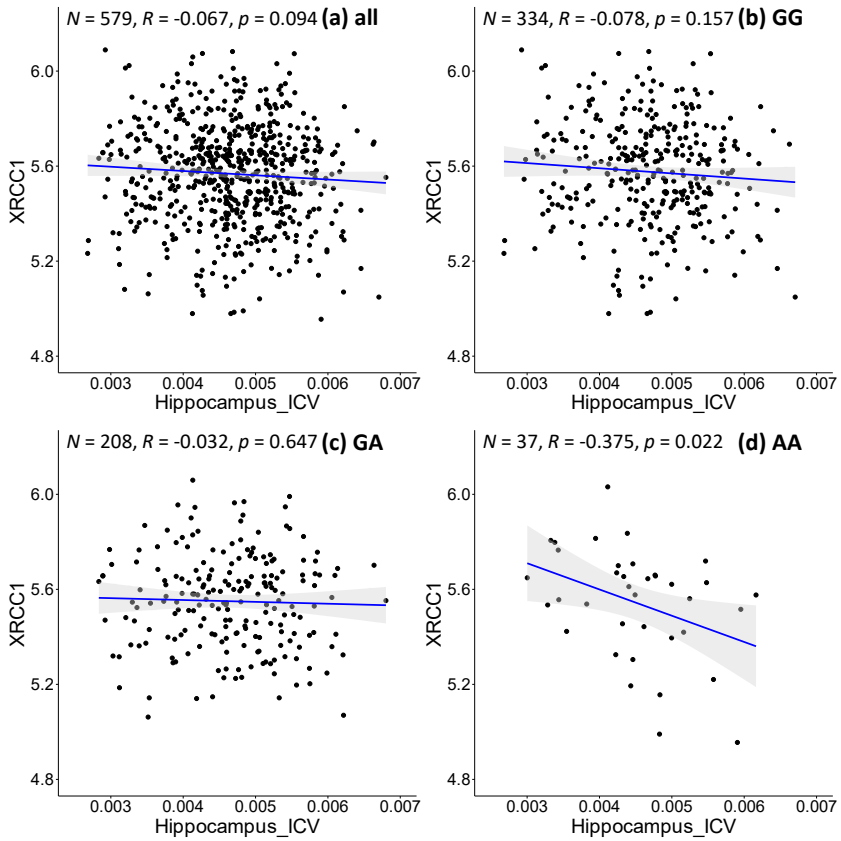

Figure 4.3.1. Correlation of Image Biomarkers and XRCC1 Gene Expression in Subpopulations Stratified by the Sample's Genotype at rs942439. (a) all samples **(b)** individuals with "GG" genotype; **(c)** those with "GA" genotype **(d)** those with "AA" genotype.

In the experiment, one of the most significant gene expression measures is for *XRCC1*, for which the *p*-value is $4.332E-03$. *XRCC1* is a gene coding for the X-ray repair cross-complementing protein; it was previously reported to be weakly associated with AD in a Turkish population (Doğru-Abbasoğlu et al., 2007).

As shown in **Figure 4.3.1**, we further adopt Pearson's correlation to evaluate the relationship between the hippocampal volume (*x*-axis) (adjusted for ICV) and *XRCC1* gene expression (*y*-axis) of each subgroup. **Figure 4.3.1 (a)** illustrates the distribution for all the samples. **Figure 4.3.1 (b), (c) and (d)** show the distribution for the samples with "GG", "GA" and "AA" genotype, respectively. Above each subfigure, *R* and *p* are the

Pearson correlation coefficient and *p*-value, and N is the number of subjects. Even so, there is always some missing information in the genotype data. Hence, before we run federated GEIDI as well as the Pearson correlation statistics, we remove the subjects without the specific genotype. Because of this, the total number N in **Figure 4.3.1 (a)** is 579 instead of 697. We find samples with an "AA" genotype had hippocampal volume negatively correlated with expression levels of *XRCC1* (N=37, R=0.37, p=0.022). In contrast, the analysis in all samples (**Figure 4.3.1 (a))** or subjects with either "GG" or

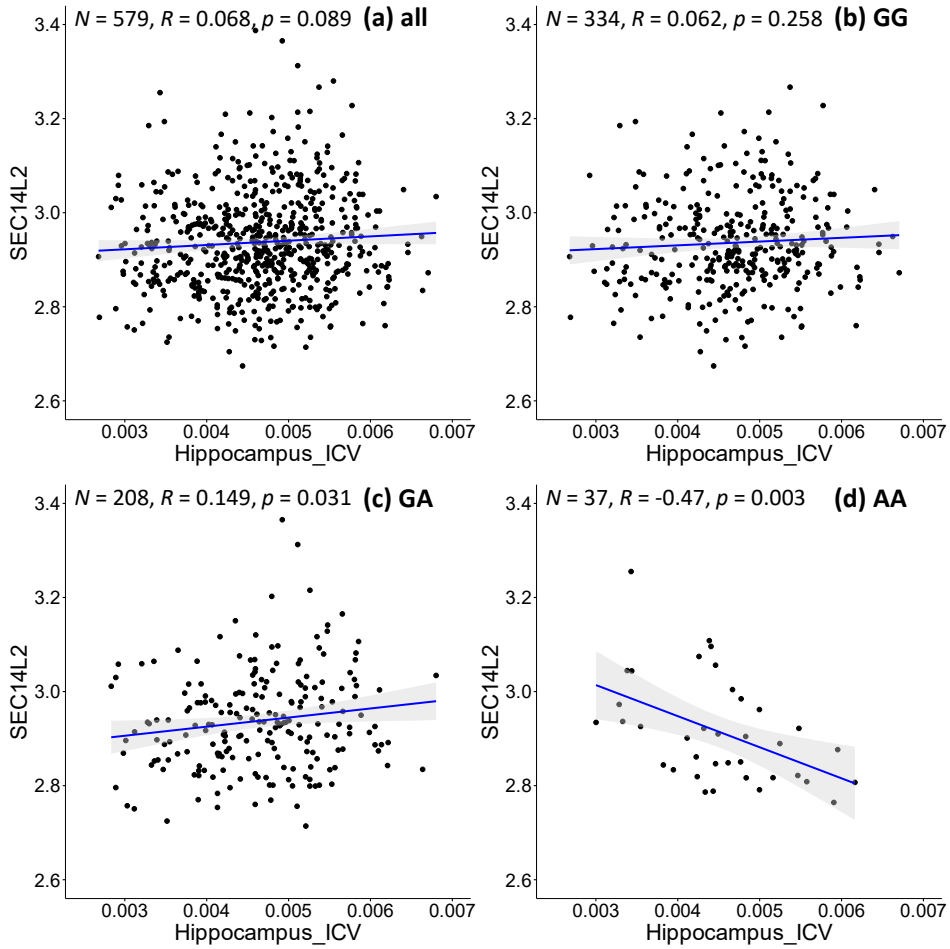

Figure 4.3.2. Correlation of Image Biomarkers and SEC14L2 Gene Expression in Subpopulation Stratified by the Sample's Genotype at rs942439. (a) all samples **(b)** those with "GG" genotype; **(c)** those with "GA" genotype; **(d)** those with "AA" genotype.

"GA" genotype (**Figure 4.3.1 (b), (c))** showed that the Pearson correlation coefficients were not significant in the overall, pooled sample. This result indicates that our method can establish associations among SNP, imaging, and gene expression data that include known AD risk factors. We further apply the above procedure to discover genes that have never been reported to be associated with AD. As shown in **Figure 4.3.2 (d)**, *SEC14L2* gene expression is negatively associated with hippocampal volume only in the subpopulation with "AA" genotype at rs942439 locus (N=37, R=-0.47, p=0.003).

Interestingly, the opposite correlation is found in a subpopulation with "GA" genotype (**Figure 4.3.2 (c)**, N=208, R=0.15, P=0.03), and when applied to all pooled subjects, the total population does not show significant correlations (**Figure 4.3.2 (a)**, N=579, R=0.07, p=0.09). The *SEC14L2* gene encodes a protein that stimulates squalene monooxygenase, a downstream enzyme in the cholesterol biosynthesis pathway. This gene has never been reported to be associated with AD, but high cholesterol levels have been linked to early-onset AD (Wingo et al., 2019). This result indicates that our method can detect strong correlations in specific subpopulations that cannot be detected in the whole population. We also observe conflicting directions in different subpopulations, as shown by "GA" and "AA" subpopulations showing opposite correlations. This also highlights the importance of individualized medicine in patient management, as the same drug may have opposing effects in different groups of samples. Thus, federated GEIDI offers a new approach to discover novel genes related to AD as potential drug targets.

4.3.4. Discovering AD-related SNPs

In the experiments of Sec. 4.3.1, we used hypergeometric statistics to evaluate the ability of our proposed model to discover AD-related gene expressions that are differentially associated with imaging measures in populations stratified by *APOE* haplotype. In this experiment, we also use hypergeometric statistics to assess the discovery rate of known AD-related genes, in the set of genes whose expression shows different correlations with imaging markers, in samples stratified according to different genotypes. Sets that are enriched in AD-related SNPs will have a more significant *p*-value in the hypergeometric test that assesses enrichment. Since the hippocampal volume measure showed superior performance for this task, among all the imaging biomarkers in Sec. 4.3.1, we adopt it as the brain imaging measure in this experiment. To illustrate the effectiveness of our GEIDI model, we perform the same experiment with the linear model in Matrix eQTL, which can evaluate the associations between SNPs and gene expression. To adjust for multiple comparisons, we will convert raw p-values to false discovery rate (FDR) and consider trios with FDR <0.05 as functionally important. When we analyze each SNP with our federated GEIDI and Matrix eQTL, we will obtain a *p*-value for each of the 20,211 expressed genes. Instead of selecting the significant gene expressions with a *p*-value < 0.05, we respectively rank the *p*-value of all the gene expressions calculated by the two methods and select the top *N* (100 and 200) gene expressions to apply the hypergeometric analysis. With the *p*-value from this hypergeometric analysis (which assesses enrichment for known AD-associated genes), we may rank the SNPs and obtain the most AD-related ones. Then, we try to prove that our GEIDI is able to detect more AD-related SNPs. From AlzGene.org, we also created a

list of 1,217 AD-related SNPs, and we randomly selected another 1,217 SNPs as the non-AD-related ones. After ranking the SNPs with the *p*-value computed by the two methods, we calculate the true positive rate (TPR) for the top *m* SNPs, which measures the percentage of AD-related SNPs in the selected top *m* SNPs. For example, the last number in **Table 4.3.4** is 0.57, which means 57% of the top 500 SNPs are AD-related ones. As the results in **Table 4.3.4**, our federated GEIDI can always achieve superior performance than Matrix eQTL.

Table 4.3.4 True Positive Rates of AD-related SNPs in the Top m SNPs.

Matrix eQTL: Linear Regression					
EXP (*N*) \SNP (*m*)	10	50	100	200	500
100	0.50	0.58	0.52	0.50	0.53
200	0.50	0.60	0.55	0.58	0.54
Matrix eQTL: ANOVA					
EXP (*N*) \SNP (*m*)	10	50	100	200	500
100	0.60	0.58	0.52	0.49	0.55
200	**0.60**	0.60	0.57	0.55	0.55
Federated GEIDI					
EXP (*N*) \SNP (*m*)	10	50	100	200	500
100	**0.60**	**0.60**	**0.60**	**0.61**	**0.60**
200	**0.60**	**0.62**	**0.61**	**0.58**	**0.57**

The SNPs are ranked with the *p-value* from hypergeometric analysis with the top *N* gene expressions as the number of samples drawn from the population

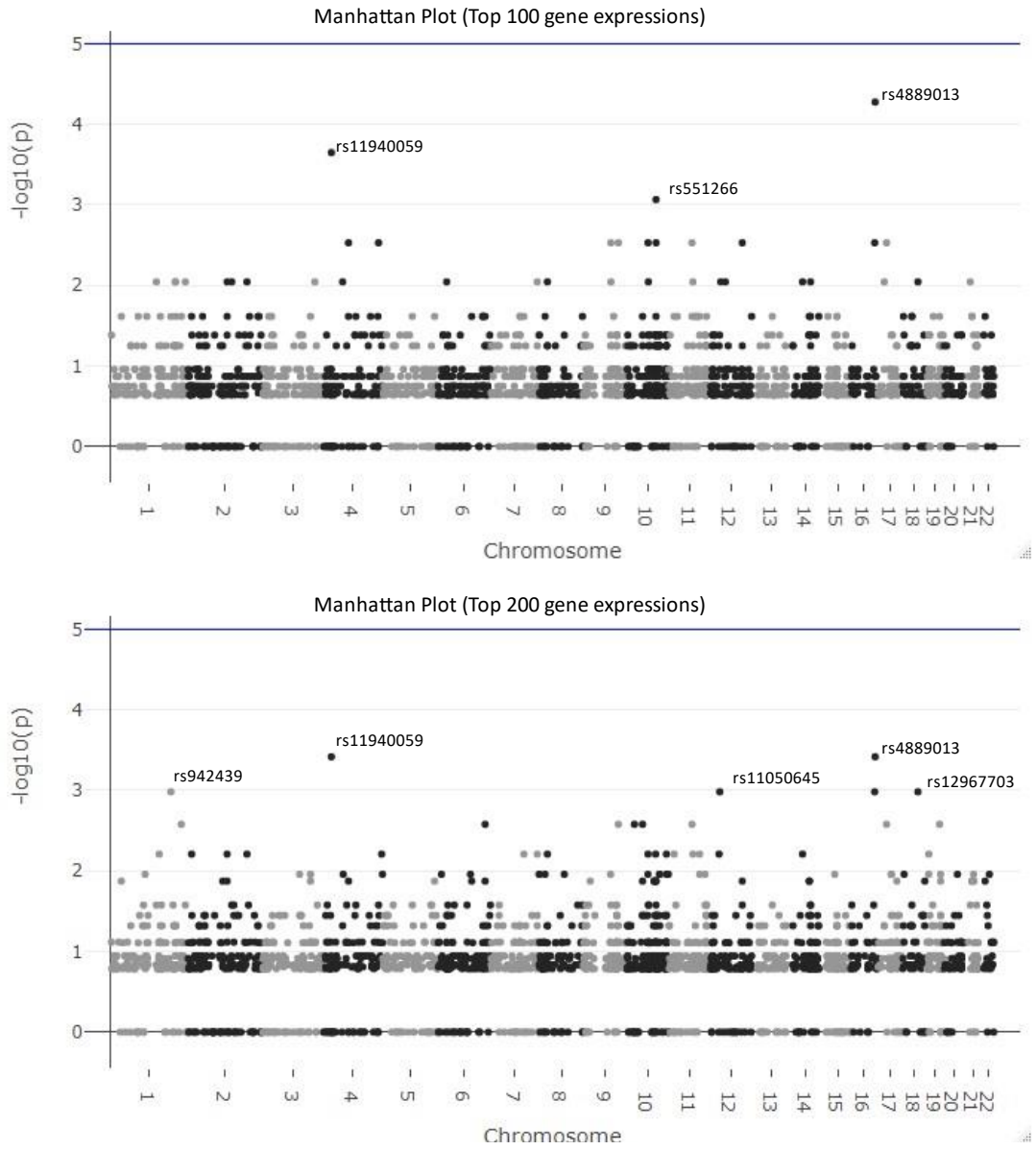

Figure 4.3.3. Manhattan Plots for the Results of Federated GEIDI. The top figure is the Manhattan plot for the results from hypergeometric analysis with the top 100 gene expressions as the number of samples drawn from the population and the bottom one is for the results with the top 200 gene expressions. The SNPs, rs4889013 and rs11940059, are the top-ranked ones for both results.

In **Figure 4.3.3,** we visualize the *p*-values of these 2,434 SNPs from hypergeometric analysis in the Manhattan plots. The top figure is the Manhattan plot for the result with the top 100 gene expressions and the bottom one is for the result of the top 200 gene expressions. The SNPs, *rs4889013* and *rs11940059*, are the top-ranked ones for

both results. When we select 100 or 200 as the number of samples drawn from the population, three parameters in of hypergeometric analysis are fixed and only the number of observed successes, *k,* varies for different SNPs. Therefore, the *p*-value from different SNPs might be the same if their numbers of observed successes are the same. This explains why results of some SNPs locate at the same horizontal position.

4.3.5. Federated Learning Stability Analysis

In this experiment, we aim to demonstrate that the performance of our federated GEIDI model is not greatly affected by different data distribution models across institutions. In practice, it would be convenient and efficient to run association tests on data that might be distributed across multiple servers without transferring it all to a centralized location. We developed this algorithm with R language and simulated the distributed condition on a cluster with several conventional x86 nodes, of which each contains two Intel Xeon E5-2680 v4 CPUs running at 2.40 GHz. Each institution is assigned one computing node. We synthesized 1,000 samples and randomly assigned them to different independent hypothetical institutions, including one institution, three institutions, five institutions and seven institutions. We compared the residuals from each linear regression model for each condition and found the residuals remained unchanged, as shown in **Table 4.3.5.** The first column is the ground truth residual, and the rest are the residuals for our federated linear model under different data distribution conditions. The residuals are the same, which means that the results of our Federated GEIDI will remain stable under different multi-site conditions. Therefore, these results demonstrate the correctness and stability of our federated GEIDI model.

Table 4.3.5 Stability Analysis of Federated GEIDI Across Different Institutional Settings.

	Ground Truth	**1-institution**	**3-institution**	**5-institution**	**7-institution**
Residual	3.9553	3.9553	3.9553	3.9553	3.9553

4.4.Results for Federated Chow Test on Hippocampal Morphometry

4.4.1. Data Description

Data for testing the performance of our proposed framework and comparable methods were obtained from the publicly available ADNI database (adni.loni.usc.edu). ADNI is the result of efforts by many co-investigators from a broad range of academic institutions and private corporations. Subjects are recruited from over 50 sites across the U.S. and Canada. From ADNI 1, ADNI 2, ADNI GO, and ADNI 3 (the different phases of ADNI), we obtained 847 pairs of MRI scans and AV1451 PET images, including 502 non-carriers (NC), 281 heterozygote subjects (HT) and 64 homozygote subjects (HM). **Table 4.4.1** shows the demographic information from the cohort that we analyzed. The PET images are reprocessed using a single pipeline consistent with the work of Sanchez et al.(Sanchez et al., 2021), so the standardized uptake value ratio (SUVR) from different ADNI study sites can be analyzed together. In this work, we analyze tau deposition in a brain region named Braak34 (S. L. Baker, Lockhart, et al., 2017).

Table 4.4.1 Demographic Information for the Subjects We Study from the ADNI.

Group	Sex (M/F)	Age	MMSE	Braak34
NC (502)	242/260	75.4±7.5	28.2±2.8	1.80±0.34
HT (281)	121/160	73.4±7.2	27.8±3.1	1.93±0.51
HM (64)	33/31	71.0±7.9	25.8±4.6	2.24±0.71

Values are mean ± standard deviation where applicable.

4.4.2. Linking Hippocampal Volume to Tau Deposition

In this experiment, we stratify 847 subjects into three subgroups based on their APOE genotype status: non-carriers (NC), heterozygotes (HT), and homozygotes (HM). And then, we use the federated chow test model to explore the difference of the changes in hippocampal volume and the measure of tau, Braak34, among these cohorts. The volume of each side of hippocampus is used as the predictors and Braak34 is used as the response. Both sides of hippocampi have significant results with our federated chow test. The *p*-value of the left side is *5.1e-16*, and the right one is *8.5e-16*, which means the changes of hippocampal volume and Braak34 are significantly different in different cohorts.

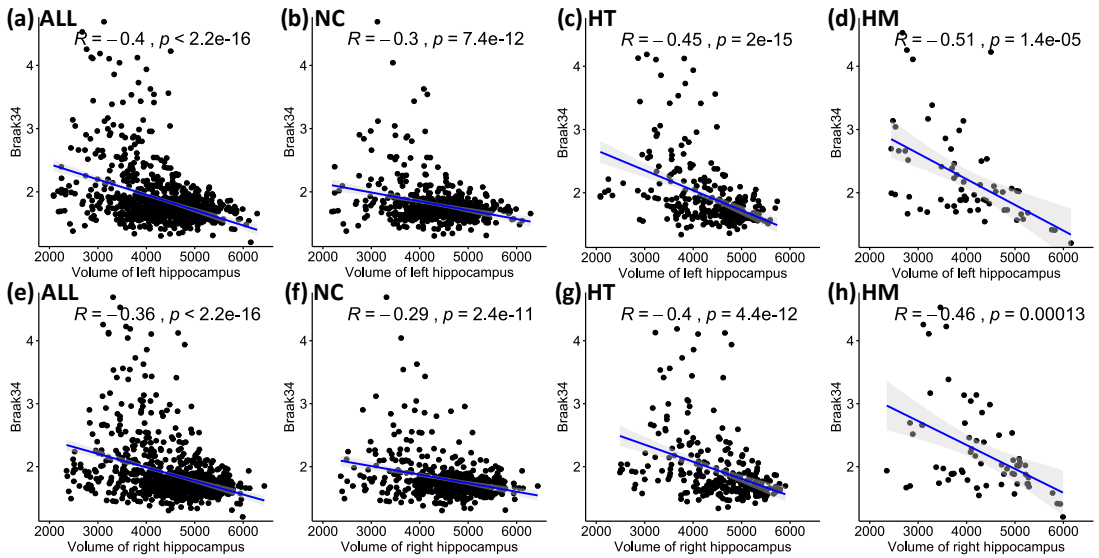

Figure 4.4.1. Correlation of Hippocampal Volume and Braak34 in Subpopulations Stratified by the Sample's APOE Genotype. The top four subfigures are the distributions for left hippocampal volume and Braak34 and the bottom four are for right hippocampus and Braak34. The first column is the results for all the 847 samples and the rest are for the cohorts of NC, HT and HM, respectively. *R* and *p* are Pearson correlation coefficient and *p*-value.

We further adopt Pearson's correlation to evaluate the relationship between the hippocampal volume (x-axis) and Braak34 (y-axis) of each subgroup, as illustrated in **Figure 4.4.1**. The top four subfigures are the distributions for left hippocampal volume and Braak34, and the bottom four are for the right hippocampus and Braak34. The first column is the result for all the 847 samples, and the rest three are for the cohorts of NC, HT, and HM, respectively. *R* and *p* in each subfigure are the Pearson correlation coefficient and *p*-value. All the *p*-values are significant and the cohort with HM genotype has the strongest negative correlation between volume and Braak34.

4.4.3. Linking Hippocampal Morphometry to Tau Deposition

We further figure out the regions where the atrophy focuses on the hippocampal surface. In this experiment, we used two morphometry features, radial distance (RD) and tensor-based morphometry (TBM) in different cohorts. Each morphometry feature on the vertex is used as predictor and Braak34 is used as response. A *p*-value is calculated with the federated chow test to evaluate effect of tau deposition to the morphometry feature on each vertex in the cohorts with different *APOE* genotype. To adjust for multiple comparisons, we will convert these raw *p*-values of all the 15,000 features on each surface to false discovery rate (FDR) (Storey, 2003) and consider trios with FDR<0.05 as functionally important. The average correct *p*-values for RD and TBM on each side of

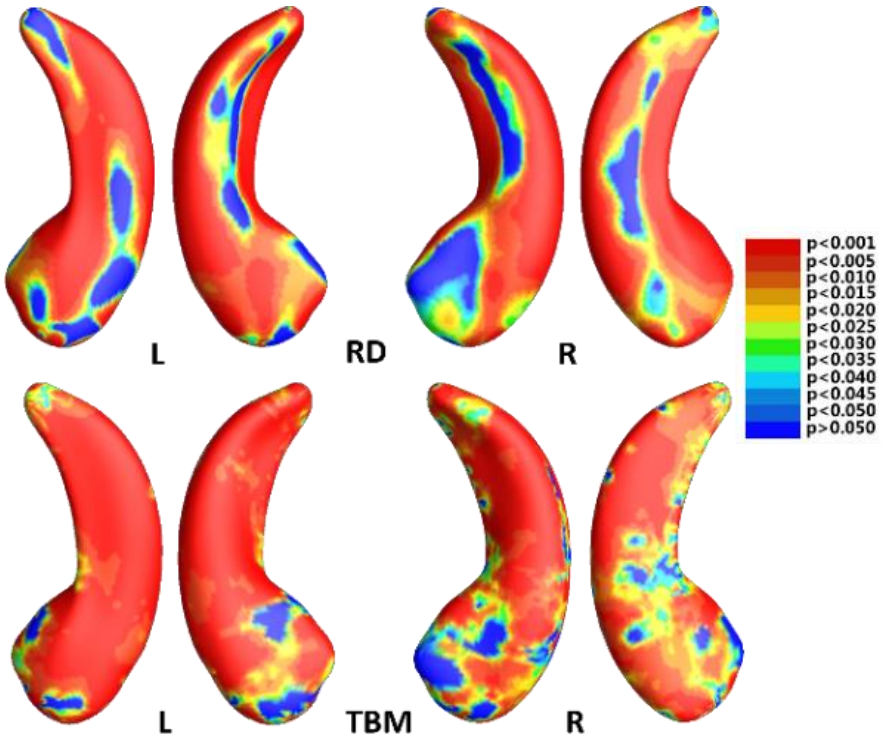

Figure 4.4.2. p-maps of Our Federated Chow Test Model. The warmer color regions have more significant p-values. The top two subfigures are the results for RD and the bottom two are for TBM.

hippocampi are all significant. The average *p*-values for RD of left and right hippocampi are *2.6e-13* and *3.6e-13*. And the *p*-values for TBM are *4.4e-13* and *6.3e-13*.

As the illustration in **Figure 4.4.2**, we further visualize the *p*-values on each hippocampal surface to figure out the atrophy regions. Since the *p*-values are very significant, we normalized the *p*-values by dividing each *p*-value by the maximal *p*-value on each surface. Then, we map the values to the color map as shown in **Figure 4.4.2**. The warmer color regions have more significant *p*-values. The top two subfigures are the results for RD and the bottom two are for TBM. The atrophy focuses on the hippocampal subregions, subiculum and cornu ammonis 1 (CA1 subfield), which is consistent with the previous studies (An et al., 2021; Hanko et al., 2019; G. Wang et al., 2021).

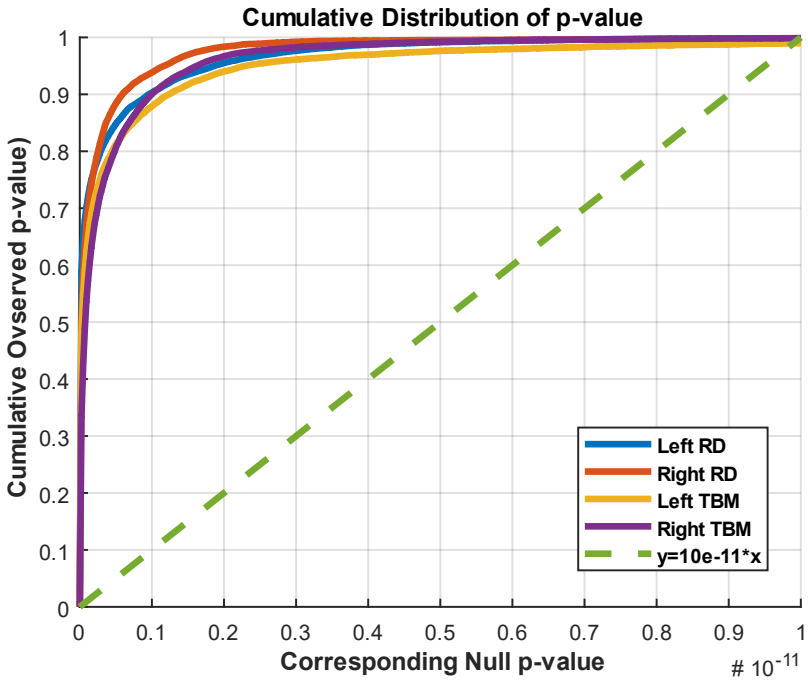

Figure 4.4.3. Cumulative Distribution Functions of the p-values.

4.4.4. Cumulative Distribution Analysis

To further compare the effect size of the morphometry features in the federated chow test model, we created cumulative distribution function (CDF) plots of the resulting uncorrected *p*-values as the previous studies (Y. Wang et al., 2011). For null distributions, the CDF of *p*-values is expected to fall approximately along the line (y = x). These empirical CDFs of *p*-values are the flip of the more common FDR PP plot; steeper CDFs show stronger effect sizes. We use $y = 10^{-11}x$ line in **Figure 4.4.3** since both RD and TBM show excellent effect size, and it is hard to figure out the better one with the null hypothesis ($y = x$). Use of the $y = 10^{-11}x$ line is related to the fact that significance is declared when the volume of suprathreshold statistics is more than 10^{11} times that expected under the null hypothesis. As shown in **Figure 4.4.3**, RD is a little bit better than TBM.

4.4.5. Predicting Clinical Decline in Participants with MCI

Table 4.4.2 Demographic Information for Participants with MCI.

Group	Sex (M/F)	Age	MMSE
MCI converter (n=63)	42/21	75.2±7.0	26.7±1.7
MCI non-converter (n=55)	38/17	74.7±7.8	27.7±1.4

Values are mean ± standard deviation, where applicable.

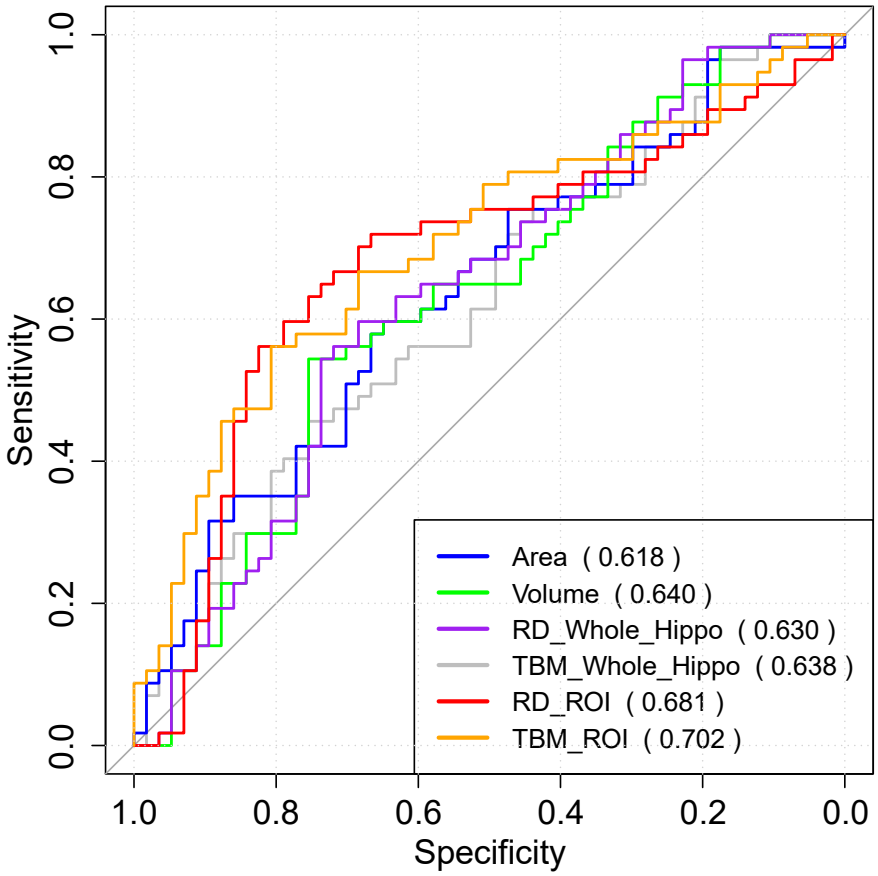

Figure 4.4.4. The ROC Analysis Results for Hippocampal Surface Area, Volume, the Whole Hippocampal Feature, and the Features on ROIs Associated with RD and TBM. The AUC for each measurement is shown in parentheses.

In this experiment, we evaluated the performance of our features on the ROI in survival analysis by using 118 MCI participants' data from a separate dataset (G. Wang et al., 2021) from ADNI (**Table 4.4.2**), including 63 MCI converters, who converted to probable AD in the next six years, and 55 MCI non-converters. We also chose 50 RD and 50 TBM from the ROIs of each hippocampal surface and 50 features on the whole hippocampal surface to predict the conversion rates from MCI to AD, respectively. For comparison, we also performed the same experiment with the surface area and volume of the hippocampus. The hippocampal volume and surface area were calculated with the

smoothed hippocampal structures after linearly registered to the MNI imaging space (Patenaude et al., 2011; Shi, Thompson, et al., 2013), and the sum of the bilateral hippocampal volume and the sum of the bilateral hippocampal surface area for each subject were used for this experiment.

To fit the univariate Cox model, we converted the features on ROIs to a single value for each subject. First, as the features on the ROIs should have stronger predictive power, we used the frequency on each vertex as a weight to multiply the RD and TBM on the vertex. And then, we respectively summed up the weighted RD and weighted TBM on the ROIs for each subject. The value for RD and the value for TBM were further reduced to a scalar with principal components analysis (PCA). PCA is an unsupervised model to reduce the dimensionality of the data while minimizing information loss. It creates new uncorrelated features which maximize the variance successively. For the randomly selected features on the whole hippocampal surface, the RD and TBM were directly summed up without multiplying the frequency and reduced to a single value with PCA.

Then, the optimal cutoffs for these measurements were determined with the maximum sensitivity and specificity for distinguishing MCI converters and non-converters based on Receiver Operating Characteristic (ROC) analysis (Robin et al., 2011). The ROC curves are illustrated in **Figure 4.4.4,** and the AUC, 95% confidence interval (CI) of AUC, and the optimal cutoffs are shown in **Table 4.4.4**.

Table 4.4.3 AUC for ROC Analysis, Optimal Cutoffs, and Estimated Hazards Ratios (HRs) for Conversion to AD in MCI Patients with High-value and Low-value Biomarkers Based on a Univariate Cox Model.

Measurements	AUC (95% CI)	Cutoff	β	HR (95% CI)	*p*-value
Area	0.62 (0.52, 0.72)	7447.9	0.76	2.1 (1.3, 3.6)	4.48E-03
Volume	0.64 (0.54, 0.74)	7814.9	0.84	2.3 (1.4, 3.9)	1.91E-03
RD_Whole_Hippo	0.63 (0.53, 0.73)	3.8	0.97	2.6 (1.6, 4.4)	2.98E-04
TBM_Whole_Hippo	0.64 (0.54, 0.74)	4.1	0.94	2.6 (1.5, 4.4)	5.36E-04
RD_ROI	0.68 (0.58, 0.78)	3.8	1.24	3.5 (1.9, 6.2)	2.89E-05
TBM_ROI	0.70 (0.61, 0.80)	4.4	1.3	3.7 (2.2, 6.2)	1.38E-06

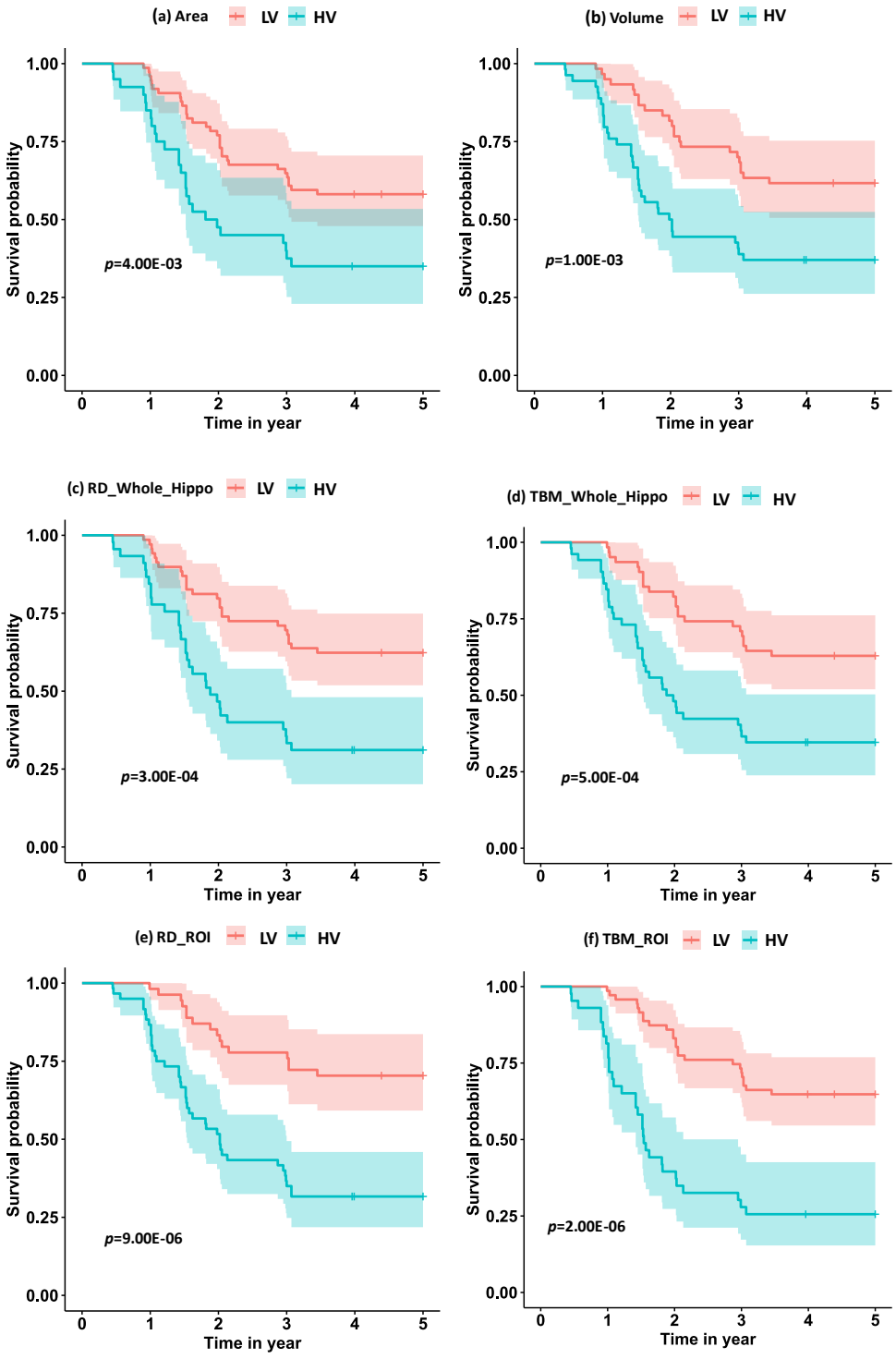

Figure 4.4.5. The Survival Probability Analysis for Progression to AD in MCI Patients Based on Hippocampal Surface Area, Volume, the Whole Hippocampal Features, and the Features on ROIs Related to Aβ, Braak12 and Braak34. The *p*-values are from the log-rank test. The red curve represents the high-value (HV) group for each measurement. and the blue one represents low-value (LV) group**.**

With the optimal cutoffs, we could divide the whole cohort into two groups with different measurements. For example, the subjects with hippocampal volume higher than 7814.9 mm^3 were assigned to a high-value (HV) group, and the rest were into a low-value (LV) group. As expected, AD may decrease the hippocampal volume as well as the other measurements. Next, we fitted a Cox proportional hazard model (Moore, 2008) with the six measurements separately, and the regression beta coefficients (β), the hazard ratios (HRs), and statistical significance (*p*-values) are shown in **Table 4.4.5**.

Moreover, we calculated the survival probabilities for conversion to AD in the HV group and the LV group by fitting Kaplan-Meier curves. The survival probabilities of the subjects based on hippocampal surface area, volume, the whole hippocampal features, and the features on ROIs related to RD and TBM are shown in **Figure 4.4.5.** Each color represents the survival curve and 95% CI of one group. Here a log-rank test was used to compare the survival group differences based on a χ^2 test, and the *p*-values are illustrated in each plot. A result with a *p*-value < 0.05 indicates that the two groups are significantly different in terms of survival time. The features from our selected ROIs tended to always yield stronger significant results than the hippocampal surface area, volume, and the whole hippocampal features.

4.4.6. Federated Learning Stability Analysis

In this experiment, we aim to demonstrate that the performance of our federated chow test model is not greatly affected by different data distribution models across institutions. We synthesized 1,000 samples and randomly assigned them to different independent hypothetical institutions, including one institution, three institutions, and

five. We compared the residuals from each linear regression model for each condition and found the residuals remained unchanged, as shown in **Table 4.4.4.** The first column is the ground truth residual and the rest are the residuals for our federated linear model under different data distribution conditions. The results demonstrate our federated chow test will remain stable under different multi-site conditions. Therefore, these results demonstrate the correctness and stability of our federated chow test model.

Table 4.4.4 Stability Analysis of Federated Linear Model Across Different Institutional Settings.

	Ground Truth	**One-institution**	**Three-institution**	**Five-institution**
Residual	3.96	3.96	3.96	3.96

Chapter 5

CONCLUSIONS

In this dissertation, I introduced three novel models I developed for the study for the A/T/N system. All these three models are exploring the association between Aβ and tau deposition and the morphometric features of hippocampus. In this final chapter, I summarize the specific contributions and future works of these studies.

- Patch Analysis-based Surface Correntropy-induced Sparse coding and max-pooling (PASCS-MP)

In this study(Wu, Dong, Gui, et al., 2021; Wu et al., 2018), I explore the association between hippocampal structures and Aβ p/tau deposition using our hippocampal MMS, PASCS-MP method. Compared to traditional hippocampal shape measures, MMS have superior performance for predicting Aβ/tau depostion in the AD continuum. Compared to similar studies, this work achieves state-of-the-art performance when predicting Aβ positivity based on MRI biomarkers. And it can enrich the understanding of the relationship between hippocampal atrophy and AD pathology, and thus help in assessing disease burden, progression, and treatment effects.

In the future, we plan to apply this proposed framework to other AD ROIs and further improve the comprehensibility of the framework by visualizing morphometry features selected in the classification.

- Federated Morphometry Feature Selection

This work proposes a novel high-dimensional federated feature selection framework, FMFS, to study the Aβ/tau burden associated with abnormalities in hippocampal subregions on two datasets (Wu, Dong, Zhang, et al., 2021; Wu et al., 2020). Experimental results showed that FMFS encoded hippocampal features at different clinical stages that were associated with Aβ/tau burden. As the clinical symptoms worsen, these ROIs appear to be more focal. Our novel proposed framework achieved superior performance in efficiency compared to a similar feature selection method. To the best of our knowledge, this is the first feature selection model to study hippocampal morphometric changes with Aβ/tau burden across the AD spectrum. More importantly, this model can visualize brain structural abnormalities affected by AD proteinopathies. The features on ROIs could provide a means for screening individuals prior to more invasive Aβ/tau burden assessments that might determine their eligibility for interventional trials. Beyond brain MRI, our framework may also be applied to any other kinds of medical data for feature selection.

- Federated Integrating Transcriptomics, Genomics, and Imaging model

In this study (Wu et al., 2022), I proposed a novel federated Genotype-Expression-Image Data Integration model. Compared to similar studies, this work achieves state-of-the-art performance in discovering downstream effects of AD-related genes and SNPs. Besides, the model provides novel insights into the relationship among image biomarkers, genotypes, and gene expression and could discover novel drug targets for precision medicine.

In the future, we will further validate our model with more datasets and more advanced imaging biomarkers. Specifically, we will introduce blood-based biomarkers into our model when such data are available.

REFERENCES

Albert, F. W., & Kruglyak, L. (2015). The role of regulatory variation in complex traits and disease. In *Nature Reviews Genetics*. https://doi.org/10.1038/nrg3891

An, N., Fu, Y., Shi, J., Guo, H. N., Yang, Z. W., Li, Y. C., Li, S., Wang, Y., Yao, Z. J., & Hu, B. (2021). Synergistic effects of APOE and CLU may increase the risk of Alzheimer's Disease: Acceleration of atrophy in the volumes and shapes of the hippocampus and amygdala. *Journal of Alzheimer's Disease*, *80*(3). https://doi.org/10.3233/JAD-201162

Ansart, M., Epelbaum, S., Gagliardi, G., Colliot, O., Dormont, D., Dubois, B., Hampel, H., & Durrleman, S. (2020). Reduction of recruitment costs in preclinical AD trials: validation of automatic pre-screening algorithm for brain amyloidosis. *Statistical Methods in Medical Research*, *29*(1), 151–164. https://doi.org/10.1177/0962280218823036

Baek, M. S., Cho, H., Lee, H. S., Lee, J. H., Ryu, Y. H., & Lyoo, C. H. (2020). Effect of APOE ε4 genotype on amyloid-β and tau accumulation in Alzheimer's disease. *Alzheimer's Research and Therapy*, *12*(1). https://doi.org/10.1186/s13195-020-00710-6

Baker, B. T., Silva, R. F., Calhoun, V. D., Sarwate, A. D., & Plis, S. M. (2015). Large scale collaboration with autonomy: Decentralized data ICA. *IEEE International Workshop on Machine Learning for Signal Processing, MLSP*. https://doi.org/10.1109/MLSP.2015.7324344

Baker, S. L., Lockhart, S. N., Price, J. C., He, M., Huesman, R. H., Schonhaut, D., Faria, J., Rabinovici, G., & Jagust, W. J. (2017). Reference tissue-based kinetic evaluation of 18F-AV-1451 for tau imaging. *Journal of Nuclear Medicine*, *58*(2). https://doi.org/10.2967/jnumed.116.175273

Baker, S. L., Maass, A., & Jagust, W. J. (2017). Considerations and code for partial volume correcting [18F]-AV-1451 tau PET data. *Data in Brief*, *15*. https://doi.org/10.1016/j.dib.2017.10.024

Barbur, V. A., Montgomery, D. C., & Peck, E. A. (1994). Introduction to Linear Regression Analysis. *The Statistician*, *43*(2). https://doi.org/10.2307/2348362

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1). https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bertram, L., McQueen, M. B., Mullin, K., Blacker, D., & Tanzi, R. E. (2007). Systematic meta-analyses of Alzheimer disease genetic association studies: The AlzGene database. *Nature Genetics*. https://doi.org/10.1038/ng1934

Boureau, Y. L., Ponce, J., & Lecun, Y. (2010). A theoretical analysis of feature pooling in visual recognition. *ICML 2010 - Proceedings, 27th International Conference on Machine Learning*.

Boyd, S., Parikh, N., & Chu, E. (2011). *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc.

Bro-Nielsen, M., & Gramkow, C. (1996). Fast fluid registration of medical images. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/bfb0046964

Brookmeyer, R., Johnson, E., Ziegler-Graham, K., & Arrighi, H. M. (2007). Forecasting the global burden of Alzheimer's disease. *Alzheimer's and Dementia*, *3*(3). https://doi.org/10.1016/j.jalz.2007.04.381

Caselli, R. J., Dueck, A. C., Osborne, D., Sabbagh, M. N., Connor, D. J., Ahern, G. L., Baxter, L. C., Rapcsak, S. Z., Shi, J., Woodruff, B. K., Locke, D. E. C., Snyder, C. H., Alexander, G. E., Rademakers, R., & Reiman, E. M. (2009). Longitudinal Modeling of Age-Related Memory Decline and the APOE ε4 Effect . *New England Journal of Medicine*. https://doi.org/10.1056/nejmoa0809437

Chauhan, G., Adams, H. H. H., Bis, J. C., Weinstein, G., Yu, L., Töglhofer, A. M., Smith, A. V., van der Lee, S. J., Gottesman, R. F., Thomson, R., Wang, J., Yang, Q., Niessen, W. J., Lopez, O. L., Becker, J. T., Phan, T. G., Beare, R. J., Arfanakis, K., Fleischman, D., … Debette, S. (2015). Association of Alzheimer's disease GWAS loci with MRI markers of brain aging. *Neurobiology of Aging*. https://doi.org/10.1016/j.neurobiolaging.2014.12.028

Chow, G. C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica*. https://doi.org/10.2307/1910133

Chung, M. K., Dalton, K. M., & Davidson, R. J. (2008a). Tensor-based cortical surface morphometry via weighted spherical harmonic representation. *IEEE Transactions on Medical Imaging*. https://doi.org/10.1109/TMI.2008.918338

Chung, M. K., Dalton, K. M., & Davidson, R. J. (2008b). Tensor-based cortical surface morphometry via weighted spherical harmonic representation. *IEEE Transactions on Medical Imaging*. https://doi.org/10.1109/TMI.2008.918338

Colom, R., Stein, J. L., Rajagopalan, P., Martínez, K., Hermel, D., Wang, Y., Álvarez-Linera, J., Burgaleta, M., Quiroga, M. Á., Shih, P. C., & Thompson, P. M. (2013). Hippocampal structure and human cognition: Key role of spatial processing and evidence supporting the efficiency hypothesis in females. *Intelligence*. https://doi.org/10.1016/j.intell.2013.01.002

Cullen, N. C., Zetterberg, H., Insel, P. S., Olsson, B., Andreasson, U., Blennow, K., Hansson, O., & Mattsson-Carlgren, N. (2020). Comparing progression biomarkers in clinical trials of early Alzheimer's disease. *Annals of Clinical and Translational Neurology*. https://doi.org/10.1002/acn3.51158

D'Agostino, E., Maes, F., Vandermeulen, D., & Suetens, P. (2003). A viscous fluid model for multimodal non-rigid image registration using mutual information. *Medical Image Analysis*. https://doi.org/10.1016/S1361-8415(03)00039-2

Dahl, M. J., Mather, M., Werkle-Bergner, M., Kennedy, B. L., Guzman, S., Hurth, K., Miller, C. A., Qiao, Y., Shi, Y., Chui, H. C., & Ringman, J. M. (2021). Locus coeruleus integrity is related to tau burden and memory loss in autosomal-dominant Alzheimer's disease. *MedRxiv*.

Davatzikos, C. (1996). Spatial normalization of 3D brain images using deformable models. *Journal of Computer Assisted Tomography*. https://doi.org/10.1097/00004728-199607000-00031

Davatzikos, C., Resnick, S. M., Wu, X., Parmpi, P., & Clark, C. M. (2008). Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2008.03.050

Doğru-Abbasoğlu, S., Aykaç-Toker, G., Hanagasi, H. A., Gülrvit, H., Emre, M., & Uysal, M. (2007). The Arg 194 Trp polymorphism in DNA repair gene XRCC1 and the risk for sporadic late-onset Alzheimer's disease. *Neurological Sciences*. https://doi.org/10.1007/s10072-007-0744-x

Dong, Q., Zhang, J., Li, Q., Wang, J., Leporé, N., Thompson, P. M., Caselli, R. J., Ye, J., & Wang, Y. (2020). Integrating Convolutional Neural Networks and Multi-Task Dictionary Learning for Cognitive Decline Prediction with Longitudinal Images. *Journal of Alzheimer's Disease : JAD*. https://doi.org/10.3233/JAD-190973

Dong, Q., Zhang, W., Wu, J., Li, B., Schron, E. H., McMahon, T., Shi, J., Gutman, B. A., Chen, K., Baxter, L. C., Thompson, P. M., Reiman, E. M., Caselli, R. J., & Wang, Y. (2019). Applying surface-based hippocampal morphometry to study APOE-E4 allele dose effects in cognitively unimpaired subjects. *NeuroImage: Clinical*, *22*. https://doi.org/10.1016/j.nicl.2019.101744

Ezzati, A., Harvey, D. J., Habeck, C., Golzar, A., Qureshi, I. A., Zammit, A. R., Hyun, J., Truelove-Hill, M., Hall, C. B., Davatzikos, C., & Lipton, R. B. (2020). Predicting Amyloid-β Levels in Amnestic Mild Cognitive Impairment Using Machine Learning Techniques. *Journal of Alzheimer's Disease : JAD*, *73*(3). https://doi.org/10.3233/JAD-191038

Feng, Y., Huang, X., Shi, L., Yang, Y., & Suykens, J. A. K. (2015). Learning with the maximum correntropy criterion induced losses for regression. *Journal of Machine Learning Research*, *16*, 993–1034.

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. *NeuroImage*. https://doi.org/10.1006/nimg.1998.0396

Fleisher, A. S., Chen, K., Liu, X., Roontiva, A., Thiyyagura, P., Ayutyanont, N., Joshi, A. D., Clark, C. M., Mintun, M. A., Pontecorvo, M. J., Doraiswamy, P. M., Johnson, K. A., Skovronsky, D. M., & Reiman, E. M. (2011). Using positron emission tomography and florbetapir F 18 to image cortical amyloid in patients with mild cognitive impairment or dementia due to Alzheimer disease. *Archives of Neurology*. https://doi.org/10.1001/archneurol.2011.150

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. https://doi.org/10.1016/0022-3956(75)90026-6

Fu, W. J. (1998). Penalized regressions: The bridge versus the lasso? *Journal of Computational and Graphical Statistics*. https://doi.org/10.1080/10618600.1998.10474784

Fu, Y., Zhang, J., Li, Y., Shi, J., Zou, Y., Guo, H., Li, Y., Yao, Z., Wang, Y., & Hu, B. (2021). A novel pipeline leveraging surface-based features of small subcortical structures to classify individuals with autism spectrum disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *104*, 109989. https://doi.org/10.1016/j.pnpbp.2020.109989

Grasby, K. L., Jahanshad, N., Painter, J. N., Colodro-Conde, L., Bralten, J., Hibar, D. P., Lind, P. A., Pizzagalli, F., Ching, C. R. K., McMahon, M. A. B., Shatokhina, N., Zsembik, L. C. P., Thomopoulos, S. I., Zhu, A. H., Strike, L. T., Agartz, I., Alhusaini, S., Almeida, M. A. A., Alnæs, D., … Medland, S. E. (2020). The genetic architecture of the human cerebral cortex. *Science*, *367*(6484). https://doi.org/10.1126/science.aay6690

Gui, J., Sun, Z., Ji, S., Tao, D., & Tan, T. (2017). Feature selection based on structured sparsity: a comprehensive study. *IEEE Transactions on Neural Networks and Learning Systems*. https://doi.org/10.1109/TNNLS.2016.2551724

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, *46*(1–3). https://doi.org/10.1023/A:1012487302797

Han, X., Xu, C., & Prince, J. L. (2003). A topology preserving level set method for geometric deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/TPAMI.2003.1201824

Hanko, V., Apple, A. C., Alpert, K. I., Warren, K. N., Schneider, J. A., Arfanakis, K., Bennett, D. A., & Wang, L. (2019). In vivo hippocampal subfield shape related to TDP-43, amyloid beta, and tau pathologies. *Neurobiology of Aging*, *74*, 171–181. https://doi.org/10.1016/j.neurobiolaging.2018.10.013

Hardy, J., & Selkoe, D. J. (2002). The amyloid hypothesis of Alzheimer’s disease: progress and problems on the road to therapeutics. *Science*, *297*(5580), 353–356. https://doi.org/10.1126/science.1072994

Hoerl, A. E., & Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, *12*(1). https://doi.org/10.1080/00401706.1970.10488634

Hoppe, H. (1996). Progressive meshes. *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1996*. https://doi.org/10.1145/237170.237216

Huang, D., Yi, X., Zhang, S., Zheng, Z., Wang, P., Xuan, C., Sham, P. C., Wang, J., & Li, M. J. (2018). GWAS4D: Multidimensional analysis of context-specific regulatory

variant for human complex diseases and traits. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gky407

Insel, P. S., Ossenkoppele, R., Gessert, D., Jagust, W., Landau, S., Hansson, O., Weiner, M. W., & Mattsson, N. (2017). Time to amyloid positivity and preclinical changes in brain metabolism, atrophy, and cognition: Evidence for emerging amyloid pathology in alzheimer's disease. *Frontiers in Neuroscience*, *11*(MAY), 1–9. https://doi.org/10.3389/fnins.2017.00281

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Feldman, H. H., Frisoni, G. B., Hampel, H., Jagust, W. J., Johnson, K. A., Knopman, D. S., Petersen, R. C., Scheltens, P., Sperling, R. A., & Dubois, B. (2016). A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology*, *87*(5), 539–547. https://doi.org/10.1212/WNL.0000000000002923

Kaissis, G., Ziller, A., Passerat-Palmbach, J., Ryffel, T., Usynin, D., Trask, A., Lima, I., Mancuso, J., Jungmann, F., Steinborn, M. M., Saleh, A., Makowski, M., Rueckert, D., & Braren, R. (2021). End-to-end privacy preserving deep learning on multi-institutional medical imaging. *Nature Machine Intelligence*, *3*(6). https://doi.org/10.1038/s42256-021-00337-8

Klunk, W. E., Koeppe, R. A., Price, J. C., Benzinger, T. L., Devous, M. D., Jagust, W. J., Johnson, K. A., Mathis, C. A., Minhas, D., Pontecorvo, M. J., Rowe, C. C., Skovronsky, D. M., & Mintun, M. A. (2015a). The Centiloid project: Standardizing quantitative amyloid plaque estimation by PET. *Alzheimer's and Dementia*, *11*(1), 1-15.e4. https://doi.org/10.1016/j.jalz.2014.07.003

Klunk, W. E., Koeppe, R. A., Price, J. C., Benzinger, T. L., Devous, M. D., Jagust, W. J., Johnson, K. A., Mathis, C. A., Minhas, D., Pontecorvo, M. J., Rowe, C. C., Skovronsky, D. M., & Mintun, M. A. (2015b). The Centiloid project: Standardizing quantitative amyloid plaque estimation by PET. *Alzheimer's and Dementia*, *11*(1), 1-15.e4. https://doi.org/10.1016/j.jalz.2014.07.003

La Joie, R., Visani, A. V, Baker, S. L., Brown, J. A., Bourakova, V., Cha, J., Chaudhary, K., Edwards, L., Iaccarino, L., Janabi, M., Lesman-Segev, O. H., Miller, Z. A., Perry, D. C., O'Neil, J. P., Pham, J., Rojas, J. C., Rosen, H. J., Seeley, W. W., Tsai, R. M., … Rabinovici, G. D. (2020). Prospective longitudinal atrophy in Alzheimer's disease correlates with the intensity and topography of baseline tau-PET. *Science Translational Medicine*, *12*(524), eaau5732. https://doi.org/10.1126/scitranslmed.aau5732

Lee, B., Wang, J., & Shen, L. (2021). Identifying precision AD biomarkers with varying prognosis effects in genetics driven subpopulations. *AAIC'21: Alzheimer's Association Int. Conf. on Alzheimer's Disease, Denver, USA and Online,, July 26-30, 2021*.

Lee, H., Battle, A., Raina, R., & Ng, A. Y. (2007). Efficient sparse coding algorithms. *Advances in Neural Information Processing Systems*. https://doi.org/10.7551/mitpress/7503.003.0105

Lee, Y. K., Hou, S. W., Lee, C. C., Hsu, C. Y., Huang, Y. S., & Su, Y. C. (2013). Increased Risk of Dementia in Patients with Mild Traumatic Brain Injury: A Nationwide Cohort Study. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0062422

Leow, A., Huang, S. C., Geng, A., Becker, J., Davis, S., Toga, A., & Thompson, P. (2005). Inverse consistent mapping in 3D deformable image registration: Its construction and statistical properties. *Lecture Notes in Computer Science*. https://doi.org/10.1007/11505730_41

Li, B., Shi, J., Gutman, B. A., Baxter, L. C., Thompson, P. M., Caselli, R. J., Wang, Y., & Neuroimaging Initiative, D. (2016). *Influence of APOE Genotype on Hippocampal Atrophy over Time-An N=1925 Surface-Based ADNI Study*. https://doi.org/10.1371/journal.pone.0152901

Li, J. Q., Wang, H. F., Zhu, X. C., Sun, F. R., Tan, M. S., Tan, C. C., Jiang, T., Tan, L., & Yu, J. T. (2017). GWAS-Linked Loci and Neuroimaging Measures in Alzheimer's Disease. *Molecular Neurobiology*. https://doi.org/10.1007/s12035-015-9669-1

Li, Q., Yang, T., Zhan, L., Hibar, D. P., Jahanshad, N., Wang, Y., Ye, J., Thompson, P. M., & Wang, J. (2016). Large-scale collaborative imaging genetics studies of risk genetic factors for Alzheimer's disease across multiple institutions. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-319-46720-7_39

Lin, B., Li, Q., Sun, Q., Lai, M.-J., Davidson, I., Fan, W., & Ye, J. (2014). *Stochastic Coordinate Coding and Its Application for Drosophila Gene Expression Pattern Annotation*.

Liu, K., Yao, X., Yan, J., Chasioti, D., Risacher, S., Nho, K., Saykin, A., & Shen, L. (2017). Transcriptome-guided imaging genetic analysis via a novel sparse CCA algorithm. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10551 LNCS*. https://doi.org/10.1007/978-3-319-67675-3_20

Loop, C. (1987). Smooth Subdivision Surfaces Based on Triangles. In *Acm Siggraph*.

Lorensen, W. E., & Cline, H. E. (1987). Marching cubes: A high resolution 3D surface construction algorithm. *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1987*. https://doi.org/10.1145/37401.37422

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. https://doi.org/10.1186/s13059-014-0550-8

Luders, E., Thompson, P. M., Kurth, F., Hong, J. Y., Phillips, O. R., Wang, Y., Gutman, B. A., Chou, Y. Y., Narr, K. L., & Toga, A. W. (2013). Global and regional alterations of hippocampal anatomy in long-term meditation practitioners. *Human Brain Mapping*. https://doi.org/10.1002/hbm.22153

Maass, A., Landau, S., Horng, A., Lockhart, S. N., Rabinovici, G. D., Jagust, W. J., Baker, S. L., & La Joie, R. (2017). Comparison of multiple tau-PET measures as biomarkers in aging and Alzheimer's disease. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2017.05.058

Mairal, J., Bach, F., Ponce, J., & Sapiro, G. (2009). Online dictionary learning for sparse coding. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/1553374.1553463

Marcus, D. S., Fotenos, A. F., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2010a). Open Access Series of Imaging Studies: Longitudinal MRI Data in Nondemented and Demented Older Adults. *Journal of Cognitive Neuroscience*, *22*(12), 2677–2684. https://doi.org/10.1162/jocn.2009.21407

Marcus, D. S., Fotenos, A. F., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2010b). Open Access Series of Imaging Studies: Longitudinal MRI Data in Nondemented and Demented Older Adults. *Journal of Cognitive Neuroscience*, *22*(12), 2677–2684. https://doi.org/10.1162/jocn.2009.21407

Marinescu, R. v., Oxtoby, N. P., Young, A. L., Bron, E. E., Toga, A. W., Weiner, M. W., Barkhof, F., Fox, N. C., Eshaghi, A., Toni, T., Salaterski, M., Lunina, V., Ansart, M., Durrleman, S., Lu, P., Iddi, S., Li, D., Thompson, W. K., Donohue, M. C., … Alexander, D. C. (2020). The Alzheimer's disease prediction of longitudinal evolution (tadpole) challenge: Results after 1 year follow-up. In *arXiv*.

Meinshausen, N., & Bühlmann, P. (2010). Stability selection. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*. https://doi.org/10.1111/j.1467-9868.2010.00740.x

Monje, M., Thomason, M. E., Rigolo, L., Wang, Y., Waber, D. P., Sallan, S. E., & Golby, A. J. (2013). Functional and structural differences in the hippocampus associated with memory deficits in adult survivors of acute lymphoblastic leukemia. *Pediatric Blood and Cancer*. https://doi.org/10.1002/pbc.24263

Moore, D. F. (2008). A Review of: " Applied Survival Analysis: Regression Modeling of Time-to-Event Data , 2nd ed., by D. W. Hosmer, S. Lemeshow, and S. May" . *Journal of Biopharmaceutical Statistics*, *18*(6). https://doi.org/10.1080/10543400802369095

Navitsky, M., Joshi, A. D., Kennedy, I., Klunk, W. E., Rowe, C. C., Wong, D. F., Pontecorvo, M. J., Mintun, M. A., & Devous, M. D. (2018). Standardization of amyloid quantitation with florbetapir standardized uptake value ratios to the Centiloid scale. *Alzheimer's and Dementia*. https://doi.org/10.1016/j.jalz.2018.06.1353

Nikolova, M., & Ng, M. K. (2006). Analysis of half-quadratic minimization methods for signal and image recovery. *SIAM Journal on Scientific Computing*. https://doi.org/10.1137/030600862

Nutini, J. (2017). *Let ' s Make Block Coordinate Descent Go Fast : arXiv : 1712 . 08859v1 [math . OC] 23 Dec 2017*. 1–49.

Olshausen, B. A., & Field, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision Research*. https://doi.org/10.1016/S0042-6989(97)00169-7

Paquette, N., Shi, J., Wang, Y., Lao, Y., Ceschin, R., Nelson, M. D., Panigrahy, A., & Lepore, N. (2017). Ventricular shape and relative position abnormalities in preterm neonates. *NeuroImage. Clinical*, *15*, 483–493. https://doi.org/10.1016/j.nicl.2017.05.025

Patenaude, B., Smith, S. M., Kennedy, D. N., & Jenkinson, M. (2011). A Bayesian model of shape and appearance for subcortical brain segmentation. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2011.02.046

Petrone, P. M., Casamitjana, A., Falcon, C., Artigues, M., Operto, G., Cacciaglia, R., Molinuevo, J. L., Vilaplana, V., & Gispert, J. D. (2019). Prediction of amyloid pathology in cognitively unimpaired individuals using voxel-wise analysis of longitudinal structural brain MRI. *Alzheimer's Research and Therapy*, *11*(1). https://doi.org/10.1186/s13195-019-0526-8

Pizer, S. M., Fritsch, D. S., Yushkevich, P. A., Johnson, V. E., & Chaney, E. L. (1999). Segmentation, registration, and measurement of shape variation via image object shape. *IEEE Transactions on Medical Imaging*. https://doi.org/10.1109/42.811263

Plis, S. M., Sarwate, A. D., Wood, D., Dieringer, C., Landis, D., Reed, C., Panta, S. R., Turner, J. A., Shoemaker, J. M., Carter, K. W., Thompson, P., Hutchison, K., & Calhoun, V. D. (2016). COINSTAC: A privacy enabled model and prototype for leveraging and processing decentralized brain imaging data. *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2016.00365

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*. https://doi.org/10.1086/519795

Qin, Z., Scheinberg, K., & Goldfarb, D. (2013). Efficient block-coordinate descent algorithms for the Group Lasso. *Mathematical Programming Computation*. https://doi.org/10.1007/s12532-013-0051-x

Rawlings, J. O., Pantula, S. G., & Dickey, D. A. (1998). Applied regression analysis: A research tool, (2nd ed.). New York, NY: In *Springer-Verlag, Inc.*

Remedios, S. W., Butman, J. A., Landman, B. A., & Pham, D. L. (2020). Federated Gradient Averaging for Multi-Site Training with Momentum-Based Optimizers. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12444 LNCS*. https://doi.org/10.1007/978-3-030-60548-3_17

Rey, D., Subsol, G., Delingette, H., & Ayache, N. (2002). Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis. *Medical Image Analysis*. https://doi.org/10.1016/S1361-8415(02)00056-7

Ritchie, J., Pantazatos, S. P., & French, L. (2018). Transcriptomic characterization of MRI contrast with focus on the T1-w/T2-w ratio in the cerebral cortex. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2018.03.027

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*. https://doi.org/10.1186/1471-2105-12-77

Rowe, C. C., Doré, V., Jones, G., Baxendale, D., Mulligan, R. S., Bullich, S., Stephens, A. W., De Santi, S., Masters, C. L., Dinkelborg, L., & Villemagne, V. L. (2017). 18F-Florbetaben PET beta-amyloid binding expressed in Centiloids. *European Journal of Nuclear Medicine and Molecular Imaging*, *44*(12), 2053–2059. https://doi.org/10.1007/s00259-017-3749-6

Sanchez, J. S., Becker, J. A., Jacobs, H. I. L., Hanseeuw, B. J., Jiang, S., Schultz, A. P., Properzi, M. J., Katz, S. R., Beiser, A., Satizabal, C. L., O'Donnell, A., DeCarli, C., Killiany, R., El Fakhri, G., Normandin, M. D., Gómez-Isla, T., Quiroz, Y. T., Rentz, D. M., Sperling, R. A., … Johnson, K. A. (2021). The cortical origin and initial spread of medial temporal tauopathy in Alzheimer's disease assessed with positron emission tomography. *Science Translational Medicine*, *13*(577). https://doi.org/10.1126/scitranslmed.abc0655

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology*. https://doi.org/10.1038/nbt.3192

Schöll, M., Lockhart, S. N., Schonhaut, D. R., O'Neil, J. P., Janabi, M., Ossenkoppele, R., Baker, S. L., Vogel, J. W., Faria, J., Schwimmer, H. D., Rabinovici, G. D., & Jagust, W. J. (2016). PET Imaging of Tau Deposition in the Aging Human Brain. *Neuron*, *89*(5). https://doi.org/10.1016/j.neuron.2016.01.028

Shabalin, A. A. (2012). Matrix eQTL: Ultra fast eQTL analysis via large matrix operations. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/bts163

Shi, J., Stonnington, C. M., Thompson, P. M., Chen, K., Gutman, B., Reschke, C., Baxter, L. C., Reiman, E. M., Caselli, R. J., & Wang, Y. (2015). Studying ventricular abnormalities in mild cognitive impairment with hyperbolic Ricci flow and tensor-based morphometry. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2014.09.062

Shi, J., Thompson, P. M., Gutman, B., & Wang, Y. (2013). Surface fluid registration of conformal representation: Application to detect disease burden and genetic influence on hippocampus. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2013.04.018

Shi, J., Thompson, P. M., & Wang, Y. (2011). Human Brain Mapping with Conformal Geometry and Multivariate Tensor-Based Morphometry. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 7012 LNCS* (pp. 126–134). https://doi.org/10.1007/978-3-642-24446-9_16

Shi, J., Wang, Y., Ceschin, R., An, X., Lao, Y., Vanderbilt, D., Nelson, M. D., Thompson, P. M., Panigrahy, A., & Leporé, N. (2013). A Multivariate Surface-Based Analysis of the Putamen in Premature Newborns: Regional Differences within the Ventral Striatum. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0066736

Silva, S., Altmann, A., Gutman, B., & Lorenzi, M. (2020). Fed-BioMed: A General Open-Source Frontend Framework for Federated Learning in Healthcare. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12444 LNCS*. https://doi.org/10.1007/978-3-030-60548-3_20

Sperling, R. A., Jack, C. R., Black, S. E., Frosch, M. P., Greenberg, S. M., Hyman, B. T., Scheltens, P., Carrillo, M. C., Thies, W., Bednar, M. M., Black, R. S., Brashear, H. R., Grundman, M., Siemers, E. R., Feldman, H. H., & Schindler, R. J. (2011). Amyloid-related imaging abnormalities in amyloid-modifying therapeutic trials: Recommendations from the Alzheimer's Association Research Roundtable Workgroup. *Alzheimer's & Dementia*, *7*(4), 367–385. https://doi.org/10.1016/j.jalz.2011.05.2351

Stein, J. L., Hua, X., Lee, S., Ho, A. J., Leow, A. D., Toga, A. W., Saykin, A. J., Shen, L., Foroud, T., Pankratz, N., Huentelman, M. J., Craig, D. W., Gerber, J. D., Allen, A. N., Corneveaux, J. J., DeChairo, B. M., Potkin, S. G., Weiner, M. W., & Thompson, P. M. (2010). Voxelwise genome-wide association study (vGWAS). *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2010.02.032

Storey, J. D. (2003). The positive false discovery rate: A Bayesian interpretation and the q-value. *Annals of Statistics*. https://doi.org/10.1214/aos/1074290335

Stripelis, D., Ambite, J. L., Lam, P., & Thompson, P. (2021). Scaling neuroscience research using federated learning. *Proceedings - International Symposium on Biomedical Imaging, 2021-April*. https://doi.org/10.1109/ISBI48211.2021.9433925

Styner, M., Lieberman, J. A., McClure, R. K., Weinberger, D. R., Jones, D. W., & Gerig, G. (2005). Morphometric analysis of lateral ventricles in schizophrenia and healthy controls regarding genetic and disease-specific factors. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.0501117102

Su, Y., Blazey, T. M., Snyder, A. Z., Raichle, M. E., Marcus, D. S., Ances, B. M., Bateman, R. J., Cairns, N. J., Aldea, P., Cash, L., Christensen, J. J., Friedrichsen, K., Hornbeck, R. C., Farrar, A. M., Owen, C. J., Mayeux, R., Brickman, A. M., Klunk, W., Price, J. C., … Benzinger, T. L. S. (2015). Partial volume correction in quantitative amyloid imaging. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2014.11.058

Su, Y., Flores, S., Wang, G., Hornbeck, R. C., Speidel, B., Joseph-Mathurin, N., Vlassenko, A. G., Gordon, B. A., Koeppe, R. A., Klunk, W. E., Jack, C. R., Farlow, M. R., Salloway, S., Snider, B. J., Berman, S. B., Roberson, E. D., Brosch, J., Jimenez-Velazques, I., van Dyck, C. H., … Benzinger, T. L. S. (2019). Comparison of Pittsburgh compound B and florbetapir in cross-sectional and longitudinal studies. *Alzheimer's and*

Dementia: Diagnosis, Assessment and Disease Monitoring. https://doi.org/10.1016/j.dadm.2018.12.008

Sun, D., van Erp, T. G. M., Thompson, P. M., Bearden, C. E., Daley, M., Kushan, L., Hardt, M. E., Nuechterlein, K. H., Toga, A. W., & Cannon, T. D. (2009). Elucidating a Magnetic Resonance Imaging-Based Neuroanatomic Biomarker for Psychosis: Classification Analysis Using Probabilistic Brain Atlas and Machine Learning Algorithms. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2009.07.019

Sun, W., Tang, Y., Qiao, Y., Ge, X., Mather, M., Ringman, J. M., & Shi, Y. (2020). A probabilistic atlas of locus coeruleus pathways to transentorhinal cortex for connectome imaging in Alzheimer's disease. *NeuroImage*, *223*. https://doi.org/10.1016/j.neuroimage.2020.117301

Ten Kate, M., Redolfi, A., Peira, E., Bos, I., Vos, S. J., Vandenberghe, R., Gabel, S., Schaeverbeke, J., Scheltens, P., Blin, O., Richardson, J. C., Bordet, R., Wallin, A., Eckerstrom, C., Molinuevo, J. L., Engelborghs, S., Van Broeckhoven, C., Martinez-Lage, P., Popp, J., … Barkhof, F. (2018). MRI predictors of amyloid pathology: Results from the EMIF-AD Multimodal Biomarker Discovery study. *Alzheimer's Research and Therapy*, *10*(1). https://doi.org/10.1186/s13195-018-0428-1

Thompson, P. M., Gledd, J. N., Woods, R. P., MacDonald, D., Evans, A. C., & Toga, A. W. (2000a). Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature*. https://doi.org/10.1038/35004593

Thompson, P. M., Gledd, J. N., Woods, R. P., MacDonald, D., Evans, A. C., & Toga, A. W. (2000b). Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature*. https://doi.org/10.1038/35004593

Thompson, P. M., Hayashi, K. M., De Zubicaray, G. I., Janke, A. L., Rose, S. E., Semple, J., Hong, M. S., Herman, D. H., Gravano, D., Doddrell, D. M., & Toga, A. W. (2004). Mapping hippocampal and ventricular change in Alzheimer disease. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2004.03.040

Tosun, D., Chen, Y.-F., Yu, P., Sundell, K. L., Suhy, J., Siemers, E., Schwarz, A. J., & Weiner, M. W. (2016). Amyloid status imputed from a multimodal classifier including structural MRI distinguishes progressors from nonprogressors in a mild Alzheimer's disease clinical trial cohort. *Alzheimer's & Dementia*, *12*(9), 977–986. https://doi.org/10.1016/j.jalz.2016.03.009

Tosun, D., Joshi, S., & Weiner, M. W. (2013). Neuroimaging predictors of brain amyloidosis in mild cognitive impairment. *Annals of Neurology*, *74*(2). https://doi.org/10.1002/ana.23921

Tosun, D., Joshi, S., & Weiner, M. W. (2014). Multimodal MRI-based imputation of the A β + in early mild cognitive impairment. *Annals of Clinical and Translational Neurology*, *1*(3), 160–170. https://doi.org/10.1002/acn3.40

Tosun, D., Veitch, D., Aisen, P., Jack, C. R., Jagust, W. J., Petersen, R. C., Saykin, A. J., Bollinger, J., Ovod, V., Mawuenyega, K. G., Bateman, R. J., Shaw, L. M., Trojanowski,

J. Q., Blennow, K., Zetterberg, H., & Weiner, M. W. (2021). Detection of β-amyloid positivity in Alzheimer's Disease Neuroimaging Initiative participants with demographics, cognition, MRI and plasma biomarkers. *Brain Communications*, *3*(2). https://doi.org/10.1093/braincomms/fcab008

Vanwinckelen, G., & Blockeel, H. (2012). On estimating model accuracy with repeated cross-validation. *21st Belgian-Dutch Conference on Machine Learning*.

Wang, G., Dong, Q., Wu, J., Su, Y., Chen, K., Su, Q., Zhang, X., Hao, J., Yao, T., Liu, L., Zhang, C., Caselli, R. J., Reiman, E. M., & Wang, Y. (2021). Developing univariate neurodegeneration biomarkers with low-rank and sparse subspace decomposition. *Medical Image Analysis*. https://doi.org/10.1016/j.media.2020.101877

Wang, J., Wonka, P., & Ye, J. (2015). Lasso Screening Rules via Dual Polytope Projection. *Journal of Machine Learning Research*.

Wang, Y., Chan, T. F., Toga, A. W., & Thompson, P. M. (2009). Multivariate tensor-based brain anatomical surface morphometry via holomorphic one-forms. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-642-04268-3_42

Wang, Y., Lui, L. M., Gu, X., Hayashi, K. M., Chan, T. F., Toga, A. W., Thompson, P. M., & Yau, S. T. (2007). Brain surface conformal parameterization using riemann surface structure. *IEEE Transactions on Medical Imaging*. https://doi.org/10.1109/TMI.2007.895464

Wang, Y., Shi, J., Yin, X., Gu, X., Chan, T. F., Yau, S. T., Toga, A. W., & Thompson, P. M. (2012). Brain surface conformal parameterization with the ricci flow. *IEEE Transactions on Medical Imaging*. https://doi.org/10.1109/TMI.2011.2168233

Wang, Y., Song, Y., Rajagopalan, P., An, T., Liu, K., Chou, Y. Y., Gutman, B., Toga, A. W., & Thompson, P. M. (2011). Surface-based TBM boosts power to detect disease effects on the brain: An N=804 ADNI study. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2011.03.040

Wang, Y., Yuan, L., Shi, J., Greve, A., Ye, J., Toga, A. W., Reiss, A. L., & Thompson, P. M. (2013). Applying tensor-based morphometry to parametric surfaces can improve MRI-based disease diagnosis. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2013.02.011

Wang, Y., Zhang, J., Gutman, B., Chan, T. F., Becker, J. T., Aizenstein, H. J., Lopez, O. L., Tamburo, R. J., Toga, A. W., & Thompson, P. M. (2010). Multivariate tensor-based morphometry on surfaces: Application to mapping ventricular abnormalities in HIV/AIDS. *NeuroImage*, *49*(3), 2141–2157. https://doi.org/10.1016/j.neuroimage.2009.10.086

Warnat-Herresthal, S., Schultze, H., Shastry, K. L., Manamohan, S., Mukherjee, S., Garg, V., Sarveswara, R., Händler, K., Pickkers, P., Aziz, N. A., Ktena, S., Tran, F., Bitzer, M., Ossowski, S., Casadei, N., Herr, C., Petersheim, D., Behrends, U., Kern, F., … Schultze,

J. L. (2021). Swarm Learning for decentralized and confidential clinical machine learning. *Nature*, *594*(7862). https://doi.org/10.1038/s41586-021-03583-3

Wingo, T. S., Cutler, D. J., Wingo, A. P., Le, N. A., Rabinovici, G. D., Miller, B. L., Lah, J. J., & Levey, A. I. (2019). Association of Early-Onset Alzheimer Disease with Elevated Low-Density Lipoprotein Cholesterol Levels and Rare Genetic Coding Variants of APOB. *JAMA Neurology*. https://doi.org/10.1001/jamaneurol.2019.0648

Woods, R. P. (2003). Characterizing volume and surface deformations in an atlas framework: Theory, applications, and implementation. *NeuroImage*. https://doi.org/10.1016/S1053-8119(03)00019-3

Wu, J., Chen, Y., Wang, P., Caselli, R. J., Thompson, P. M., Wang, J., & Wang, Y. (2022). Integrating Transcriptomics, Genomics, and Imaging in Alzheimer's Disease: A Federated Model. *Frontiers in Radiology*, *1*. https://doi.org/10.3389/fradi.2021.777030

Wu, J., Dong, Q., Gui, J., Zhang, J., Su, Y., Chen, K., Thompson, P. M., Caselli, R. J., Reiman, E. M., Ye, J., & Wang, Y. (2021). Predicting Brain Amyloid Using Multivariate Morphometry Statistics, Sparse Coding, and Correntropy: Validation in 1,101 Individuals From the ADNI and OASIS Databases. *Frontiers in Neuroscience*, *15*, 985.

Wu, J., Dong, Q., Zhang, J., Su, Y., Wu, T., Caselli, R. J., Reiman, E. M., Ye, J., Lepore, N., Chen, K., Thompson, P. M., & Wang, Y. (2021). Federated Morphometry Feature Selection for Hippocampal Morphometry Associated Beta-Amyloid and Tau Pathology. *Frontiers in Neuroscience*, *15*, 1585. https://www.frontiersin.org/article/10.3389/fnins.2021.762458

Wu, J., Zhang, J., Li, Q., Su, Y., Chen, K., Reiman, E., Wang, J., Lepore, N., Ye, J., Thompson, P., & Wang, Y. (2020). Patch-based surface morphometry feature selection with federated group lasso regression. *The 16th International Symposium on Medical Information Processing and Analysis*, *11583*.

Wu, J., Zhang, J., Shi, J., Chen, K., Caselli, R. J., Reiman, E. M., & Wang, Y. (2018). Hippocampus morphometry study on pathology-confirmed Alzheimer's disease patients with surface multivariate morphometry statistics. *Proceedings - International Symposium on Biomedical Imaging*. https://doi.org/10.1109/ISBI.2018.8363870

Yao, Z., Fu, Y., Wu, J., Zhang, W., Yu, Y., Zhang, Z., Wu, X., Wang, Y., & Hu, B. (2018). Morphological changes in subregions of hippocampus and amygdala in major depressive disorder patients. *Brain Imaging and Behavior*. https://doi.org/10.1007/s11682-018-0003-1

Yeganeh, Y., Farshad, A., Navab, N., & Albarqouni, S. (2020). Inverse Distance Aggregation for Federated Learning with Non-IID Data. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12444 LNCS*. https://doi.org/10.1007/978-3-030-60548-3_15

Yip, S. H., Sham, P. C., & Wang, J. (2018). Evaluation of tools for highly variable gene discovery from single-cell RNA-seq data. *Briefings in Bioinformatics*. https://doi.org/10.1093/bib/bby011

Yip, S. H., Wang, P., Kocher, J. P. A., Sham, P. C., & Wang, J. (2017). Linnorm: improved statistical analysis for single cell RNA-seq expression data. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkx828

Yuan, M., & Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*. https://doi.org/10.1111/j.1467-9868.2005.00532.x

Zhang, A., Zhao, Q., Xu, D., & Jiang, S. (2018). Brain APOE expression quantitative trait loci-based association study identified one susceptibility locus for Alzheimer's disease by interacting with APOE ϵ4. *Scientific Reports*. https://doi.org/10.1038/s41598-018-26398-1

Zhang, J., Fan, Y., Li, Q., Thompson, P. M., Ye, J., & Wang, Y. (2017). Empowering cortical thickness measures in clinical diagnosis of Alzheimer's disease with spherical sparse coding. *Proceedings - International Symposium on Biomedical Imaging*. https://doi.org/10.1109/ISBI.2017.7950557

Zhang, J., Li, Q., Caselli, R. J., Thompson, P. M., Ye, J., & Wang, Y. (2017). *Multi-source Multi-target Dictionary Learning for Prediction of Cognitive Decline* (pp. 184–197). Springer, Cham. https://doi.org/10.1007/978-3-319-59050-9_15

Zhang, J., Shi, J., Stonnington, C., Li, Q., Gutman, B. A., Chen, K., Reiman, E. M., Caselli, R., Thompson, P. M., Ye, J., & Wang, Y. (2016). Hyperbolic space sparse coding with its application on prediction of Alzheimer's disease in mild cognitive impairment. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-319-46720-7_38

Zhang, J., Stonnington, C., Li, Q., Shi, J., Bauer, R. J., Gutman, B. A., Chen, K., Reiman, E. M., Thompson, P. M., Ye, J., & Wang, Y. (2016). Applying sparse coding to surface multivariate tensor-based morphometry to predict future cognitive decline. *Proceedings - International Symposium on Biomedical Imaging*. https://doi.org/10.1109/ISBI.2016.7493350

Zhang, J., Wu, J., Li, Q., Caselli, R. J., Thompson, P. M., Ye, J., & Wang, Y. (2021). Multi-Resemblance Multi-Target Low-Rank Coding for Prediction of Cognitive Decline with Longitudinal Brain Images. *IEEE Transactions on Medical Imaging*. https://doi.org/10.1109/TMI.2021.3070780

Zhang, L., Mak, E., Reilhac, A., Shim, H. Y., Ng, K. K., Ong, M. Q. W., Ji, F., Chong, E. J. Y., Xu, X., Wong, Z. X., Stephenson, M. C., Venketasubramanian, N., Tan, B. Y., Zhou, J. H., & Brien, J. T. O. (2020). *Longitudinal trajectory of Amyloid-related hippocampal subfield atrophy in nondemented elderly*. *January*, 1–11. https://doi.org/10.1002/hbm.24928

Zhang, Q., & Li, B. (2015). Dictionary learning in visual computing. *Synthesis Lectures on Image, Video, and Multimedia Processing*, *18*. https://doi.org/10.2200/S00640ED1V01Y201504IVM018