
ORIGINAL ARTICLE**Journal Section**

Tracking the Progression & Influence of Beta-Amyloid Plaques Using Percolation Centrality and Collective Influence Algorithm: A Study using PET images

Gautam Kumar Baboo^{1†} | Raghav Prasad^{1†} | Pranav Mahajan^{1†} | Veeky Baths^{1†}

¹Cognitive Neuroscience Lab, Department of Biological Sciences, BITS, Pilani - K. K. Birla Goa Campus, South Goa, Goa, 403726, India

Correspondence

Veeky Baths, Cognitive Neuroscience Lab, Department of Biological Science, BITS, Pilani - K.K. Birla Goa Campus, South Goa, Goa, 403726, India
Email: veeky@goa.bits-pilani.ac.in

Present address

†Cognitive Neuroscience Lab, Department of Biological Sciences, BITS, Pilani - K. K. Birla Goa Campus, South Goa, Goa, 403726, India

Funding information

Data usage and analysis for this study is enabled by the Alzheimer's Disease Neuroimaging Initiative - National Institutes of Health Grant U01 AG024904, DOD ADNI, Department of Defense award number W81XWH-12-2-0012

Network analysis allows investigators to explore the many facets of brain networks, particularly the proliferation of disease, using graph theory to model the disease movement. One of the hypotheses behind the disruption in brain networks in Alzheimer's disease (AD) is the abnormal accumulation of beta-amyloid plaques and tau protein tangles. In this study, the potential use of percolation centrality to study the movement of beta-amyloids, as a feature of given PET image-based networks, is studied. The PET image-based network construction is possible using a public access database - Alzheimer's Disease Neuroimaging Initiative, which provided 551 scans. For each image, the Julich atlas provides 121 regions of interest, which are the network nodes. Besides, the influential nodes for each scan are calculated using the collective influence algorithm. Analysis

Abbreviations: AD, Alzheimer's Disease; ANOVA, Analysis of Variance; AV45, Florbetapir (18F-AV-45); CI, Collective Influence; CN, Cognitively Normal; CSF, Cerebrospinal Fluid; EEG, Electroencephalography; FAB, Frontal Assessment Battery; FDG, Fluorodeoxyglucose (18F-FDG); fMRI, functional Magnetic Resonance Imaging; FSL, FMRIB Software Library; GNU, GNU's Not Unix; PCv, Percolation Centrality Value PET, Positron Emission tomography; PIB, Pittsburgh compound B (11C-PIB); RMSE = Root Mean Square Error; MCI, Mild Cognitive Impairment; MEG, Magnetoencephalography; MLR, Multivariate Linear Regression; MMSE, Mini-Mental State Examination; NPIQ, Neuropsychiatric Inventory Questionnaire; OMST, Orthogonal Minimum Spanning Tree.

[†]Equally contributing authors.

of variance ($p<0.05$) yields the region of interest GM Superior parietal lobule 7A L, for which percolation centrality is significant irrespective of the tracer type. Pairwise variance analysis between the clinical groups provides five and twelve candidates for AV45 and PiB. Multivariate linear regression between the percolation centrality values for nodes and psychometric assessment scores reveals Mini-Mental State Examination is a reliable metric. Finally, a ranking of the regions of interest is made based on the collective influence algorithm to indicate the anatomical areas strongly influencing the beta-amyloid network. Through this study, it is possible to use percolation centrality values to indicate the regions of interest that reflect the disease's spread.

KEY WORDS

Brain Connectivity, PET, Percolation Centrality, Alzheimer's Disease, Graph theory, Collective Influence

1 | INTRODUCTION

Alzheimer's disease predominantly stands out when it comes to neurodegenerative diseases—affecting the middle-age (early-onset Alzheimer's disease (AD)) and the old-age (Late-onset AD) human population. Current projections are estimated to cost about 2 trillion US Dollars by 2030[1], affecting 75 million individuals by the same year. The indirect costs are estimated to be about 244 billion US Dollars[2]. With no sight of a cure for AD and with increasing cases, early diagnosis and active management is the key to tackling this disease for now. The ability to predict the disease's progression with high accuracy helps design a suitable treatment regime at an early stage, thereby bringing the disease's management to an affordable cost range.

Current methods of diagnosis of the disease include both non-invasive and invasive techniques of investigations ranging from Positron Emission Tomography (PET) scans or Cerebrospinal Fluid (CSF) analysis to bedside pen and paper-based questionnaires; each with its pros

and cons. The ability to accurately determine that the dementia is due to Alzheimer's is of utmost importance, followed by the ability to indicate the severity of the disease, which is a unique challenge.

A combination of techniques or criteria is currently employed to detect and determine the extent of dementia due to AD. Methods include family history, psychiatric history for cognitive and behavioral changes, which is then followed by psychometric assessments such as Mini-Mental State Examination (MMSE) [3], Frontal Assessment Battery[4], and the Neuropsychiatric Inventory Questionnaire (NPIQ)[5]. Others include Genetic testing[6] for markers of AD, the apolipoprotein-e4 (APOE-e4)[7], or the use of blood testing or brain imaging to rule out dementia due to other factors. The use of PET imaging[8] and lumbar puncture[9] to determine the levels of beta amyloids in either of them beyond the normal levels is the current standard of practice for the determination of dementia due to AD[10, 11].

Positron Emission Tomography or PET imaging involves the use of radiopharmaceuticals such as 2-[¹⁸F],

florbetapir-fluorine-18 (AV45), or 11C-Pittsburgh compound B (PiB). AV45 and PiB[12] are comparatively newer and different in terms of the image construction mechanism. Both AV45 and PiB bind to beta-amyloid but vary in their half-life. AV45 has a half-life of 109.75 minutes and PiB, 20 minutes[13]. A comparison between PiB and AV45 varies in the fact that AV45 shows uptake within the white matter region[14].

The application of network analysis/graph theory to anatomical neural networks has proved useful in understanding the brain connectivity[15, 16] deviations under various psychological and neurological disease states. Network analysis on neuroimaging data such as EEG, MEG, fMRI, and PET scans proves to be useful to show the variation between a cognitively normal population versus other diagnostic states using various graph-theoretic metrics[17, 18].

Network analysis on AD is a practical application wherein it describes the Alzheimer's brain network's behavior. Connectivity analysis using fMRI and EEG data reports provides mixed responses; when comparing AD patients and the control group[19], there is an increase or decrease in the network's connectivity. A reduction in connectivity could explain the cortical atrophy/disruption of the network. An increase could explain the compensatory mechanism[20].

Network Analysis on PET images related to AD mainly revolve around learning models or are limited to tracers that focus on the metabolic networks and the associated deviations of these networks[21]. Other methods include applying algorithms to the raw PET images to recognize patterns to resolve differences between healthy controls and patients with neurodegeneration [22]. Recent methods include genetic and protein markers to improve predicting the course of the disease[23]. These methods rely on a considerable amount of data points and equally reliable computing hardware; this is currently a challenge. Given that AD diagnosis is a global challenge, a method that works well in a spectrum of nations, from developed countries such as the United States to rural hospitals of southeast Asia or Africa[2] is a basic necessity. Methods such as principal component analysis have a few drawbacks; for instance, choosing

the number of principal components and data standardization for multiple PET scans of patients with different tracers leads to controlling multiple variables. Regression analysis is based on the assumption that there are cause and effect in place. Furthermore, a relationship that is present within a limited data set might get overturned with an exhaustive data set.

To understand the propagation of beta-amyloids, we propose applying graph theoretic methods on PET images to understand the advancement of beta-amyloids. The main benefit of adding this method is that

1. This does not introduce any new steps for data collection from the patient and, at the same time
2. adds value to the existing data by computing the percolation centrality of a given node at a given time

Network topology offers insights into the evolution of the network in a clinical setting. Studying such an evolution provides a possibility to understand the weak links within the said anatomical neural networks. Such networks' structural connectivity information yields the source and sink of neurodegeneration with the brain architecture.

Percolation centrality is defined as the proportion of 'percolated paths' that pass through that node; this measure quantifies the relative impact of nodes based on their topological connectivity, as well as their percolated states. In other words, it is one such graph metric that looks at the extent to which a given node within a network has percolated information or can percolate information. The volume of information transmitted via a given node is provided by the values ranging from 0.0 to 1.0[24, 25]. Prior exploration of percolation centrality on disease networks[26, 27, 28, 29] and percolation centrality in disease networks of the brain[30] have shown this as a promising metric for brain network investigation.

The knowledge on the application of percolation centrality on human PET-image based networks is scarce at present. This work aims at adding knowledge to the gap. On the other hand, collective influence provides a minimum set of nodes or regions of the in-

terest that can transfer information or spread disease with ease with optimal spread[31] based on the optimal percolation theory. By examining the network for the minimum set of nodes, this set will provide the regions of interest within the brain that optimally move beta-amyloid, disrupting the normal functioning of the existing neural networks. Thus, the ability to detect the disease and predict the rate of progression of the disease at an early stage is imperative. To this end, the study aims to answer two main questions: 1) Can percolation centrality measure be used to determine the percolation of beta-amyloids within the brain? 2) Can the collective influence algorithm provide a minimum set of nodes that are vital to the AD network?

2 | MATERIALS AND METHODS

2.1 | Current Ethics Statement

2.1.1 | IRB Waiver Statement

Informed consent from the patients is obtained prior to the assessment carried out by ADNI study team (See ADNI website for details), and this study is a secondary data analysis of the ADNI data collection, which aims at providing a simplified metric to an already diagnosed patient. The data access and usage is within the ADNI data use agreements

2.2 | Patient Distribution

Based on the tracer agents used for acquiring the PET images, each diagnostic state subset of the data set is divided into the two available tracers; AV45[32] and PiB[13].

The patients are categorized as Cognitively normal, with Mild Cognitive Impairment, or having Alzheimer's Disease (AD) based on the ADNI study's psychometric assessments. Next, the PET image is matched with the patient's diagnostic state at the time of the imaging procedure. This provides us with a set of observations for each type of tracer for each patient condition clinical group (see Table 1). Finally, the resulting set of patients

are matched with the table containing demographic information providing a total of 531 patients.

3 | NETWORK CONSTRUCTION AND PROCESSING

3.1 | PET Image preprocessing

Image preprocessing is carried out in two steps:

1. Combining individual frames of the PET image to form a 4D raw activity image. This is done using the fslmerge utility included in FSL[33].
2. The 4D raw activity image is converted to a 4D SUV image using the following formula:

$$SUV = \frac{c_{img}}{c_{inj}} \quad (1)$$

c_{img} (MBq ml^{-1}) is given by the raw activity image and $c_{inj} = \frac{ID}{BW}$. ID (MBq) is the injection dose[34], and BW (g) is the bodyweight of the patient, considering the equivalency $1\text{g} = 1\text{ml}$

3. Coregistering the 4D SUV image from subject space to MNI[35] space. This is done using FreeSurfer[36]. The image used for coregistration is the MNI152_T1_2mm_brain.
4. Spatially realigning the PET frames to correct for motion. This was done using MCFLIRT.[37] The motion correction occurs with 6 DOF. The PET frames are realigned using the mean image as a template. The mean image is obtained by applying the motion correction parameters to the time series and averaging the volumes.

Given the large volume of data, carrying out computation in a sequential manner would be highly time-consuming. Thus, to parallelize this operation, GNU Parallel[38] was used. This could potentially give a maximum speedup of up to 12 times on the system we used for computation (6-core, hyperthreaded Intel i7).

3.2 | PET Image-based Network Construction

The network is constructed using the regions of interest (ROIs) from the Julich Atlas[39, 40, 41]. This atlas provides 121 ROIs, which translates to 121 nodes or vertices in the network. Building networks from the pre-processed images requires the generation of adjacency matrices. The adjacency matrix is computed by calculating the combined functional connectivity (combinedFC).

Bivariate Pearson correlation performs poorly in cases of "confounding" or "chain" interactions. In such cases, partial correlation measures the direct connectivity between two nodes by estimating their correlation after regressing out effects from all the other nodes in the network, hence avoiding spurious effects in network modeling. Whereas in cases of "colliding" interactions, a partial correlation may induce a spurious correlation. Thus, Sanchez-Romero and Cole have introduced a combined multiple functional connectivity method[42].

The network is constructed by computing the pairwise partial correlation values of voxel intensities in the PET images to produce an initial adjacency matrix ($matrix_{part}$). A second matrix ($matrix_{bivar}$) is constructed by computing the bivariate correlation values of voxel intensities in the PET images. Now, $matrix_{part}$ is modified using $matrix_{bivar}$ as follows:

$$matrix_{part}(i,j) = \begin{cases} 0 & \text{if } matrix_{bivar}(i,j) = 0 \\ \text{no change} & \text{otherwise} \end{cases} \quad (2)$$

where $matrix_{part}(i,j)$ and $matrix_{bivar}(i,j)$ is the element at (i,j) in the respective matrices. $matrix_{part}$ is now the combinedFC adjacency matrix that defines the network.

The partial correlation is calculated using the correlation between two residuals; the values are computed using $N - 2$ ROIs as co-factors for every pair of ROIs[43]. The partial correlation values serve as the edge weights and constitute the values in the adjacency matrices

Partial correlations are computed as correlation of residuals. The first order partial correlation ($\rho_{ij,k}$) of x_i

and x_j , controlling for x_k is given by [44]

$$corr(resid(i|k), resid(j|k)) = \frac{c_{ij} - c_{ik}v_kc_{kj}}{\sqrt{v_i - c_{ik}v_kc_{ki}}\sqrt{v_j - c_{jk}v_kc_{kj}}} \quad (3)$$

where $c_{ij} = cov(x_i, x_j)$ and $v_k = var(x_k)$

Further,

$$\rho_{ij,k} = \frac{\rho_{ij} - \rho_{ik}\rho_{jk}}{\sqrt{1 - \rho_{ik}^2}\sqrt{1 - \rho_{jk}^2}} \quad (4)$$

Since we are controlling for $(N - 2)$ ROIs for each pair of ROIs ROI_i and ROI_j , we calculate the $(N - 2)^{th}$ order partial correlation. This is calculated recursively as

For each $ROI_k \in ROIs$

$$\rho_{ij,ROIs} = \frac{\rho_{ij,ROIs\setminus\{k\}} - \rho_{ik,ROIs\setminus\{k\}}\rho_{kj,ROIs\setminus\{k\}}}{\sqrt{1 - \rho_{ik,ROIs\setminus\{k\}}^2}\sqrt{1 - \rho_{kj,ROIs\setminus\{k\}}^2}} \quad (5)$$

The base case of this recursive algorithm is given by Equation 4.

The estimation of partial correlations[45] is a computationally intensive task, mainly due to the pre-calculation of residuals before computing cross-correlation, because the number of covariates is large; this calculation is done in a time-optimized manner using the R package ppcor[44].

Next, the adjacency matrices are thresholded using a data-driven thresholding scheme based on Orthogonal Minimal Spanning Trees (OMSTs)[46, 47, 22]. Network thresholding serves to remove inconsequential (or low-impact) edges and reduce the network complexity.

The Networkx[48] Python library is used for network construction from the thresholded adjacency matrices and subsequent percolation centrality computation.

3.3 | Percolation Centrality Computation

Percolation centrality is a nodal metric and is calculated for each node. The percolation centrality for each node v at time t is calculated as shown below:

$$PC^t(v) = \frac{1}{(N-2)} \sum_{s \neq v \neq r} \frac{\sigma_{s,r}(v)}{\sigma_{s,r}} \frac{x_s^t}{[\sum x_v^t] - x_v^t} \quad (6)$$

Where $\sigma_{s,r}$ is the number of shortest paths between nodes s and r pass-through node v ,

x_i^t is the percolation state of node i at time t ,
 $x_i^t = 0$ indicates a non-percolated node and,
 $x_i^t = 1$ indicates a fully percolated node.

The percolation centrality value is calculated for each network using the inbuilt function of Networkx. This has a worst-case time complexity of $O(n^3)$, where n is the number of nodes in the network. Using a modified form of Brandes' fast algorithm for betweenness centrality[49], the complexity can be reduced to $O(nm)$, where m is the number of edges. However, percolation centrality calculation with target nodes cannot take advantage of this optimization and has a worst-case time complexity of $O(n^3)$ (see supplementary data)

3.4 | Collective Influence Algorithm

The algorithm is on the basis that, given a network: the flow of information within the network is optimal with a minimum number of nodes that weigh heavily on the flow of information through the said network[31]. In the context of this investigation, the small sets of nodes/ROIs would prove to be vital in the movement of beta-amyloid plaques.

The core idea is that the overall functioning of a network in terms of the spread of information (or in our case, movement of beta-amyloid plaques) hinges on a specific set of nodes called influencers. This idea of finding the most influential nodes has been previously used in other contexts, for example, activating influential nodes in social networks to spread information[50] or de-activating or immunizing influential nodes to prevent large scale pandemics[27, 51]. In recent applica-

tions to neuroscience, this method has been used to find nodes essential for global integration of a memory network in rodents[30]. Our work is the first to apply it to study the progression of AD, to the best of our knowledge. In the context of this investigation, these small sets of influential nodes/ROIs would prove to be vital in the movement of beta-amyloid plaques.

With the implementation of Collective Influence (CI) algorithm, it facilitates to pinpoint the most influential nodes, more efficiently than previously known heuristic techniques. CI is an optimization algorithm that aims to find the minimal set of nodes that could fragment the network in optimal percolation, or in a sense, their removal would dismantle the network in many disconnected and non-extensive components. In percolation theory, if we remove nodes randomly, the network would undergo a structural collapse at a critical fraction where the probability that the giant connected component exists is $G = 0$. The optimal percolation is an optimization problem that attempts to find the minimal fraction of influencers q to achieve the result $G(q) = 0$.

4 | STATISTICAL ANALYSIS

For this study the null hypothesis is that percolation centrality value does not indicate the propagation of beta-amyloids within the brain network.

To determine the impact the percolation value has over each PET scan, a comparison with the regions of interest from the brain atlas is done using the Multiple linear regression analysis.

This study is exploratory in nature, and that the multiplicity problem is significant. And implementation of multiple test procedures does not solve the problem of making valid statistical inference for hypotheses that were generated by the data. But it does assist in describing the possible mechanism.

4.1 | Pairwise Analysis of Variance

To obtain pairwise group differences, we carry out a posteriori (post hoc) analysis using scikits-posthoc pack-

age; the Student T-test pairwisely gives us the respective p values. The ANOVA test is performed for each node in the network with the null hypothesis that the mean percolation centrality of that node is the same across the three stages. To test the null hypothesis, Analysis of variance with significance level (α) of 0.05 is used.

4.2 | Error Correction

To control for multiple comparisons of 121 nodes, the use Scheffe Test and control for Experiment-wise Error Rate (EER) is carried out. It is a single-step procedure that calculates the simultaneous confidence intervals for all pairwise differences between means.

4.3 | Multivariate Linear Regression

A correlation between the percolation centrality values for all 121 nodes and psychometric test scores - MMSE and NPIQ - is computed to identify the regions of interest that can be used as reliable predictors. Instead of performing multiple correlations across all three diagnoses, a multivariate regression analysis using regularization techniques, wherein the features are the nodal percolation centrality values, and the target variable is the MMSE or NPIQ score. The goal is not to build a predictive model but to use it to quantify each node's influence in distinguishing between the clinical conditions for interpretation purposes. Had the purpose been building a machine learning model, it would imply the need to develop elaborate features sets (more than just percolation centrality) and utilize complex machine learning architectures (which provide less room for interpretability)

4.4 | Regularization and Cross-Validation

We use regularization in our multivariate linear regression (MLR) to make sure our regression model generalizes better to unseen data. Regularization is necessary to control for overfitting. Here, both Lasso regression (L1 penalty) as well as Ridge regression (L2 penalty)

are tested, and both provide similar root mean squared errors (RMSE) and similar desired results. We choose Lasso with $\alpha = 0.1$, for reporting our results (see Figure 9). To quantify the robustness and reliability of our model, before and after regularization, we perform a leave one out cross-validation (LOOCV). We choose this cross-validation strategy because it is unbiased and better suited to our smaller sample sizes (especially in PiB tracer subset). We observe an improvement in validation RMSE with an increase in regularization (parameter α), but we also observe that excessive penalization of weights at very high values of α can result in the regression model converging to the mean of the output MMSE/NPIQ scores. To take this into account, we also plot the standard deviation in predicted MMSE/NPIQ outputs and choose $\alpha = 1$ for sufficient but not excessive regularization (Further details in supplementary figures).

5 | RESULTS

5.1 | Demographics

On the basis of the selection criteria, 531 patients were available for this study. Of this, 48% of the females were of the Cognitively Normal group, 25% with Mild cognitive impairment, and 27% with Alzheimer's disease.

43% of the patients received more than 12 years of education as opposed to only 16% who received less than 12 years of education, 31% received more than 12 years of education in the MCI group as opposed to 69% with less than 12 years of education.

47% of the Left-handed patients were in the AD clinical group as opposed to 26% in the right-handed patients. One patient in the MCI, two in CN, and four in AD groups were multilingual.

5.2 | Pairwise ANOVA

The student t-test provides the following results, following ROIs are suited to discern between CN and MCI AV45 tracer:

- GM Superior parietal lobule 7A L/7P L(p-value= 0.040 & 0.032),
- GM Medial geniculate body L,(p-value = 0.049)

for PiB tracer, following are better at distinguishing between CN and AD clinical groups:

- GM Anterior intra-parietal sulcus hIP3 R
- GM Insula Ig2 R

For the CN-MCI pairwise p-value <0.05 between MCI and AD, the following are identified:

- GM Superior parietal lobule 7A L
- WM Superior occipito-frontal fascicle R

Similarly, for AV45 tracer:

- GM Anterior intra-parietal sulcus hIP3 R
- GM Insula Ig2 R
- GM Broca's area BA44 R

and for the PiB tracer:

- GM Broca's area BA44 R
- GM Superior parietal lobule 7A L
- WM Superior longitudinal fascicle L
- WM Superior occipito-frontal fascicle R
- GM Amygdala-laterobasal group L
- GM Amygdala-laterobasal group R

have CN-AD clinical group pairwise p-value <0.05. The regions of interest(ROI) - both the AV45 and PiB tracers that reject the null hypothesis on the basis of F-value and p-value are listed in Table 3. The five ROIs for AV45 are:

- GM Superior parietal lobule 7P L
- GM Medial geniculate body L
- GM Anterior intra-parietal sulcus hIP3 R
- **GM Superior parietal lobule 7A L**
- GM Superior parietal lobule 5L L

and the nine ROIs for PiB are:

- GM Broca's area BA44 R
- GM Amygdala-laterobasal group L
- GM Amygdala-laterobasal group R
- WM Superior occipito-frontal fascicle R
- **GM Superior parietal lobule 7A L**
- GM Visual cortex V3V R
- GM Hippocampus hippocampal-amygdaloid transition area R
- WM Superior longitudinal fascicle L
- GM Primary auditory cortex TE1.1 L

Error Correction using the Scheffe test for the tracers are as follows: AV45:

- GM Anterior intra-parietal sulcus hIP3 R(p-value = 0.042334) between CN & AD.
- GM Superior parietal lobule 7A L (p-value = 0.041955) between CN & MCI (p-value = 0.024760) between CN & AD.

PiB:

- GM Broca's area BA44 R (p-value = 0.005371) between MCI & AD.
- GM Superior parietal lobule 7A L (p-value = 0.041955) between CN & MCI (p-value = 0.024760) between CN & AD.
- WM Superior occipito-frontal fascicle R (p-value = 0.019056) between CN & MCI (p-value = 0.049955) between CN & AD.

A one-way between clinical groups ANOVA was conducted to compare the effect of beta-amyloid accumulation on percolation centrality values in the cognitive normal, mild cognitive impairment and Alzheimer's disease patients.

There was a significant effect of the beta-amyloid accumulation on the percolation centrality values at $p<0.05$ level for the three clinical groups[F(3, 454) = 3.002 for AV45 and F(3, 97) = 3.027 for PiB] (see Table 2).

5.3 | Cross-Validation

Increasing regularization (α) improves the validation RMSE (see Supplementary), making it more robust and generalize to unseen data. But at higher values of α , we see that the standard deviation of predicted MMSE scores decreases to less than < 2 , irrespective of clinical condition. Which could mean that it saturates to predicting the mean MMSE value when regression weights are extremely penalized. Thereby choosing a reasonably small yet effective α value (less than 2), for which the validation RMSE and the standard deviation in output predicted MMSE.

5.4 | Multivariate Linear regression

A linear regression model between the percolation centrality values for all 121 nodes and psychometric test scores - MMSE and NPIQ - is computed to identify the regions of interests that can be used as reliable predictors. Instead of performing multiple correlations across all three diagnoses, a multivariate regression analysis using leave one out cross validation is carried out, wherein the features are the nodal percolation centrality values and the target variable are the psychological assessment scores. For the AV45 tracer with alpha = 0.1, the validation RMSE = 7.895 & standard deviation = 5.491 and training RMSE = 3.734 & standard deviation of 3.631 and with a alpha of 1.0, the validation RMSE = 6.786 & standard deviation = 3.906 and training RMSE = 3.836 & standard deviation of 3.049 and in the case of PiB tracer alpha = 0.1, the validation RMSE = 8.993 & standard deviation = 7.525 and training RMSE = 0.714 & standard deviation of 4.633 and with a alpha of 1.0, the validation RMSE = 6.404 & standard deviation = 3.789 and training RMSE = 2.037 & standard deviation of 3.544

5.5 | Collective Influence Ranking

The collective influence algorithm ranks the ROIs; here, the rank list is generated for the two tracers- AV45 and PiB. When a comparison of the rank is carried out between the clinical groups and tracers in the case of PiB,

the ranking increases when moving from CN clinical condition to MCI, and then ranking decreases going from MCI to AD. Overall the ranking increases by 50% from cognitively normal condition to Alzheimer's disease condition. A comparison of the ROIs across clinical conditions and tracers does not provide any new information at this stage regarding a common or group of common ROIs across the data(see Table 6).

6 | DISCUSSION

This study's demographic does not provide enough knowledge into the possible advantages or disadvantages that environmental or lifestyle provide for beta-amyloid percolation. This is mainly because the sample size is small, and the distribution of data points within the clinical groups is asymmetric(see Figures 4-6).

The student t-test provides nodes that are capable of distinguishing between CN and MCI, whereas some are better at distinguishing between MCI and AD, and most nodes (Table 4) are better at distinguishing between CN and AD clinical groups. These are good indicators of distinguishing between the two extreme clinical conditions, but are not necessarily good indicators of distinguishing between CN and MCI. It is also observed that in the case of PiB tracer, GM Broca's area BA44 R is particularly good at distinguishing between MCI (p-value = 0.044) and AD (p-value = 0.003) (Table 4).

It is observed that percolation centrality values of certain areas of the brain, such as inferior and superior parietal lobules, are reliable for the tracer PiB, whereas for most other cases, the brain areas differ for each tracer considerably. The variation due to the tracers could be because AV45 and PiB bind to the amyloids, it is also observed that the percolation centrality of Broca's area is a reliable differentiator between CN and AD clinical conditions, which validates previous findings that cognitive impairment affects speech production[52].

Graph metrics such as characteristic path length, clustering coefficient, modularity, and hubs have been studied and have provided insights into the brain networks of AD patients and control groups. Some studies

have tried to map the progression of MCI to dementia due to AD [53, 19]; thus, network analysis and the various graph metrics have shown potential as a tool to investigate the brain networks.

The reliability of these graph metrics is questionable since they tend to be influenced by other factors such as genetic predisposition, lifestyle, etc. Here, based on the variance analysis and multivariate regression testing, and the percolation centrality graph metric computed using the PET images, it is possible to show the Alzheimer's disease progressing through the beta-amyloid networks.

The intensity of a voxel/node is used to determine if the node is percolated or not; a higher intensity signifies a percolated node, and zero or lower intensity signifies an ability to permeate with ease (See Figure 2).

The collective influence algorithm yields nine rank lists. These tables contain the ranking of the nodes, from most influential to least, for each patient condition and type of tracer used(see Figure 5). Moreover, since the influential nodes are identified using optimal percolation theory, it validates percolation centrality as a potential metric for diagnosis. Furthermore, ROIs that could be a potential indicator of Alzheimer's disease progression across the clinical conditions and tracers can be identified by increasing the number samples.

The results from the Scheffe test provides a means to validate and increase the confidence in the results (Table 5).

The leave one out cross-validation (LOOCV) strategy to test the robustness and reliability of our regression. Here cross-validation strategy is implemented because it is unbiased and better suited to our smaller sample size.

By using the regularization (L1 - Lasso or L2 - Ridge) to control for overfitting, and it is observed that the effect of increasing regularization on validation RMSE.

The ROIs obtained from the pairwise t-test for between the clinical conditions sheds some light and the percolation of beta-amyloids within the brain. Previous studies show that the seeding of amyloid-beta occurs in neocortical and subcortical regions[54], from this study it is observed that for PiB the following ROI - WM Su-

perior occipito-frontal fascicle R is part of both the neocortical and subcortical regions of the brain. Apart from this, AV45 tracer has GM Medial geniculate body L ROI in the subcortical region and the following in the neocortical region -GM Superior parietal lobule 7P L, GM Anterior intra-parietal sulcus hIP3 R, GM Superior parietal lobule 7A L and GM Superior parietal lobule 5L L.

Prior research shows that damage to the parietal lobe is common in AD, which can lead to apraxia[52, 55], which is attested by these results. AD is associated with atrophy of the cornu ammonis, the subfield of the hippocampus, and deficits in episodic memory and spatial orientation[56, 57, 58].

Whereas with PiB the following ROIs are in neocortical region - GM Broca's area BA44 R, GM Superior parietal lobule 7A L, WM Superior longitudinal fascicle L and GM Primary auditory cortex TE 1.1 L.

And the following in the the subcortical region

- GM Amygdala-laterobasal group L, GM Amygdala-laterobasal group R and GM Hippocampus hippocampal-amygdaloid transition area R.

Age factor not so important but the presence of beta amyloid deposits is[59], Since these ROIs stand out irrespective of the clinical condition or demographic backgrounds, the percolation centrality has a potential to be a reliable value for AD diagnosis.

6.1 | Limitations

This study does not give any evidence regarding the disease progression in terms of the ROIs or patient clinical group. However, this can be addressed by increasing the number of observations within each patient clinical group.

The PET tracers used for acquiring the images, Pittsburgh Compound B (PiB) and Florbetapir (AV45), are compared to check for which among the two tracers provide a more consistent or reliable PCv. Here, the AV45 tracer binds with a high affinity to the beta-amyloid plaque, whereas PiB binds to oligomers or protofibrils. A possible explanation for the difference in PCv generated using these tracers would be their binding targets. The use of second-generation tracers can help improve the

accuracy and test the applicability of percolation centrality on other neurodegenerative diseases and the possibility of using it in metastatic cancer scenarios.

By expanding the dataset to include more patients and with comprehensive data that factors in healthy aging shrinkage of the brain, which results in a decrease of the distances of the brain networks, can help improve the reliability of the Percolation Centrality value. This can then provide a setting for testing out other psychological assessments that can be used as early indicators for dementia due to Alzheimer's disease, thereby tailoring it to specific demographics or population subsets.

The current pipeline is built for tracers such as AV45 and PiB, which indicate beta-amyloid plaque concentrations directly. However, the pipeline can work with a second generation tracers and tracers like FDG with some appropriate modifications, namely: taking the multiplicative inverse of the percolation states of each of the ROIs to reflect the behavior of the FDG tracer.

7 | CONCLUSION

This study aims to show that percolation centrality is a reliable predictor and identifies the nodes that regulate the movement of beta-amyloid plaque and use them to track the disease.

This work demonstrates that using the existing neuroimaging method, PET-CT, this work can add value with relatively short computation time provided sufficient hardware capability is present. The ability to provide a metric to the extent of the disease state is advantageous to the current world of Alzheimer's. Prolonging life with modern-day medicine pushes patients to a world of medical experiences that deviate from the normal. Being able to show the deviation with a value such as percolation centrality has potential applications.

The reliability of percolation centrality can be improved by addressing the concerns that arise by the factors such as the number of patients and the number of patients within each clinical group, time points of data collection, demographic, and the PET tracers used were the limiting factors. Thus, this study provides the usabil-

ity of percolation centrality value to determine the state of the patient and also sets the stage for studying other neurodegenerative diseases.

Unlike measures such as hub centrality or betweenness centrality, which provide information regarding a vital vertex/node within a network, the collective influence algorithm provides a minimum set of nodes of the network that are key to the beta-amyloid plaque movement, which can provide information regarding a particular pathway that is susceptible to the disease.

Author Contribution

Gautam Kumar conceived the hypothesis and the design of the study. Pranav explored the potential of collective influence on the network. Both Raghav and Pranav carried out the analysis, and the necessary scripting for the computing of PC values, listing of the influential nodes and statistical analyses, and Veeky Baths conceptualized the idea and helped in the manuscript preparation.

Acknowledgements

ADNI, for providing access to the PET images.

Conflict of interest

No competing financial interests exist.

references

- [1] El-Hayek YH, Wiley RE, Khoury CP, Daya RP, Ballard C, Evans AR, et al. Tip of the Iceberg: Assessing the Global Socioeconomic Costs of Alzheimer's Disease and Related Dementias and Strategic Implications for Stakeholders. *Journal of Alzheimer's Disease* 2019;70:1–19.
- [2] 2020 Alzheimer's disease facts and figures. *Alzheimer's and Dementia* 2020;16(3):391–460.
- [3] Tombaugh TN, McDowell I, Kristjansson B, Hubley AM. Mini-Mental State Examination (MMSE) and the Modified MMSE (3MS): A psychometric comparison and normative data. *Psychological Assessment* 1996;8(1):48–59.

- [4] Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: a Frontal Assessment Battery at bedside. *Neurology* 2000 dec;55(11):1621–1626. [//pubmed.ncbi.nlm.nih.gov/24647577/](https://pubmed.ncbi.nlm.nih.gov/24647577/)<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4055504/>.
- [5] Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology* 1994 dec;44(12):2308–2314.
- [6] Loy CT, Schofield PR, Turner AM, Kwok JBJ. Genetics of dementia. *Lancet (London, England)* 2014 mar;383(9919):828–840.
- [7] Holtzman DM, Herz J, Bu G. Apolipoprotein E and apolipoprotein E receptors: normal biology and roles in Alzheimer disease. *Cold Spring Harbor perspectives in medicine* 2012 mar;2(3):a006312.
- [8] Johnson KA, Minoshima S, Bohnen NI, Donohoe KJ, Foster NL, Herscovitch P, et al. Appropriate use criteria for amyloid PET: a report of the Amyloid Imaging Task Force, the Society of Nuclear Medicine and Molecular Imaging, and the Alzheimer's Association. *Alzheimer's & dementia : the journal of the Alzheimer's Association* 2013 jan;9(1):e-1–16.
- [9] Shaw LM, Arias J, Blennow K, Galasko D, Molinuevo JL, Salloway S, et al. Appropriate use criteria for lumbar puncture and cerebrospinal fluid testing in the diagnosis of Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association* 2018 nov;14(11):1505–1521.
- [10] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CRJ, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association* 2011 may;7(3):263–269.
- [11] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association* 2011 may;7(3):280–292.
- [12] Landau SM, Thomas BA, Thurfjell L, Schmidt M, Margolin R, Mintun M, et al. Amyloid PET imaging in Alzheimer's disease: a comparison of three radio-tracers. *European journal of nuclear medicine and molecular imaging* 2014 jul;41(7):1398–1407. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4055504/>
- [13] Yamin G, Teplow DB. Pittsburgh Compound-B (PiB) binds amyloid β -protein protofibrils. *Journal of neurochemistry* 2017 jan;140(2):210–215. <https://pubmed.ncbi.nlm.nih.gov/27943341/><https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5225051/>.
- [14] Su Y, Flores S, Wang G, Hornbeck RC, Speidel B, Joseph-Mathurin N, et al. Comparison of Pittsburgh compound B and florbetapir in cross-sectional and longitudinal studies. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 2019;11(1):180–190. <https://doi.org/10.1016/j.jad.2018.12.008>.
- [15] Reijneveld JC, Ponten SC, Berendse HW, Stam CJ. The application of graph theoretical analysis to complex networks in the brain. *Clinical Neurophysiology* 2007;118(11):2317–2331.
- [16] Stam CJ, Reijneveld JC, Graph theoretical analysis of complex networks in the brain; 2007.
- [17] Sakkalis V. Modern electroencephalographic assessment techniques: Theory and applications; 2014.
- [18] Vecchio F, Miraglia F, Maria Rossini P. Connectome: Graph theory application in functional brain network architecture. *Clinical Neurophysiology Practice* 2017;2:206–213. <https://doi.org/10.1016/j.cnp.2017.09.003>.
- [19] Tijms BM, Wink AM, de Haan W, van der Flier WM, Stam CJ, Scheltens P, et al. Alzheimer's disease: connecting findings from graph theoretical studies of brain networks. *Neurobiology of Aging* 2013;34(8):2023–2036. <http://dx.doi.org/10.1016/j.neurobiolaging.2013.02.020>.
- [20] Wook Yoo S, Han CE, Shin JS, Won Seo S, Na DL, Kaiser M, et al. A network flow-based analysis of cognitive reserve in normal ageing and Alzheimer's disease. *Scientific Reports* 2015;5:1–14. <http://dx.doi.org/10.1038/srep10057>.
- [21] Chung J, Yoo K, Kim E, Na DL, Jeong Y, Glucose Metabolic Brain Networks in Early-Onset vs. Late-Onset Alzheimer's Disease ; 2016. <https://www.frontiersin.org/article/10.3389/fnagi.2016.00159>.
- [22] Klyuzhin IS, Fu JF, Hong A, Sacheli M, Shenkov N, Matarazzo M, et al. Data-driven, voxel-based analysis

- of brain PET images: Application of PCA and LASSO methods to visualize and quantify patterns of neurodegeneration. *PLoS ONE* 2018;13(11):1–20.
- [23] Gupta Y, Lama RK, Kwon GR, Weiner MW, Aisen P, Weiner M, et al., Prediction and Classification of Alzheimer's Disease Based on Combined Features From Apolipoprotein-E Genotype, Cerebrospinal Fluid, MR, and FDG-PET Imaging Biomarkers ; 2019. <https://www.frontiersin.org/article/10.3389/fncom.2019.00072>.
- [24] Newman MEJ, Watts DJ. Scaling and percolation in the small-world network model. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* 1999;60(6):7332–7342.
- [25] Piraveenan M, Prokopenko M, Hossain L. Percolation Centrality: Quantifying Graph-Theoretic Impact of Nodes during Percolation in Networks. *PLOS ONE* 2013 jan;8(1):e53095. <https://doi.org/10.1371/journal.pone.0053095>.
- [26] Moore C, Newman MEJ. Epidemics and percolation in small-world networks. *Physical Review E* 2000 may;61(5):5678–5682. <https://link.aps.org/doi/10.1103/PhysRevE.61.5678>.
- [27] Newman MEJ. Spread of epidemic disease on networks. *Physical Review E* 2002 jul;66(1):16128. <https://link.aps.org/doi/10.1103/PhysRevE.66.016128>.
- [28] Piraveenan M, Prokopenko M, Hossain L. Percolation Centrality: Quantifying Graph-Theoretic Impact of Nodes during Percolation in Networks. *PLOS ONE* 2013;8(1).
- [29] Sander LM, Warren CP, Sokolov IM, Simon C, Koopman J. Percolation on heterogeneous networks as a model for epidemics. *Mathematical biosciences* 2002;180:293–305.
- [30] Del Ferraro G, Moreno A, Min B, Morone F, Pérez-Ramírez Ú, Pérez-Cervera L, et al. Finding influential nodes for integration in brain networks using optimal percolation theory. *Nature Communications* 2018;9(1). <http://dx.doi.org/10.1038/s41467-018-04718-3>.
- [31] Morone F, Min B, Bo L, Mari R, Makse HA. Collective Influence Algorithm to find influencers via optimal percolation in massively large social media. *Scientific reports* 2016 jul;6:30062.
- [32] Chopra A, Shan L, Eckelman WC, Leung K, Latterner M, Bryant SH, et al. Molecular Imaging and Contrast Agent Database (MICAD): evolution and progress. *Molecular imaging and biology* 2012 feb;14(1):4–13. <https://europapmc.org/articles/PMC3259264>.
- [33] Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. *NeuroImage* 2012 aug;62(2):782–790.
- [34] Initiative ADN, PET Acquisition; 2017. <https://adni.loni.usc.edu/methods/pet-analysis-method/pet-analysis/>.
- [35] Vizza P, Tradigo G, Messina D, Cascini GL, Veltri P. Methodologies for the analysis and classification of PET neuroimages. *Network Modeling Analysis in Health Informatics and Bioinformatics* 2013;2(4):191–208. <https://doi.org/10.1007/s13721-013-0035-9>.
- [36] Reuter M, Schmansky NJ, Rosas HD, Fischl B. Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage* 2012 jul;61(4):1402–1418. <https://pubmed.ncbi.nlm.nih.gov/22430496https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3389460/>.
- [37] Jenkinson M, Bannister P, Brady M, Smith S. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage* 2002;17(2):825 – 841. <http://www.sciencedirect.com/science/article/pii/S1053811902911328>.
- [38] Tange O, GNU Parallel 20200722 ('Privacy Shield'). Zenodo; 2020. <https://doi.org/10.5281/zenodo.3956817>, GNU Parallel is a general parallelizer to run multiple serial command line programs in parallel without changing them.
- [39] Kötter R, Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, et al. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 2001 aug;356(1412):1293–1322. <https://doi.org/10.1098/rstb.2001.0915>.
- [40] Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, et al. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage* 2005 may;25(4):1325–1335.

- [41] Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 2006 jul;31(3):968–980.
- [42] Sanchez-Romero R, Cole MW. Combining multiple functional connectivity methods to improve causal inferences. *bioRxiv* 2020;https://www.biorxiv.org/content/early/2020/02/06/841890.
- [43] Madigan D. Graphical models in applied multivariate statistics, by J. Whittaker, John Wiley & Sons, New York, 1990, 448 pp. Price: \$59.95. *Networks* 1994 mar;24(2):125. https://doi.org/10.1002/net.3230240213.
- [44] Kim S. ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients. *Communications for Statistical Applications and Methods* 2015 nov;22(6):665–674. http://http://www.csam.or.kr/journal/view.html?doi=10.5351/CSAM.2015.22.6.665.
- [45] Son SJ, Kim J, Seo J, min Lee J, Park H. Connectivity analysis of normal and mild cognitive impairment patients based on FDG and PiB-PET images. *Neuroscience Research* 2015;98:50–58.
- [46] Dimitriadis SI, Salis C, Tarnanas I, Linden DE. Topological Filtering of Dynamic Functional Brain Networks Unfolds Informative Chronnectomics: A Novel Data-Driven Thresholding Scheme Based on Orthogonal Minimal Spanning Trees (OMSTs); 2017. https://www.frontiersin.org/article/10.3389/fninf.2017.00028.
- [47] Dimitriadis SI, Antonakakis M, Simos P, Fletcher JM, Papanicolaou AC. Data-Driven Topological Filtering Based on Orthogonal Minimal Spanning Trees: Application to Multigroup Magnetoencephalography Resting-State Connectivity. *Brain connectivity* 2017 dec;7(10):661–670. https://pubmed.ncbi.nlm.nih.gov/28891322https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6435350/.
- [48] Hagberg AA, Schult DA, Swart PJ. Exploring Network Structure, Dynamics, and Function using NetworkX. In: Varoquaux G, Vaught T, Millman J, editors. *Proceedings of the 7th Python in Science Conference* Pasadena, CA USA; 2008. p. 11–15.
- [49] Brandes U. A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology* 2001 jun;25(2):163–177. https://doi.org/10.1080/0022250x.2001.9990249.
- [50] Richardson M, Domingos P. Mining Knowledge-Sharing Sites for Viral Marketing. In: *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD '02*, New York, NY, USA: Association for Computing Machinery; 2002. p. 61–70. https://doi.org/10.1145/775047.775057.
- [51] Pastor-Satorras R, Vespignani A. Epidemic Spreading in Scale-Free Networks. *Phys Rev Lett* 2001 Apr;86:3200–3203. https://link.aps.org/doi/10.1103/PhysRevLett.86.3200.
- [52] Fernández M, Gobart AL, Balañá M, Group CS. Behavioural symptoms in patients with Alzheimer's disease and their association with cognitive impairment. *BMC neurology* 2010 sep;10:87. https://pubmed.ncbi.nlm.nih.gov/20920205https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2955564/.
- [53] de Haan W, Mott K, van Straaten ECW, Scheltens P, Stam CJ. Activity Dependent Degeneration Explains Hub Vulnerability in Alzheimer's Disease. *PLoS Computational Biology* 2012;8(8).
- [54] Jucker M, Walker LC. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* 2013;501(7465):45–51.
- [55] Kueper JK, Speechley M, Lingum NR, Montero-Odasso M. Motor function and incident dementia: a systematic review and meta-analysis. *Age and Ageing* 2017 sep;46(5):729–738. https://doi.org/10.1093/ageing/afx084.
- [56] Adler DH, Wisse LEM, Ittyerah R, Pluta JB, Ding SL, Xie L, et al. Characterizing the human hippocampus in aging and Alzheimer's disease using a computational atlas derived from ex vivo MRI and histology. *Proceedings of the National Academy of Sciences of the United States of America* 2018;115(16):4252–4257.
- [57] Masurkar AV. Towards a circuit-level understanding of hippocampal CA1 dysfunction in Alzheimer's disease across anatomical axes. *Journal of Alzheimer's disease & Parkinsonism* 2018;8(1):412. https://pubmed.ncbi.nlm.nih.gov/29928558https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6005196/.
- [58] Ogawa M, Sone D, Beheshti I, Maikusa N, Okita K, Takano H, et al. Association between subfield volumes of the medial temporal lobe and cognitive assessments. *Heliyon* 2019;5(6):e01828. http://www.sciencedirect.com/science/article/pii/S2405844018361486.

- [59] Hamaguchi T, Eisele YS, Varvel NH, Lamb BT, Walker LC, Jucker M. The presence of A β seeds, and not age per se, is critical to the initiation of A β deposition in the brain. *Acta Neuropathologica* 2012 Jan;123(1):31–37. <https://doi.org/10.1007/s00401-011-0912-1>.

8 | FIGURES AND TABLES

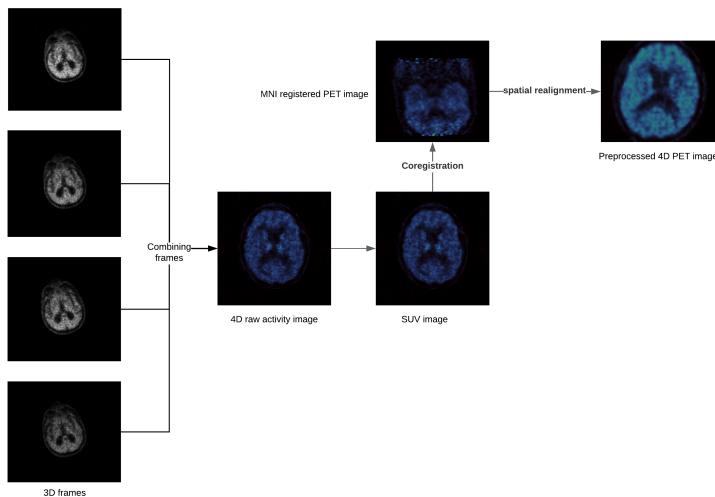


FIGURE 1 PET image preprocessing flowchart.

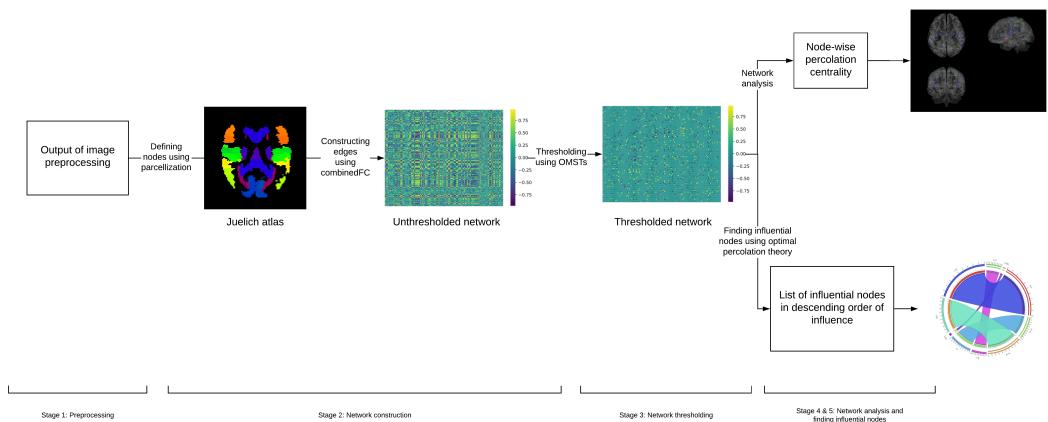


FIGURE 2 The pipeline of the PET-image processing and network construction.

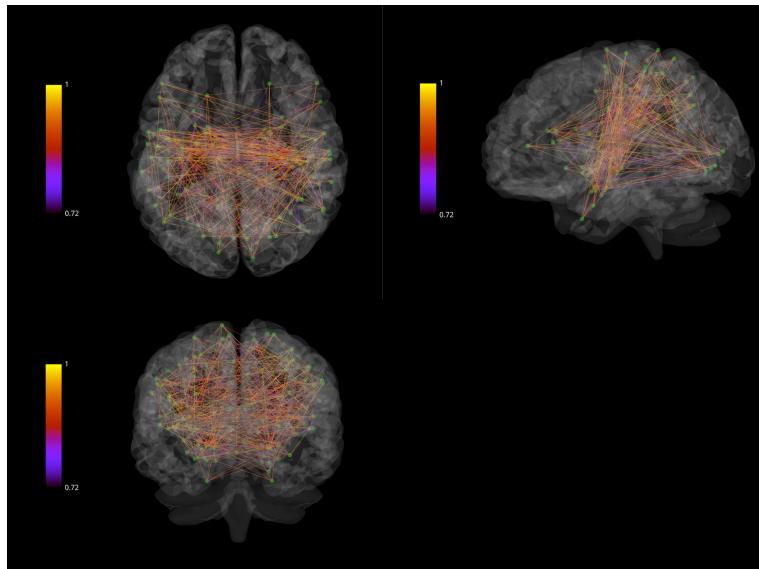


FIGURE 3 A connected network of all the nodes using the Julich Atlas. Green circles indicate the ROIs, the connecting lines indicate the edges with their weights as denoted by the accompanying color bar



FIGURE 4 Pie chart of gender distribution of patients within each category: M-male, F-female



FIGURE 5 Distribution of patients with (R)Right or (L)Left Handedness

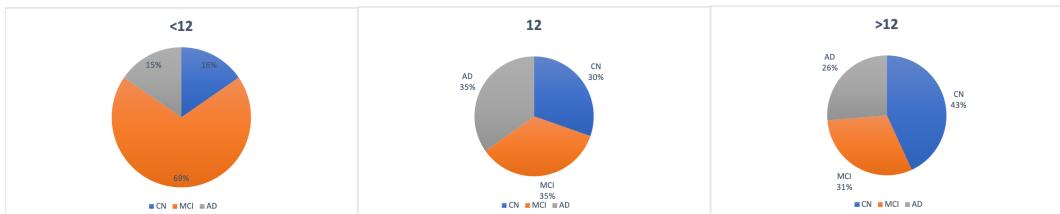


FIGURE 6 Patients with number of years of education received within each category, less than 12 years(<12), 12 years(12) and greater than 12 years(>12)

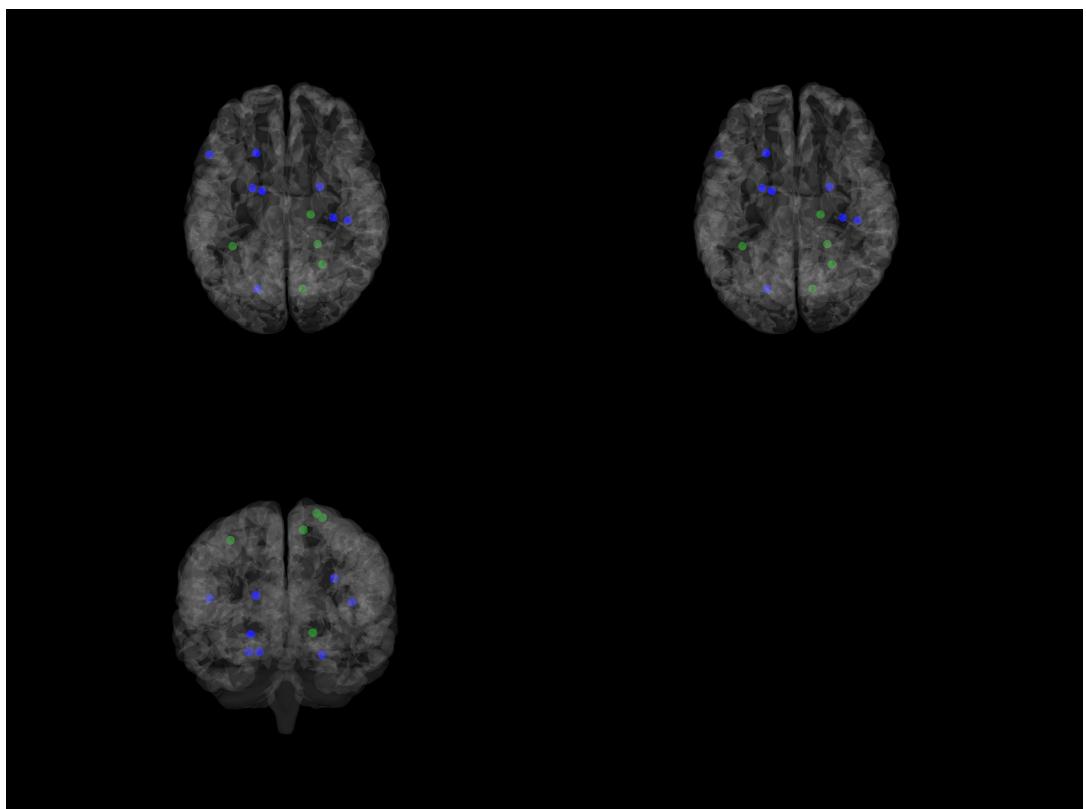


FIGURE 7 Illustrates the findings of the ANOVA. The green circles represent ROIs from AV45 scans with an f-value greater than the AV45 critical f-value and the blue circles represent ROIs from PiB scans with an f-value greater than the PiB critical f-value. The critical f-values are tabulated in Table 2.

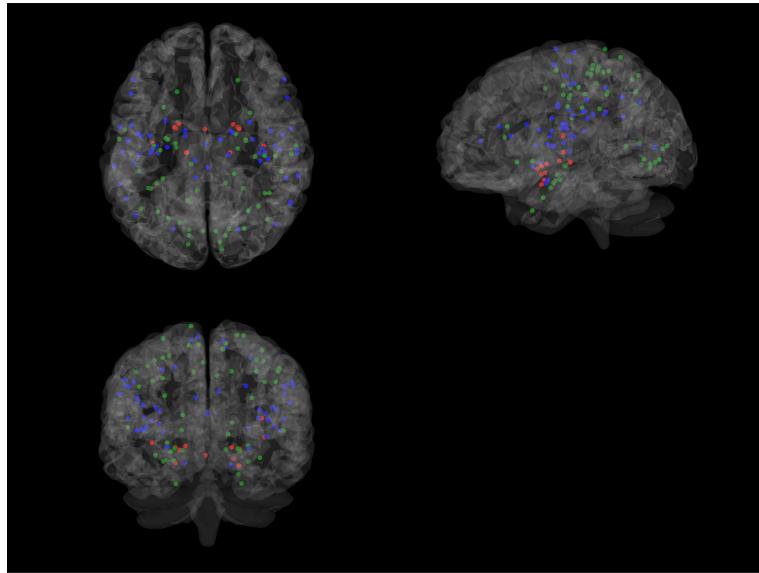


FIGURE 8 Illustrates the ROIs that corresponds to MMSE and NPIQ. The green circles represent ROIs associated with the MMSE psychometric assessment, the red circles represent ROIs associated with the NPIQ psychometric assessment, and the blue circles represent ROIs associated with both MMSE and NPIQ

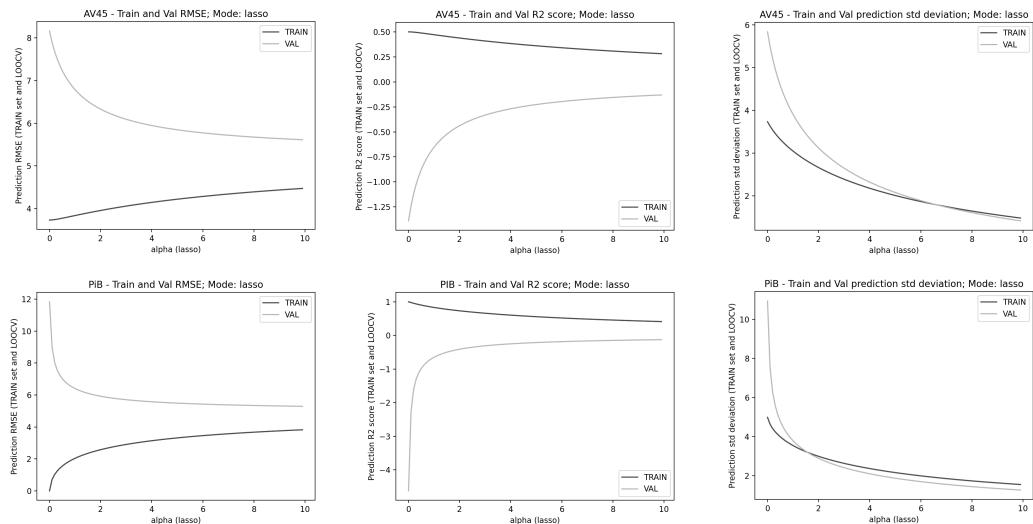


FIGURE 9 Regularization using Lasso regression with L1 penalty.

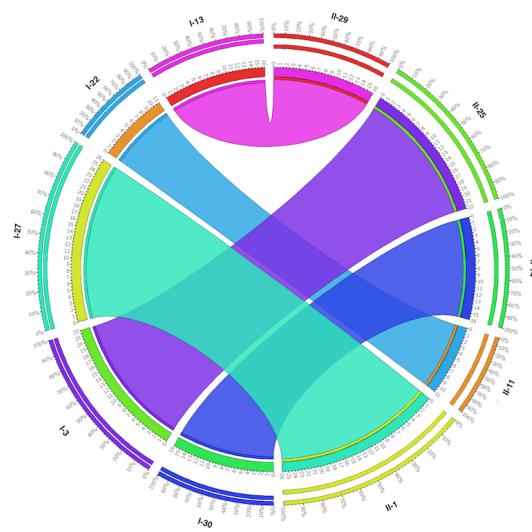


FIGURE 10 CIRCOS plot comparing ROI ranklists of AD-AV45 and MCI-PiB scans

TABLE 1 Distribution of patients

CN		MCI		AD	
AV45	PiB	AV45	PiB	AV45	PiB
262	13	76	65	116	19
M - 122	M - 8	M - 54	M - 45	M - 67	M - 11
F - 140	F - 5	F - 22	F - 20	F - 49	F - 8
Total - 275; M - 130, F - 145		Total - 141; M - 99, F - 42		Total - 135; M - 78, F - 57	

TABLE 2 Number of Scans per tracer type and corresponding critical F-values

Tracer	AV45	PiB
No. of scans	454	97
Critical F-value	3.002	3.027

Supporting Information

TABLE 3 ROIs that reject the ANOVA null hypothesis, i.e. they have F-values greater than the critical F-values tabulated in Table 2 and they have p-values less than 0.05. **BG** - Between Groups, **WG** - Within Groups

AV45						
ROI		Sum of Squares	df	F-value	p-value	
GM Superior parietal lobule 7P L	BG	0.295529	2	3.7881	0.0234	
	WG	17.592231	451			
GM Medial geniculate body L	BG	0.314216	2	3.5756	0.0288	
	WG	19.816397	451			
GM Anterior intra-parietal sulcus hIP3 R	BG	0.255406	2	3.2166	0.0410	
	WG	17.905028	451			
GM Anterior intra-parietal sulcus hIP3 R	BG	0.235190	2	3.2131	0.0412	
	WG	16.505845	451			
GM Anterior intra-parietal sulcus hIP3 R	BG	0.227484	2	3.1184	0.0452	
	WG	16.449730	451			
PiB						
ROI		Sum of Squares	df	F-value	p-value	
GM Broca's area BA44 R	BG	0.580235	2	5.6906	0.0046	
	WG	4.792265	94			
GM Amygdala_laterobasal group L	BG	0.403694	2	4.6474	0.0151	
	WG	4.082629	94			
GM Amygdala_laterobasal group R	BG	0.364432	2	4.3872	0.0119	
	WG	3.904149	94			
WM Superior occipito-frontal fascicle R	BG	0.418668	2	4.3465	0.0157	
	WG	4.527190	94			
WGM Superior parietal lobule 7A L	BG	0.472869	2	4.2163	0.0176	
	WG	5.271143	94			
GM Visual cortex V3V R	BG	0.287708	2	4.0442	0.0206	
	WG	3.343623	94			
GM Hippocampus hippocampal-amygdaloid transition area R	BG	0.416872	2	3.7953	0.0260	
	WG	5.162439	94			
WM Superior longitudinal fascicle L	BG	0.301459	2	3.7325	0.0276	
	WG	3.796017	94			
GM Primary auditory cortex TE1.1 L	BG	0.246789	2	3.5855	0.0316	
	WG	3.235020	94			

TABLE 4 Pairwise p-values from the Analysis of Variance. The values in bold represent significant p-values (<0.05), indicating that the ROI is suited to distinguish between that particular pair of clinical conditions

AV45			
ROI	CN-MCI p-value	CN-AD p-value	MCI-AD p-value
GM Anterior intra-parietal sulcus HIP3 R	0.67875	0.030838	0.27162
GM Superior parietal lobule 7A L	0.040561	0.886276	0.060147
GM Superior parietal lobule 7P L	0.032889	0.17607	0.364166
GM Medial geniculate body L	0.049724	0.12863	0.529499
GM Insula Ig2 R	0.537545	0.04853	0.413302
PiB			
ROI	CN-MCI p-value	CN-AD p-value	MCI-AD p-value
GM Broca's area BA44 R	0.932158	0.044271	0.003756
GM Hippocampus hippocampal-amygdaloid transition area R	0.273329	0.602867	0.032379
GM Superior parietal lobule 7A L	0.030896	0.006372	0.412819
GM Visual cortex V3V R	0.366099	0.182049	0.020161
WM Superior longitudinal fascicle L	0.065172	0.023275	0.249212
WM Superior occipito-frontal fascicle R	0.015984	0.019919	0.934997
GM Amygdala_laterobasal group L	0.216778	0.007312	0.032816
GM Amygdala_laterobasal group R	0.077198	0.006742	0.129689

TABLE 5 Scheffe Test results. The values in bold represent significant p-values (<0.05)

AV45				
ROI	CN-MCI p-value	CN-AD p-value	MCI-AD p-value	
GM Anterior intra-parietal sulcus hIP3 R	0.916342	0.042334	0.307361	
GM Superior parietal lobule 7A L	0.041955	0.024760	0.700914	
PiB				
ROI	CN-MCI p-value	CN-AD p-value	MCI-AD p-value	
GM Broca's area BA44 R	0.996228	0.070505	0.005371	
GM Hippocampus hippocampal-amygdaloid transition area R	0.301853	0.873924	0.045042	
GM Superior parietal lobule 7A L	0.041955	0.024760	0.700914	
GM Visual cortex V3V R	0.613399	0.490895	0.022726	
WM Superior longitudinal fascicle L	0.096172	0.029781	0.503721	
WM Superior occipito-frontal fascicle R	0.019056	0.049955	0.996449	
GM Amygdala_laterobasal group L	0.430360	0.018984	0.054526	
GM Amygdala_laterobasal group R	0.096886	0.015237	0.312719	

TABLE 6 Ranking of the top five ROIs under the AD clinical condition with other clinical conditions for both, AV45 and PiB tracers. First column within each clinical condition indicates the number of occurrence of the ROI and the second column provides the rank.

ROI	Tracer	CN		MCI		AD	
WM Uncinate fascicle L	AV45	142	28	35	30	63	24
	PiB	4	20	25	37	10	30
GM Superior parietal lobule 7P L	AV45	145	29	35	29	57	25
	PiB	10	42	35	28	12	31
GM.Inferior parietal lobule PFop R	AV45	150	30	34	29	67	26
	PiB	9	24	33	35	6	50
GM Insula Ig2 R	AV45	130	31	44	32	51	26
	PiB	3	29	32	33	11	15
GM Mamillary body	AV45	136	30	47	35	58	27
	PiB	5	21	32	32	10	30