



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

의학박사 학위논문

**Distinct subtypes of spatial brain
metabolism patterns in Alzheimer's
disease identified by deep learning-
based FDG PET clusters**

딥러닝 기반 군집화 방법을 이용하여 FDG
PET에서 알츠하이머병의 공간적 뇌 대사
패턴의 특징적 아형 분류

2022년 2월

서울대학교 융합과학기술대학원

분자의학 및 바이오제약학과

유 현 지

Distinct subtypes of spatial brain metabolism patterns in Alzheimer's disease identified by deep learning-based FDG PET clusters

지도 교수 이 동 수

이 논문을 의학박사 학위논문으로 제출함
2021년 11월

서울대학교 융합과학기술대학원
분자의학 및 바이오제약학과
유 현 지

유현지의 의학박사 학위논문을 인준함
2022년 01월

위 원 장 윤 형 진

부위원장 이 동 수

위 원 이 동 영

위 원 이 재 성

위 원 진 경 환

ABSTRACT

Distinct subtypes of spatial brain metabolism patterns in Alzheimer's disease identified by deep learning-based FDG PET clusters

Hyun Gee Ryoo

Department of Molecular Medicine and Biopharmaceutical Science,
The Graduate School of Convergence Science and Technology,
Seoul National University

Alzheimer's disease (AD) presents a broad spectrum of clinicopathologic profiles, despite common pathologic features including amyloid and tau deposition. Here, we aimed to identify AD subtypes using deep learning-based clustering on FDG PET images to understand distinct spatial patterns of neurodegeneration. We also aimed to investigate clinicopathologic features of subtypes defined by spatial brain metabolism patterns.

A total of 3620 FDG brain PET images with AD, mild cognitive impairment (MCI), and cognitively normal controls (CN) at baseline and

follow-up visits were obtained from Alzheimer's Disease Neuroimaging Initiative (ADNI) database. In order to identify representations of brain metabolism patterns different from disease progression in AD, a conditional variational autoencoder (cVAE) was used, followed by clustering using the encoded representations. FDG brain PET images with AD (n=838) and Clinical Demetria Rating Scale Sum of Boxes (CDR-SB) scores were used as inputs of cVAE model and the k-means algorithm was applied for the clustering. The trained deep learning model was also transferred to FDG brain PET image with MCI (n=1761) to identify differential trajectories and prognosis of subtypes. Statistical parametric maps were generated to visualize spatial patterns of clusters, and clinical and biological characteristics were compared among the clusters. The conversion rate from MCI to AD was also compared among the subtypes.

Four distinct subtypes were identified by deep learning-based FDG PET clusters: (i) S1 (angular), showing prominent hypometabolism in the angular gyrus with a diffuse cortical hypometabolism pattern; frequent in males; more amyloid; less tau; more hippocampal atrophy; cognitive decline in the earlier stage. (ii) S2 (occipital), showing prominent hypometabolism in the occipital cortex with a posterior-predominant hypometabolism pattern; younger age; more tau; less hippocampal atrophy; lower executive and visuospatial scores; faster

conversion from MCI to AD. (iii) S3 (orbitofrontal), showing prominent hypometabolism in the orbitofrontal cortex with an anterior-predominant hypometabolism pattern; older age; less amyloid; more hippocampal atrophy; higher executive and visuospatial scores. (iv) S4 (minimal), showing minimal hypometabolism; frequent in females; less amyloid; more tau; less hippocampal atrophy; higher cognitive scores.

In conclusion, we could identify distinct subtypes in AD with different brain pathologies and clinical profiles. Also, our deep learning model was successfully transferred to MCI to predict the prognosis of subtypes for conversion from MCI to AD. Our results suggest that distinct AD subtypes on FDG PET may have implications for the individual clinical outcomes and provide a clue to understanding a broad spectrum of AD in terms of pathophysiology.

Keywords: Alzheimer's disease, subtype, FDG PET, deep learning, conditional variational autoencoder

Student Number: 2018-36289

CONTENTS

ABSTRACT	i
CONTENTS	iv
List of Figures	vi
List of Tables.....	ix
1. Introduction	1
1.1 Heterogeneity of Alzheimer’s disease	1
1.2 FDG PET as a biomarker of Alzheimer’s disease	1
1.3 Biologic subtypes of Alzheimer’s disease	2
1.4 Dimensionality reduction methods	5
1.5 Variational autoencoder for clustering	8
1.6 Final goal of the study	10
2. Methods	11
2.1 Subjects.....	11
2.2 FDG PET data acquisition and preprocessing	12
2.3 Deep learning-based model for representations of FDG PET in AD ...	12
2.4 Clustering method for AD subtypes on FDG PET	17
2.5 Transfer of deep learning-based FDG PET cluster model for MCI subtypes.	17
2.6 Visualization of subtype-specific spatial brain metabolism pattern	21
2.7 Clinical and biological characterization	21
2.8 Prognosis prediction of MCI subtypes	22
2.9 Generation of subtype-specific FDG PET images.....	22

2.10 Statistical analysis.....	23
3. Results.....	24
3.1 Deep learning-based FDG PET clusters	24
3.2 Spatial brain metabolism pattern in AD subtypes	27
3.3 Clinical and biological characterization in AD subtypes.....	32
3.4 Subtype-specific spatial metabolism patterns resemble in MCI.....	43
3.5 Clinical and biological characterization in MCI subtypes	50
3.6 Prognosis prediction of subtypes for conversion from MCI to AD.....	56
3.7 Generating FDG PET images of AD subtypes	61
4. Discussion	66
4.1 Limitations of previous subtyping approach	68
4.2 Interpretation of results.....	68
4.3 Strength of our deep learning-based clustering approach.....	73
4.4 Strength of our deep learning-based AD subtypes	77
4.5 Limitations and future directions.....	82
5. Conclusion	83
References.....	84
Supplementary Figures	99
국문 초록	101

List of Figures

Figure 1. cVAE architecture of deep learning-based model for representations of FDG PET in AD	16
Figure 2. Deep learning-based FDG PET cluster model for AD subtypes	19
Figure 3. Transfer of deep learning-based FDG PET cluster model for MCI subtypes	20
Figure 4. t-SNE visualization of the latent feature of FDG PET in AD	25
Figure 5. Elbow method for choosing the number of clusters	26
Figure 6. Spatial metabolism pattern on FDG PET with cerebellar normalization of AD subtypes showing the differences between subtypes and controls.....	29
Figure 7. Spatial metabolism pattern on FDG PET with cerebellar normalization of AD subtypes showing the differences between one subtype and all other subtypes	30
Figure 8. Spatial metabolism pattern on FDG PET with cerebellar normalization of AD subtypes at baseline visit showing the differences between subtypes and controls.....	31
Figure 9. Heatmap of clinical and biological characteristics of AD subtypes	42

Figure 10. Frequency of subtypes in AD and MCI	46
Figure 11. Spatial metabolism pattern on FDG PET with cerebellar normalization of MCI subtypes showing the differences between subtypes and controls	47
Figure 12. Spatial metabolism pattern on FDG PET with cerebellar normalization of MCI subtypes showing the differences between one subtype and all other subtypes.....	48
Figure 13. Spatial metabolism pattern on FDG PET with cerebellar normalization of MCI subtypes with the exclusion of overlapping individuals showing the differences between subtypes and controls.....	49
Figure 14. Heatmap of clinical and biological characteristics of MCI subtypes	55
Figure 15. Frequency of MCI converter across subtypes.....	57
Figure 16. Time to conversion from MCI to AD.....	59
Figure 17. Kaplan-Meier survival curve for the conversion from MCI to AD.....	60
Figure 18. Generation of subtype-specific FDG brain PET images according to latent features with different CDR-SB scores by using a deep learning-based model.....	65
Figure 19. Theoretical model summarizing AD subtypes on FDG PET	67

Figure 20. Spatial metabolism of clusters using VAE without CDR-	
SB condition	76
Figure 21. The trajectory of spatial brain metabolism pattern across	
the subtypes	81

List of Tables

Table 1. Key features of dimension reduction methods 7

Table 2. cVAE architecture of deep learning-based model for
representations of FDG PET in AD 15

Table 3. Clinical and biological characteristics of subtypes in AD... 34

Table 4. Clinical and biological characteristics of subtypes in AD at
baseline visit 38

Table 5. Clinical and biological characterizations of subtypes in MCI
..... 51

Table 6. Time to conversion from MCI to AD 58

1. Introduction

1.1 Heterogeneity of Alzheimer's disease

Alzheimer's disease (AD) is a heterogeneous disease that presents a broad spectrum of clinicopathologic profiles, despite common pathologic features including amyloid and tau deposition [1]. AD is commonly regarded as an amnesic disorder, but it often has a non-amnesic clinical presentation [2, 3]. Variability in age onset, clinical presentation, tau-related pathology, and brain atrophy has been reported [4-8]. Due to its heterogeneous biologic pathology, many studies have been reported to reveal biological subtypes of AD on structural magnetic resonance imaging (MRI) and tau positron emission tomography (PET) [9-11].

1.2 FDG PET as a biomarker of Alzheimer's disease

[¹⁸F]fluorodeoxyglucose (FDG) PET has been used to support the differential diagnosis of neurodegenerative disease including dementia and parkinsonism [12-14]. FDG PET is also used to predict the progression of cognitive dysfunction in neurodegenerative disease as well as mild cognitive impairment (MCI) as an imaging-derived biomarker [15-18]. Since regional glucose hypometabolism precedes macroscopic brain atrophy detectable on MRI, FDG PET may indicate

an earlier stage of neurodegeneration [19, 20]. Previous studies have found a relationship between spatial pattern of hypometabolism on FDG PET and clinical symptoms of AD [21-23], and in addition, recently a few attempts have been reported to differentiate spatial hypometabolic patterns of AD subtypes in a data-driven or hypothesis-driven manner [24, 25].

1.3 Biologic subtypes of Alzheimer's disease

Because of the variability of AD, various attempts have been made to explain heterogeneity within AD for several decades. In 1992, Ritchie and Touchon [26] proposed 3 descriptive hypotheses for the explanation of heterogeneity of AD: phase, compensation, and subtype models. The phase model refers to the stage of the disease that is associated with the severity of the symptom. The compensation model hypothesizes that the disease progression can be counter-balanced by a compensatory biological mechanism that is different for each patient. The sub-type model hypothesizes separate subtypes at functional, clinical, and biological levels. Ritchie and Touchon suggested that these 3 models together result in the heterogeneity of the AD population.

Over the last few years, many studies have been reported to determine the biological subtypes of AD in a hypothesis and a data-driven manner. Most of the early studies were performed on post-mortem

data using data-driven clustering methods on pathologic measures of senile plaques, neurofibrillary tangles, and cerebral amyloid angiopathy [27-29]. Murray et al. classified 3 subtypes with distinct clinical characteristics from the distribution of neurofibrillary tangle: typical, hippocampal sparing, and limbic predominant subtypes [30]. Hippocampal sparing AD was younger and had a higher proportion of males, whereas limbic predominant AD was older and higher in females. Late studies focused on neuroimaging biomarkers, mainly tau pathology (tau PET) and neurodegeneration (MRI), for subtyping AD based on a hypothesis-driven or data-driven method. Whitwell et al. first proposed 3 subtypes based on entorhinal and cortical uptake on tau PET [31]: low entorhinal and low cortical uptake (characterized with older age), high entorhinal and high cortical uptake (worst memory impairment), and low entorhinal and high cortical uptake (worst impairment in nonmemory domains). In addition to 3 subtypes identified on postmortem and tau PET analyses, various MRI studies have consistently identified the fourth subtype with minimal atrophy [11, 32, 33]. Combining various neuropathology and neuroimaging studies of AD subtypes, Ferreira et al. proposed 4 biological subtypes of AD (typical, hippocampal sparing, limbic predominant, and minimal atrophy) distinguished by 2 core dimensions (typicality and severity) [9].

Few studies have been reported to identify subtypes based on the

pattern of hypometabolism on FDG PET. Bittner et al. proposed 2 AD groups (temporoparietal area vs. no hypometabolism) [34], and Collette et al. proposed 2 AD groups (frontal and posterior cortex vs. only in posterior cortex) [35] with a hypothesis-driven approach. Meyer et al. also proposed 2 groups (perisylvian areas vs. typical posterior brain areas) of AD based on the pattern of hypometabolism [36]. Recently, Levin et al. proposed 3 data-driven FDG PET subtypes of AD (typical, limbic-predominant, and cortical-predominant) [24]. Groot et al. proposed AD subgroups based on the impairments across cognitive domains (memory, executive function, language, visuospatial function, multiple domains, and no domains) which showed subgroup-specific hypometabolism patterns and trajectories [25]. However, the proposed number and patterns of hypometabolism of AD subtypes on FDG PET across the studies are not consistent. In addition, the hypometabolism patterns identified in the studies on FDG PET cannot easily be combined with subtypes in other modalities (postmortem, MRI, and tau PET). For this reason, a biologic subtype of AD on FDG PET is not yet been clarified.

Above all, the concern of previous approaches in defining AD subtypes on various modalities was difficulty in removing the effect of disease progression. The clusters based on the traditional data-driven method may have a risk of representing the dimension of disease severity rather than the intrinsic biologic subtype of AD.

1.4 Dimensionality reduction methods

For the extraction of key features and classification of the subtypes from the high-dimensional data, we can use dimensionality reduction methods. There are several different types of traditional methods for dimensionality reduction depending on their purpose and strategy [37, 38]. One type of method only keeps the most important features without transformation of the set of the features: backward elimination, forward selection [39, 40], and Random forests [41]. The other type of method applies appropriate transformation in the set of features and finds a combination of new features. This method can be divided into linear methods and non-linear methods, also known as manifold learning, depending on which type of activation function is used. Linear method uses linear activation function for dimensionality reduction: principal component analysis (PCA) [42, 43], factor analysis (FA) [44, 45], linear discriminant analysis (LDA) [46] and truncated singular value decomposition (SVD) [47, 48]. Non-linear method uses non-linear activation function to reduce dimension: kernel PCA [49], t-distributed stochastic neighbor embedding (t-SNE) [50], multidimensional scaling (MDS) [51], isometric mapping (Isomap) [52, 53], and autoencoder [54, 55].

Among these methods, PCA was the most commonly used traditional method for dimensionality reduction for its fast and simple

application. To compare PCA and autoencoder for dimensionality reduction, the biggest difference is that the autoencoder can utilize a non-linear activation function whereas the PCA is a linear transformation [38]. The use of the non-linear function for dimensionality reduction in the autoencoder could make superior performance in learning patterns from the high-dimensional image data and accurate image generation [56-58]. However, the autoencoder is prone to an overfitting problem. To avoid overfitting, the variational autoencoder (VAE) has been introduced [59]. Instead of encoding an input as a single point, VAE encodes it as a distribution over the latent space. Key features of dimension reduction methods are summarized in **Table 1**.

Table 1. Key features of dimension reduction methods

Methods	Key features
PCA	<ul style="list-style-type: none">• Utilize linear activation function• Limitation with non-linear data
Autoencoder	<ul style="list-style-type: none">• Utilize non-linear activation function• Superior performance in learning patterns from the high-dimensional image data and accurate image generation• Prone to an overfitting problem
VAE	<ul style="list-style-type: none">• Consist of encoder and decoder component• Single input• Reduce the dimension in a probabilistic manner• Avoid overfitting
cVAE	<ul style="list-style-type: none">• Consist of encoder and decoder component• Conditioned variable as additional input• Latent features represent other data-specific features different from conditions

1.5 Variational autoencoder for clustering

VAE is an unsupervised learning method that can generate images from some representation [60]. This generative model has advantages in a flexible application in image generation and a model transfer to different disease domains [61, 62].

VAE consists of two components, encoder and decoder (generator). The encoder can reduce the high-dimensional image (X) into low-dimensional latent features (z) which are hidden representations of X , and the decoder can generate an image from values of latent features. Compared to a traditional simple autoencoder, VAE reduces the dimension in a probabilistic manner. For the simple explanation of the model, an input (x) is encoded into the latent distribution in a probabilistic manner $p(z|x)$ followed by sampling. Sampled latent representation $z \sim p(z|x)$ is then decoded for the input reconstruction $d(z)$. In practice, generator input was resampled by the encoded latent feature z assuming normal distribution $z_{\text{resampled}} = z_{\text{encoder}} + z_{\text{sd}} \times \epsilon$, where ϵ represents a random variable. The probabilistic generator can be defined by $p_{\theta}(x|z)$, where θ represents the parameters of the generator. The posterior distribution $p_{\theta}(z|x)$ can be obtained by prior distribution $p(z)$, $p_{\theta}(z|x) \sim p(z) p_{\theta}(x|z)$. Variational Bayes learns both parameters, $p_{\theta}(x|z)$ and an approximation $q_{\phi}(z|x)$ to the intractable true posterior $p_{\theta}(z|x)$.

The VAE model is trained to minimize the loss function,

$$L(\phi, \theta) = -\mathbb{E}_{z \sim q_\phi(z|x)}(\log p_\theta(x|z)) + \text{KL}(q_\phi(z|x) \parallel p_\theta(z))$$

The first term represents the reconstruction loss of the autoencoder. The second term, Kullback-Leibler (KL) is a divergence between a learned distribution and prior distribution. KL divergence act as a regularization term to ensure our learned distribution q is similar to the true prior distribution [60].

The conditional VAE (cVAE) is an extension of VAE. While VAE uses single high-dimensional input data for the dimension reduction, cVAE uses the conditioned variable as additional input labels. Whereas VAE generates data from the latent variable directly, the cVAE model is aimed to generate data from both conditioned and latent variables [63]. The probabilistic generator (decoder) and the encoder can be defined as $p_\theta(x|y, z)$ and $q_\phi(z|x, y)$, respectively. The loss function is changed to,

$$L(\phi, \theta) = -\mathbb{E}_{z \sim q_\phi(z|x,y)}(\log p_\theta(x|y, z)) + \text{KL}(q_\phi(z|x, y) \parallel p_\theta(z))$$

where the first and second term represents reconstruction loss and KL divergence respectively. By giving conditioned variables as additional input information, the latent features are expected to reflect hidden information other than given conditions. As shown in the experiment of cVAE on Modified National Institute of Standards and Technology (MNIST) dataset (large database of handwritten digits) with conditioned

input labels, cVAE learns handwriting styles for the latent feature which is hidden information other than given condition [63].

In this study, we utilize cVAE on FDG PET of AD subjects with the corresponding condition of AD severity score. To find AD subtypes using brain images, clustering based on unsupervised learning could be used to define specific clusters. Nonetheless, decreased FDG uptake patterns in the cortex mostly depend on disease severity, there has been a difficulty in defining subtypes removing the effect of disease progression. As a cVAE model directly uses conditional variables to generate images from latent features, latent features are expected to represent other data-specific features different from conditions. We aimed to reduce FDG PET image dimension into a lower-dimensional feature for clustering, which is expected to have hidden representation of images other than given condition (e.g. disease severity).

1.6 Final goal of the study

We aimed to identify AD subtypes using the deep learning-based clustering on FDG PET images to understand distinct spatial patterns of neurodegeneration. We also aimed to investigate clinicopathologic features of subtypes defined by spatial brain metabolism patterns.

2. Methods

2.1 Subjects

We obtained 3620 FDG brain PET images from 1607 participants at enrollment (baseline) and follow-up visits from Alzheimer's Disease Neuroimaging Initiative (ADNI) database. We defined cohorts of Alzheimer's disease (AD), mild cognitive impairment (MCI), and cognitive normal controls (CN) based on clinical diagnosis at each visit for this study. Based on the criteria, we included data of 838 AD, 1761 MCI, and 1021 CN for the present study. The detailed inclusion criteria for the different diagnostic categories are available on the ADNI website (<http://adni.loni.usc.edu/methods/documents/>). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD, VA Medical Center and University of California San Francisco. ADNI recruited subjects from over 50 sites across the US and Canada. The primary purpose of ADNI has been to test whether serial imaging and biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. For up-to-date information, see <http://www.adni-info.org>. Written informed consent to cognitive testing and neuroimaging prior to participation was obtained, approved by the institutional review boards of all participating institutions.

2.2 FDG PET data acquisition and preprocessing

FDG PET data were acquired in a pre-processed image file. 185 MBq of FDG were intravenously injected and a dynamic 3D scan of six 5-min frames was acquired for 30 minutes after 30-60 minutes of injection. Each extracted frame was co-registered followed by averaging the six 5-minute frames. Scans were reoriented into a standard $160 \times 160 \times 96$ voxel image grid, having 1.5 mm cubic voxel. The intensity of each PET image was normalized using a subject-specific mask so that the average of voxels within the mask is exactly one. Each image set was filtered to produce images of a uniform isotropic resolution of 8 mm full width at half maximum [64, 65]. The detailed PET acquisition and preprocessing protocols are available on the ADNI website (<http://adni.loni.usc.edu/methods/pet-analysis-method/pet-analysis/>).

For the present deep learning study, FDG PET images were further spatially normalized to a customized PET template and reshaped into $91 \times 109 \times 91$ matrices using and smoothed with a Gaussian smoothing kernel of 10 mm full-width at half maximum (FWHM).

2.3 Deep learning-based model for representations of FDG PET in AD

To find hidden representations of FDG PET image patterns in AD, a cVAE was used [63]. 838 FDG brain PET images with AD were used as

input images. Clinical Demetria Rating Scale Sum of Boxes (CDR-SB) score was used as the corresponding condition vector of each image. CDR-SB is commonly used in clinical and research settings to stage dementia severity [66]. Since removing the effect from the disease severity dimension was the biggest concern for previous AD subtype studies, we chose CDR-SB as an input condition to remove the effect of disease severity of AD in latent features and find hidden representation. CDR-SB score was rescaled to the range of 0 to 1, which was divided by a maximum CDR-SB score of 18, and used as input. To encode 3-dimensional PET volume, we used multiple 3D convolutional neural network. Four sequential convolution layers and rectified linear units were applied, and the features were fed into the fully connected layer. The sizes of the convolutional filters were $(5 \times 5 \times 5)$, $(5 \times 5 \times 5)$, $(5 \times 5 \times 5)$, and $(3 \times 3 \times 3)$ in the sequence, and the stride size of 2, 3, and 3 voxels were applied for the first, second, and third convolution filters. After multiple convolutional filters, 3D feature volumes were reduced to a 1-dimensional feature with a size of 32 followed by a rectified linear unit. These features were merged into CDR-SB information of each subject followed by sampling layer and finally connected to the layer of 10 latent features. Conversely, the decoder consisted of convolutional and upsampling layers with input variables of 10 latent features and CDR-SB information to generate PET volume. We chose a latent

dimension size of 10 that gives the best performance in dimension reduction of FDG PET, as shown in the previous cVAE study [67]. The detail of the cVAE architecture is summarized in **Table 2 and Figure 1**. We trained the cVAE model using an optimization algorithm, Adam [68], to minimize the loss function, and the learning rate was set to 0.0001. It took 300 epochs with a batch size of 32 for the training. The cVAE was implemented using a deep learning library, Keras with TensorFlow backend (version 2.5.0). Ten percent of PET data were used for the internal validation to determine hyperparameters and find out the best model.

Table 2. cVAE architecture of deep learning-based model for representations of FDG PET in AD

Layer type	Output size	Kernel size	Stride	Activation
Encoder				
Input 1 (x; image)	$(91 \times 109 \times 91)$			
Cropping3D	$(90 \times 108 \times 90)$			
Conv3D	$(45 \times 54 \times 45)$, 16	$5 \times 5 \times 5$	2	ReLU
Conv3D	$(15 \times 18 \times 15)$, 16	$5 \times 5 \times 5$	3	ReLU
Conv3D	$(5 \times 6 \times 5)$, 64	$5 \times 5 \times 5$	3	ReLU
Conv3D	$(3 \times 4 \times 3)$, 64	$3 \times 3 \times 3$		ReLU
Flatten	2304			
Dense	32			ReLU
Input 2 (y; condition)	1			
Concatenate	33			
Dense	10			
Dense	10			
Lambda (Sampling)	10			
Latent variables: z				
Decoder				
Concatenate (z and y)	11			
Dense	32			ReLU
Dense	2304			ReLU
Reshape	$(3 \times 4 \times 3)$, 64			
Conv3D Transpose	$(5 \times 6 \times 5)$, 64	$3 \times 3 \times 3$		ReLU
Conv3D Transpose	$(15 \times 18 \times 15)$, 16	$5 \times 5 \times 5$	3	ReLU
Conv3D Transpose	$(45 \times 54 \times 45)$, 16	$5 \times 5 \times 5$	3	ReLU
Conv3D Transpose	$(90 \times 108 \times 90)$	$5 \times 5 \times 5$	2	Sigmoid
Zero padding	$(91 \times 109 \times 91)$			

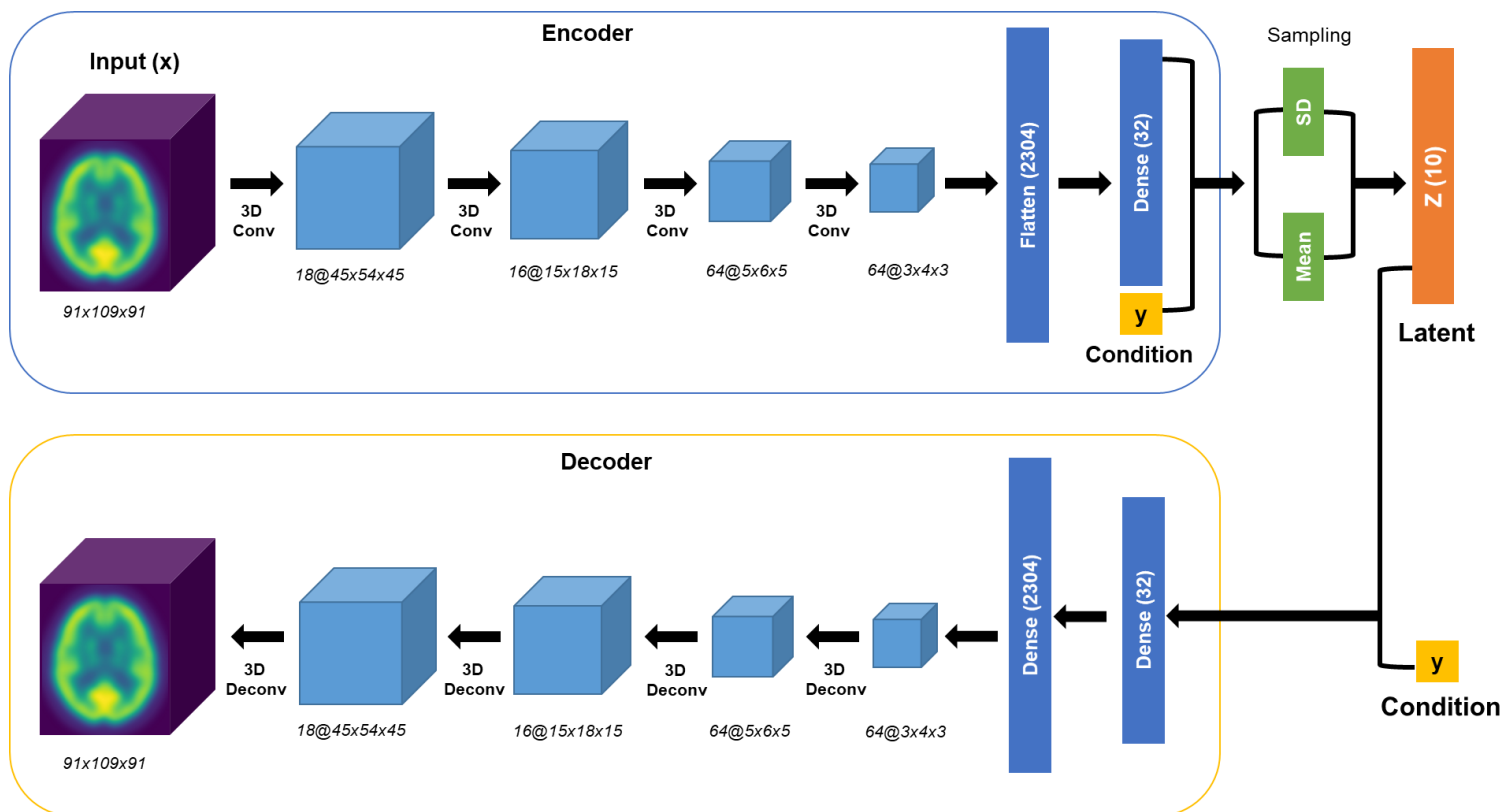


Figure 1. cVAE architecture of deep learning-based model for representations of FDG PET in AD

2.4 Clustering method for AD subtypes on FDG PET

To identify AD subtypes from the hidden representations of FDG PET image patterns, the k-means algorithm was applied to generate clusters from the 10 latent features [69, 70]. The study design of our deep learning-based FDG PET cluster model for AD subtypes is summarized in **Figure 2**. The number of clusters was determined by using the elbow method [71, 72]. The elbow method is a heuristic method to determine the optimal number of clusters in a dataset by calculating the sum of squared distance between each point and the centroid in a cluster. The t-SNE algorithm was applied using random initialization and perplexity of 30 to visualize latent features on a 2-dimensional plot [50]. The Scikit-learn library (version 0.24.2) was implemented for k-means and t-SNE algorithms.

2.5 Transfer of deep learning-based FDG PET cluster model for MCI subtypes

The trained deep learning-based FDG PET cluster model in AD was transferred to the MCI cohort to predict their subtypes and to identify differential trajectories and prognosis of subtypes in MCI. The study design for the transfer of our deep learning-based FDG PET cluster model for MCI subtypes is summarized in **Figure 3**. The cVAE model and k-means clustering model used in the AD cohort were directly

transferred to the MCI cohort. FDG PET images and CDR-SB scores from the MCI subjects were used as input images and conditions of the transferred cVAE model to extract latent features of FDG PET with MCI. And then, the k-means clustering model was transferred to identify MCI subtypes from the latent features.

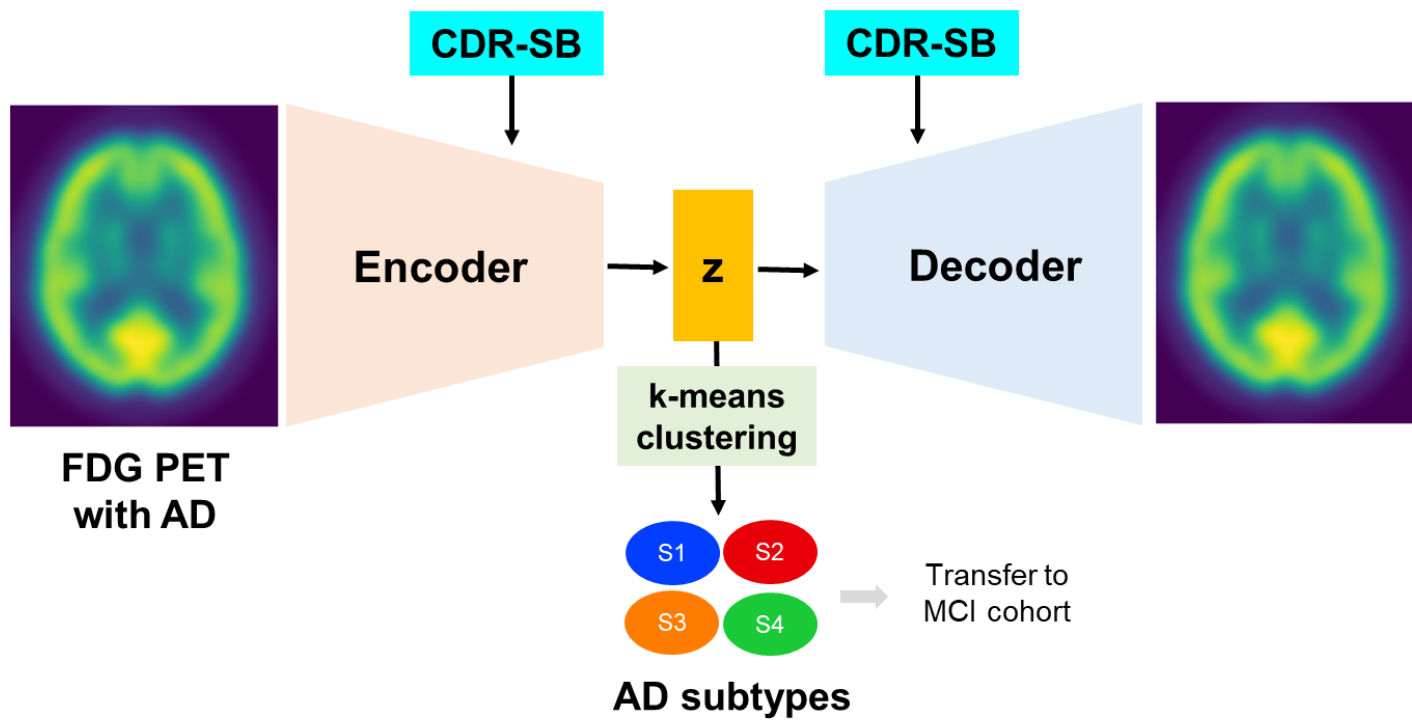


Figure 2. Deep learning-based FDG PET cluster model for AD subtypes

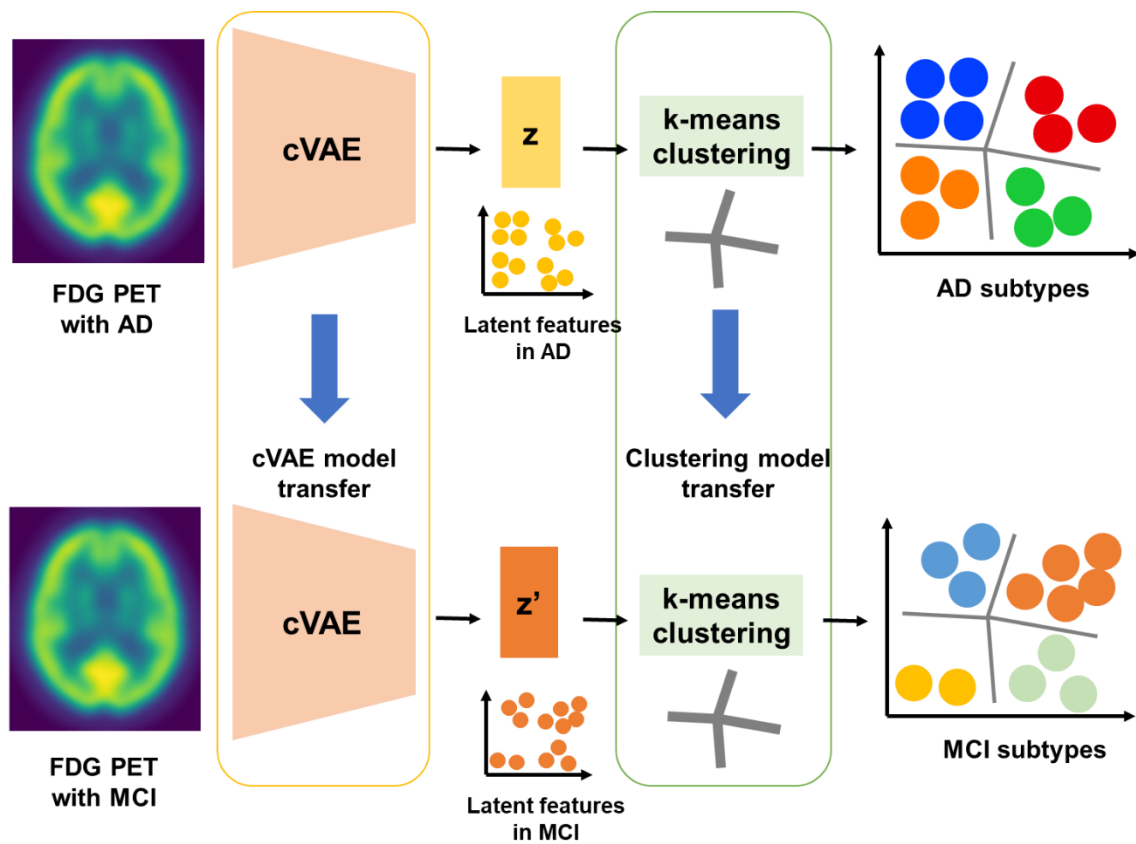


Figure 3. Transfer of deep learning-based FDG PET cluster model for MCI subtypes

2.6 Visualization of subtype-specific spatial brain metabolism pattern

To visualize spatial brain metabolism patterns of identified subtypes, a voxel-wise two-sample t-test was performed on FDG PET images using SPM12 software. Intensities of FDG PET images were normalized to the cerebellum using the automated anatomical labeling (AAL) template for the cerebellar normalization [73]. FDG PET images with global normalization were also used for the additional analysis, and overall grand mean scaling was applied for the global normalization. Statistical parametric maps (SPM) were generated to visualize spatial patterns of the cluster. The significance threshold was $P = 0.01$ with family-wised error correction, and the extent threshold was 50 voxels. The statistical analyses were performed comparing FDG PET of each subtype against controls as well as one subtype against all other subtypes.

2.7 Clinical and biological characterization

We compared demographic, cognitive, and biomarker variables among clusters for clinical and biologic characterization of subtypes. All of these variables were downloaded from the LONI Image and Data Archive (IDA). For the cognition scores, we used Alzheimer's Disease Sequencing Project (ADSP) Phenotype Harmonization Consortium (PHC) - composite cognitive scores: harmonized composite memory

score (PHC-MEM), executive function score (PHC-EXF), language score (PHC-LAN), and visuospatial score (PHC-VSP) [74]. Individual's standardized uptake value ratio (SUVR) of [^{18}F]florbetapir (AV45) PET was downloaded on LONI, and florbetapir mean of whole cerebellum was used as reference regions. The regions were defined by Freesurfer and details are available on the documents on LONI. The hippocampal volume to intracranial volume ratio (HV/ICV) was also calculated as an indicator of hippocampal atrophy.

2.8 Prognosis prediction of MCI subtypes

We defined MCI converters for those who converted from MCI to AD within 2 years from the baseline visit, and non-converters for those who did not convert to AD during a follow-up of at least 2 years from the baseline visit. Time to conversion from MCI to AD was calculated for the subjects who have a change of diagnosis from MCI to AD at any time point. Kaplan-Meier survival curves were generated to evaluate the risk of conversion from MCI to AD across the subtypes.

2.9 Generation of subtype-specific FDG PET images

Decoder layers of our cVAE model were used for the generation of subtype-specific FDG PET images. The centroids of latent features of 4 subtypes and rescaled CDR-SB scores were used for the input of the

decoder for image generation.

2.10 Statistical analysis

Values are expressed as percentages or mean with standard deviation (SD). Group differences in demographic and clinical variables in subtypes were evaluated using the one-way ANOVA with posthoc analysis and chi-square test. Z-scores of demographic, cognitive, and biomarker variables were calculated, and averaged Z-score of each subtype were used for the heatmap generation to visualize clinical and biological characteristics of each subtype. Kaplan-Meier survival analysis was used to test for subtype differences in conversion from MCI to AD, and curves were compared using the log-rank test. Statistical analyses were performed using a JAMOVİ statistical software version 1.6, and a P-value lower than 0.05 was considered statistically significant.

3. Results

3.1 Deep learning-based FDG PET clusters

Of the 838 FDG brain PET images with AD, four distinct AD subtypes were identified by deep-learning-based FDG PET clusters (**Figure 4**).

We set 4 as a subtype number since the inflection point appeared when the number was 4 on the elbow method plot (**Figure 5**).

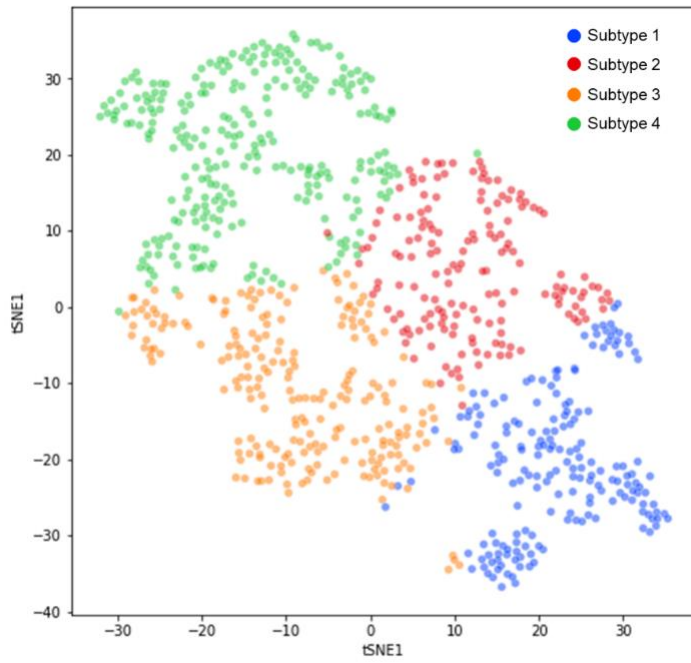


Figure 4. t-SNE visualization of the latent feature of FDG PET in AD

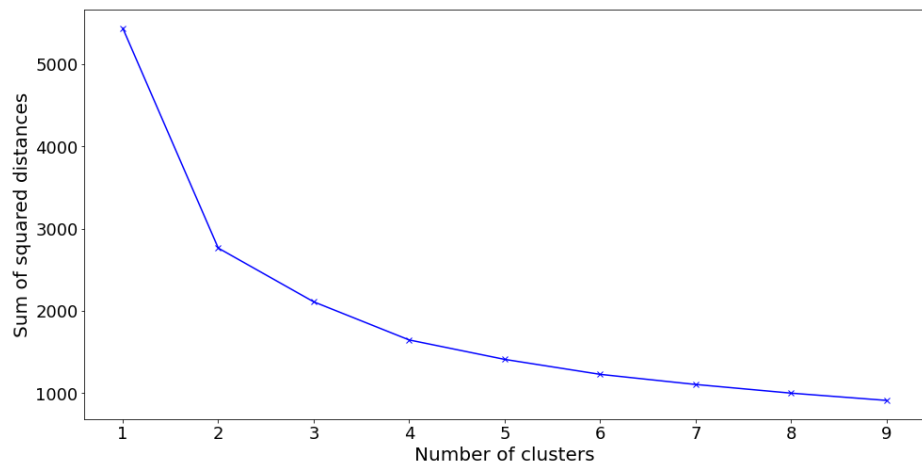


Figure 5. Elbow method for choosing the number of clusters

3.2 Spatial brain metabolism pattern in AD subtypes

Spatial brain metabolism patterns on FDG PET with cerebellar normalization of AD subtypes were compared with those of the CN, and the regions showing significant differences were revealed using a voxelwise t-test. All AD subtypes shared a common pattern of hypometabolism involving frontal, parietal, temporal lobes, precuneus, and posterior cingulate, which is typically observed in AD. In addition to typical AD patterns shared across all subtypes, subtype-specific regions of hypometabolism were observed in each subtype (**Figure 6**). Subtype 1 (S1: angular) included 185 (22%) of AD and showed prominent hypometabolism in the angular gyrus with a diffuse hypometabolism pattern involving parietotemporal, frontal, limbic, occipital, and cingulate cortices. Subtype 2 (S2: occipital) included 161 (19%) of AD and showed prominent hypometabolism in the occipital cortex with a posterior-predominant hypometabolism pattern involving occipital, posterior-parietal cortices, and precuneus. Subtype 3 (S3: orbitofrontal) included 224 (27%) of AD and showed prominent hypometabolism in the orbitofrontal cortex with an anterior-predominant hypometabolism pattern involving frontal, limbic, and anterior cingulate cortices. Subtype 4 (S4: minimal) included 268 (32%) of AD and showed no additional hypometabolic region. Subtype-specific spatial

metabolism patterns also corresponded to the regions observed in the comparisons between one subtype and all other subtypes (**Figure 7**). In the additional analysis using FDG PET with global normalization, the subtype-specific hypometabolism patterns were consistent with those observed in the analysis using FDG PET with cerebellar normalization (**Supplementary figure 1 and 2**).

To exclude the effect of overlapping images from the same subject at different visits, FDG PET from individuals with AD at baseline visit were only used for the additional spatial metabolism pattern analysis. For the FDG PET with AD at baseline visit, S1 (angular), S2 (occipital), S3 (orbitofrontal), and S4 (minimal) included 58 (20%), 52 (18%), 83 (28%), and 99 (34%), respectively. FDG PET images of AD subtypes at baseline visit were compared with those of CN, and regions showing significant differences were revealed using a voxelwise t-test. The subtype-specific hypometabolism patterns observed in AD at baseline visit resembled those observed in the analysis using whole AD subjects but showed a lower degree of t-value (**Figure 8**).

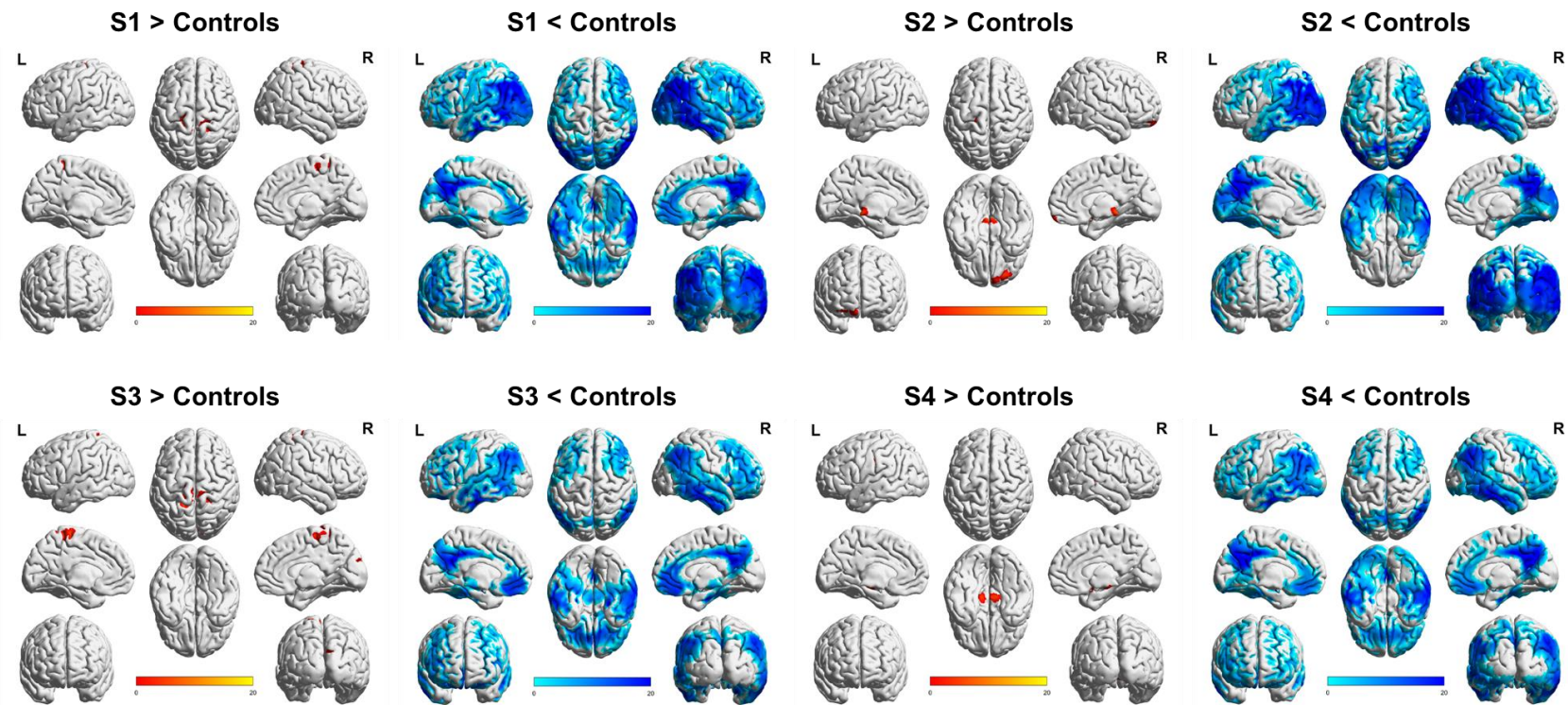


Figure 6. Spatial metabolism pattern on FDG PET with cerebellar normalization of AD subtypes showing the differences between subtypes and controls

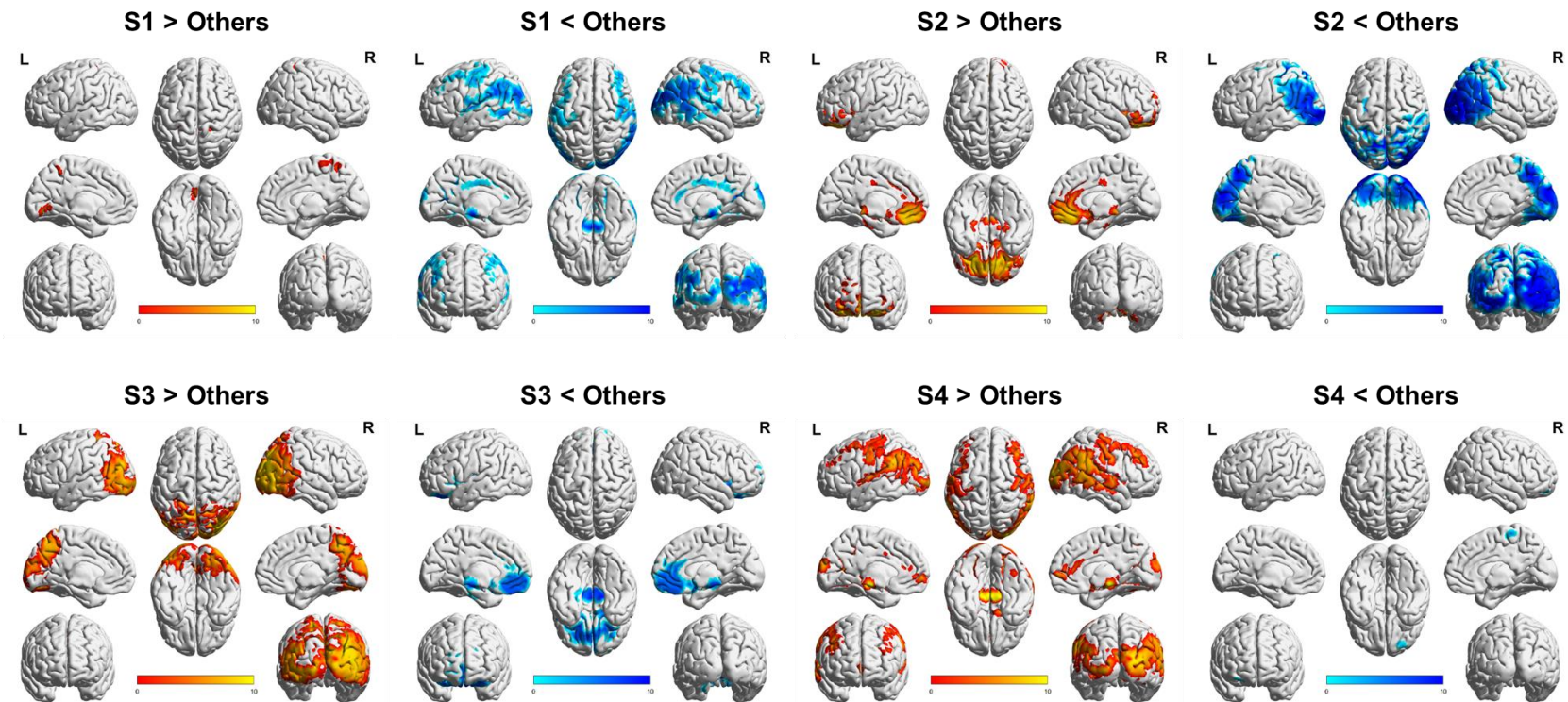


Figure 7. Spatial metabolism pattern on FDG PET with cerebellar normalization of AD subtypes showing the differences between one subtype and all other subtypes

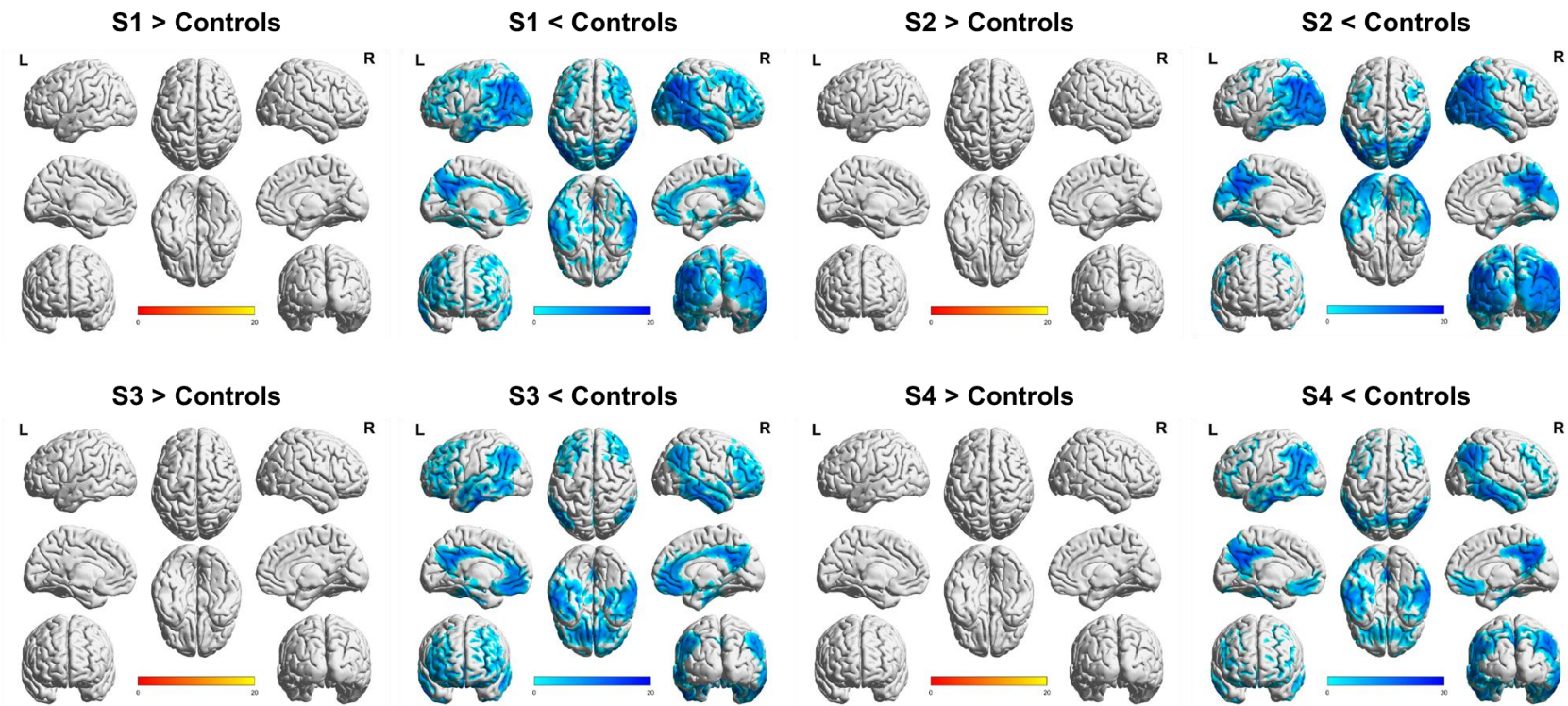


Figure 8. Spatial metabolism pattern on FDG PET with cerebellar normalization of AD subtypes at baseline visit showing the differences between subtypes and controls

3.3 Clinical and biological characterization in AD subtypes

We compared demographic, cognitive, and biomarker variables between the AD subtypes (**Table 3-4 and Figure 9**). Characteristics of subtypes were compared both in groups with all AD subjects (n=838) and AD subjects at baseline visit (n=292). CDR-SB scores did not show a significant difference across AD subtypes ($P = 0.609$) as we used CDR-SB as the input variable of our cVAE model.

Compared to other subtypes, S1 (angular) included the highest frequency of male and more educated individuals. Individuals in S1 had more amyloid burden (i.e. less cerebrospinal fluid (CSF) amyloid-beta (A β) and higher AV45-PET SUVR) but less overall tau burden (i.e. less CSF p-tau and t-tau level). S1 individuals tended to have lower HV/ICV (i.e. more hippocampal atrophy). S2 (occipital) individuals were younger and had more overall tau burden but less hippocampal atrophy. S2 individuals tended to have better relative memory and language, but worse relative executive and visuospatial scores. S3 (orbitofrontal) individuals were older and had less amyloid burden but more hippocampal atrophy. S3 individuals tended to have better executive and visuospatial scores. Finally, S4 (minimal) included the highest frequency of female and less educated individuals. S4 individuals had a more tau burden but less hippocampal atrophy. S4 individuals

tended to have better executive, language, and visuospatial scores.

Table 3. Clinical and biological characteristics of subtypes in AD

	CN	AD (n=838)				P-value, global comparison (S1, S1, S3, and S4)
		S1	S2	S3	S4	
Demographics						
n (%)	1021	185 (22%)	161 (19%)	224 (27%)	268 (32%)	
Age, years (n=838)	75.7 (6.2)	76.4 (7.2)	75.0 (7.5)	78.2 (6.5)	75.2 (7.9)	< 0.001 ***
Sex, female (%) (n=838)	43%	5%	32%	41%	72%	< 0.001 ***
Education, years (n=838)	16.3 (2.8)	16.1 (3.3)	15.5 (2.6)	15.3 (2.7)	15.1 (2.9)	0.009 **
Cognition						
CDR-SB (n=838)	0.1 (0.4)	5.5 (2.7)	5.2 (2.7)	5.4 (2.6)	5.3 (2.3)	0.609
MMSE (n=837)	29.0 (1.2)	21.9 (4.0)	22.0 (4.4)	22.3 (3.8)	22.2 (3.9)	0.699
MOCA (n=354)	25.8 (2.5)	16.7 (4.9)	16.0 (5.2)	17.1 (4.8)	16.8 (5.0)	0.645
PHC-MEM (n=833)	0.86 (0.51)	-0.92 (0.49)	-0.81 (0.53)	-0.89 (0.46)	-0.90 (0.45)	0.217
PHC-EXF (n=832)	0.74 (0.48)	-0.55 (0.72)	-0.71 (0.74)	-0.36 (0.65)	-0.45 (0.70)	< 0.001 ***
PHC-LAN (n=833)	0.80 (0.49)	-0.29 (0.59)	-0.24 (0.63)	-0.27 (0.62)	-0.22 (0.64)	0.63
PHC-VSP (n=659)	0.10 (0.31)	-0.36 (0.61)	-0.60 (0.74)	-0.19 (0.52)	-0.26 (0.62)	< 0.001 ***
Biomarker						

AV45 PET, SUVR (n=318)	1.12 (0.19)	1.40 (0.24)	1.37 (0.22)	1.35 (0.25)	1.39 (0.21)	0.585
APOE4 carrier (%) (n=817)	27%	68%	64%	68%	66%	0.249
CSF Abeta, pg/ml (n=321)	1211 (439)	580 (276)	677 (402)	681 (345)	664 (280)	0.159
CSF t-tau, pg/ml (n=321)	245 (91)	293 (108)	382 (158)	358 (124)	423 (173)	< 0.001 ***
CSF p-tau, pg/ml (n=321)	23 (10)	29 (12)	38 (17)	35 (14)	41 (19)	< 0.001 ***
HV/ICV, cm ³ /mm ³ (n=547)	4.78 (0.68)	3.37 (0.56)	3.80 (0.59)	3.50 (0.57)	3.85 (0.73)	< 0.001 ***

The data are expressed as percentages or means with standard deviation in parentheses. Missing values are excluded.

* p < 0.05, ** p < 0.01, *** p < 0.001

(continued)

	P-value, pair-wise comparisons					
	S1 vs S2	S1 vs S3	S1 vs S4	S2 vs S3	S2 vs S4	S3 vs S4
Demographics						
Age, years	0.296	0.078	0.308	< 0.001 ***	0.995	< 0.001 ***
Education, years	0.294	0.023 *	0.002 **	0.800	0.442	0.929
Cognition						
CDR-SB	0.618	0.897	0.635	0.934	0.998	0.962
MMSE	0.994	0.684	0.892	0.854	0.977	0.967
MOCA	0.889	0.949	0.997	0.562	0.748	0.979
PHC-MEM	0.165	0.967	0.994	0.318	0.197	0.996
PHC-EXF	0.148	0.039 *	0.512	< 0.001 ***	0.002 **	0.461
PHC-LAN	0.901	0.993	0.632	0.967	0.979	0.770
PHC-VSP	0.009 **	0.056	0.399	< 0.001 ***	< 0.001 ***	0.667
Biomarker						
AV45 PET, SUVR	0.944	0.632	1.000	0.933	0.949	0.571
CSF Abeta, pg/ml	0.328	0.249	0.357	1.000	0.994	0.984
CSF t-tau pg/ml	0.004 **	0.045 *	< 0.001 ***	0.751	0.289	0.015 *
CSF p-tau, pg/ml	0.009 **	0.066	< 0.001 ***	0.818	0.508	0.067

HV/ICV, cm ³ /mm ³	< 0.001 ***	0.382	< 0.001 ***	0.001 **	0.913	< 0.001 ***
--	-------------	-------	-------------	----------	-------	-------------

* p < 0.05, ** p < 0.01, *** p < 0.001

Table 4. Clinical and biological characteristics of subtypes in AD at baseline visit

	CN	AD at baseline (n=292)				P-value, global comparison (S1, S1, S3, and S4)
		S1	S2	S3	S4	
Demographics						
n (%)	385	58 (20%)	52 (18%)	83 (28%)	99 (34%)	
Age, years (n=292)	73.6 (5.9)	75.1 (7.7)	72.1 (8.5)	77.2 (6.7)	73.7 (8.4)	0.001 **
Sex, female (%) (n=292)	51%	3%	31%	39%	69%	< 0.001 ***
Education, years (n=292)	16.4 (2.7)	16.4 (3.1)	15.4 (2.4)	15.3 (2.8)	14.9 (2.9)	0.038 *
Cognition						
CDR-SB (n=292)	0.05 (0.15)	4.6 (1.6)	4.2 (1.6)	4.6 (1.8)	4.5 (1.7)	0.557
MMSE (n=292)	29.0 (1.2)	23.1 (2.1)	23.4 (2.3)	23.2 (2.3)	23.2 (2.2)	0.912
MOCA (n=188)	25.8 (2.4)	16.9 (4.7)	17.2 (3.9)	17.3 (4.7)	17.3 (4.7)	0.985
PHC-MEM (n=287)	0.81 (0.46)	-0.75 (0.40)	-0.72 (0.30)	-0.80 (0.32)	-0.80 (0.35)	0.347
PHC-EXF (n=287)	0.69 (0.47)	-0.49 (0.75)	-0.55 (0.69)	-0.33 (0.61)	-0.42 (0.67)	0.247
PHC-LAN (n=287)	0.76 (0.46)	-0.25 (0.52)	-0.11 (0.44)	-0.26 (0.68)	-0.16 (0.55)	0.292
PHC-VSP (n=230)	0.13 (0.27)	-0.41 (0.56)	-0.44 (0.71)	-0.10 (0.48)	-0.28 (0.60)	0.007 **
Biomarker						

AV45 PET, SUVR (n=159)	1.12 (0.18)	1.44 (0.22)	1.39 (0.22)	1.29 (0.24)	1.42 (0.19)	0.026 *
APOE4 carrier (%) (n=271)	28%	68%	75%	61%	67%	0.356
CSF Abeta, pg/ml (n=177)	1228 (439)	546 (241)	698 (390)	720 (370)	686 (257)	0.034 *
CSF t-tau, pg/ml (n=177)	238 (91)	287 (104)	398 (155)	358 (126)	415 (166)	< 0.001 ***
CSF p-tau, pg/ml (n=177)	22 (9)	29 (11)	40 (17)	36 (14)	41 (18)	< 0.001 ***
HV/ICV, cm ³ /mm ³ (n=201)	4.96 (0.64)	3.64 (0.46)	3.98 (0.54)	3.69 (0.53)	3.99 (0.70)	0.002 **

The data are expressed as percentages or means with standard deviation in parentheses. Missing values are excluded.

* p < 0.05, ** p < 0.01, *** p < 0.001

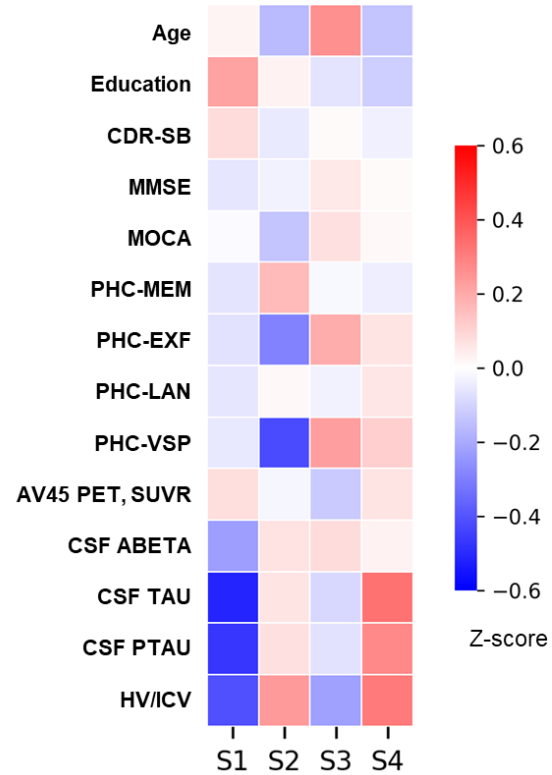
(continued)

	P-value, pair-wise comparisons					
	S1 vs S2	S1 vs S3	S1 vs S4	S2 vs S3	S2 vs S4	S3 vs S4
Demographics						
Age, years	0.194	0.392	0.741	0.002 **	0.601	0.018 *
Education, years	0.286	0.122	0.010 *	0.996	0.723	0.783
Cognition						
CDR-SB	0.688	0.999	0.988	0.534	0.797	0.947
MMSE	0.893	0.996	0.995	0.946	0.946	1.000
MOCA	0.993	0.987	0.981	1.000	1.000	1.000
PHC-MEM	0.951	0.820	0.819	0.491	0.479	1.000
PHC-EXF	0.969	0.488	0.929	0.246	0.694	0.775
PHC-LAN	0.550	0.999	0.787	0.411	0.940	0.640
PHC-VSP	0.996	0.038 *	0.610	0.020 *	0.458	0.290
Biomarker						
AV45 PET, SUVR	0.771	0.029 *	0.984	0.242	0.871	0.016 *
CSF Abeta, pg/ml	0.222	0.080	0.185	0.989	0.998	0.938
CSF t-tau pg/ml	0.012 *	0.145	< .001 ***	0.592	0.948	0.156
CSF p-tau, pg/ml	0.033 *	0.222	0.002 **	0.694	0.968	0.270

HV/ICV, cm ³ /mm ³	0.067	0.969	0.013 *	0.122	0.999	0.024 *
--	-------	-------	---------	-------	-------	---------

* p < 0.05, ** p < 0.01, *** p < 0.001

A



B

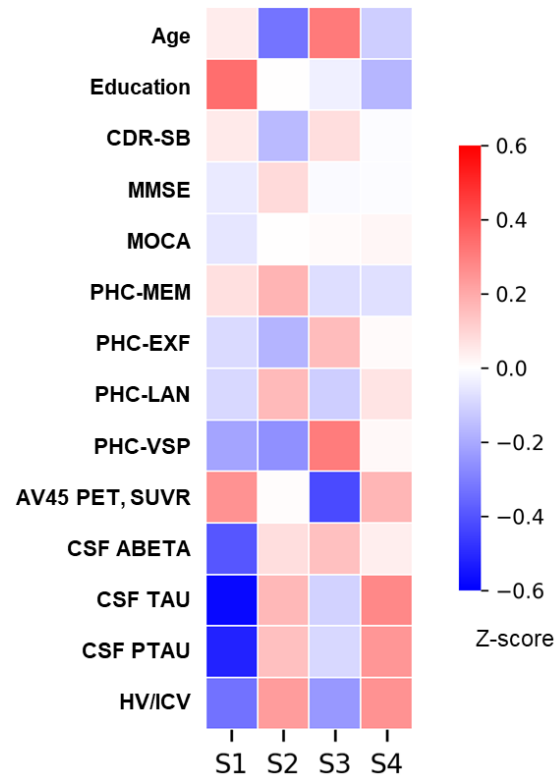


Figure 9. Heatmap of clinical and biological characteristics of AD subtypes

Comparisons in groups with (A) all AD subjects and (B) AD subjects at baseline visit. Averaged Z-scores of demographic, cognitive, and biomarker variables in each subtype are visualized in the heatmap.

3.4 Subtype-specific spatial metabolism patterns resemble in MCI

We transferred our trained cVAE model to FDG PET of cohorts with MCI (n=1761) and their subtypes (MCI-S) were predicted. Different from the distribution of subtypes in AD, individuals were less distributed in S1 and S2, but more distributed in S3 and S4 in the MCI cohort (**Figure 10**): S1, S2, S3, and S4 were 252 (14%), 202 (11%), 596 (34%), and 711 (40%), respectively.

Spatial brain metabolism patterns on FDG PET with cerebellar normalization of MCI subtypes were compared with those of the CN using a voxel-wise t-test (**Figure 11**). MCI-S1, MCI-S2, and MCI-S3 shared a typical AD pattern of hypometabolism involving frontal, parietal, temporal lobes, precuneus, and posterior cingulate which were similar to those observed in AD subtypes. Unlike S4 in AD, MCI-S4 did not share a typical AD pattern but only showed a minimal region of hypometabolism compared to CN. In addition to the shared pattern of hypometabolism, subtype-specific regions of hypometabolism were observed in MCI-S1, MCI-S2, and MCI-S3 which were similar to those observed in AD. As shown in AD subtypes, MCI-S1 (angular) showed prominent hypometabolism in the angular gyrus with a diffuse hypometabolism pattern. MCI-S2 (occipital) showed prominent hypometabolism in the occipital cortex with a posterior-predominant

hypometabolism pattern, and MCI-S3 (orbitofrontal) showed prominent hypometabolism in the orbitofrontal cortex with an anterior-predominant hypometabolism pattern. MCI-S4 (minimal) did not show additional subtype-specific regions of hypometabolism. Subtype-specific spatial metabolism patterns in MCI also corresponded to the regions observed in the comparisons between one subtype and all other subtypes (**Figure 12**).

To exclude the effect of overlapping individuals included in both MCI and AD cohorts due to their conversion from MCI to AD in follow-up visits, we also compared the spatial brain metabolism patterns of MCI subtypes with the exclusion of overlapping individuals from the MCI cohort. With exclusion of overlapping individuals, MCI-S1 (angular), MCI-S2 (occipital), MCI-S3 (orbitofrontal), and MCI-S4 (minimal) included 183 (13%), 140 (10%), 494 (36%), and 543 (40%), respectively. The subtype-specific regions showing significant differences in MCI compared to CN were revealed using a voxelwise t-test. Subtype-specific hypometabolism patterns in MCI with the exclusion of overlapping individuals were similar to those observed in the whole MCI cohort but showed a slightly lower degree of t-value (**Figure 13**). In addition, typical AD patterns (frontal, parietal, temporal lobes, precuneus, and posterior cingulate) were less involved when the overlapping individuals

were excluded for the comparison.

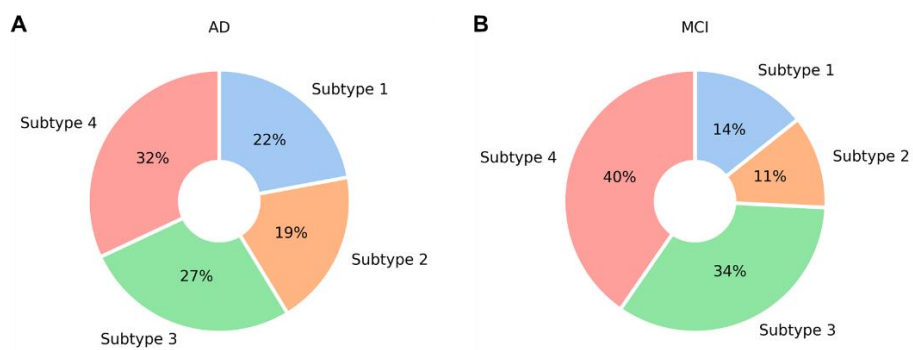


Figure 10. Frequency of subtypes in AD and MCI
(A) AD, (B) MCI

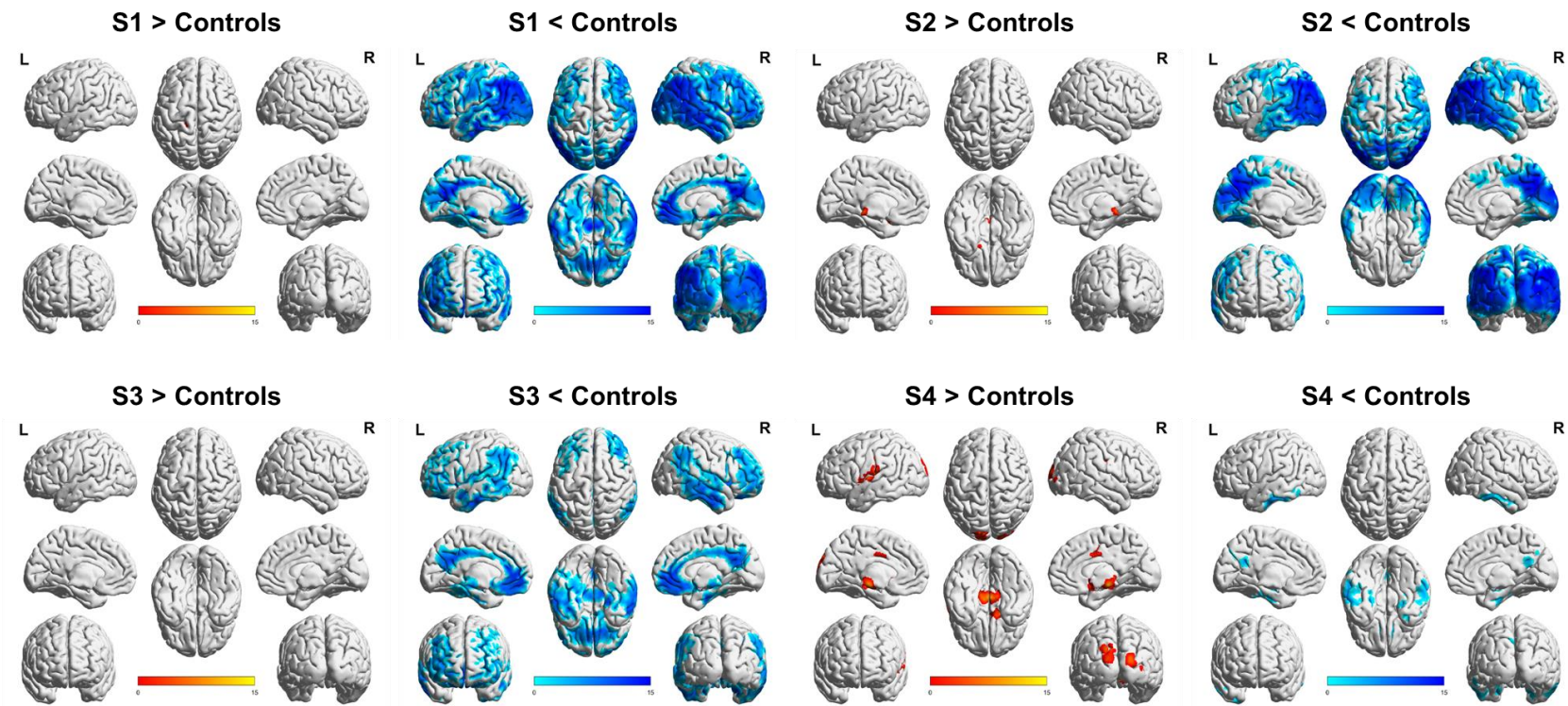


Figure 11. Spatial metabolism pattern on FDG PET with cerebellar normalization of MCI subtypes showing the differences between subtypes and controls

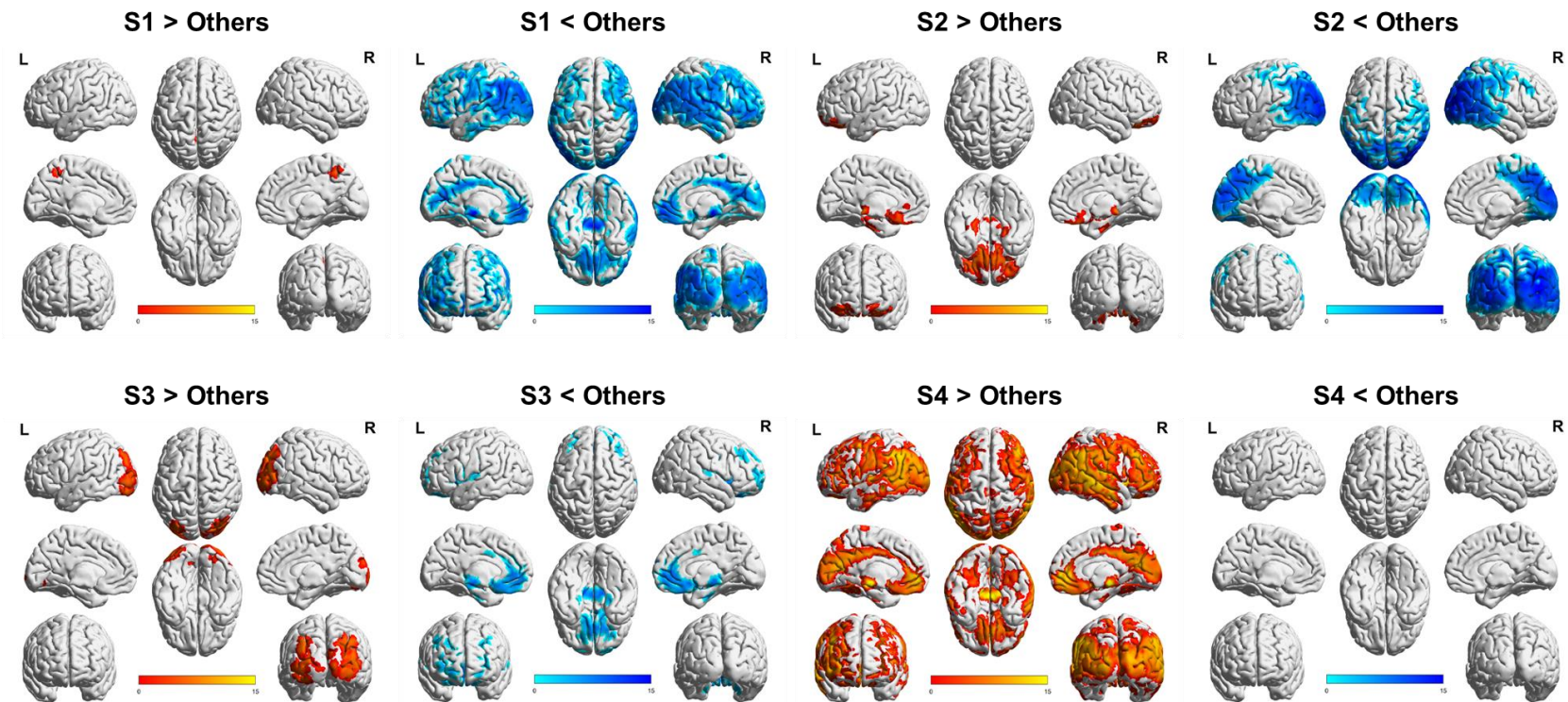


Figure 12. Spatial metabolism pattern on FDG PET with cerebellar normalization of MCI subtypes showing the differences between one subtype and all other subtypes

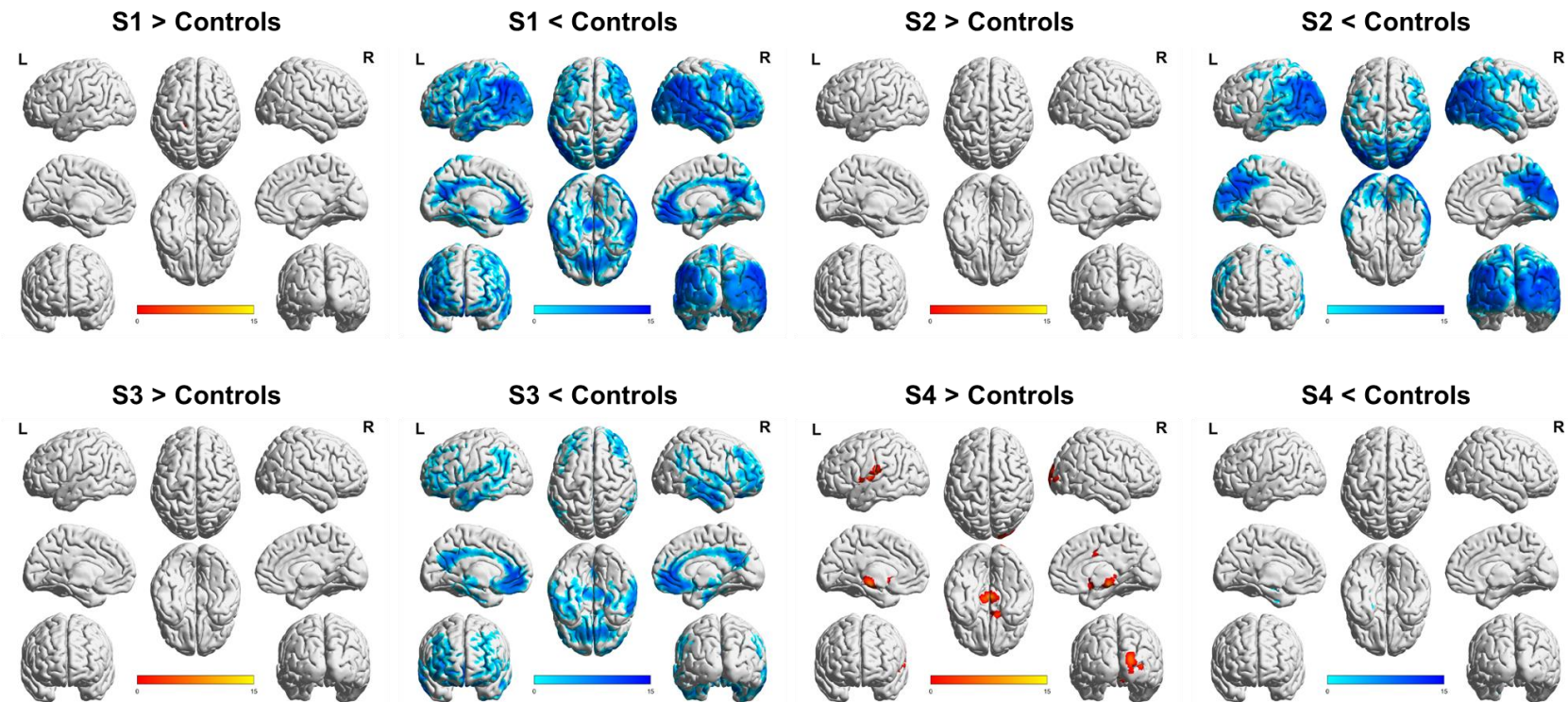


Figure 13. Spatial metabolism pattern on FDG PET with cerebellar normalization of MCI subtypes with the exclusion of overlapping individuals showing the differences between subtypes and controls
 Overlapping MCI individuals, who were included in both MCI and AD cohorts due to their conversion from MCI to AD in follow-up visits, were excluded.

3.5 Clinical and biological characterization in MCI subtypes

We compared demographic, cognitive, and biomarker variables between the subtypes in MCI (**Table 5 and Figure 14**). CDR-SB scores were significantly different across the MCI subtypes ($P < 0.001$); CDR-SB score was highest in MCI-S1 but lowest in MCI-S3. MCI-S1 included the highest frequency of males, and the individuals in MCI-S1 had more amyloid burden, more hippocampal atrophy, but less overall tau burden, which was a similar pattern observed in AD subtypes. In addition, individuals in MCI-S1 tended to have lower memory, executive, language, and visuospatial scores. Individuals in MCI-S2 were younger and more educated. MCI-S2 tended to have more overall tau burden and had better language, but worse executive and visuospatial scores, which was similarly observed in S2 of AD. Individuals in MCI-S3 tended to have a less amyloid burden and higher executive and visuospatial scores like S3 in AD. Finally, MCI-S4 included the highest frequency of female and less educated individuals as observed in S4 of AD. Individuals in MCI-S4 had the most favorable clinical presentation: higher Montreal Cognitive Assessment (MOCA), Mini-Mental State Examination (MMSE), memory, executive, language, and visuospatial scores. MCI-S4 had the lowest amyloid deposit and less hippocampal atrophy among the subtypes but had more tau burden.

Table 5. Clinical and biological characterizations of subtypes in MCI

	CN	MCI (n=1761)				P-value, global comparison (S1, S1, S3, and S4)
		S1	S2	S3	S4	
Demographics						
n (%)	1021	252 (14%)	202 (11%)	596 (34%)	711 (40%)	
Age, years (n=1759)	75.7 (6.2)	76.7 (6.6)	73.7 (7.7)	75.3 (7.7)	73.7 (7.9)	< 0.001 ***
Sex, female (%) (n=1761)	43%	8%	18%	30%	63%	< 0.001 ***
Education, years (n=1761)	16.3 (2.8)	16.4 (2.8)	16.6 (2.7)	16.0 (2.8)	15.7 (2.8)	< 0.001 ***
Cognition						
CDR-SB (n=1761)	0.1 (0.4)	1.8 (1.1)	1.7 (1.2)	1.5 (0.9)	1.6 (1.0)	< 0.001 ***
MMSE (n=1758)	29.0 (1.2)	27.4 (2.2)	27.3 (2.3)	27.6 (2.0)	27.7 (2.0)	0.037 *
MOCA (n=965)	25.8 (2.5)	23.3 (2.7)	23.1 (3.5)	23.4 (3.1)	23.7 (3.5)	0.455
PHC-MEM (n=1741)	0.86 (0.51)	0.01 (0.48)	0.08 (0.64)	0.11 (0.52)	0.23 (0.63)	< 0.001 ***
PHC-EXF (n=1738)	0.74 (0.48)	0.23 (0.56)	0.21 (0.61)	0.35 (0.54)	0.41 (0.55)	< 0.001 ***
PHC-LAN (n=1741)	0.80 (0.49)	0.33 (0.50)	0.37 (0.52)	0.36 (0.52)	0.44 (0.53)	0.009 **
PHC-VSP (n=1000)	0.10 (0.31)	-0.02 (0.43)	-0.10 (0.50)	0.04 (0.35)	0.02 (0.36)	0.032 *
Biomarker						

AV45 PET, SUVR (n=841)	1.12 (0.19)	1.20 (0.23)	1.23 (0.23)	1.20 (0.23)	1.19 (0.22)	0.487
APOE4 carrier (%) (n=1703)	27%	44%	52%	44%	48%	0.026 *
CSF Abeta, pg/ml (n=742)	1211 (439)	763 (315)	886 (442)	941 (441)	1052 (449)	< 0.001 ***
CSF t-tau, pg/ml (n=742)	245 (91)	245 (101)	291 (127)	274 (106)	312 (152)	< 0.001 ***
CSF p-tau, pg/ml (n=742)	23 (10)	24 (11)	29 (14)	26 (12)	30 (17)	< 0.001 ***
HV/ICV, cm ³ /mm ³ (n=1228)	4.78 (0.68)	3.85 (0.67)	4.22 (0.71)	4.19 (0.72)	4.73 (0.87)	< 0.001 ***

The data are expressed as percentages or means with standard deviation in parentheses. Missing values are excluded.

* p < 0.05, ** p < 0.01, *** p < 0.001

(continued)

	P-value, pair-wise comparisons					
	S1 vs S2	S1 vs S3	S1 vs S4	S2 vs S3	S2 vs S4	S3 vs S4
Demographics						
Age, years	< 0.001 ***	0.076	< 0.001 ***	0.054	1.000	< 0.001 ***
Education, years	0.814	0.160	0.002 **	0.017 *	< 0.001 ***	0.204
Cognition						
CDR-SB	0.532	< 0.001 ***	< 0.001 ***	0.081	0.239	0.854
MMSE	1.000	0.254	0.076	0.277	0.097	0.905
MOCA	0.991	0.973	0.657	0.892	0.542	0.760
PHC-MEM	0.585	0.137	< 0.001 ***	0.954	0.008 **	< 0.001 ***
PHC-EXF	0.948	0.037 *	< 0.001 ***	0.010 *	< 0.001 ***	0.132
PHC-LAN	0.800	0.871	0.024 *	0.985	0.418	0.034 *
PHC-VSP	0.414	0.361	0.558	0.006 **	0.015 *	0.970
Biomarker						
AV45 PET, SUVR	0.853	1.000	0.904	0.748	0.441	0.865
CSF Abeta, pg/ml	0.244	0.003 **	< 0.001 ***	0.771	0.017 *	0.012 *
CSF t-tau, pg/ml	0.103	0.235	< 0.001 ***	0.761	0.587	0.003 **
CSF p-tau, pg/ml	0.102	0.347	< 0.001 ***	0.634	0.840	0.009 **

HV/ICV, cm ³ /mm ³	< 0.001 ***	< 0.001 ***	< 0.001 ***	0.979	< 0.001 ***	< 0.001 ***
--	-------------	-------------	-------------	-------	-------------	-------------

* p < 0.05, ** p < 0.01, *** p < 0.001

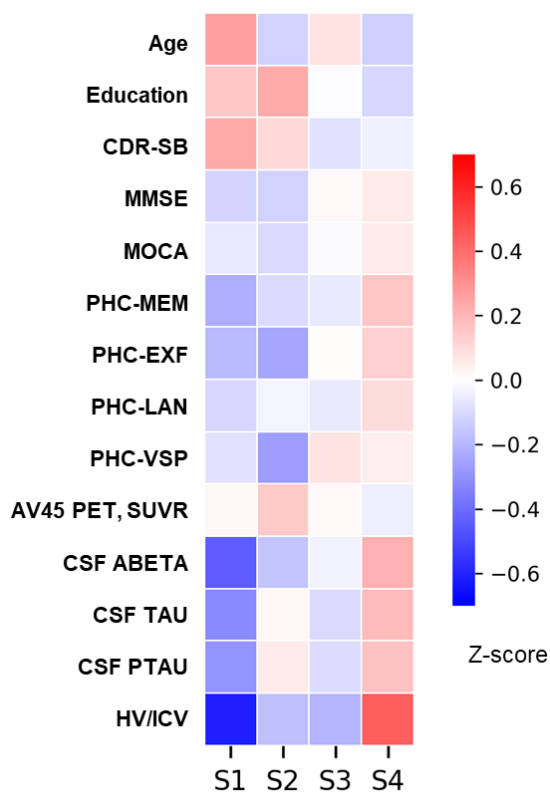


Figure 14. Heatmap of clinical and biological characteristics of MCI subtypes
Averaged Z-scores of demographic, cognitive, and biomarker variables in each subtype are visualized in the heatmap.

3.6 Prognosis prediction of subtypes for conversion from MCI to AD

Conversion rate and time to conversion from MCI to AD were analyzed across the MCI subtypes. There was a significant difference in the frequency of MCI to AD conversion within a 2-year follow-up across the subtypes ($P < 0.001$). MCI to AD conversion was observed more frequently in S1 (23.4%, 18/77) and S2 (35.7%, 25/70) compared to S3 (16.7%, 35/210) and S4 (12.7%, 31/245) (**Figure 15**). Individuals in S2 also had a significantly faster conversion from MCI to AD (S2 vs S3, 24.0 vs 35.3 months, $P = 0.047$; S2 vs S4, 24.0 vs 34.0 months, $P = 0.037$; **Table 6 and Figure 16**). Kaplan-Meier survival curves were constructed to compare the rate of conversion from MCI to AD across the subtypes (**Figure 17**). There was a significant difference in the risk of conversion across the subtypes (Log-rank $P < 0.0001$). S2 showed the highest risk of conversion from MCI to AD, followed by S1.

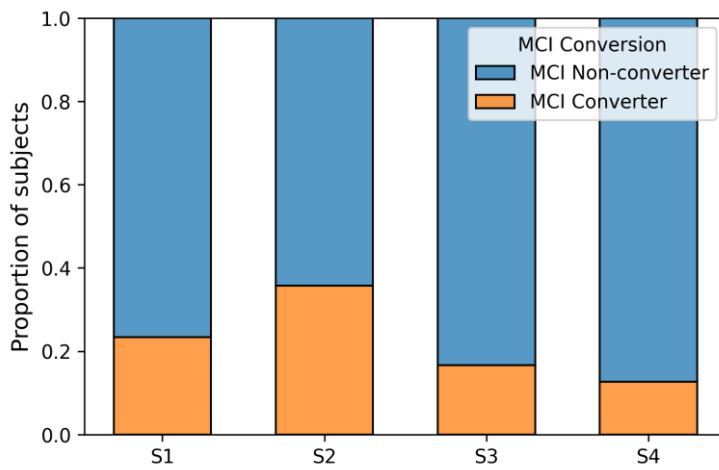


Figure 15. Frequency of MCI converter across subtypes

Table 6. Time to conversion from MCI to AD

	MCI at baseline				P-value, global	P-value, pair-wise comparisons					
	S1 (n=39)	S2 (n=40)	S3 (n=78)	S4 (n=83)		S1 vs S2	S1 vs S3	S1 vs S4	S2 vs S3	S2 vs S4	S3 vs S4
Time to conversion, months	34.2 (27.7)	24.0 (16.9)	35.3 (29.5)	34.0 (22.9)	0.019	0.213	0.997	1.000	0.047 *	0.037 *	0.991

The data are expressed as percentages or means with standard deviation in parentheses.

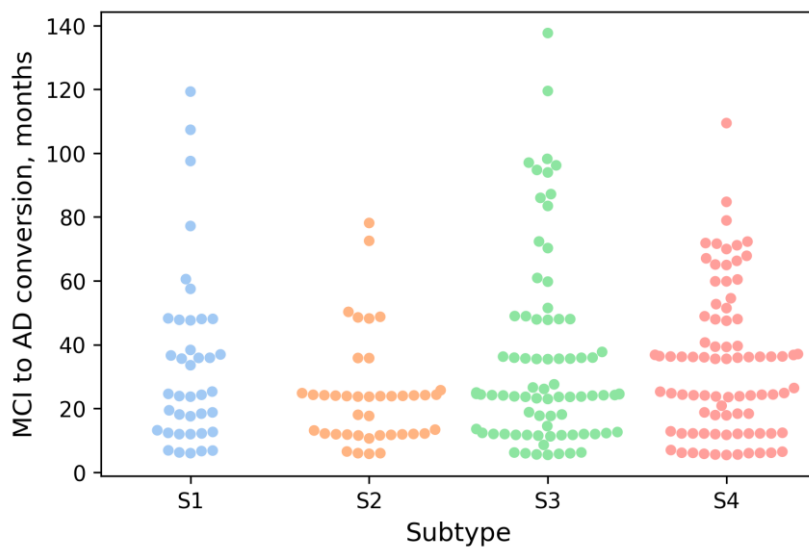


Figure 16. Time to conversion from MCI to AD

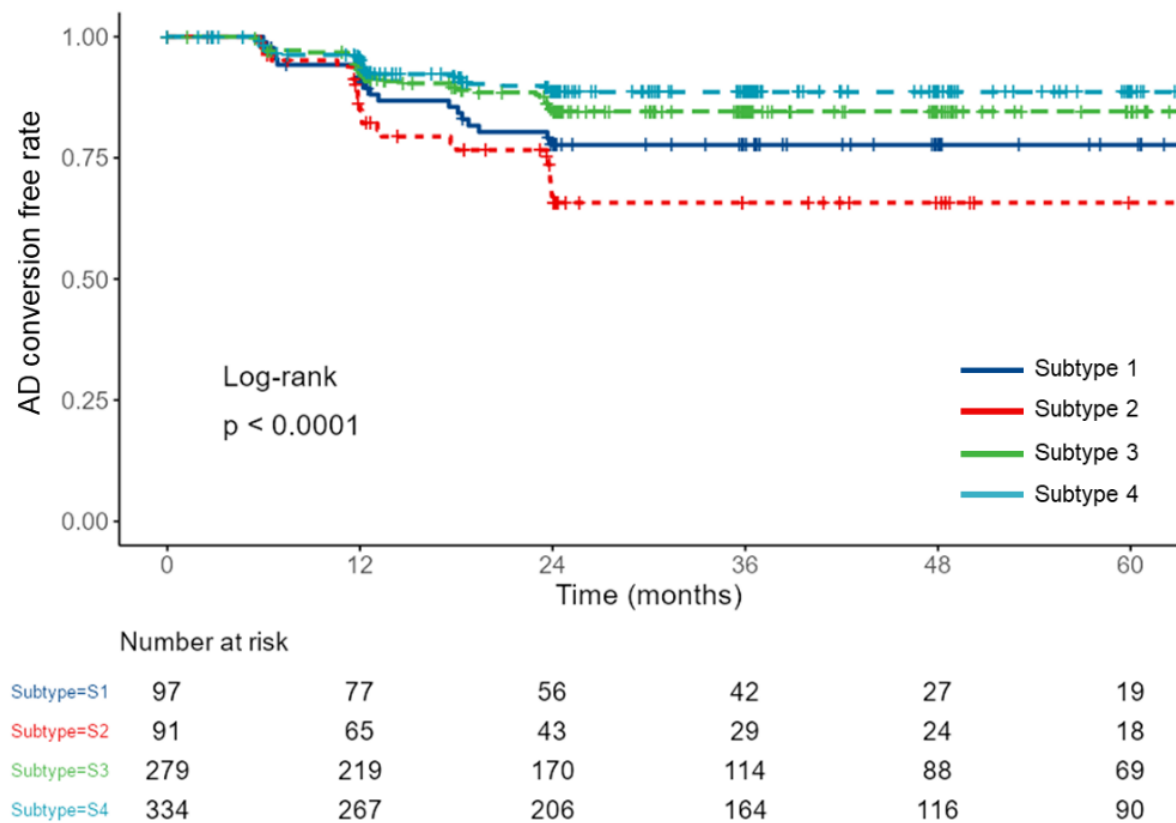
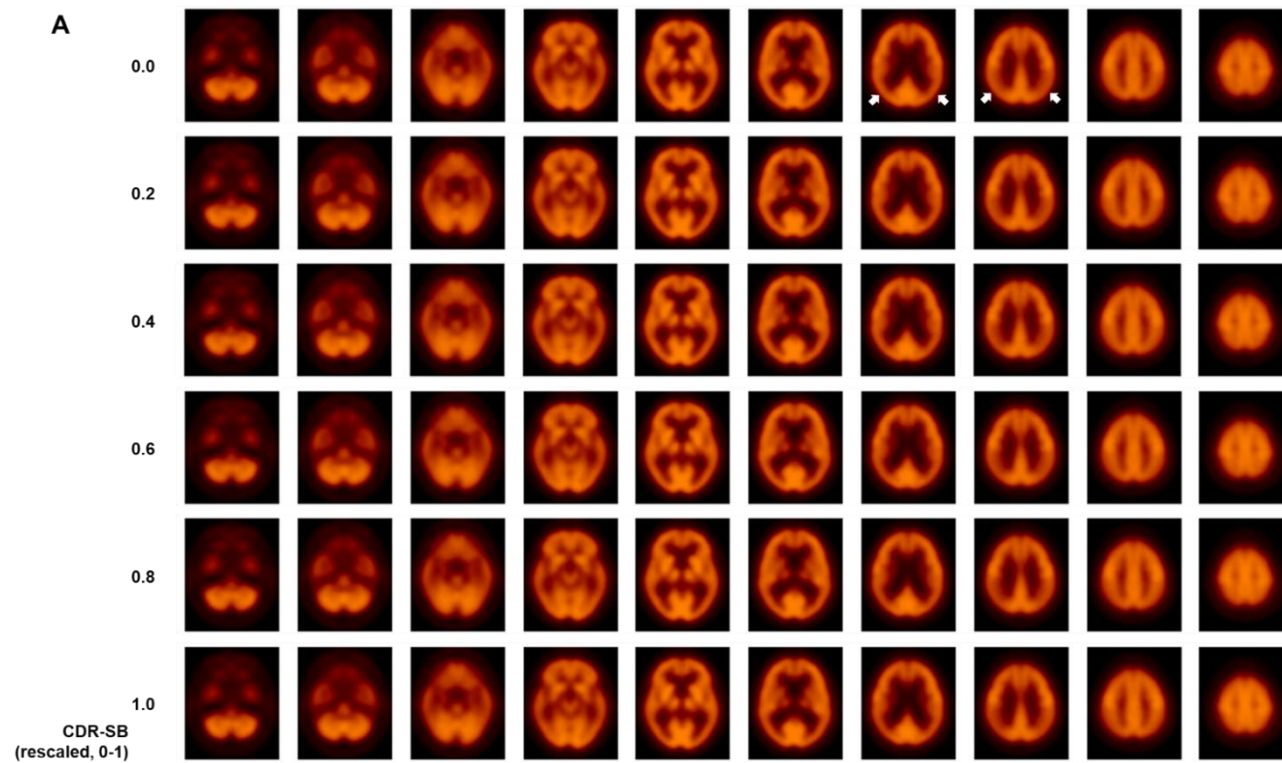
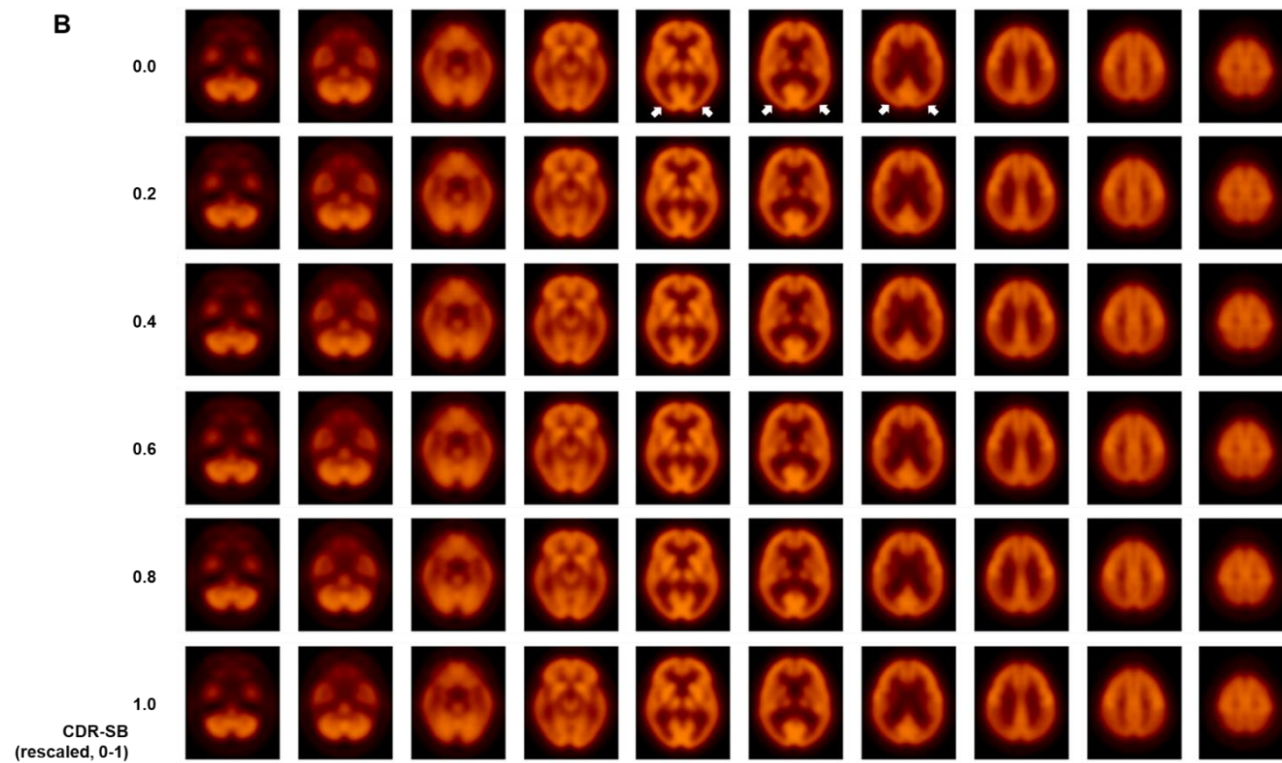


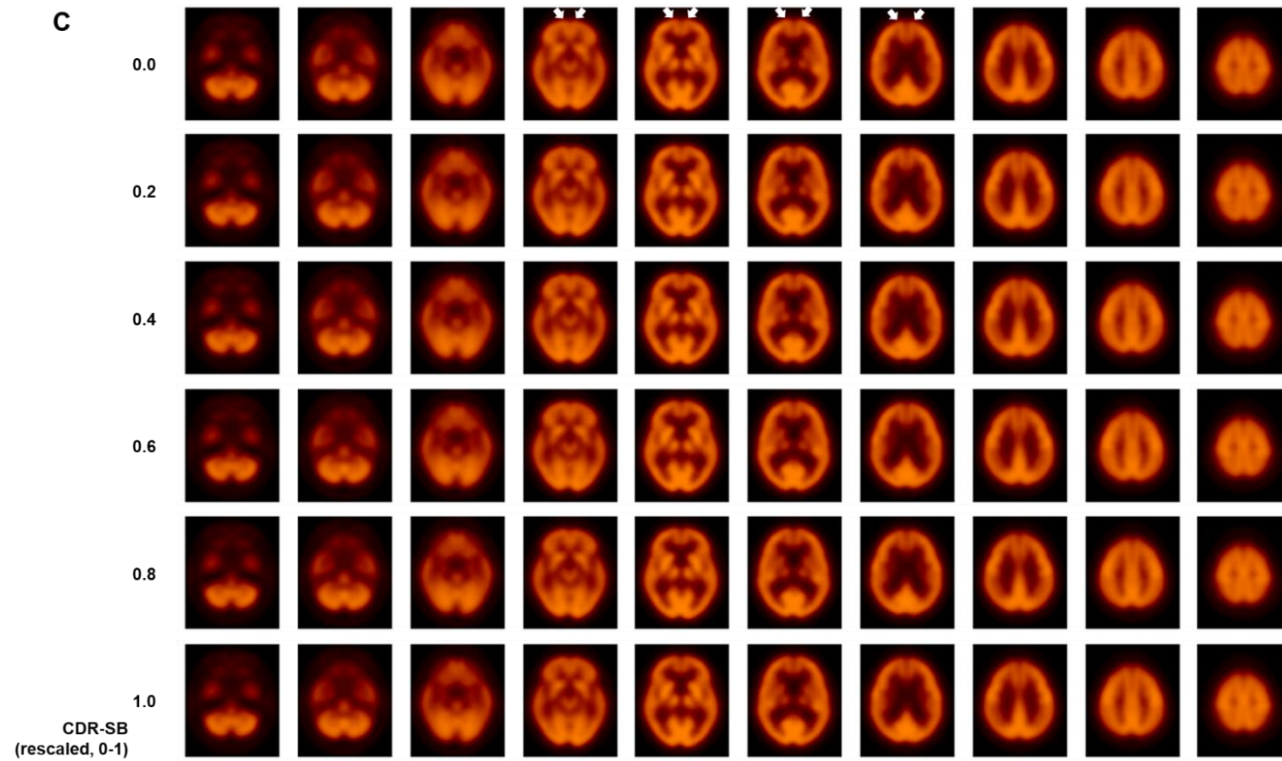
Figure 17. Kaplan-Meier survival curve for the conversion from MCI to AD

3.7 Generating FDG PET images of AD subtypes

We applied decoder layers of our cVAE model to generate FDG PET images and to compare subtype-specific metabolism patterns from the latent features across AD subtypes. The representative FDG brain PET images generated from the centroids of latent features of each subtype showed subtype-specific hypometabolism patterns as identified in our study (**Figure 18**): S1 (angular), S2 (occipital), S3 (orbitofrontal), and S4 (minimal). The intensities of voxels in subtype-specific hypometabolism regions decreased as the rescaled CDR-SB scores increased.







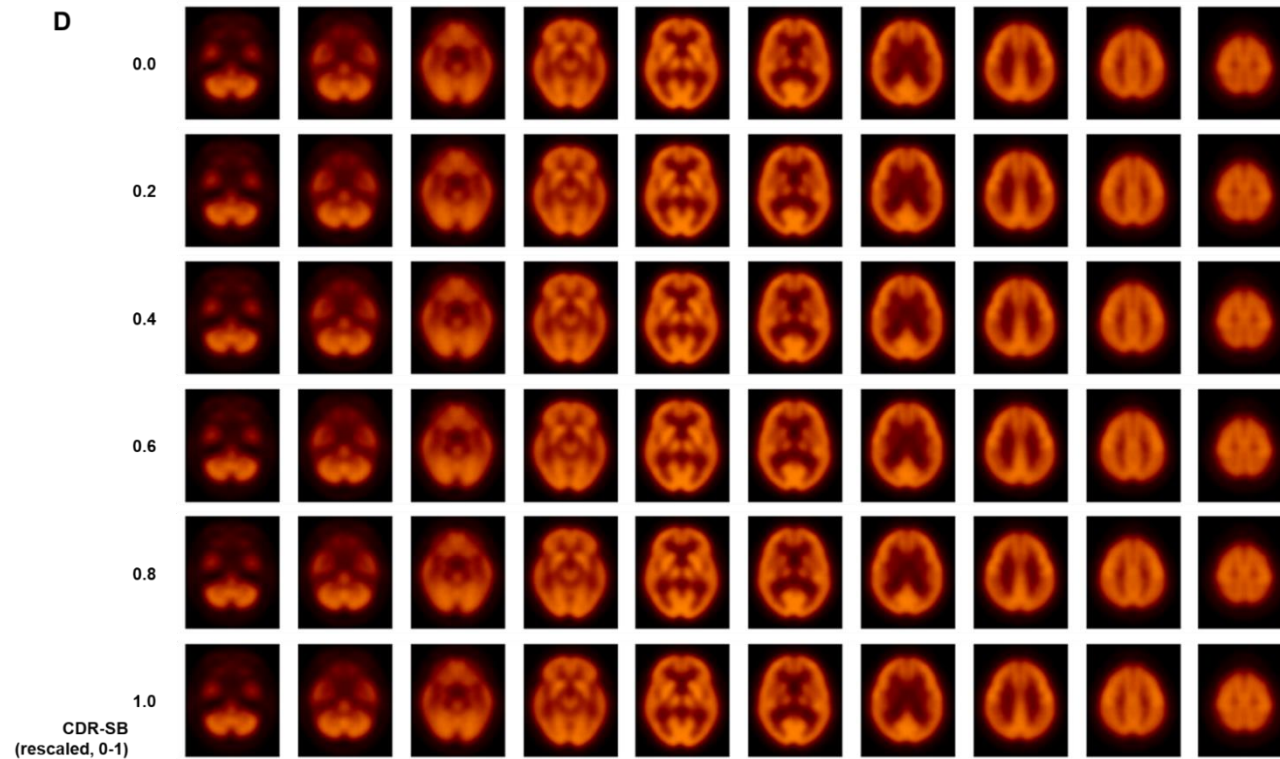


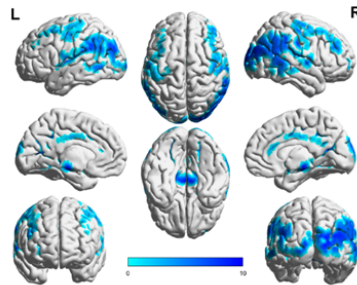
Figure 18. Generation of subtype-specific FDG brain PET images according to latent features with different CDR-SB scores by using a deep learning-based model

Representative generated images of (A) S1 (angular), (B) S2 (occipital), (C) S3 (orbitofrontal), and (D) S4 (minimal). Subtype-specific regions of hypometabolism are annotated with white arrows in the first row of (A)-(C).

4. Discussion

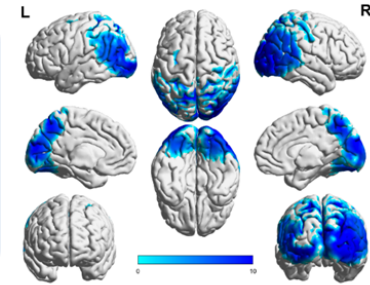
In this study, we identified distinct subtypes of spatial brain metabolism patterns with different clinicopathologic features in AD using deep learning-based FDG PET clusters (summarized in **Figure 19**). Our deep learning model was also successfully transferred to predict the prognosis of subtypes for conversion from MCI to AD.

S1: Angular



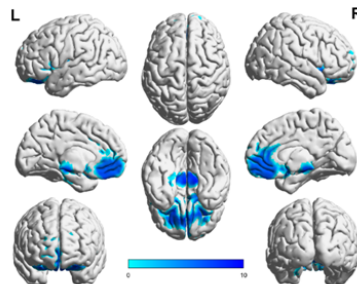
- Frequent in males
- More amyloid
- Less tau
- More hippocampal atrophy
- Cognitive decline in the earlier stage

S2: Occipital



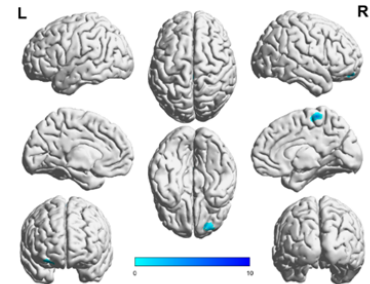
- Younger age
- More tau
- Less hippocampal atrophy
- Lower executive and visuospatial scores
- Faster conversion from MCI to AD

S3: Orbitofrontal



- Older age
- Less amyloid
- More hippocampal atrophy
- Higher executive and visuospatial scores

S4: Minimal



- Frequent in females
- Less amyloid
- More tau
- Less hippocampal atrophy
- Higher cognitive scores

Figure 19. Theoretical model summarizing AD subtypes on FDG PET

4.1 Limitations of previous subtyping approach

AD has been recognized as a heterogeneous disease with variable clinicopathologic profiles, such as age onset, clinical presentation, tau-related pathology, and brain atrophy [4-8]. For that reason, there were many efforts to reveal biologic subtypes of AD. Previous studies using multiple brain imaging, mostly brain MRI and tau PET, tried to identify AD subtypes in a hypothesis-driven or data-driven manner [9-11, 75]. However, the hypothesis-driven study design was limited to the prior definition of neuropathologic subtypes. As a result, the hypothesis-driven approach on neuroimaging may not provide a comprehensive description and cannot reflect heterogeneous clinicopathologic profiles of AD. Data-driven subtyping in previous studies was also limited by the confound of disease stage, demonstrating disease severity rather than a heterogeneous biologic profile.

4.2 Interpretation of results

We defined 4 distinct subtypes of hypometabolism pattern in AD: S1 (angular), S2 (occipital), S3 (orbitofrontal), and S4 (minimal). Our proposed AD subtypes on FDG PET show different patterns in regions of involvement from the previously reported 4 biologic subtypes (typical, hippocampal sparing, limbic predominant, and minimal atrophy AD)

based on neurodegeneration and tau pathology. Spatial patterns in S1 (angular), S2 (occipital), and S3 (orbitofrontal) obviously differ from previously described AD subtypes in other modalities, but only the S4 (minimal hypometabolism) showed similarity with minimal atrophy AD. In addition, whereas the previously described MRI-defined subtypes separate hippocampal sparing AD as an independent subgroup, both S2 (occipital) and S4 (minimal) showed less hippocampal atrophy than other subtypes, which implies both subtypes share the characteristics of hippocampal sparing. From this point of view, it can be considered that our AD subtypes based on spatial metabolism patterns belongs to a domain different from the previously-defined subtypes reported on other modalities.

Subtype-specific spatial metabolism pattern in the current study notably differs from previously reported FDG PET subtypes. In the study of Levin et al. [24], hypometabolic regions of their 3 subtypes (typical, limbic predominant, and cortical-predominant) overlapped in parietotemporal areas and frontal lobe with different effect sizes, and the “minimal” subtype was not included. They employed a hierarchical clustering approach for voxel-wise FDG PET profile which has a limitation in removing the effect of disease progression from clustering. The clusters defined in the studies of Levin et al. may have rather

represented a severity dimension, and their subtypes actually showed a significant difference in MMSE scores across the subtypes. In contrast, AD subtypes in the current study shared typical AD patterns but showed distinct subtype-specific spatial metabolism patterns which have few overlapping regions. Our deep learning-based approach is intended to remove the severity dimension from clustering, and CDR-SB and MMSE scores were not significantly different across the subtypes.

The considerable disagreement in subtypes between previous reports and the current study is likely due to the presence or absence of severity dimension along with the AD subtypes. In the recent subtype study on tau PET, Vogel et al. [10] applied an algorithm by combining clustering with disease progression modeling [76] to overcome the limitation from the disease severity dimension, and their identified 4 subtypes (limbic, medial-temporal lobe-sparing, posterior, and lateral temporal) showed a completely different spatial pattern of tau deposition compared to the previously identified subtypes. From this point of view, we can reasonably infer that our AD subtypes represent distinct spatial metabolism patterns independent from the dimension of disease severity.

In the previous reports, female patients more frequently had limbic-predominant AD, and male patients more frequently had hippocampal-sparing AD [9]. In our study, S1 (angular) and S3

(orbitofrontal) were more hypometabolic in the limbic area, and S2 (occipital) and S4 (minimal) had more hippocampal-sparing patterns compared to other subtypes. In our result, male patients were more frequent in S1 (95%), and female patients were more frequent in S4 (72%). The difference of dimension across the AD subtypes between our result (hidden representation without severity dimension) and previous studies (with severity dimension) may be a careful explanation for disagreement of a demographic pattern between previous and current studies.

For the S2 (occipital), we showed that posterior-predominant hypometabolism and visuospatial impairment are prominent, which is in line with the “posterior” tau deposition subtype characterized by visuospatial impairment on recent tau PET study [10]. In addition, we showed that the S2 has the faster conversion from MCI to AD, and this result corresponds to a faster decline in metabolism in the AD subgroup of visuospatial function impairment reported in a recent trajectory study on FDG PET [25].

In our study, the orbitofrontal region was classified as an independent group with distinct spatial metabolism patterns. The orbitofrontal cortex is associated with apathy, non-memory-related behavior, and agitation in AD patients [77-79]. Even though previous

studies in AD subtypes did not focus on the orbitofrontal region, S3 (orbitofrontal) might be an independent biologic subtype with distinct spatial metabolism patterns in AD.

Minimal subtype-specific hypometabolism was observed in S4 (minimal). S4 had the lowest level of hippocampal atrophy, but CSF p-tau and t-tau levels were higher than in other subtypes. In addition, S4 had the lowest level of education and a frequent prevalence of females, and the MCI-S4 had younger age than in other subtypes. This clinicopathologic pattern (tau, education) is in line with the results observed in minimal atrophy AD defined in the previous studies [9]. Based on these observations, we may propose a hypothesis that the neurodegeneration at a molecular level (tau-related pathology) may be sufficient to give dementia symptoms with minimal hypometabolism and atrophy [80]. Co-existing risk factors such as female sex and lower level of education may counteract with a compensatory biologic mechanism of AD resulting lowest level of cognitive reserve in the “minimal” subtype, which is related to the concepts of brain resilience and resistance [26, 81, 82].

In our study, we used both FDG PET images with cerebellar normalization and global normalization for the generation of SPM images to visualize spatial metabolism patterns. Subtype-specific

patterns were consistent in both analyses, but commonly shared regions of hypometabolism across the subtypes, which is a typical AD pattern, were extended in the analysis with cerebellar normalization. More reduced regions of hypermetabolism were also observed in cerebellar normalization. Cerebellar and global normalization are frequently used methods for the detection and differentiation of dementia in FDG PET studies. Earlier studies showed that the glucose metabolism in the cerebellum is less affected in AD patient compared to global cerebral glucose metabolism [83], which result in a difference in SPM images between cerebellar and global normalization methods. While global normalization is superior for differential diagnosis between types of dementia, cerebellar normalization is known to be superior in identifying dementia patients in comparison to control subjects and in early detection of dementia [84].

4.3 Strength of our deep learning-based clustering approach

To overcome the limitations of the previous subtyping approach, in this study, we applied the deep learning-based clustering method to reveal biologic subtypes of AD on FDG PET by using cVAE. VAE could generate images in an unsupervised manner and reduce image dimension into a lower-dimensional feature [60]. The unsupervised approach could

help to make new discoveries of unappreciated subtypes and it is expected to overcome the limitations of prior definition on hypothesis-driven studies. Compared to early studies using traditional dimensionality reduction methods such as PCA, autoencoder has its advantage in the use of a non-linear function that makes superior performance in learning patterns from high-dimensional image data such as neuroimaging. Moreover, VAE could overcome the overfitting problem from the traditional autoencoder.

The greatest strength of our study is the application of the cVAE model which has a structure of additional input information for the condition of each image [63, 85]. For instance, cVAE applied on MNIST database with condition information of number label, the latent vector contained the image feature of handwriting style, which could be applied for the number image generation with different handwriting style. In addition, cVAE was also successfully applied on FDG brain PET for predicting aging of brain metabolic topography [67]. From this idea, we hypothesized that when we give AD severity information (CDR-SB) as an input condition of FDG PET images on the cVAE model, the latent features could reflect hidden heterogeneity information of AD other than disease severity. When VAE without an additional input condition was applied for the subjects in our study, FDG PET images were clustered

across the disease severity domain and showed spatial brain metabolism patterns of disease progression (**Figure 20**). When cVAE with CDR-SB condition was applied, there was no significant difference in CDR-SB among the AD subtypes ($P = 0.609$), and the brain metabolism pattern showed distinct features with different clinicopathologic profiles.

The deep learning-based model has the advantage of being easy to transfer trained models to other cohorts, as shown in our study. Since our deep learning network is a generative model, it can also be used to generate a new image from the input image and information. This approach is also expected to be readily applied to the other disease domain with heterogeneous clinicopathologic profiles.

Our deep learning-based approach has a unique novelty. First, this is the first study to remove the effect of disease progression (severity dimension) to reveal the hidden representations of FDG PET spatial metabolism pattern in AD. Second, this is the first approach to apply cVAE in AD subtype studies not only on FDG PET but also on other modalities. Third, AD subtype study on FDG PET is very rare.

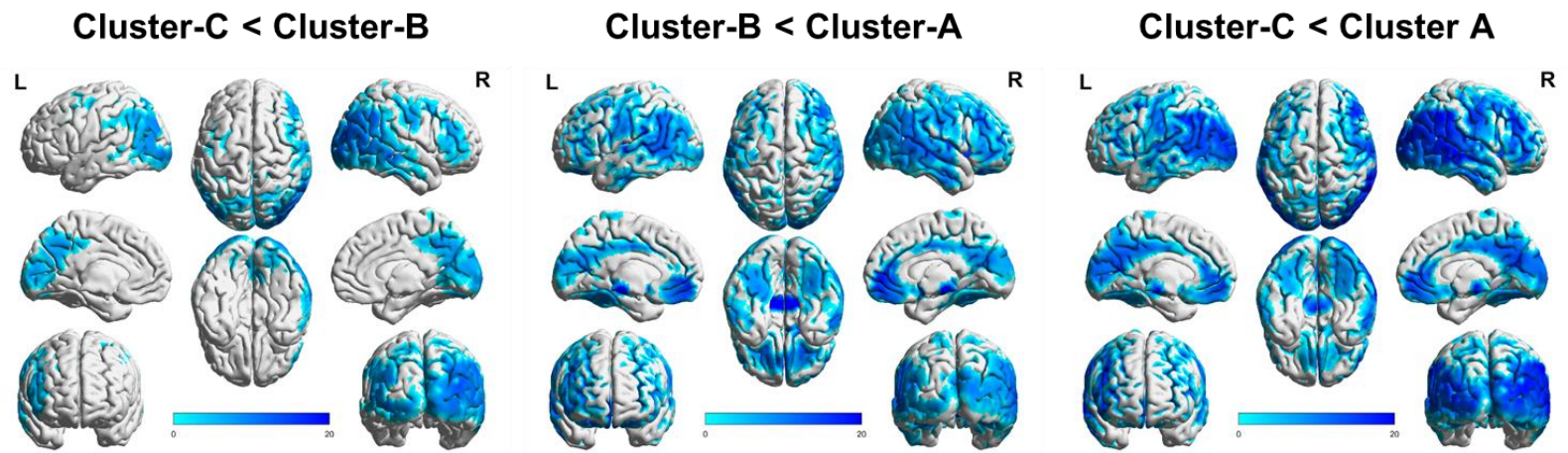


Figure 20. Spatial metabolism of clusters using VAE without CDR-SB condition

4.4 Strength of our deep learning-based AD subtypes

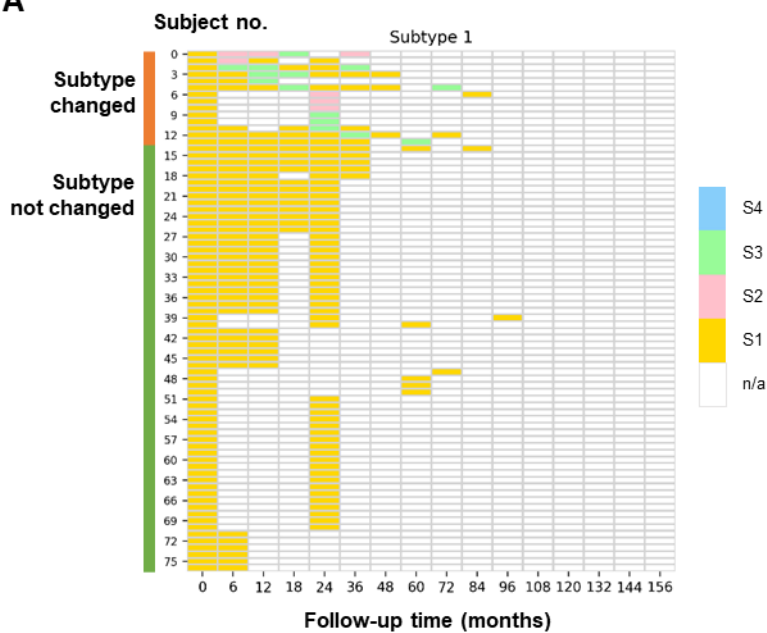
In our study, the deep learning model could predict the distinct subtype of brain metabolism patterns in AD with different clinicopathologic features. The model was also successfully transferred to MCI subjects. The spatial metabolism pattern in MCI subtypes was almost similar to those of AD. The demographic, cognitive, and biomarker profiles of MCI subtypes were similar but with some differences from those of AD. These findings imply that subtypes may have distinct trajectories of disease progression and different features in earlier and late stages.

Until now there is no curative treatment for AD, but only to postpone the disease progression. Therefore, a biomarker for the early detection of MCI converter and the prediction of prognosis and disease course is very important to allow patients to have early treatment before the stage of AD. Predicting who will convert from MCI to AD has been an important problem, and there were many previous studies for the development of biomarkers and models to predict MCI to AD conversion [86-90]. In the clinic, it is difficult to differentiate biologic subtypes in AD patients and to predict the prognosis of individual patients since the symptoms and clinical characteristics of AD patients are diverse and complex. In our study, the deep learning model was transferred to FDG PET of MCI patients, and the predicted MCI subtypes revealed

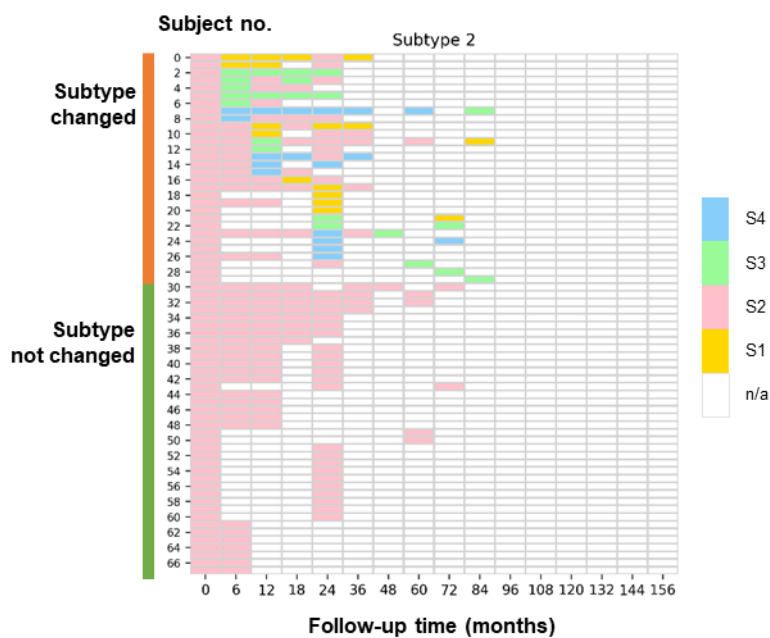
significantly different prognoses of MCI to AD conversion. Our deep learning-based models for the prediction of AD and MCI subtypes on FDG PET have great novelty and strong clinical implications for the early and individualized diagnosis and treatment that could be applied in precision medicine. In addition, according to our preliminary result, most AD and MCI patients maintained their subtypes from baseline to follow-up, but some showed changes and fluctuation of their subtypes (**Figure 21**). The trajectory of spatial metabolism pattern across the subtypes in MCI and AD and their relationship to patient prognosis should be investigated in the future.

In clinical routine, differentiating subtype-specific metabolism patterns on FDG PET only with visual interpretation is very difficult which is subject to inter-observer variability and requires highly trained experts. Because of diffuse overlapping regions of hypometabolism (typical AD pattern) across the subtypes, it is difficult to distinguish subtype-specific hypometabolic regions only with visual interpretation as shown in our generated images and SPM images with cerebellar normalization. Since our deep learning model could differentiate distinct subtypes, we suggest that our model be used to help physicians' visual interpretation of FDG PET to make a precise and accurate diagnosis of subtypes, which is a great advantage of the deep learning model.

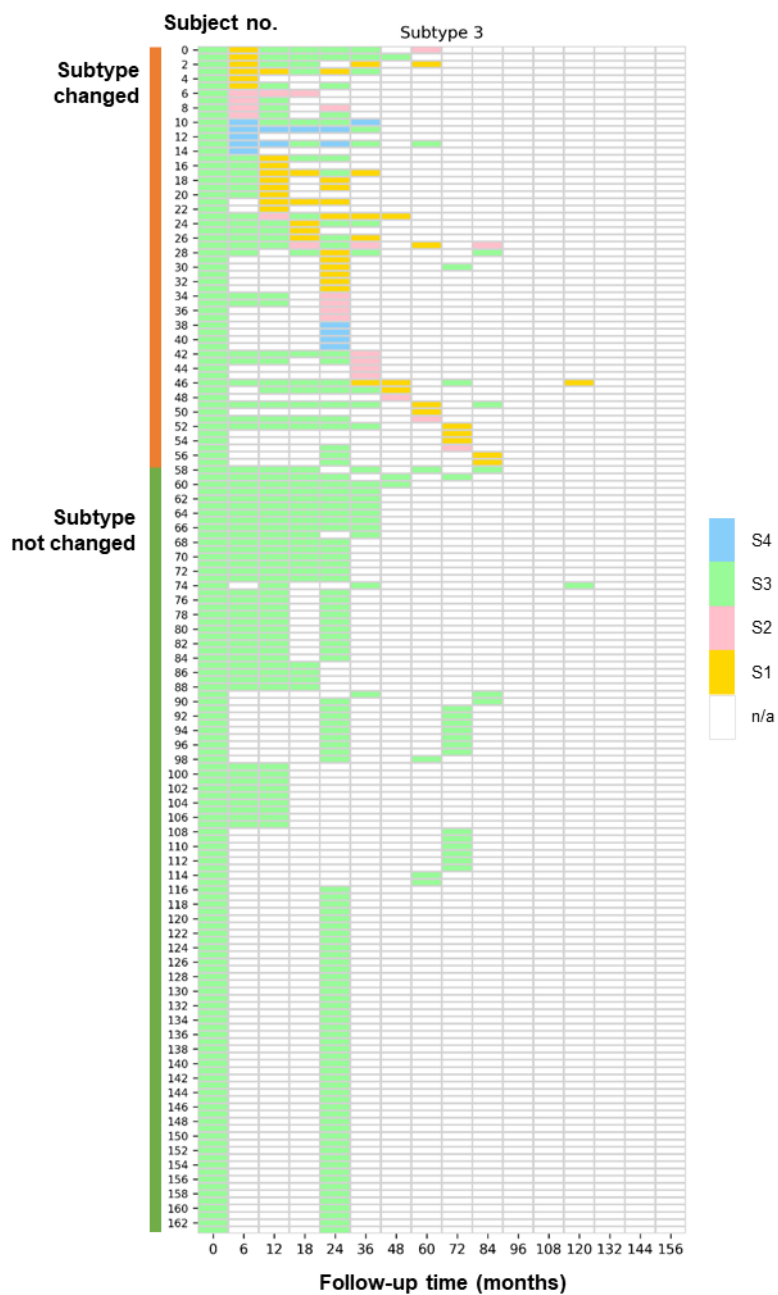
A



B



C



D

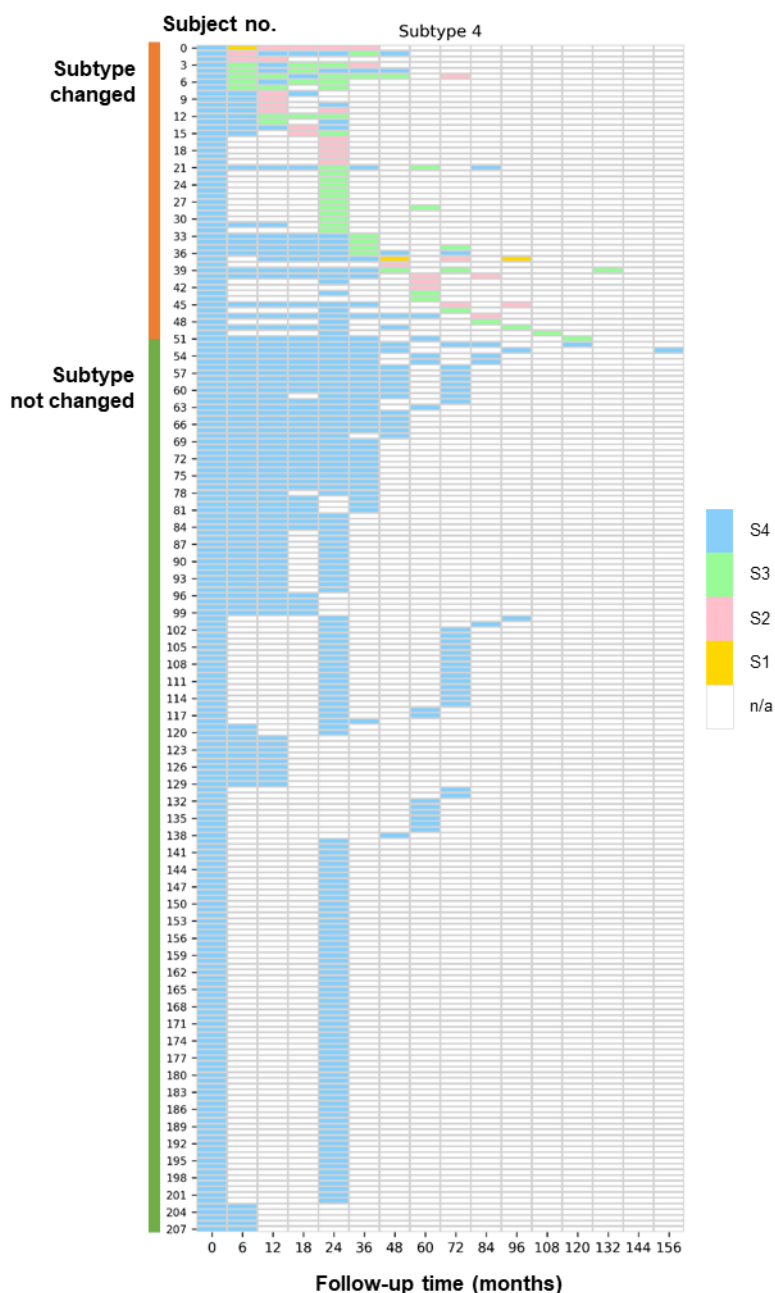


Figure 21. The trajectory of spatial brain metabolism pattern across the subtypes Deep learning-based model predicted subtypes of individual subjects with AD and MCI at baseline and follow-up visits are visualized in the heatmap. Subjects were (A) S1, (B) S2, (C) S3, and (D) S4 on baseline FDG PET. Each subtype is marked with a different color. Subjects are sorted in the order in which the subtypes changed or not during the follow-up. n/a, not available PET

4.5 Limitations and future directions

Some limitations should be noted. First, the deep learning model was trained by using FDG PET with AD from the single database (ADNI). A validation study using external AD datasets is mandatory. Second, the deep learning model trained in the AD cohort was directly transferred to the MCI cohort. However, MCI is a heterogeneous syndrome resulting from AD as well as non-AD and non-neurodegenerative conditions [91-97], which implies that some of the patients in our MCI cohort may be originated from other disease populations different from AD. Therefore, a validation study using external MCI datasets is also mandatory. Third, a longitudinal study may be warranted to investigate the subtype-specific trajectory of spatial metabolism changes and clinicopathologic profiles. Finally, quantitative analysis of regions of interest of subtypes could give more evidence to our results of spatial metabolism patterns. We are planning the next step for the longitudinal study and quantitative analysis.

5. Conclusion

We could identify distinct subtypes in AD with different brain pathologies and clinical profiles. Also, our model was successfully transferred to predict the prognosis of subtypes for conversion from MCI to AD. Our results suggest that distinct AD subtypes on FDG PET may have implications for the individual clinical outcomes and provide a clue to understanding a broad spectrum of AD in terms of pathophysiology.

References

1. Blennow, K. and H. Zetterberg, *Biomarkers for Alzheimer's disease: current status and prospects for the future*. J Intern Med, 2018. **284**(6): p. 643-663.
2. Crane, P.K., et al., *Incidence of cognitively defined late-onset Alzheimer's dementia subgroups from a prospective cohort study*. Alzheimers Dement, 2017. **13**(12): p. 1307-1316.
3. Scheltens, N.M.E., et al., *Cognitive subtypes of probable Alzheimer's disease robustly identified in four cohorts*. Alzheimers Dement, 2017. **13**(11): p. 1226-1236.
4. Friedland, R.P., et al., *NIH conference. Alzheimer disease: clinical and biological heterogeneity*. Ann Intern Med, 1988. **109**(4): p. 298-311.
5. Noh, Y., et al., *Anatomical heterogeneity of Alzheimer disease: based on cortical thickness on MRIs*. Neurology, 2014. **83**(21): p. 1936-1944.
6. Dujardin, S., et al., *Tau molecular diversity contributes to clinical heterogeneity in Alzheimer's disease*. Nature medicine, 2020. **26**(8): p. 1256-1263.
7. Wesseling, H., et al., *Tau PTM profiles identify patient heterogeneity and stages of Alzheimer's disease*. Cell, 2020.

- 183(6):** p. 1699-1713. e13.
8. Holzer, M., et al., *Abnormally phosphorylated tau protein in Alzheimer's disease: heterogeneity of individual regional distribution and relationship to clinical severity*. Neuroscience, 1994. **63(2):** p. 499-516.
 9. Ferreira, D., A. Nordberg, and E. Westman, *Biological subtypes of Alzheimer disease: A systematic review and meta-analysis*. Neurology, 2020. **94(10):** p. 436-448.
 10. Vogel, J.W., et al., *Four distinct trajectories of tau deposition identified in Alzheimer's disease*. Nat Med, 2021. **27(5):** p. 871-881.
 11. Ferreira, D., et al., *Distinct subtypes of Alzheimer's disease based on patterns of brain atrophy: longitudinal trajectories and clinical applications*. Scientific reports, 2017. **7(1):** p. 1-13.
 12. Mosconi, L., et al., *Multicenter standardized 18F-FDG PET diagnosis of mild cognitive impairment, Alzheimer's disease, and other dementias*. Journal of nuclear medicine, 2008. **49(3):** p. 390-398.
 13. Juh, R., et al., *Different metabolic patterns analysis of Parkinsonism on the 18F-FDG PET*. European journal of radiology, 2004. **51(3):** p. 223-233.

14. Meyer, P.T., et al., *18F-FDG PET in parkinsonism: differential diagnosis and evaluation of cognitive impairment*. Journal of Nuclear Medicine, 2017. **58**(12): p. 1888-1898.
15. Herholz, K., et al., *Evaluation of a calibrated 18F-FDG PET score as a biomarker for progression in Alzheimer disease and mild cognitive impairment*. Journal of Nuclear Medicine, 2011. **52**(8): p. 1218-1226.
16. Landau, S., et al., *Comparing predictors of conversion and decline in mild cognitive impairment*. Neurology, 2010. **75**(3): p. 230-238.
17. Huang, C., et al., *Metabolic abnormalities associated with mild cognitive impairment in Parkinson disease*. Neurology, 2008. **70**(16 Part 2): p. 1470-1477.
18. Huang, C., et al., *Metabolic brain networks associated with cognitive function in Parkinson's disease*. Neuroimage, 2007. **34**(2): p. 714-723.
19. Shaffer, J.L., et al., *Predicting cognitive decline in subjects at risk for Alzheimer disease by using combined cerebrospinal fluid, MR imaging, and PET biomarkers*. Radiology, 2013. **266**(2): p. 583-591.
20. Kljajevic, V., et al., *Distinct pattern of hypometabolism and*

- atrophy in preclinical and predementia Alzheimer's disease.* Neurobiology of aging, 2014. **35**(9): p. 1973-1981.
21. Besson, F.L., et al., *Cognitive and brain profiles associated with current neuroimaging biomarkers of preclinical Alzheimer's disease.* Journal of Neuroscience, 2015. **35**(29): p. 10402-10411.
 22. Laforce Jr, R., et al., *Molecular imaging in dementia: Past, present, and future.* Alzheimer's & Dementia, 2018. **14**(11): p. 1522-1552.
 23. Vanhoutte, M., et al., *18F-FDG PET hypometabolism patterns reflect clinical heterogeneity in sporadic forms of early-onset Alzheimer's disease.* Neurobiology of aging, 2017. **59**: p. 184-196.
 24. Levin, F., et al., *Data-driven FDG-PET subtypes of Alzheimer's disease-related neurodegeneration.* Alzheimers Res Ther, 2021. **13**(1): p. 49.
 25. Groot, C., et al., *Differential trajectories of hypometabolism across cognitively-defined Alzheimer's disease subgroups.* Neuroimage Clin, 2021. **31**: p. 102725.
 26. Ritchie, K. and J. Touchon, *Heterogeneity in senile dementia of the Alzheimer type: individual differences, progressive deterioration or clinical sub-types?* J Clin Epidemiol, 1992.

- 45(12): p. 1391-8.
27. Armstrong, R.A. and L. Wood, *The identification of pathological subtypes of Alzheimer's disease using cluster analysis*. Acta Neuropathol, 1994. **88**(1): p. 60-6.
 28. Armstrong, R.A., et al., *The use of multivariate methods in the identification of subtypes of Alzheimer's disease: a comparison of principal components and cluster analysis*. Dementia, 1996. **7**(4): p. 215-20.
 29. Armstrong, R.A., D. Nochlin, and T.D. Bird, *Neuropathological heterogeneity in Alzheimer's disease: a study of 80 cases using principal components analysis*. Neuropathology, 2000. **20**(1): p. 31-7.
 30. Murray, M.E., et al., *Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study*. Lancet Neurol, 2011. **10**(9): p. 785-96.
 31. Whitwell, J.L., et al., *[(18) F]AV-1451 clustering of entorhinal and cortical uptake in Alzheimer's disease*. Ann Neurol, 2018. **83**(2): p. 248-257.
 32. Poulakis, K., et al., *Heterogeneous patterns of brain atrophy in Alzheimer's disease*. Neurobiology of aging, 2018. **65**: p. 98-108.
 33. Persson, K., et al., *MRI-assessed atrophy subtypes in Alzheimer's*

- disease and the cognitive reserve hypothesis*. PLoS One, 2017. **12**(10): p. e0186595.
34. Bittner, D., et al., *[18F]FDG-PET in patients with Alzheimer's disease: marker of disease spread*. Dement Geriatr Cogn Disord, 2005. **19**(1): p. 24-30.
 35. Collette, F., et al., *Frontal hypometabolism does not explain inhibitory dysfunction in Alzheimer disease*. Alzheimer Dis Assoc Disord, 2002. **16**(4): p. 228-38.
 36. Meyer, F., et al., *Characterization of a temporoparietal junction subtype of Alzheimer's disease*. Human brain mapping, 2019. **40**(14): p. 4279-4286.
 37. Fodor, I.K., *A survey of dimension reduction techniques*. 2002, Lawrence Livermore National Lab., CA (US).
 38. Anowar, F., S. Sadaoui, and B. Selim, *Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE)*. Computer Science Review, 2021. **40**: p. 100378.
 39. Mao, K.Z., *Orthogonal forward selection and backward elimination algorithms for feature subset selection*. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2004. **34**(1): p. 629-634.

40. Sutter, J.M. and J.H. Kalivas, *Comparison of forward selection, backward elimination, and generalized simulated annealing for variable selection*. Microchemical journal, 1993. **47**(1-2): p. 60-66.
41. Breiman, L., *Random forests*. Machine learning, 2001. **45**(1): p. 5-32.
42. Wold, S., K. Esbensen, and P. Geladi, *Principal component analysis*. Chemometrics and intelligent laboratory systems, 1987. **2**(1-3): p. 37-52.
43. Abdi, H. and L.J. Williams, *Principal component analysis*. Wiley interdisciplinary reviews: computational statistics, 2010. **2**(4): p. 433-459.
44. Yong, A.G. and S. Pearce, *A beginner's guide to factor analysis: Focusing on exploratory factor analysis*. Tutorials in quantitative methods for psychology, 2013. **9**(2): p. 79-94.
45. Hermansen, F. and A.A. Lammertsma, *Linear dimension reduction of sequences of medical images: III. Factor analysis in signal space*. Physics in medicine and biology, 1996. **41**(8): p. 1469-1481.
46. Balakrishnama, S. and A. Ganapathiraju, *Linear discriminant analysis-a brief tutorial*. Institute for Signal and information

- Processing, 1998. **18**(1998): p. 1-8.
47. Baker, K., *Singular value decomposition tutorial*. The Ohio State University, 2005. **24**.
 48. Hansen, P.C., *Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank*. SIAM Journal on Scientific and Statistical Computing, 1990. **11**(3): p. 503-518.
 49. Schölkopf, B., A. Smola, and K.-R. Müller. *Kernel principal component analysis*. in *International conference on artificial neural networks*. 1997. Springer.
 50. Van der Maaten, L. and G. Hinton, *Visualizing data using t-SNE*. Journal of machine learning research, 2008. **9**(11).
 51. Cox, M.A. and T.F. Cox, *Multidimensional scaling*, in *Handbook of data visualization*. 2008, Springer. p. 315-347.
 52. Tenenbaum, J.B., V. De Silva, and J.C. Langford, *A global geometric framework for nonlinear dimensionality reduction*. science, 2000. **290**(5500): p. 2319-2323.
 53. Samko, O., A.D. Marshall, and P.L. Rosin, *Selection of the optimal parameter value for the Isomap algorithm*. Pattern Recognition Letters, 2006. **27**(9): p. 968-979.
 54. Kramer, M.A., *Nonlinear principal component analysis using*

- autoassociative neural networks*. AIChE journal, 1991. **37**(2): p. 233-243.
55. Baldi, P. *Autoencoders, unsupervised learning, and deep architectures*. in *Proceedings of ICML workshop on unsupervised and transfer learning*. 2012. JMLR Workshop and Conference Proceedings.
 56. Manning-Dahan, T., *PCA and Autoencoders*. Montreal: Concordia University, INSE, 2018. **6220**.
 57. Fournier, Q. and D. Aloise. *Empirical comparison between autoencoders and traditional dimensionality reduction methods*. in *2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*. 2019. IEEE.
 58. Almotiri, J., K. Elleithy, and A. Elleithy. *Comparison of autoencoder and Principal Component Analysis followed by neural network for e-learning using handwritten recognition*. in *2017 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*. 2017. IEEE.
 59. Mahmud, M.S., J.Z. Huang, and X. Fu, *Variational autoencoder-based dimensionality reduction for high-dimensional small-sample data classification*. International Journal of Computational Intelligence and Applications, 2020. **19**(01): p.

2050002.

60. Kingma, D.P. and M. Welling, *Auto-encoding variational bayes*. arXiv preprint arXiv:1312.6114, 2013.
61. Pu, Y., et al., *Variational autoencoder for deep learning of images, labels and captions*. Advances in neural information processing systems, 2016. **29**: p. 2352-2360.
62. An, J. and S. Cho, *Variational autoencoder based anomaly detection using reconstruction probability*. Special Lecture on IE, 2015. **2**(1): p. 1-18.
63. Kingma, D.P., et al. *Semi-supervised learning with deep generative models*. in *Advances in neural information processing systems*. 2014.
64. Jagust, W.J., et al., *The Alzheimer's Disease Neuroimaging Initiative positron emission tomography core*. Alzheimer's & Dementia, 2010. **6**(3): p. 221-229.
65. Jagust, W.J., et al., *The Alzheimer's disease neuroimaging initiative 2 PET core: 2015*. Alzheimer's & Dementia, 2015. **11**(7): p. 757-771.
66. Hughes, C.P., et al., *A new clinical scale for the staging of dementia*. The British journal of psychiatry, 1982. **140**(6): p. 566-572.

67. Choi, H., H. Kang, and D.S. Lee, *Predicting Aging of Brain Metabolic Topography Using Variational Autoencoder*. Front Aging Neurosci, 2018. **10**: p. 212.
68. Kingma, D.P. and J. Ba, *Adam: A method for stochastic optimization*. arXiv preprint arXiv:1412.6980, 2014.
69. Lloyd, S., *Least squares quantization in PCM*. IEEE transactions on information theory, 1982. **28**(2): p. 129-137.
70. MacQueen, J. *Some methods for classification and analysis of multivariate observations*. in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*. 1967. Oakland, CA, USA.
71. Kodinariya, T.M. and P.R. Makwana, *Review on determining number of Cluster in K-Means Clustering*. International Journal, 2013. **1**(6): p. 90-95.
72. Marutho, D., S.H. Handaka, and E. Wijaya. *The determination of cluster number at k-mean using elbow method and purity evaluation on headline news*. in *2018 International Seminar on Application for Technology of Information and Communication*. 2018. IEEE.
73. Tzourio-Mazoyer, N., et al., *Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation*

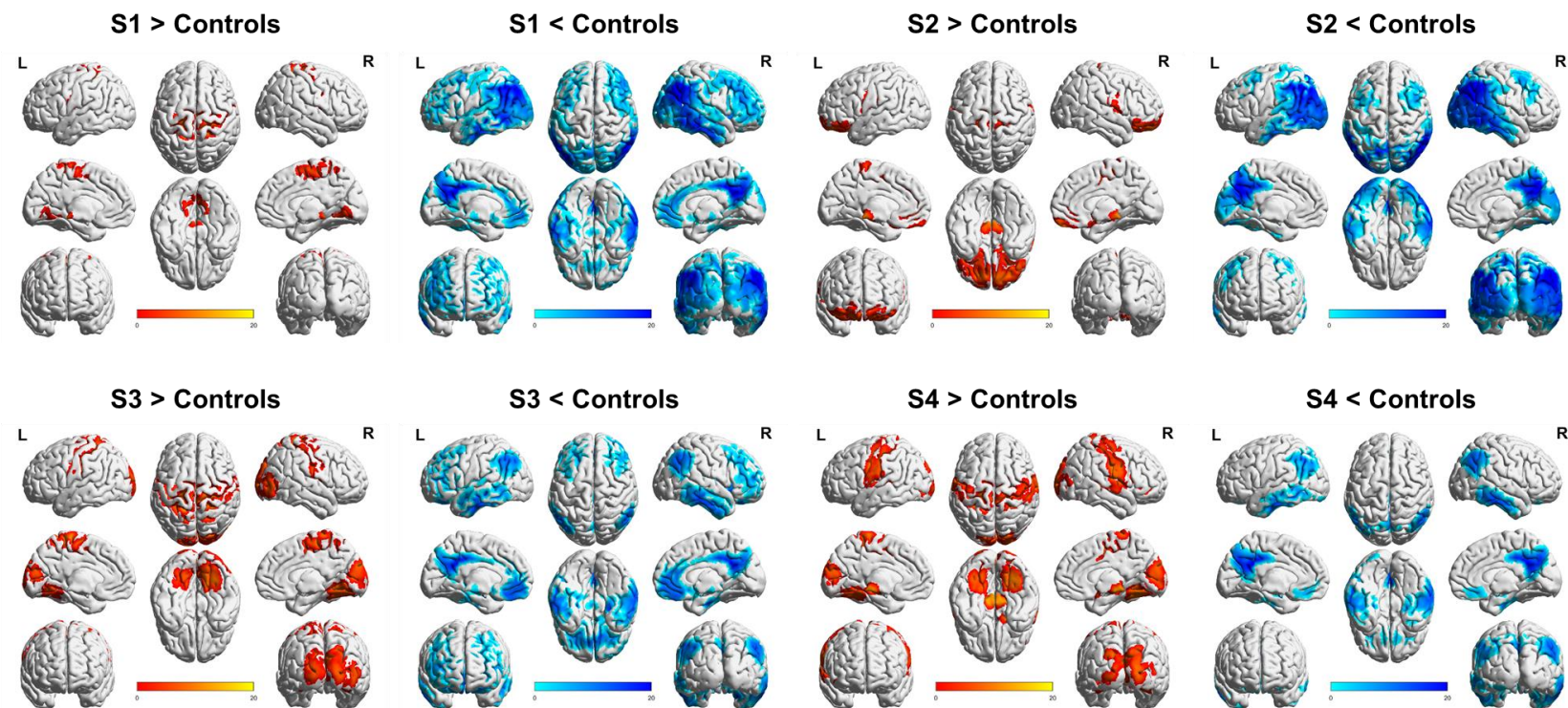
- of the MNI MRI single-subject brain*. Neuroimage, 2002. **15**(1): p. 273-289.
74. Shubhabrata Mukherjee, S.-E.C., Michael L. Lee, Phoebe Scollard, Emily H. Trittschuh, Andrew J. Saykin, R. Elizabeth Sanders, Laura E. Gibbons, Logan C. Dumitrescu, Michael L. Cuccaro, Timothy J. Hohman, Paul K. Crane *ADSP Phenotype Harmonization Consortium – Derivation of Cognitive Composite Scores*.
 75. Habes, M., et al., *Disentangling Heterogeneity in Alzheimer's Disease and Related Dementias Using Data-Driven Methods*. Biol Psychiatry, 2020. **88**(1): p. 70-82.
 76. Young, A.L., et al., *Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference*. Nature communications, 2018. **9**(1): p. 1-16.
 77. Marshall, G.A., et al., *Positron emission tomography metabolic correlates of apathy in Alzheimer disease*. Archives of neurology, 2007. **64**(7): p. 1015-1020.
 78. Van Hoesen, G.W., J. Parvizi, and C.-C. Chu, *Orbitofrontal cortex pathology in Alzheimer's disease*. Cerebral Cortex, 2000. **10**(3): p. 243-251.
 79. Tekin, S., et al., *Orbitofrontal and anterior cingulate cortex*

- neurofibrillary tangle burden is associated with agitation in Alzheimer disease. Ann Neurol*, 2001. **49**(3): p. 355-61.
80. Ferreira, D., L.-O. Wahlund, and E. Westman, *The heterogeneity within Alzheimer's disease. Aging (Albany NY)*, 2018. **10**(11): p. 3058.
81. Arenaza-Urquijo, E.M. and P. Vemuri, *Resistance vs resilience to Alzheimer disease: clarifying terminology for preclinical studies. Neurology*, 2018. **90**(15): p. 695-703.
82. Stern, Y., et al., *Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. Alzheimer's & Dementia*, 2020. **16**(9): p. 1305-1311.
83. Kushner, M., et al., *Cerebellar glucose consumption in normal and pathologic states using fluorine-FDG and PET. J Nucl Med*, 1987. **28**(11): p. 1667-70.
84. Dukart, J., et al., *Differential effects of global and cerebellar normalization on detection and differentiation of dementia in FDG-PET studies. Neuroimage*, 2010. **49**(2): p. 1490-5.
85. Sohn, K., H. Lee, and X. Yan, *Learning structured output representation using deep conditional generative models. Advances in neural information processing systems*, 2015. **28**: p. 3483-3491.

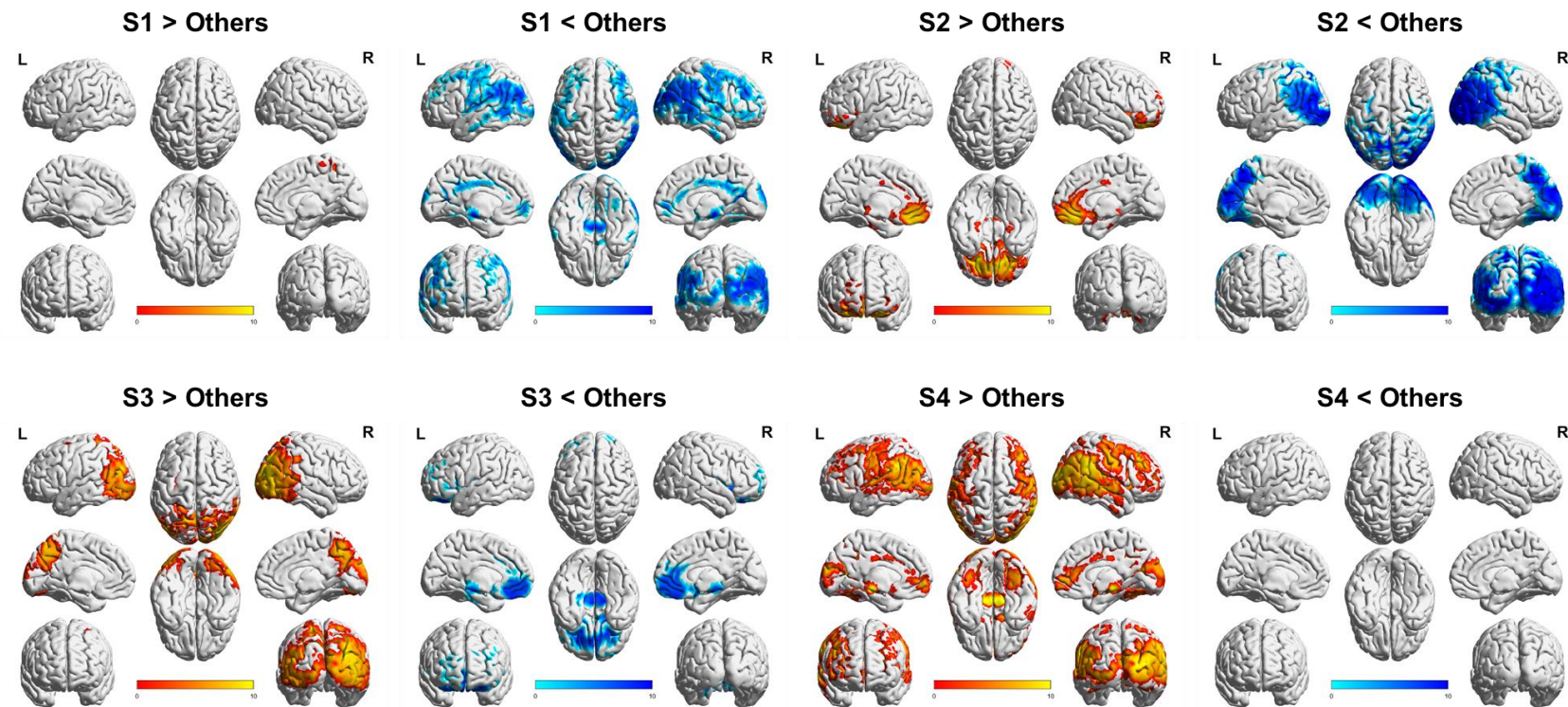
86. Misra, C., Y. Fan, and C. Davatzikos, *Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI*. Neuroimage, 2009. **44**(4): p. 1415-1422.
87. Davatzikos, C., et al., *Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification*. Neurobiology of aging, 2011. **32**(12): p. 2322. e19-2322. e27.
88. Moradi, E., et al., *Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects*. Neuroimage, 2015. **104**: p. 398-412.
89. Gao, F., et al., *AD-NET: Age-adjust neural network for improved MCI to AD conversion prediction*. NeuroImage: Clinical, 2020. **27**: p. 102290.
90. Cheng, B., et al., *Domain transfer learning for MCI conversion prediction*. IEEE Transactions on Biomedical Engineering, 2015. **62**(7): p. 1805-1817.
91. Petersen, R.C., *Mild cognitive impairment as a diagnostic entity*. J Intern Med, 2004. **256**(3): p. 183-94.
92. Hornberger, M., et al., *How preserved is episodic memory in behavioral variant frontotemporal dementia?* Neurology, 2010. **74**(6): p. 472-9.

93. Ferrer, I., G. Santpere, and F.W. van Leeuwen, *Argyrophilic grain disease*. Brain, 2008. **131**(Pt 6): p. 1416-32.
94. Tolnay, M., et al., *Argyrophilic grain disease: widespread hyperphosphorylation of tau protein in limbic neurons*. Acta Neuropathol, 1997. **93**(5): p. 477-84.
95. Jack, C.R., Jr., et al., *An operational approach to National Institute on Aging-Alzheimer's Association criteria for preclinical Alzheimer disease*. Ann Neurol, 2012. **71**(6): p. 765-75.
96. Wisse, L.E.M., et al., *Suspected non-AD pathology in mild cognitive impairment*. Neurobiol Aging, 2015. **36**(12): p. 3152-3162.
97. Jack, C.R., Jr., et al., *Suspected non-Alzheimer disease pathophysiology--concept and controversy*. Nat Rev Neurol, 2016. **12**(2): p. 117-24.

Supplementary Figures



Supplementary figure 1. Spatial metabolism pattern on FDG PET with global normalization of AD subtypes showing the differences between subtypes and controls



Supplementary figure 2. Spatial metabolism pattern on FDG PET with global normalization of AD subtypes showing the differences between one subtype and all other subtypes

국문 초록

딥러닝 기반 군집화 방법을 이용하여 FDG PET에서 알츠하이머병의 공간적 뇌 대사 패턴의 특징적 아형 분류

유현지

분자의학 및 바이오제약학과

서울대학교 융합과학기술대학원

알츠하이머병은 아밀로이드와 타우 침착과 같은 병리학적 특징을 공유함에도 불구하고 광범위한 임상병리학적 특성을 보인다. 본 연구에서는 딥러닝 기반 군집화 방법을 이용하여 FDG PET 영상에서 알츠하이머병 특징적 아형을 분류하여 신경 퇴행의 공간적 뇌 대사 패턴을 이해하고자 하였으며, 공간적 뇌 대사 패턴에 의해 정의된 아형의 임상병리학적 특징을 밝히고자 하였다.

Alzheimer's Disease Neuroimaging Initiative (ADNI) 데이터베이스로부터 첫번째 방문 및 추적 방문을 포함한 알츠하이머병, 경도인지장애, 인지 정상군의 총 3620개의 FDG 뇌 양전자단층촬영 (PET) 영상을 수집하였다. 알츠하이머병에서

질병의 진행 외의 뇌 대사 패턴을 나타내는 표현(representation)을 찾기 위하여, 조건부 변이형 오토인코더(conditional variational autoencoder)를 사용하였으며, 인코딩된 표현으로부터 군집화(clustering)를 시행하였다. 알츠하이머병의 뇌 FDG PET (n=838)과 CDR-SB(Clinical Demetria Rating Scale Sum of Boxes) 점수가 cVAE 모델의 입력값으로 사용되었으며, 군집화에는 k-means 알고리즘이 사용되었다. 훈련된 딥러닝 모델은 경도인지장애군 (n=1761)의 뇌 FDG PET에 전이(transfer)되어 각 아형의 서로 다른 궤적(trajecory)과 예후를 밝히고자 하였다. 통계적 파라미터 지도작성법(Statistical Parametric Mapping, SPM)을 이용하여 각 군집의 공간적 패턴을 시각화 하였으며, 각 군집의 임상적 및 생물학적 특징을 비교하였다. 또한 아형 별 경도인지장애로부터 알츠하이머병으로 전환되는 비율을 비교하였다.

딥러닝 기반 군집화 방법으로 4개의 특징적 아형이 분류되었다. (i) S1 (angular): 모이랑(angular gyrus)에서 현저한 대사 저하를 보이며 분산된 피질의 대사 저하 패턴, 남성에서 빈도 높음, 더 많은 아밀로이드 침착, 더 적은 타우 침착, 더 심한 해마 위축, 초기 단계의 인지 저하의 특징을 보였다. (ii) S2 (occipital): 후두엽(occipital) 피질에서 현저한 대사 저하를 보이며 후부 우세한 대사 저하 패턴, 더 적은 연령, 더 많은 타우, 더 적은 해마 위축, 더 낮은 집행 및 시공간 점수, 경도인지장애로부터 알츠하이머병으로의 빠른 전환의 특징을 보였다. (iii) S3(orbitofrontal): 안와전두(orbitofrontal) 피질에서 현저한 대사 저하를 보이며 전방 우세한 대사 저하 패턴,

더 높은 연령, 더 적은 아밀로이드 침착, 더 심한 해마 위축, 더 높은 집행 및 시공간 점수의 특징을 보였다. (iv) S4(minimal): 최소의 대사 저하를 보임, 여성에서 빈도 높음, 더 적은 아밀로이드 침착, 더 많은 타우 침착, 더 적은 해마 위축, 더 높은 인지기능 점수의 특징을 보였다.

결론적으로, 본 연구에서 우리는 서로 다른 뇌 병리 및 임상 특성을 가진 알츠하이머병의 특징적 아형을 분류하였다. 또한 우리 딥러닝 모델은 경도인지장애군에 성공적으로 전이되어 아형 별 경도인지장애로부터 알츠하이머병으로 전환되는 예후를 예측할 수 있었다. 본 결과는 FDG PET에서 알츠하이머병의 특징적 아형은 개인의 임상 결과에 영향을 미칠 수 있고, 병태생리학 측면에서 알츠하이머병의 광범위한 스펙트럼을 이해하는데 단서를 제공할 수 있음을 시사한다.

Keywords: 알츠하이머병, 아형, FDG 양전자단층촬영, 딥러닝, 조건부 변이형 오토인코더

Student Number: 2018-36289