Published in final edited form as:

Mol Biosyst. 2017 March 28; 13(4): 725–735. doi:10.1039/c6mb00815a.

Predicting brain network changes in Alzheimer's disease with link prediction algorithms†,‡

Sadegh Sulaimany[a], **Mohammad Khansari**[b], **Peyman Zarrineh**[a], **Madelaine Daianu**[c], **Neda Jahanshad**[c], **Paul M. Thompson**[c], and **Ali Masoudi-Nejad***[a]

[a]Laboratory of Systems Biology and Bioinformatics (LBB), Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran. amasoudin@ut.ac.ir; Web: http://LBB.ut.ac.ir; Fax: +98-21-6640-4680; Tel: +98-21-6695-9256

[b]Department of New Sciences and Technologies (FNST), University of Tehran, Tehran, Iran

[c]Imaging Genetics Center, Institute for Neuroimaging and Informatics, University of Southern California, Marina del Rey, CA, USA

Abstract

Link prediction is a promising research area for modeling various types of networks and has mainly focused on predicting missing links. Link prediction methods may be valuable for describing brain connectivity, as it changes in Alzheimer's disease (AD) and its precursor, mild cognitive impairment (MCI). Here, we analyzed 3-tesla whole-brain diffusion-weighted images from 202 participants in the Alzheimer's Disease Neuroimaging Initiative (ADNI) – 50 healthy controls, 72 with earlyMCI (eMCI) and 38 with lateMCI (lMCI) and 42 AD patients. We introduce a novel approach for Mixed Link Prediction (MLP) to test and define the percent of predictability of each heightened stage of dementia from its previous, less impaired stage, in the simplest case. Using well-known link prediction algorithms as the core of MLP, we propose a new approach that predicts stages of cognitive impairment by simultaneously adding and removing links in the brain networks of elderly individuals. We found that the optimal algorithm, called "Adamic and Adar", had the best fit and most accurately predicted the stages of AD from their previous stage. When compared to the other link prediction algorithms, that mainly only predict the added links, our proposed approach can more inclusively simulate the brain changes during disease by both adding and removing links of the network. Our results are also in line with computational neuroimaging and clinical findings and can be improved for better results.

†Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

‡Electronic supplementary information (ESI) available: There is a folder containing the original matrices for the Normal, eMCI, lMCI and AD groups. Another folder resulting from the computation is named according to the type of transition between Alzheimer's disease stages. For each folder we have named the calculated matrices according to one of the five prediction algorithms (AA, JC, RA, PA and CN) and one of the MLP approaches (Selective, NLPFirst or PLPFirst). See DOI: 10.1039/c6mb00815a

1. Introduction

Alzheimer's disease (AD) is the most common form of dementia, accounting for around 50–70% of all dementia cases.[1] To optimize potential treatments and care, early detection and understanding how the brain changes in AD is of great clinical importance. The most common assessments of cognitive impairment and decline including standardized cognitive evaluations, such as the Mini Mental State Examination (MMSE), General Practitioner Assessment of Cognition (GPCOG), Alzheimer's Disease Assessment Scale-cognitive (Adas-COG), Clinical Dementia Rating and other cognitive tests performed in the clinic, show relatively low sensitivity and specificity. Meanwhile, neuroimaging techniques, such as magnetic resonance imaging (MRI) and positron emission tomography (PET) for evaluating brain structure and metabolism, have improved early diagnosis of AD, as they can directly monitor brain changes and often complement conventional cognitive measures.[2–4]

Diffusion weighted imaging (DWI) can reveal microstructural and organizational patterns of the brain's white matter that cannot otherwise be seen through more conventional anatomical MRI. Tractography algorithms can be used with DWI to trace the pattern and density of white matter connections and determine cortical regions that are connected. As shown in Fig. 1, this type of analysis represents key components of the brain as a network or graph of nodes consisting of, for example, cortical regions. The edges or "links" in this case represent the pathways that connect them, traced by tractography;[5–7] this data representation may be stored in a 2D connectivity matrix, allowing mathematical analysis of brain network topology. In addition, graph theory can be applied to study topological changes in the organization of the brain described as a network of connections.[7,8] Fig. 1 shows the general process of the whole-brain structural network reconstruction with diffusion-weighted and structural MRI.

DWI is increasingly used and recent studies show how brain structural networks change in AD.[2,9–13] Identifying typical changes in connectivity patterns between the different stages of cognitive impairment, from early to late-stage mild cognitive impairment (MCI; eMCI and lMCI respectively) to full dementia, will help to understand the disease and its progression.

Graph theoretical analysis of Alzheimer's disease network has been investigated widely so far. For example, Daianu *et al.*,[7] Xie and He,[14] Seo *et al.*[15] and Jalili[16] have investigated the AD networks based on the cost/sparsity, degree, clustering coefficient, characteristic path length, global efficiency, local efficiency, betweenness centrality, small-world and scale-free properties, *etc.*

Link prediction is a research area that aims to find missing links in a network or predict links that may appear in the near future of an evolving network, graph. It can help better understand network evolution processes. To the best of our knowledge, no prior studies used link prediction algorithms to predict AD stages of impairment using brain connectivity networks. The main goal of this paper is to find the eligibility of applying link prediction to predict the brain changes during Alzheimer's disease. In order to achieve this goal we propose a novel link prediction based approach to determine connectivity changes based on adding and removing edges from brain networks in healthy controls, eMCI, lMCI and AD

participants. To do this, we interpreted the brain networks as described by the connectivity matrices in the form of Boolean networks, with the matrix having entries 0 for unconnected pairs of brain regions and 1 for connected pairs.

1.1. Background

1.1.1. Link prediction problem.—Given a snapshot of a graph at time t, which new edges among its vertices are likely to be established at time $t'(t' > t)$? We can formulate this question as a link prediction problem.[17] Link prediction has been applied to many networks such as biological and social ones.[18–20] There are few papers about using link prediction in brain networks also.[21,22] The diversity of link prediction applications is motivating to test its usability for predicting brain network changes during Alzheimer's disease.

There have been some surveys of publications on link prediction, and some classifications of link prediction algorithms. For example, Lü and Zhou[18] divided link prediction algorithms into three categories: similarity based, maximum-likelihood, and probabilistic models. Also, Al Hasan and Zaki[19] used another similar categorization that divides the methods into three main categories: feature based, probabilistic and linear algebraic algorithms. A more abstract method by Chawla and Yang[23] classifies the algorithms into unsupervised (using topological features of the network such as node neighborhood similarity) and supervised algorithms – using classification and probabilistic methods.

1.1.2. Link prediction approaches.—The most commonly used link prediction approaches are known as Positive Link Prediction (PLP)[24] and they predict future stages of the network by adding links to an earlier stage. In contrast, only a few papers[24,25] proposed Negative Link Prediction (NLP), which predicts the links most likely to be removed from a graph in the near future. Formally, we define the NLP problem as follows:[24] given a network structure at time t_D, we seek to predict the status of the network at time $t\,(t > t_D)$ by predicting the links that are expected to be removed or dropped from the network. NLP works similarly to PLP and applies PLP on the complement of the Boolean adjacency matrix. Complement of a Boolean matrix is created by converting the 0 entries to 1 and *vice versa*. Edges found to be more probably established in the complement matrix by the PLP method are the edges most likely to be removed from the real adjacency matrix.

As brain structure changes during AD, we need to consider both new edge formation and edge loss. In degenerative brain disease, there is little evidence for new connections being formed, but at least in theory there could be compensatory responses. In any case the more general theory (where connections are added) is needed to model functional connectivity (synchrony) of the brain in AD. As far as we know, no studies have proposed an explicit approach to predict stages of disease by both adding and removing links from the brain's network. We will refer to this type of link prediction as the Mixed Link Prediction (MLP). Differences between PLP, NLP and MLP are shown in Fig. 2. There are also prediction methods that simultaneously add and remove edges in biological networks like PPI[26] and GRN.[27] For these, a score is computed between all pairs of edges, with existing or non-existing edges of the network. Then, a threshold is used to retain or establish the edges with greatest scores and to remove those with scores below the threshold. We name these implicit

approaches as the number of edges to be removed (or established) is not defined at first and can be varied according to the problem condition and threshold selection, while in the explicit approach of link prediction it is necessary to define the number of added and removed edges before. Therefore, implicit approaches have only been applied to the networks with one state. In contrast the explicit approach, MLP, is suitable for networks that have at least two states. For example a before and after state could represent different stages of disease progression.

Finding the best edge to add or remove via the MLP approach needs a suitable formula, algorithm. We will use a special set of algorithms, node neighborhood similarity based algorithms, in our proposed approach as they have low complexity, low time requirements and good prediction accuracy, making them one of the most common and effective link prediction algorithms.[21,26,28] Another benefit of using these algorithms is that they use their topology dependent parameters that do not need prior feature selection, as do the supervised algorithms defined in ref. 23. Their limitation is that they only accept the Boolean network as input for their basic version used here. There are five popular node neighborhood similarity based algorithms: Common Neighbors (CN), Preferential Attachments (PA), Adamic and Adar (AA), Resource Allocation (RA) and Jaccard (JC). They all use a similar set of operations, their only difference being their prediction score function. A general algorithm to compute node neighborhood similarity is illustrated in Fig. 3. This algorithm gets a Boolean Brain Connectivity Matrix and returns an edge, $\langle x, y \rangle$, and its rank, TopRank. TopRank will be used to accurately determine whether to add or remove an edge in the Selective version of MLP implementation (Types of MLP subsection under the section titled Mixed link prediction approaches). It only computes calculations on half of the matrix, as the matrix is not directional (i.e., it is symmetric; all connections are considered bidirectional). We will use this algorithm iteratively in our MLP approach to construct a new matrix, the prediction matrix.

Score functions are defined as described in Table 1 and ref. 17. For a vertex x, $\Gamma(x)$ is the set of its neighbors and $|\Gamma(x)|$ is its size and $|\Gamma(x)|$ is the number of common neighbors of x and y vertices. While CN just considers the network transitivity property and gives a higher score to not connected nodes that have more common neighbors, JC normalizes the score based on the selection of the two nodes with more common neighbors with respect to the total number of their neighbors. The idea of PA in a network is that nodes with higher neighborhood size have more probability to be connected in near future. RA and AA refer to prediction of an edge between two nodes based on their similarity to each other but give the common neighbors with smaller degree a higher rank.

1.2. Paper structure

In the next section we explain the data processing, details of our proposed approach for link prediction in brain networks as AD progresses, and the evaluation metric we used to determine matching percentage of prediction. In Section 3, we run the MLP variations to find the best prediction coefficient and algorithm for the transition from cognitively healthy (Controls) to eMCI, eMCI to lMCI, lMCI to AD and direct prediction of the transition to AD from Controls. In the last section we discuss the results including comparison with

clinical and neurobiological evidence of brain network changes in AD. Finally, we propose some extensions and future work.

2. Materials and methods

2.1. Data

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5 year public–private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. The principal investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California – San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 subjects but ADNI has been followed by ADNI-GO and ADNI-2. To date these three protocols have recruited over 1500 adults, ages 55 to 90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow-up duration of each group is specified in the protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

We analyzed diffusion-weighted images (DWI) from 202 participants scanned as part of the ADNI. Table 2 shows the demographics of the participants included here, including age, sex, and mini-mental state exam (MMSE) scores, broken down by diagnosis. All 202 participants underwent whole-brain MRI on 3-tesla GE Medical Systems scanners, at 16 sites across North America. Standard anatomical T1-weighted IR-FSPGR (inverse recovery fast spoiled gradient recalled echo) sequences were collected (256 × 256 matrix; voxel size = 1.2 × 1.0 × 1.0 mm^3; TI = 400 ms, TR = 6.984 ms; TE = 2.848 ms; flip angle = 11°) in the same session as the DWI (128 × 128 matrix; voxel size: 2.7 × 2.7 × 2.7 mm^3; scan time = 9 min). 46 separate images were acquired for each DWI scan: 5 T_2-weighted images with no diffusion sensitization (b_0 images) and 41 diffusion-weighted images (b = 1000 s mm^{-2}). Image preprocessing was performed as in ref. 7. In this dataset we do not observe the same number of participants in all states, but there is no bias in selecting them and the participants in each diagnostic group are representative of that diagnosis.

2.1.1. Connectivity matrix creation. Briefly, we performed whole-brain tractography using the Hough transform method[29] to recover ~10 000 fibers (3D curves) per participant. The Hough method uses a constant solid angle orientation distribution function (CSA-ODF) to

model the local diffusion propagator. For each hemisphere, 34 cortical labels, or regions of interest (ROIs), were automatically extracted from all aligned T_1-weighted structural MRI scans with FreeSurfer (http://surfer.nmr.mgh.harvard.edu/) using the Desikan-Killiany atlas.[30] Then, a 70 × 70 connectivity matrix was created for each participant in which each element represented the total number of detected fibers that passed through each pair of ROIs; the corpus callosum was indirectly eliminated from the study by replacing rows 4 and 39 by 0s, *i.e.* the region is still included in the study but no connections exist with the others. After that we computed a matrix for each diagnostic group by calculating the mean matrix for the group and we created Boolean matrices by converting non-zero elements to one because link prediction works on the Boolean matrix in its basic form selected for this paper. The main diagonal of each matrix was set to zero as selflooping is not permitted in link prediction algorithms.[18] The preliminary information is listed in Table 3.

Subtracting the matrix for each diagnostic group from the one that immediately preceded it provides the input parameters for our approach (Fig. 4). Each entry in the subtraction matrix can be 0, 1 or –1. Zero entries indicate an unchanged state for edges, while entries of one indicate that new edges were added to the next state and negative entries indicate that edges were removed in the new state. For example, for transitioning from the normal state into the eMCI state, 99 additional edges were needed and 79 edges had to be removed from the connectivity matrix of normal participants. Of course, given the finite sampling of a population, some of these changes may reflect noise, as there are no known processes that add structural connections as the brain degenerates. Also, the success rate of our prediction approach relies on how accurately eMCI participants were diagnosed. Furthermore, we also subtracted the first step, normal state, from the AD state (Fig. 4), to find its predictability and compare to known clinical information.

2.2. Mixed link prediction approaches

Given the connectivity changes between stages of cognitive impairment presented in Fig. 4, the algorithm must both add and remove edges to predict one state from its precursor. It can be done in different manners. In this study, we propose 3 MLP approaches, described next.

2.2.1. Types of MLP.—

MLP can be implemented in three ways. We ran all three approaches to find the most suitable one for predicting the brain network state in AD.

1. PLP First: First, we predict the edges that need to be added and add them accordingly. Then, we predict and remove the edges that need to be removed.
2. NLP First: First, we predict and remove the edges that need to be removed. Then, we will perform the same task for edges that need to be added.
3. Selective: At each step we will calculate a score to find the best action, either to remove or add edges; the action with the highest score is selected accordingly.

We show flowcharts for the described approaches in Fig. 5 and 6. All MLP variations use a node neighborhood similarity based Link Prediction Algorithm, described in the second subsection of introduction of the paper, as their core. We can easily calculate A, the number of edges to be removed, and B, the edges to be added, by subtracting the source state matrix

from the target state matrix (Fig. 4). Selective MLP should not select a link for addition or removal more than once. To prevent an edge from being selected more than once, we constructed an ordered list of candidate links in our implementation to select from, to pick another alternative predicted edge.

Before performing the computations, we define the success criteria as follows.

2.3. Evaluation metric

For the prediction results, different evaluation metrics can be applied to measure the performance of the approaches and find the best fit algorithm. MLP prediction can be treated as a binary predictor, which means that it has either a positive or a negative prediction (*i.e.*, the link will either be formed or removed). Similar to a previous study,[24] we used the confusion matrix[31] as an evaluation metric to measure the accuracy of matching the predicted edges. Table 4 shows the confusion matrix. Accuracy is defined as the ratio of the sum of all data points classified correctly over the number of all possible data points. Matching accuracy here refers to successful prediction of edge changes while usually the word accuracy in Alzheimer's imaging works refers to how well some method can predict one stage *vs.* the other. So for avoiding misunderstanding of the meaning of accuracy, here we use 'matching percentage' or shortly 'matching'. By knowing the true matrix at each disease stage, matching can be easily described and computed for the transition between AD stages in the following sections.

$$\text{Matching} = (\text{TP} + \text{TN}) \Big/ (\text{TP} + \text{FP} + \text{FN} + \text{TN}) = (\text{True Additions} + \text{True Removals}) \Big/ (\text{All Additions and Removals})$$

To assess the significance of the prediction results, we simulated the changes using a randomization method. For each transition we calculated the prediction matching in cases where the same numbers of edges, the same A and B used in MLP, were added and removed randomly. This calculation was repeated 10 000 times for each change between AD stages. Then, we reported the p-value for the actual matching of best predictions as output from the approaches. In other words, here the p-value represents the probability of obtaining a matching score, under random prediction, higher than the best one provided by the MLP. The smaller the p-value, the higher the improvement of the MLP compared to random prediction. To compute it, we performed random prediction 10 000 times. Then for each time, we computed the associated matching score and obtained the p-value as the relative frequency of samples in which the matching score from random prediction is greater than the best matching score from MLP.

3. Results

In this section, we apply all variations of our MLP approaches to the clinical diagnostic groups to see how accurately we can predict the next disease state. We compute the difference between the next real state and the predicted one, including the random prediction, and calculate accuracy metric for each prediction as described in the previous section.

3.1 Predicting eMCI from normal

According to Fig. 4, 99 edges should be added to the *Normal* matrix and 79 edges should be removed correctly from it by our MLP prediction approaches to reach the *eMCI state*. If we name the prediction result *eMCIP*, and define Distance as *Distance* = *eMCI* – *eMCIP*, the worst possible score value for Distance is 356 = 2 × (99 + 79), when all the added and removed edges are false, and the best value for Distance is 0, when *eMCI* is equal to *eMCIP*. So the Matching percentage or shortly Matching is calculated as follows (Table 5):

$$\text{Matching} = (356 - \text{Distance}) / 356$$

3.2. Predicting lMCI from eMCI

According to Fig. 4, 44 edges should be added to the *eMCI* matrix and 174 edges should be removed from it by our MLP prediction approaches to reach the *lMCI*. If we name the prediction result *lMCIP*, and define *Distance* = *lMCI* – *lMCIP*, the worst value for Distance is 436 = 2 × (44 + 174) and the best value is 0. So the Matching is calculated as follows (Table 6):

$$\text{Matching} = (436 - \text{Distance}) / 436$$

3.3. Predicting AD from lMCI

According to Fig. 4, 74 edges should be added to the *lMCI* matrix and 90 edges should be removed from it by our MLP prediction approaches to reach the AD. If we name the prediction result *ADP*, and define *Distance* = *AD* – *ADP*, the worst value for Distance is 328 = 2 × (74 + 90) and the best value is 0. So the Matching is calculated as follows (Table 7):

$$\text{Matching} = (328 - \text{Distance}) / 328$$

3.4. Predicting AD directly from normal

According to Fig. 4, 46 edges should be added to the *Normal* matrix and 172 edges should be removed from it by our MLP prediction approaches to reach the *AD* directly. If we name the prediction result *ADP*, we can compare it with *AD* as follows. We define Distance as *Distance* = *AD* – *ADP*. The worst value for Distance is 436 = 2 × (46 + 172) and the best value is 0. So the Matching is calculated as follows (Table 8):

$$\text{Matching} = (436 - \text{Distance}) / 436$$

3.5. Best predictor

The AA and RA had similar results and outperformed the other node neighborhood similarity based algorithms used in the MLP approach with p-value < 0.0001 as compared to random prediction. These algorithms also perform better for other commonly analyzed link prediction problems. We simply refer to the AA algorithm as the best predictor because of its superior performance in other link prediction research studies.[19,32,33] Selective MLP performed with the highest matching percentage. However, the performance of the Selective

Prediction algorithm was comparable (often similar) to the performance of the NLP First and PLP First algorithms. Based on the AA results, Fig. 7 shows the best matching percentage of prediction from one disease stage to the other.

4. Discussion

In this work, we applied link prediction algorithms to brain network connections to determine the patterns of connections gained and lost in the stages thought to lead up to Alzheimer's disease. Link prediction research has not been applied to Alzheimer's disease progression before for several reasons. First, sample sizes with DWI in AD neuroimaging studies have been insufficient to obtain sufficient power to assess link prediction algorithms. Learning about the likely network changes via link predicting models offers a new direction in neuroscience as the application of graph analytical methods to these matrices is proving quite successful. However, how best to create a connectivity matrix from neuroimaging techniques is still a much studied research topic, suggesting that with improved matrix development techniques, the opportunities for link prediction can be greatly expanded. Second, the link prediction algorithms almost always aim to predict missing links in the network while in the case of examining stages of cognitive impairment for Alzheimer's disease progression we need to predict both link addition and removal.

Despite the meaningful results for the prediction values of the state change during the disease, maybe they seem low at first glance. This is because of two reasons. First, we are going to test the usability of link prediction for the task in its basic form with most simplifications here and this research tries to show that results are promising for future customization and improvement and can be a base for related future works and research. Second, the results show that not all the states are predictable equally from their predecessor; some of the predictions are more difficult such as predicting lMCI from eMCI.

Predicting lMCI from eMCI is a challenge. Based on the results from the previous section, Fig. 7 shows a greater predictability from Normal to eMCI than from Normal to Alzheimer's and so on. From a computational point of view, link prediction may work best when the network has few changes and can be used for removing noise or finding missing links. Clearly, the Normal state is similar to eMCI and lMCI is similar to AD and eMCI is similar to lMCI. But we often combine eMCI and lMCI – naming it MCI – as they are harder to distinguish. Another Alzheimer's prediction paper with different neuroimaging approach supports these findings.[34] It proposes a new graph construction algorithm to predict AD severity based on neuroimaging predictors using 758 subjects from ADNI including 180 AD subjects, 160 cMCI (MCI subjects converted to AD in a determined time) subjects, 214 ncMCI (non-converted MCI subjects to AD) subjects and 204 normal aging subjects. According to their findings the prediction of cMCI was the most difficult task compared to other state transition predictions and most of the misclassifications occur between cMCI and ncMCI patients. In another study,[35] over 1 year, a small proportion (0–20%) of participants with eMCI progressed to more severe cognitive impairment, lMCI. Somewhat larger proportions (6–53%) of participants improved or reverted to normal. Most participants (29–92%) remained stable.

Many papers treat MCI as a single state, viewing eMCI and lMCI as a single state, MCI, between normal cognition and dementia. We explored this breakdown of ADNI patients as well; we have represented this by a dashed oval in Fig. 7. Based on ref. 36, those with MCI were more likely to experience development of AD than Normal elderly subjects and our results confirm this; we also indicate a greater likelihood for developing AD from MCI than AD from Normal (Fig. 7).

According to the interpretation of the link prediction algorithms, certain algorithms work better for particular types of networks, such as scale-free or small-world networks. For example, the PA is suitable for the prediction of scale-free networks, where according to its formula the probability that a new link is connected to the node x and y is proportional to the degree of x and y.[28] On the other hand, because of the small-world nature of social networks, the AA algorithm is one of the best predictors for this type of network. In the networks that satisfy the small-world property, like the brain networks studied here, most nodes are not neighbors of one another, but most nodes can be reached from every other by a small number of hops or steps. Our results show that brain network differences, across the different stages of Alzheimer's disease, may be more predictable with the AA than PA algorithms, which may be due in part to the small-world property of brain networks. The small world property of human structural brain networks is well known in the literature and is confirmed by our optimal choice in the link prediction algorithm.[37]

Link prediction generally works on social network data where the relations are deterministic and the way of network construction is explicit and based on the human contacts. But for the brain, we use experiments and implicit results, and the network is constructed after several steps and approximation. So this can be an error prone process with different results. As per our knowledge, we do not currently have individuals who change state from healthy to disease (Normal, eMCI, lMCI and Alzheimer's); therefore we are working with population means across groups, where we aim to predict the brain links that are gained and lost between one state and the other. As longitudinal neuroimaging studies of Alzheimer's disease increase, we may be able to capture more individuals as they degenerate through the different stages of disease longitudinally and compute the algorithm on individual level networks and gain more accurate results.

Nonetheless, the success rate of our prediction approach relies on how accurately participants were diagnosed. For instance, we observed that there were more edges in eMCI than controls in Table 3 and participants of each stage of the disease are not necessarily the same. This design was specifically chosen for ADNI for identifying early progression and intervention approaches for dementia.

There are statistical models for dynamic/temporal network modeling like Temporal Exponential Random Graphs (TERGM) or Stochastic Actor Oriented Models (SAOM), which could be supposed similar to MLP to test and compare, directly or with modifications, to address the same link prediction goal. These models have been used for various applications including sociology, economics, political science, industry, *etc.*, but as per ref. 38 neither of TERGM and SAOM has been adopted for the brain study yet.

The network modeling processes in SAOM and TERGM are very similar;[39] they both use the snapshots $G_1, \cdots, G_T$ of a network evolving over time to model its dynamics. TERGM works based on the Markov-chain model and predicts the next state of the network given the proceeding states.[40] On the other hand, SAOM process ends at the target network simulating the network changes during the transition from the previous stage to the current one.

It is not easily possible to use or modify SAOM and TERGM to verify our method's results. This is because they both cannot predict eMCI from Normal without any previous knowledge and prior steps. They need prior observations of the network snapshots to estimate the network parameters and predict its dynamics. Furthermore, the next state predicted by the TERGM or SAOM is not essentially what we want. For example, it may jump directly to a state that is more similar to Alzheimer's instead of lMCI by starting from eMCI! In other words, its stopping criterion is unclear.

5. Conclusion and future work

In this study, we tested the idea of customizing the link prediction problem to apply it to brain network changes in Alzheimer's disease. Our novel approach, called Mixed Link Prediction, was able to simulate the dynamics of brain network changes in the ADNI study. The link prediction methods assess the likelihood of brain network elements changing in each successive stage of AD. Prediction matching percentage is not equal for each transition. Therefore we would expect to see different accuracies for different transitions. Future work may not lead to a transition prediction diagram, like Fig. 7, with coefficients near to 1 for all transitions.

Basic link prediction algorithms only accept the network in the Boolean form. Converting all non-zero elements to one for creating a connectivity matrix will give the same weights to all the connections and produce some false positive connections. An idea is including the weight of connections in the calculations; this may require the use of a special version of link prediction and customizing it for the MLP approach and improve the results considerably.

On the other hand connectivity matrices extracted from neuroimaging and their accuracy depend on the calculation method, including threshold selection for converting the weighted matrix into Boolean, which also affects the input to our approach. Therefore extensions of our approach may be used as a metric to find the best threshold in converting the weighted matrix of the brain connectivity into Boolean form.

We inserted topological neighborhood based link prediction algorithms into our mixed approach. The method can be extended to other categories of link prediction algorithms such as topological path-based algorithms or machine-learning classification based ones to gain better prediction accuracies. Future works considering the edge weight can determine edge changes rather than loss or gain.

In general our Selective MLP approach ranks edges to be removed or added similarly. Therefore, for sparse matrices, candidate edges to be removed will gain a greater rank and our Selective MLP will behave like NLP First and give preference for removing edges, while

for dense matrices, Selective MLP will behave like PLP First and give preference for adding edges. We will further explore alterations in the Selective MLP algorithm in future works to rank add and remove edges based on a more biologically informative criterion or at least a normalized number of related edges to make a fair ranking, which will be proportional to matrix edge numbers.

One interesting future work is predicting the next state of the disease using only the current state information. After proving the predictability of the changes during Alzheimer's disease stages using link prediction algorithms, one can expect to develop a modified version of the MLP to estimate the next stage of the network without using information from the target network to be predicted. This need is more consistent with reality.

Finally it is possible to show the order of the link changes between clinical stages based on the calculation sequence of the MLP. Eventually we aim to compare the function of the edges gained or lost with other biological and cognitive change sequences to home in on biological mechanisms underlying the AD etiology.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. The first author (S. S.) would like to appreciate being partly supported by the Cognitive Science and Technologies Council (CSTC) of Iran.

References

1. Kukull WA and Bowen JD Med. Clin. North Am, 2002, 86, 573–590. [PubMed: 12168560]
2. Huang S, Li J, Ye J, Wu T, Chen K, Fleisher A and Reiman E Presented in part at the Advances in Neural Information Processing Systems 24 (NIPS 2011), 2011.
3. Coimbra A, Williams DS and Hostetler ED Curr. Top. Med. Chem, 2006, 6, 629–647. [PubMed: 16712496]
4. Jensen JR, Cisek K, Funk KE, Naphade S, Schafer KN and Kuret J J. Alzheimer's Dis, 2011, 26, 147–157. [PubMed: 21971459]
5. Sporns O, Networks of the Brain, MIT press, 2011.
6. van den Heuvel MP and Sporns O Trends Cognit. Sci, 2013, 17, 683–696. [PubMed: 24231140]

7. Daianu M, Dennis EL, Jahanshad N, Nir TM, Toga AW, Jack CR, et al., Alzheimer's disease disrupts rich club organization in brain connectivity networks, in IEEE 10th International Symposium on Biomedical Imaging (ISBI), 2013, pp. 266–269.
8. Rubinov M and Sporns O NeuroImage, 2010, 52, 1059–1069. [PubMed: 19819337]
9. Huang S, Li J, Sun L, Liu J, Wu T, Chen K, et al., Learning Brain Connectivity of Alzheimer's Disease from Neuroimaging Data, in NIPS, 2009, pp. 808–816.
10. Huang S, Li J, Ye J, Fleisher A, Chen K, Wu T, et al., Brain effective connectivity modeling for alzheimer's disease by sparse gaussian bayesian network, in Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, 2011, pp. 931–939.
11. Dennis EL and Thompson PM Neuropsychology Review, 2014, 24, 49–62. [PubMed: 24562737]
12. Toga AW and Thompson PM Brain, 2014, 137, 3104–3106. [PubMed: 25413934]
13. Braskie MN and Thompson PM Trends Cognit. Sci, 2013, 17, 510–516. [PubMed: 24029445]
14. Xie T and He Y Magnetic resonance imaging of disturbed brain connectivity in psychiatric illness, 2011, p. 21.
15. Seo EH, Lee DY, Lee J-M, Park J-S, Sohn BK, Lee DS, Choe YM and Woo JI PLoS One, 2013, 8, e53922. [PubMed: 23335980]
16. Jalili M, Inf. Sci, 2017, 384, 145–156.
17. Liben-Nowell D and Kleinberg J J. Am. Soc. Inf. Sci. Technol, 2007, 58, 1019–1031.
18. Lü L and Zhou T Phys. A, 2011, 390, 1150–1170.
19. Al Hasan M and Zaki MJ Social network data analytics, Springer, 2011, pp. 243–275.
20. Wang P, Xu B, Wu Y and Zhou X Science China Information Sciences, 2014, 1–38.
21. Cannistraci CV, Alanis-Lobato G and Ravasi T Sci. Rep, 2013, 3, 1613. [PubMed: 23563395]
22. Yang Y, Guo H, Tian T and Li H Tsinghua Sci. Technol, 2015, 20, 306–315.
23. Chawla NV and Yang Y, in Encyclopedia of Social Network Analysis and Mining, ed. Alhajjand R Rokne J Springer, 2012.
24. Almansoori W, Gao S, Jarada TN, Elsheikh AM, Murshed AN, Jida J, Alhajj R and Rokne J Network Modeling Analysis in Health Informatics and Bioinformatics, 2012, 1, 27–36.
25. Hong W, Hua YW and Qian Z Frontier and Future Development of Information Technology in Medicine and Education, Springer, 2014, pp. 2031–2035.
26. Lei C and Ruan J, Bioinformatics, 2013, 29, 355–364. [PubMed: 23235927]
27. Barzel B and Barabási A-L Nat. Biotechnol, 2013, 31, 720–725. [PubMed: 23851447]
28. Liu H, Hu Z, Haddadi H and Tian H EPL, 2013, 101, 18004.
29. Aganj I, Lenglet C, Jahanshad N, Yacoub E, Harel N, Thompson PM and Sapiro G Med. Image Anal, 2011, 15, 414–425. [PubMed: 21376655]
30. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP and Hyman BT NeuroImage, 2006, 31, 968–980. [PubMed: 16530430]
31. Kohavi R and Provost F Machine Learning, 1998, 30, 271–274.
32. Lü L and Zhou T Role of weak ties in link prediction of complex networks, in Proceedings of the 1st ACM international workshop on Complex networks meet information & knowledge management, 2009, pp. 55–58.
33. Srinivas V and Mitra P Link Prediction in Social Networks, Springer, 2016, pp. 15–25.
34. Liu S, Cai W, Wen L and Feng D Neuroimaging biomarker based prediction of Alzheimer'S disease severity with optimized graph construction, in IEEE 10th International Symposium on Biomedical Imaging (ISBI), 2013, pp. 1336–1339.
35. Ganguli M, Snitz BE, Saxton JA, Chang C-CH, Lee C-W, Vander Bilt J, Hughes TF, Loewenstein DA, Unverzagt FW and Petersen RC Arch. Neurol, 2011, 68, 761–767. [PubMed: 21670400]
36. Manly JJ, Tang MX, Schupf N, Stern Y, Vonsattel JPG and Mayeux R Ann. Neurol, 2008, 63, 494–506. [PubMed: 18300306]
37. Kaiser M, NeuroImage, 2011, 57, 892–907. [PubMed: 21605688]
38. Simpson SL, Bowman FD and Laurienti PJ Statistics surveys, 2013, 7, 1. [PubMed: 25309643]
39. Leifeld P and Cranmer SJ arXiv preprint arXiv:1506. 06696, 2015.

40. Hanneke S, Fu W and Xing EP Electronic Journal of Statistics, 2010, 4, 585–605.

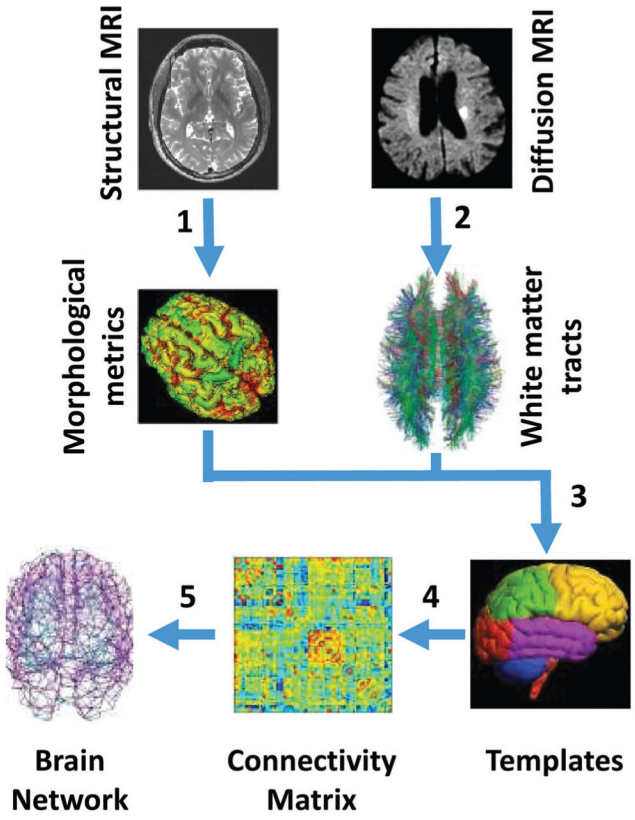

Fig. 1.

Construction of the brain network. (1) Calculate morphological metrics such as cortical thickness and gray matter volume. (2) Estimate white matter fiber tracts using tractography. (3) Regional information extraction from the original vertex- or voxel-based MRI data based on templates. (4) When integrating anatomical MRI and diffusion MRI, the connectivity matrix can be a matrix including numbers of fibers connecting the various regions or the connectivity strength (proportion) of fibers detected between the regions. (5) Create the brain network using extra changes of the connectivity matrix, such as using thresholds and making it Boolean.

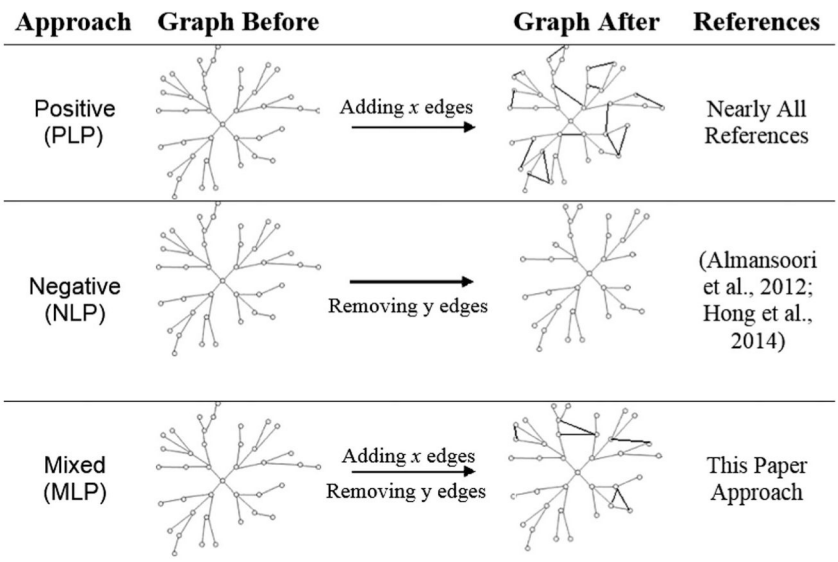

Fig. 2.

Comparison of three different link prediction approaches: PLP, NLP and MLP. While PLP only predicts adding links and NLP only predicts removing links, MLP prediction is done by implementing both addition and removal of links in the graph.

Algorithm: Link Prediction
Input: Adjacency Matrix
Output: x,y,TopRank

```
MaxIndex=70;          /* Number of the regions in Connectivity Matrix */
TopRank=-1;
for i=1 to MaxIndex-1 do
   for j=i+1 to MaxIndex do
      if Matrix(i,j)==0 then
         TempRank=Score Function Matrix(i,j) ;
            if TempRank >TopRank then
                  x=i;
                  y=j;
                  TopRank=TempRank;
         end if
      end if
   end for
end for
```

Fig. 3.

The general pseudocode of node neighborhood similarity based link prediction algorithms. The score function can be calculated based on different available algorithms like CN, JC, PA, AA and RA.

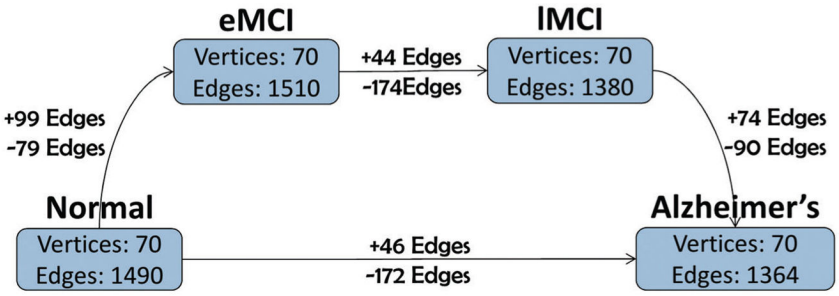

Fig. 4.

Edge changes in transitions between Alzheimer's disease stages. Each box is representative of an Alzheimer's disease stage described by vertexes and edges from the associate connectivity matrix. Arc labels show the number of added and removed edges in conversion from one state to the next.

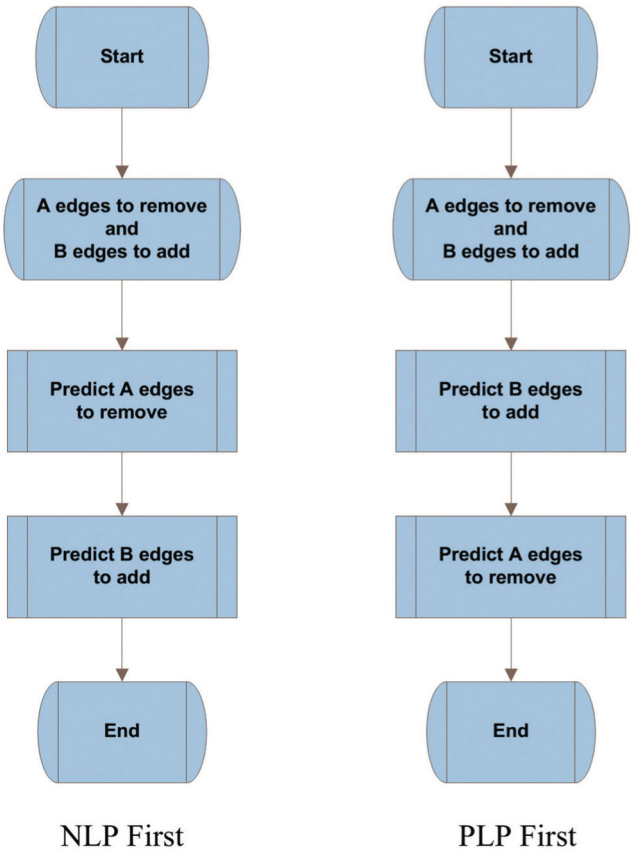

Fig. 5.

Flowchart for NLP First and PLP First variations of the MLP Link Prediction Approach. A and B are the input parameters in each approach; A describes the number of predicted edges to be removed and B describes the number of predicted edges to be added.

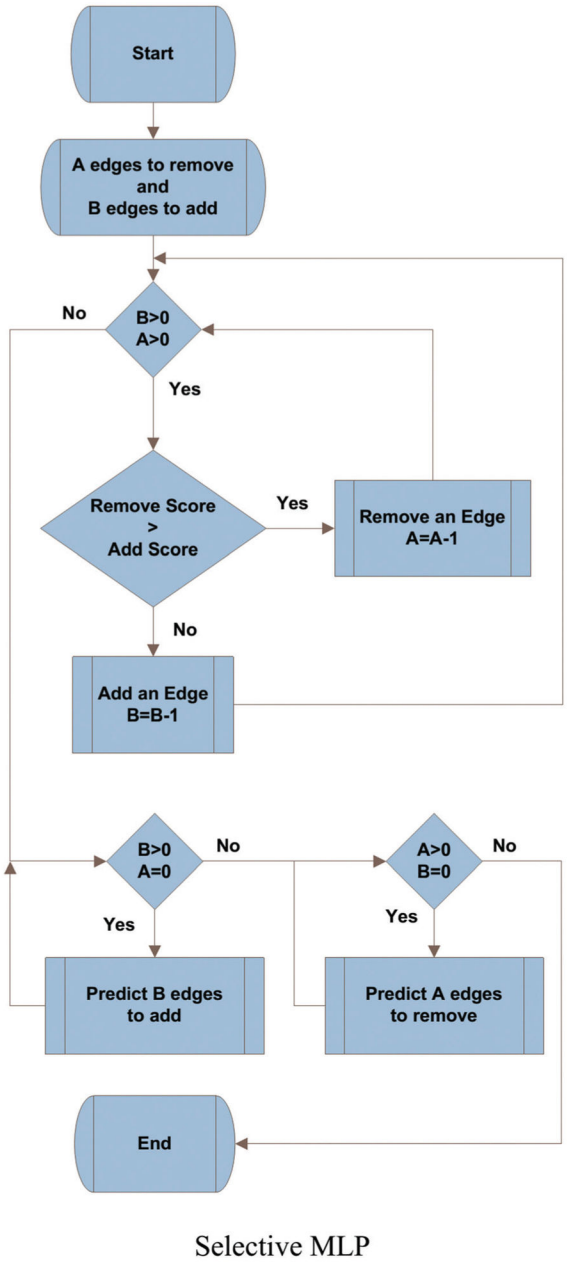

Fig. 6.

Flowchart for the Selective variation of the MLP Link Prediction Approach. Each iteration in the approach predicts whether to add or remove edges. *A* and *B* are the input parameters in each approach; *A* describes the number of predicted edges to be removed.

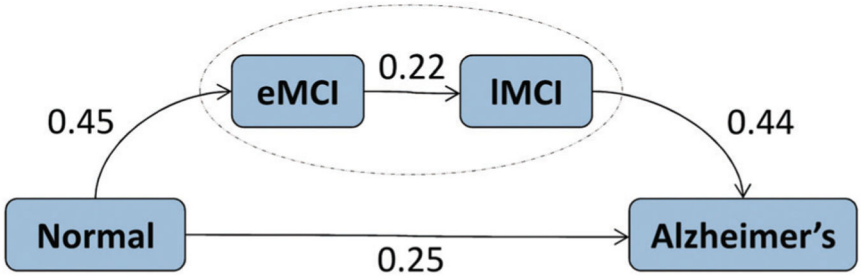

Fig. 7.

Best prediction matching percentage between different Alzheimer's disease states using the Selective variation of the MLP approach obtained by the core Adamic and Adar algorithm.

Table 1

Score functions used by the node neighborhood similarity based link prediction algorithm

Link prediction algorithms	Score function
Common neighbors (CN)	$\mid \Gamma(x) \cap \Gamma(y) \mid$
Jaccard (JC)	$\frac{\mid \Gamma(x) \cap \Gamma(y) \mid}{\mid \Gamma(x) \cup \Gamma(y) \mid}$
Preferential attachments (PA)	$\mid \Gamma(x) \mid \cdot \mid \Gamma(y) \mid$
Adamic and Adar (AA)	$\sum_{s \in \Gamma(x) \cap \Gamma(y)} \frac{1}{\log_2\left(\mid \Gamma(s) \mid \right)}$
Resource allocation (RA)	$\sum_{s \in \Gamma(x) \cap \Gamma(y)} \frac{1}{\mid \Gamma(s) \mid}$

Table 2

Demographic information for 50 controls, 72 eMCI, 38 lMCI and 42 AD participants scanned with diffusion MRI as part of the ADNI project. Mean age and mini mental state exam (MMSE) scores are listed for each diagnostic group

	Controls	eMCI	lMCI	AD	Total
N	50	72	38	42	202
Age (mean ± SD in years)	72.6 ± 6.1	72.4 ± 7.9	72.6 ± 5.6	75.5 ± 8.9	73.1 ± 7.4
MMSE (mean ± SD)	28.9 ± 1.4	28.1 ± 1.5	26.9 ± 2.1	23.3 ± 1.9	27.1 ± 2.7
Sex	22M/28F	45M/27F	25M/13F	28M/14F	120M/82F

Table 3

The average number of edges for each diagnostic group computed using the 70 × 70 connectivity matrices is listed

Matrix	Normal	eMCI	lMCI	AD
Number of edges	1490	1510	1380	1364

Table 4

Confusion matrix representation

	Predicted positive	**Predicted negative**
Actual positive	True positive (TP)	False negative (FN)
Actual negative	False positive (FP)	True negative (TN)

Table 5

Prediction results of the eMCI matrix from the Normal matrix with all 3 variations of the MLP approach (PLP first, MLP First and Selective MLP) and the random method. AA is the most predictive method (p-value < 0.0001) relative to random

	PLP First		NLP First		Selective		Best result	
Method	**Total difference**	**Matching (%)**	**Total difference**	**Matching (%)**	**Total difference**	**Matching (%)**	**Method**	**Matching (%)**
CN	304	15	314	12	304	15	Selective	15
JC	312	12	270	24	270	24	Selective	24
PA	320	10	312	12	320	10	NLP first	12
RA	312	12	196	45	196	45	**Selective**	**45**
AA	312	12	196	45	196	45	Selective	45
Random								12

Table 6

Prediction results for the transition to the lMCI matrix from the eMCI matrix with all 3 variations of the MLP approach (PLP First, MLP First, and Selective MLP) and the random method. AA is proved to be the best method with p-value < 0.0001 relative to random

	PLP First		NLP First		Selective		Best result	
Method	**Total difference**	**Matching (%)**	**Total difference**	**Matching (%)**	**Total difference**	**Matching (%)**	**Method**	**Matching (%)**
CN	376	14	378	13	376	14	Selective	14
JC	418	4	394	10	392	10	Selective	10
PA	356	18	370	15	356	18	Selective	18
RA	378	13	338	22	338	22	Selective	22
AA	378	13	338	22	338	22	**Selective**	**22**
Random								12

Table 7

Prediction results of the AD matrix from the lMCI matrix with all 3 variations of the MLP approach (PLP First, MLP First and Selective MLP) and the random method. AA is proved to be the best method with p-value < 0.0001 relative to random

	PLP First		NLP First		Selective		Best result	
Method	**Total difference**	**Matching (%)**	**Total difference**	**Matching (%)**	**Total difference**	**Matching (%)**	**Method**	**Matching (%)**
CN	294	10	296	10	294	10	Selective	10
JC	312	5	300	9	296	10	Selective	10
PA	282	14	286	13	282	14	Selective	14
RA	276	16	184	44	184	44	Selective	44
AA	276	16	184	44	184	44	**Selective**	**44**
Random								10

Table 8

Prediction results of the AD matrix from the Normal matrix with all 3 variations of the MLP approach (PLP First, MLP First and Selective MLP) and the random method. AA is proved to be the best method with p-value < 0.0001 relative to random

	PLP First		NLP First		Selective		Best result	
Method	**Total difference**	**Matching (%)**	**Total difference**	**Matching (%)**	**Total difference**	**Matching (%)**	**Method**	**Matching (%)**
CN	382	12	386	11	382	12	Selective	12
JC	424	3	404	7	396	9	Selective	9
PA	388	11	382	12	388	11	NLP First	12
RA	368	16	328	25	328	25	Selective	25
AA	368	16	328	25	328	25	**Selective**	**25**
Random								12