RESEARCH ARTICLE

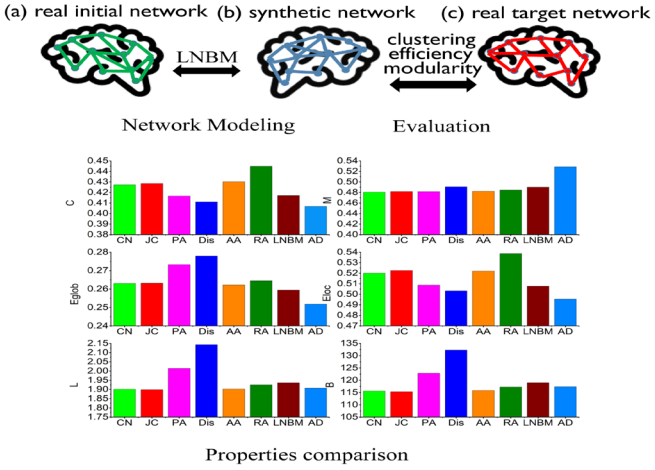
Brain networks modeling for studying the mechanism underlying the development of Alzheimer's disease

Shuai-Zong Si, Xiao Liu*, Jin-Fa Wang, Bin Wang*, Hai Zhao
School of Computer Science and Engineering, Northeastern University, Shenyang, China

Funding: *This study was supported in part by Fundamental Research Funds for the Central Universities in China, No. N161608001 and No. N171903002.*

Graphical Abstract

A local naïve Bayes brain network model for uncovering the wiring changes in human brain networks from normal control to Alzheimer's disease



Correspondence to:
Xiao Liu, Doctoral candidate, liu.xiao.xiao.1881@gmail.com;
Bin Wang, Doctoral candidate, bin.wang.brilliant@gmail.com.

orcid:
0000-0002-4143-7757 (Bin Wang)

doi: *10.4103/1673-5374.257538*

Received: *August 21, 2018*
Accepted: *December 28, 2018*

Abstract

Alzheimer's disease is a primary age-related neurodegenerative disorder that can result in impaired cognitive and memory functions. Although connections between changes in brain networks of Alzheimer's disease patients have been established, the mechanisms that drive these alterations remain incompletely understood. This study, which was conducted in 2018 at Northeastern University in China, included data from 97 participants of the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset covering genetics, imaging, and clinical data. All participants were divided into two groups: normal control (n = 52; 20 males and 32 females; mean age 73.90 ± 4.72 years) and Alzheimer's disease (n = 45, 23 males and 22 females; mean age 74.85 ± 5.66). To uncover the wiring mechanisms that shaped changes in the topology of human brain networks of Alzheimer's disease patients, we proposed a local naïve Bayes brain network model based on graph theory. Our results showed that the proposed model provided an excellent fit to observe networks in all properties examined, including clustering coefficient, modularity, characteristic path length, network efficiency, betweenness, and degree distribution compared with empirical methods. This proposed model simulated the wiring changes in human brain networks between controls and Alzheimer's disease patients. Our results demonstrate its utility in understanding relationships between brain tissue structure and cognitive or behavioral functions. The ADNI was performed in accordance with the Good Clinical Practice guidelines, US 21CFR Part 50–Protection of Human Subjects, and Part 56–Institutional Review Boards (IRBs)/Research Good Clinical Practice guidelines Institutional Review Boards (IRBs)/ Research Ethics Boards (REBs).

Key Words: *nerve regeneration; Alzheimer's disease; graph theory; functional magnetic resonance imaging; network model; link prediction; naïve Bayes; topological structures; anatomical distance; global efficiency; local efficiency; neural regeneration*

Chinese Library Classification No. *R447; R741*

Introduction

Alzheimer's disease (AD) is the most widespread progressive neurodegenerative disorder among older people (Alzheimer's Association, 2018). According to reports from the World Health Organization, approximately 29.8 million people worldwide have AD, which contributes to 60–70% of cases of dementia (World Health Organization, 2017). In contrast to healthy people, AD patients show a decline in cognitive and memory functions. This decline has mainly been attributed to dynamic alterations in structural and functional connections of brain networks in AD patients (Sperling et al., 2014; Frere and Slutsky, 2018). However, despite an impressive increase of knowledge about AD, little is known about the intrinsic causes underlying changes of con-

nections or wiring mechanisms that shape the topologies of brain networks in AD patients. This problem has profound significance for our understanding of how brain regions interact with each other, and is crucial for early diagnosis and treatment of AD.

Recently, resting-state functional magnetic resonance imaging (fMRI) has emerged as a powerful and efficient tool for investigating changes in functional brain organization in AD patients (Harrison et al., 2016; Dimitriadis et al., 2018). Simultaneously, graph theory provides a series of methods to quantify dynamic changes of brain network properties, such as network connectivity, clustering coefficient, modularity, efficiency, and small-worldness between normal control (NC) and AD patients (Zhao et al., 2012; Brier et al., 2014; Stam, 2014; Avena-Koenigsberger et al., 2017; Jalili, 2017; Tan et al., 2017; Wang et al., 2018). Moreover, network models of graph theory *i.e*. Watts–Strogatz (Watts and Strogatz, 1998) and Barabási–Albert (Barabasi and Albert, 1999) have been widely used to generate synthetic networks with the same features as those encountered in real-world networks (Guimerà and Sales-Pardo, 2009; Zhang et al., 2017; Muldoon, 2018). Through network modeling, we can anticipate mechanisms underlying the existence of links among nodes, and thus explain the mechanisms of network reorganization (Muldoon, 2018). Therefore, network modeling has been viewed as a promising method for investigating the dynamic mechanism of interconnections in complex networks. Using this tool, we can produce desired network topologies using appropriate network models.

With the rapid development of network science, network modeling has been applied to brain network simulation (Bullmore and Sporns, 2009; Sporns, 2011; Bassett et al., 2018). Both network topology and anatomical distance are considered as important impactors for forming synthetic topologies similar to observed brain networks (Vértes et al., 2012; Betzel et al., 2016; Cheng et al., 2016). However, it is worth mentioning that models proposed in previous studies are generative models, which means nodes are isolated and no edges exist in the initial simulation networks. Instead, edges are continually added into the initial simulation networks to generate desired synthetic networks. Obviously, these generative models are not suitable for simulating the dynamic process of brain networks from one existing state to another, *i.e*. NC to AD. Simultaneously, new connections will appear, and old links will be cut off in observed brain networks from NC to AD (Tijms et al., 2013; Toussaint et al., 2014). Therefore, models should take both the addition and deletion of edges into account to realize brain network simulation.

Considering the factors mentioned above, this study proposed a brain network model based on a novel link-prediction algorithm named local naïve Bayes (LNB) (Liu et al., 2011) to simulate wiring changes in human brain networks between NC and AD groups. In all cases, we aimed to uncover wiring mechanisms that produce synthetic networks with properties similar to those of real observed brain networks topologies, as shown in **Figure 1**. Data used in this analysis are from imaging and clinical data of NC and AD patients in the Alzheimer's Disease Neuroimaging Initiative dataset.

Materials and Methods

Data acquisition

This study, which was conducted at Northeastern University in 2018, contained a total of 97 participants (aged 61–85 years; average 74 years; 54 males) divided into two groups: NC and AD. All participants were recruited from public resting-state fMRI datasets of the Alzheimer's disease Neuroimaging Initiative (http://adni.loni.ucla.edu). AD patients had a Mini-Mental State Examination score of 17–26 (Tombaugh and McIntyre, 1992) and a Clinical Dementia Rating (Morris, 1993) of 0.5 or 1.0. Participants in the NC group were non-demented, non-depressed, and had an Mini-Mental State Examination score of 24–30 and Clinical Dementia Rating of 0. Demographic and clinical characteristics of the two groups are shown in **Table 1**. Each group included the same number of participants, and no significant differences were found in age distribution among the two groups.

Table 1 Demographic and clinical characteristics of normal control and AD groups

	Normal control group	AD group
N	52	45
Sex (male/female, *n*)	20/32	23/22
Age (year)	73.90±4.72	74.85±5.66
MMSE score	28.69±1.46	23.24±4.02*
CDR score	0.00±0.00	0.92±0.33*

Data of age, MMSE score, and CDR score are expressed as the mean ± SD. *$P < 0.05$, *vs*. normal control group (two sample *t*-test). AD: Alzheimer's disease; MMSE: Mini-Mental State Examination; CDR: Clinical Dementia Rating.

All participants underwent a scan session for brain imaging on a Philips MRI scanner with a 3.0-Tesla field strength. An echo-planar imaging sequence was used to collect resting fMRI scans with the following parameters: flip angle = 80.0°; axial slices = 48; slice thickness = 3.313 mm; slice acquisition order = sequential ascending; echo time = 30 ms; and repetition time = 3000 ms.

Data preprocessing

Acquired images from each participant were preprocessed before constructing real brain networks. A standard data preprocessing strategy was performed using the Data Processing Assistant for Resting-State fMRI software (Yan and Zang, 2010). Data Processing Assistant for Resting-State fMRI software was developed based on the well-known Statistical Parametric Mapping software package (SPM8, Wellcome Department of Cognitive Neurology, Institute of Neurology, London, UK) (http: //www.fil.ion.ucl.ac.uk/spm). Specifically, the steps of data preprocessing for each resting-state scan were as follows: (1) the first seven slice time points were discarded to allow for magnetization equilibrium. (2) Slice timing correction was performed to ensure

all time points in the time domain. (3) Head motion correction was subsequently executed to eliminate movement artifacts in the BOLD time series. Either participants whose head rotation was more than 2.5° or participants whose head translation exceeded 2.5 mm were discarded. (4) Functional volumes were normalized to the standard echo-planar imaging template and re-sliced to 3 × 3 × 3 mm^3 resolution in Montreal Neurological Institute space. (5) Spatial smoothing with a standard 4-mm full width at half-maximum Gaussian kernel was performed for fMRI data to increase the signal-to-noise ratio. (6) Temporal band-pass filtering was performed over each smoothed image in the frequency range from 0.06–0.11 Hz. (7) Both linear and quadratic trends were removed. (8) Other nuisance covariates involving white matter, cerebrospinal fluid, whole-brain signal, and six head motion parameters were regressed out from the preprocessed data. (9) The cerebrum of each participant was further functionally parceled into 90 regions (45 regions for each hemisphere) according to the automated anatomical labeling atlas in Montreal Neurological Institute space (Tzourio-Mazoyer et al., 2002).

Real brain network construction

An effort was made to examine the wiring mechanisms that shape brain network topologies of AD from graph theory. The real brain network of each participant after preprocessing was expressed by a graph that consisted of a constant number of nodes and connections. Nodes represented functional brain regions and connections indicated the internal-relationship amongst these regions. For each participant, 90 regions of interest in the cerebrum on the automated anatomical labeling template were selected to define nodes in brain networks. Functional connectivity then was defined by calculating the Pearson correlation coefficient between whole-run BOLD time courses of each pair of nodes. As a result, each of participant obtained one 90 × 90 inter-regional symmetric correlation matrix. Fisher's r-to-z transformation was performed to improve the normality of correlation coefficients in the matrix. Then, each inter-regional correlation matrix was thresholded to retain only a fraction of the strongest connections for statistical significance. If the corresponding correlation coefficients between two nodes were above the threshold θ, the element of the matrix was set to 1; otherwise, the element of the matrix was set to zero. Finally, a binary adjacent matrix g was obtained for each participant in both NC and AD groups to represent the real functional brain network. Different connectivity densities of the network were generated according to various θ, and θ was set as 15%.

Brain network models

In this subsection, we provided a detailed illustration of our proposed brain network model for generating synthetic brain networks with properties similar to the real brain networks, as obtained from original images of the AD group. Both local topology-based link prediction methods and naïve Bayes classifiers were adopted in calculating connection probabilities between two nodes. We introduced some previous models that had been intensively investigated over recent years before presenting our model.

In the past, simulations of brain networks were based on simple models for cost minimization (Kaiser and Hilgetag, 2004; Zhao et al., 2007). We named this type of model as "Dis" in the current study. The connection probability $p(u,v)$ between any pair of regions in Dis is defined as a function of anatomical distance, given by:

$$p(u,v) = E(u,v)^{-\eta} \quad (1)$$

where $E(u,v)$ represents the anatomical distance between nodes u and v, and η is the parameter for distance penalization. Dis can generate network topologies with minimal cost.

At present, not only anatomical distance but also network topology are considered essential impactors for brain network modeling (Vértes et al., 2012; Brier et al., 2014; Cheng et al., 2016). Vértes et al. (2012) established a hemisphere-brain network using an economical clustering model based on a famous link-prediction algorithm of local topologies named common neighbors (Liben-Nowell and Kleinberg, 2007) to simulate the formation mechanism of human brain networks. Edges were added to the initial network step by step until a total number of connections were built. Connections are created according to the probability given by:

$$p(u,v) = E(u,v)^{-\eta} * S_{(u,v)}^{CN\ \gamma} \quad (2)$$

$$S_{(u,v)}^{CN} = |\Gamma(u) \cap \Gamma(v)| \quad (3)$$

where $\Gamma(u)$ and $\Gamma(v)$ are the neighbors set of nodes u and v, respectively. $S_{(u,v)}^{CN}$ represents the number of common neighbors between nodes u and v, and denotes the contribution of topological similarity to the connection probability. The other term, $E(u,v)$, is the anatomical distance between nodes u and v. η is the parameter of anatomical distance penalization, while γ represents the parameter of topological similarity. Simulation results showed that networks modeled by the economical clustering model could capture an impressive range of topological properties of real functional brain networks.

It should be noted that the economical clustering model gave each common neighbor the same contribution of topological similarity between nodes u and v. However, different common neighbors might have different degrees, thus playing various roles during the formation of a connection. Therefore, LNB was proposed to better classify these common neighbors (Liu et al., 2011). In LNB, two pairs of nodes with the same number of neighbors in common might get different probabilities during connection establishment.

Taking the mechanism of economical clustering model and naïve Bayes classifiers as a basis, this study presented a novel brain network model named LNBM to investigate changes in the wiring mechanism of human brain networks between NC and AD patients. We assumed that gradual changes of local topologies in NC networks resulted the topologies of AD. Thus, LNBM took local topologies of common neighbors into consideration. Notably, different from economical clustering models, our proposed model considered both adding and deleting connections during the

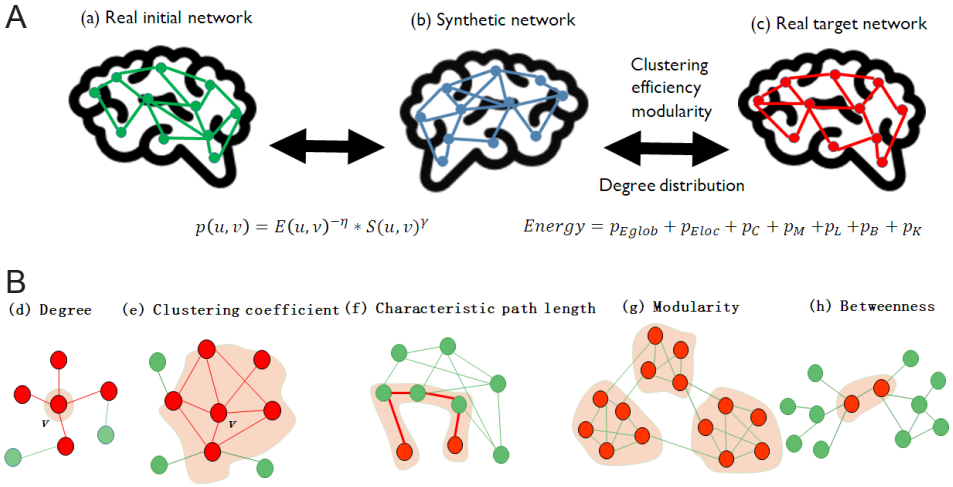


Figure 1 Network modeling and evaluation.
(A) Three different states of brain networks including (a) real initial network, (b) synthetic network, and (c) real target network. Brain network modeling starts from the initial network; then connections are added and deleted continuously according to the probabilities of $p(u,v)$ to form the synthetic network; $E(u,v)$ is the anatomical distance between nodes u and v, $S(u,v)$ represents the topological similarity, η is the parameter of anatomical distance penalization, γ represents the parameter of topological similarity. We aimed to reproduce synthetic networks with similar properties of real target networks. For each property, P-values of the two samples was examined by t-test between synthetic networks and real target networks. P_c is the P-value of clustering coefficient; P_M represents the P-value of modularity; P_{Eglob} is the P-value of global efficiency; P_{Eloc} represents the P-value of modularity; P_L and P_B indicate P-values of characteristic path length and betweenness, respectively; P_K represents the P-value of Kolmogorov-Smirnoff test for the difference in degree distributions between synthetic networks and real target networks. Finally, all P-values of these network properties were considered as an *Energy* function to evaluate different models. (B) Descriptions of interested network properties. Degree (K) indicates the number of edges connecting with a node. Clustering coefficient is a measure of the degree to which nodes in a network exhibit a tendency to cluster together. Characteristic path length is defined as the average shortest path length between all node pairs in the network; it is a measure to quantify the efficiency for information transmission. Modularity is used to detect the strength of the division of a network into communities or modules. Betweenness is the number of shortest paths in the network that pass through the node; it is designed for quantifying the control of information flow.

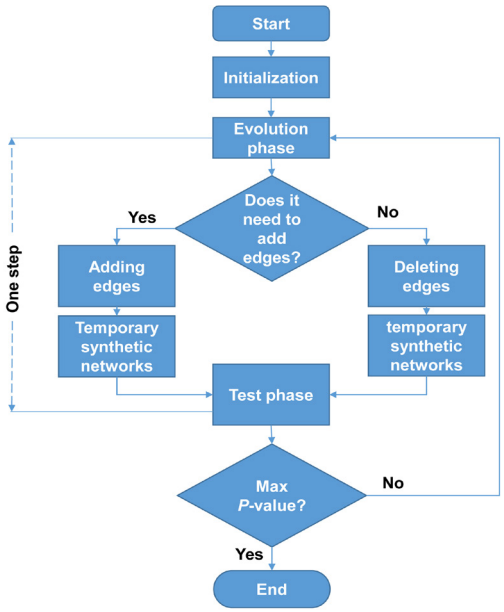


Figure 2 Construction process for brain networks.

construction of synthetic networks.

As shown in **Figure 2**, the construction process of brain networks was simulated step by step. Each step was dynamically performed in two phases, commencing with an evolution phase that was then followed by a two-sample t-test phase.

(1) Initialization: At the beginning of the simulation, all participants in NC and AD groups were preprocessed and the corresponding initial brain networks $G_c=(g_{c1}, g_{c2},\ldots, g_{cn})$, $G_a=(g_{a1}, g_{a2},\ldots, g_{an})$ were obtained from real fMRI data. Each initial network consisted of a fixed number of nodes $|V|$ = 90 and edges sets were represented by $E_c=(e_{c1}, e_{c2},\ldots, e_{cn})$, $E_c=(e_{c1}, e_{c2},\ldots, e_{cn})$, respectively.

(2) Evolution: After initialization, the evolution phase in the first step started. First, the connection probabilities of any node pair were calculated for each participant. Second, each node pair was sorted according to its connection probability. Next, we set a constant number α of connections to establish and another constant number β of connections to delete in every step. Here, α < β because a reduced number of edges was found in real brain networks of AD. Specially, the node pair with the largest connection probability could establish a link if its two nodes disconnected from each other. Meanwhile, the node pair with the smallest connection probability could cut off its link if there was a connection between its two nodes. It is noted that a new pair of nodes must be chosen based on the connection probability of LNBM, if the degree of either of the two selected nodes was less than 2 when deleting the edge between them. The simulation could not proceed to the t-test phase until α connections were established and β connections were successfully deleted for each participant.

(3) Two-sample *t*-test: After evolution, a *t*-test phase starts, and a *P*-value of a two-sample *t*-test for the difference in number of edges between the synthetic network and real target network, *i.e.* brain networks of AD group, was calculated. The construction process ended when the *P*-value stopped increasing, and the synthetic network with the max *P*-value in the current step was treated as the final synthetic network.

The definition of connection probabilities of LNBM is given by:

$$p(u,v) = E(u,v)^{-\eta} * S_{(u,v)}^{LNB\,\gamma} \tag{4}$$

where $S_{(u,v)}^{LNB}$ is the topological similarity calculated by common neighbors and a naïve Bayes classifier. We provide a comprehensive description of how to calculate this similarity, as follows.

We first gave an illustration of the naïve Bayes classifier. Naïve Bayes classifiers are a family of simple probabilistic classifiers based on Bayes' theorem that are particularly suited for assumptions of strong independence among features. Abstractly, the conditional probability model of a classifier is written as $p(C_k|\, x_1,\ldots, x_n)$, in which C_k represents a dependent class, and $X=(x_1, \ldots, x_n)$ denotes the features. Using Bayes' theorem, the posterior probability $p(C|X)$, given by equation (5), can be calculated from $p(C)$, $p(X)$ and $P(X|C)$.

$$p(C_k|x_1, \ldots, x_n) = \frac{p(C_k)p(x_1,\ldots,x_n|C_k)}{p(x_1,\ldots,x_n)} \tag{5}$$

Taking the "naïve" conditional independence assumptions into consideration, each feature x_i of a given class C_k is conditionally independent of every other feature x_j for $i \neq j$. Therefore, the posterior probability can be rewritten as:

$$p(C_k|x_1, \ldots, x_n) = \frac{p(C_k)p(x_1|C_k)\cdots p(x_n|C_k)}{p(x_1,\cdots,x_n)} = \frac{p(C_k)\prod_{i=1}^{n} p(x_i|C_k)}{p(x_1,\cdots,x_n)} \tag{6}$$

Given a graph *g*, $S_{(u,v)}^{LNB}$ aims to calculate the probability of a connection existence between two nodes. The probability for one connection establishment between nodes *u* and *v* conditionally depends on its local topological similarity, *i.e.*, number of common neighbors. Therefore, according to Bayesian theory, the posterior probability that nodes *u* and *v* are connected and disconnected are respectively given by:

$$p(C_{u,v}|O_{u,v}) = \frac{p(C_{u,v})}{p(O_{u,v})} \cdot \prod_{\xi\in O_{u,v}} p(\xi|C_{u,v}) \tag{7}$$

$$p(\hat{C}_{u,v}|O_{u,v}) = \frac{p(\hat{C}_{u,v})}{p(O_{u,v})} \cdot \prod_{\xi\in O_{u,v}} p(\xi|\hat{C}_{u,v}) \tag{8}$$

where $O_{u,v}$ denotes the common neighbors set of two nodes *u* and *v*. $p(C_{u,v})$ and $p(\hat{C}_{u,v})$ denote the prior probabilities of connection and disconnection between nodes *u* and *v*, respectively, and both can be calculated according to a given graph matrix *g* as follows:

$$p(C_{u,v}) = \frac{|E|}{|V|(|V|-1)/2} \tag{9}$$

$$p(\hat{C}_{u,v}) = 1 - \frac{|E|}{|V|(|V|-1)/2} \tag{10}$$

where $|E|$ and $|V|$ represent existing edges number and nodes number in *g*, respectively.

For each node ξ, $p(\xi|C_{u,v})$ denotes the conditional probability that two connected nodes *u* and *v* have a neighbor node ξ in common, and $p(\xi|\hat{C}_{u,v})$ is the conditional probability that two disconnected nodes *u* and *v* have a neighbor node ξ in common. According to Bayesian theory, these two conditional probabilities are given by:

$$p(\xi|C_{u,v}) = \frac{p(\xi)\cdot p(C_{u,v}|\xi)}{p(C_{u,v})} \tag{11}$$

$$p(\xi|\hat{C}_{u,v}) = \frac{p(\xi)\cdot p(\hat{C}_{u,v}|\xi)}{p(\hat{C}_{u,v})} \tag{12}$$

where $p(\xi)$ represents the probability that a selected node ξ is a common neighbor of one node pair. As a result, the topological similarity $S_{(u,v)}^{LNB}$ between node *u* and node *v* is defined as the ratio of $p(C_{u,v}|O_{u,v})$ and $p(\hat{C}_{u,v}|O_{u,v})$ in equation (13).

$$S_{(u,v)}^{LNB} = \frac{p(C_{u,v})}{p(\hat{C}_{u,v})} \cdot \prod_{\xi\in O_{u,v}} \frac{p(\hat{C}_{u,v})\cdot p(C_{u,v}|\xi)}{p(C_{u,v})\cdot p(\hat{C}_{u,v}|\xi)} \tag{13}$$

$p(C_{u,v}|\xi)$ can be obtained by calculating the clustering coefficient of ξ, given by:

$$p(C_{u,v}|\xi) = \frac{2|E_\xi|}{|N_\xi|(|N_\xi|-1)} \tag{14}$$

where N_ξ denotes neighbors set of ξ, $|N_\xi|$ represents existing edges number in N_ξ and $|N_\xi|(|N_\xi|-1)/2$ is the edges number could possibly exist within N_ξ. As $p(C_{u,v}|\xi) + p(\hat{C}_{i,v}|\xi) = 1$, we have:

$$p(\hat{C}_{u,v}|\xi) = 1 - \frac{2|E_\xi|}{|N_\xi|(|N_\xi|-1)} \tag{15}$$

Therefore, the topological similarity can be written as:

$$S_{(u,v)}^{LNB} = s^{-1}\prod_{\xi\in O_{u,v}} s\cdot R_\xi \tag{16}$$

$$s = \frac{p(\hat{C}_{u,v})}{p(C_{u,v})} = \frac{|V|(|V|-1)-2|E|}{2|E|} \tag{17}$$

$$R_\xi = \frac{2|E_\xi|}{|N_\xi|(|N_\xi|-1)-2|E_\xi|} \tag{18}$$

From equation (16), we observe that $S_{(u,v)}^{LNB}$ is equal to $S_{(u,v)}^{CN}$ in equation (2) if R_ξ=1. Clearly, each node has its own local topology, *i.e.*, degree, connections. Thus, different common neighbors $\xi \in O_{u,v}$ can achieve different R_ξ, thereby differentially contributing to topological similarity in calculating the connecting probability.

Additionally, to achieve a linear formula for the topological similarity of connection establishment, a logarithmic function is then used on both sides of equation (16). Thus, the topological similarity $S_{(u,v)}^{LNB\,\prime}$ is given by:

$$S_{(u,v)}^{LNB\,\prime} = s^{-1}\{|O_{c,v}|\log(s) + \textstyle\sum_{\xi\in O_{u,v}}\log(R_\xi)\} \tag{19}$$

Evaluation of simulation networks

For different η and γ, an *Energy* function was defined as the performance of evaluation based on *P*-values for the difference in degree distribution, characteristic path length (L), clustering coefficient (C), modularity (M), betweenness (B), global efficiency (Eglob), and local efficiency (Eloc) between synthetic networks and real target networks. These properties were selected because they could reflect the most important

aspects of network performance. Thus, the evaluation can be more convincing and reasonable considering these properties. A detailed description of these properties is shown in **Figure 1B**. Simulated annealing is performed on an energy function to find the optimal η and γ in parameter space that maximizes *Energy*. We aimed to find the optimal synthetic networks that best fit real data networks of the target group (*i.e.*, AD group). The definition of *Energy* is given by:

$$Energy = p_{Eglob} + p_{Eloc} + p_C + p_M + p_L + p_B + p_K \quad (20)$$

where p_{Eglob} and p_{Eloc} are *P*-values associated with a two-samples *t*-test for differences in global efficient and local efficient, respectively, between a set of 45 simulated model networks and set of 52 real observed brain networks of the AD group. Similarly, p_{C}, p_{M}, p_{L}, and p_{B} are *P*-values of a two-sample *t*-test for differences between clustering coefficients, modularity, characteristic path length, and betweenness in modeled *vs.* real target networks of AD group, respectively; and, p_{K} is the *P*-value of the Kolmogorov-Smirnoff test for the difference in degree distributions of synthetic networks *vs.* real target networks of the AD group. Larger values of *Energy* indicate that the simulated model network is more similar to the real target network of AD.

Results

In this study, to simultaneously take both network topology information and anatomical distance into account, a novel brain network model named LNBM was proposed for better exploration of the dynamic mechanism of human functional brain networks changing from normal control to AD. Definitions of the compared models are listed in **Table 2**. **Table 3** records the maximum *Energy* of each model and its corresponding optimal parameters η and γ in the parameter space.

Table 2 Comparison of link prediction index

Link prediction index	Index abbreviation	Mathematical definition
Common neighbor	CN	$S^{CN}_{(u,v)} = \lvert\Gamma(u) \cap \Gamma(v)\rvert$
Jaccard	JC	$S^{JC}_{(u,v)} = \frac{\lvert\Gamma(u) \cap \Gamma(v)\rvert}{\lvert\Gamma(u) \cup \Gamma(v)\rvert}$
Preferential attachment	PA	$S^{PA}_{(u,v)} = \lvert\Gamma(u)\rvert \times \lvert\Gamma(v)\rvert$
Adamic-Adar	AA	$S^{AA}_{(u,v)} = \sum_{\xi \in \lvert\Gamma(u) \cap \Gamma(v)\rvert} \frac{1}{\log \lvert\xi\rvert}$
Resource allocation	RA	$S^{RA}_{(u,v)} = \sum_{\xi \in \lvert\Gamma(u) \cap \Gamma(v)\rvert} \frac{1}{\lvert\xi\rvert}$

$\Gamma(u)$ represents the neighbor set of node u; $|\Gamma(u)|$ represents the number of neighbors of node u.

Topological differences in brain networks between normal control and AD patients

In this subsection, topological differences in brain networks were investigated between NC and AD patients. As shown in **Figure 3**, six topological features are presented. The results indicate that clustering coefficient, global efficiency, local efficiency, characteristic path length, and betweenness of brain networks in the AD group were decreased compared with the normal control group; whereas, modularity of AD *vs.* NC indicated an increase. Analysis of these topological differences guided the construction of brain networks and also helped us to evaluate the performance of our proposed brain network model.

Evaluation of anatomical distance based model

The one-parameter model, Dis, was first investigated to analyze the effects of anatomical distance on brain network modeling. Dis did not consider topological structures compared with other models. We found that at a particular value of the model parameter η = 3, as shown in **Table 3**, simulated networks based on Dis had an average clustering coefficient that exactly matched the clustering coefficient in brain networks of AD patients (P = 0.8625), followed by PA (P = 0.6855) and LNBM (P = 0.6680). Whereas, the clustering coefficient of networks simulated by resource allocation and Adamic-Adar exhibited a worse fitness of the real data compared with other models, with P values of 0.1085 and 0.3338, respectively. Moreover, Dis generated synthetic networks that best fitted the real networks of AD with regard to properties of modularity (P = 0.3353) and local efficiency (P = 0.8045), followed by LNBM (P = 0.2742 and 0.6973, respectively). However, the results shown in **Table 3** demonstrated that, although synthetic networks generated by the Dis model were best-fitting for clustering coefficient, local

Table 3 Minimum *Energy* values and *P*-values of various models

Index	*Energy*	γ	η	P_C	P_M	P_{Eglob}	P_{Eloc}	P_L	P_B	P_K
CN	3.9337	2	2.5	0.4000	0.1705	0.6896	0.4368	0.9726	0.8886	0.3756
JC	3.7404	1	3	0.3663	0.1831	0.6858	0.3851	0.9617	0.8759	0.2823
PA	3.4843	2	6	0.6855	0.1894	0.4491	0.6731	0.5375	0.6674	0.2822
Dis	3.0910	–	3	0.8625	0.3353	0.3654	0.8045	0.1859	0.2551	0.2823
AA	3.8017	8	3.5	0.3338	0.1930	0.7106	0.3965	0.9789	0.9066	0.2823
RA	3.4295	4.5	7.5	0.1085	0.2169	0.6513	0.1615	0.9217	0.9941	0.3756
LNBM	4.5811	1	7	0.6680	0.2742	0.7845	0.6973	0.8762	0.9054	0.3756

P_c represents the *P*-value for the difference in clustering coefficient between synthetic networks and real target networks; P_M is the *P*-value of modularity; P_{Eglob} is the *P*-value of global efficiency; P_{Eloc} represents the *P*-value of modularity; P_L and P_B indicate the *P*-values of characteristic path length and betweenness, respectively; P_K represents the *P*-value of Kolmogorov-Smirnoff test for the difference in degree distributions between synthetic networks and real target networks. CN: Common neighbor; JC: Jaccard; PA: preferential attachment; Dis: a function of anatomical distance; AA: Adamic–Adar; RA: resource allocation; LNBM: local naïve Bayes model.

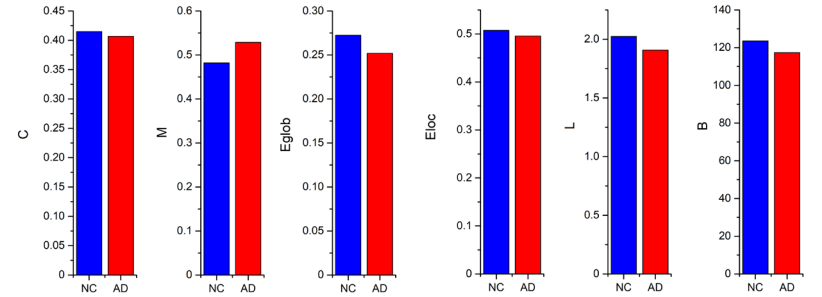


Figure 3 Differences of topological properties between real brain networks of NC and AD.
NC represents the real brain networks of the normal control group; AD represents the real brain networks of the Alzheimer's disease group; C represents the clustering coefficient; M is the modularity; Eglob is the global efficiency; Eloc indicates the local efficiency; L represents the characteristic path length; and B indicates betweenness.

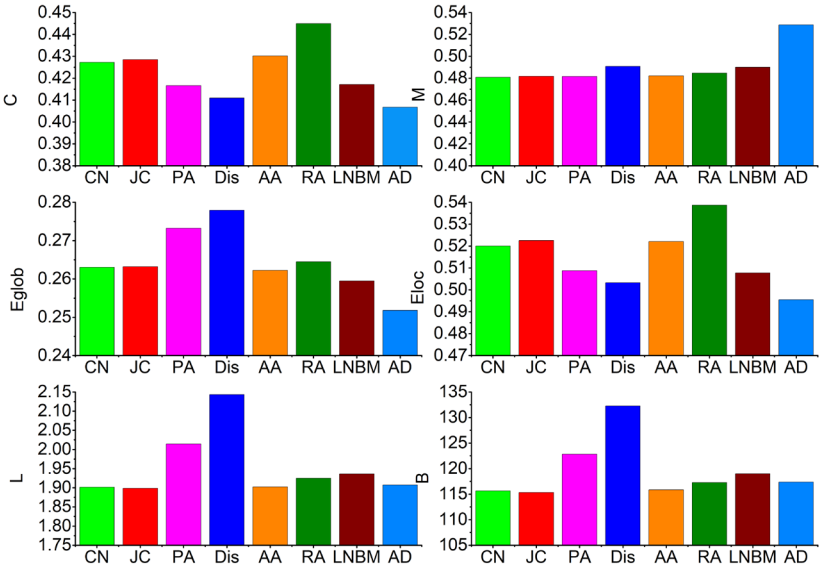


Figure 4 Topological properties of real target brain networks and synthetic brain networks.
C is the clustering coefficient; M represents the modularity; Eglob is the global efficiency; Eloc indicates the local efficiency; L is the characteristic path length and B represents betweenness. CN: Common neighbor; JC: Jaccard; PA: preferential attachment; Dis: a function of anatomical distance; AA: Adamic–Adar; RA: resource allocation; LNBM: local näive Bayes model; AD: Alzheimer's disease.

efficiency, and modularity, they could not simultaneously match observed networks of AD in terms of global efficiency, characteristic path length, and betweenness. Indeed, Dis fitting was the worst in terms of characteristic path length (P = 0.1859) and betweenness (P = 0.2551), which resulted in the smallest *Energy* = 3.0636 among all models. We achieved similar results from a further finding, as shown in **Figure 4**.

Evaluation of models considering both network topology and anatomical distance

In this subsection, models were evaluated by taking both network topological similarity and anatomical distance into account. As shown in **Table 3**, with the best-fitting parameters, models considering both network topologies and anatomical distance outperformed the pure anatomical distance-based model, generating synthetic networks with significantly larger *Energy*. Importantly, our proposed model LNBM achieved the largest *Energy* = 4.5811, with optimal parameters η = 1 and γ = 7, followed by common neighbors (*Energy* = 3.9337, η = 2 and γ = 2.5) and Adamic–Adar (*Energy* = 3.8017, η = 8 and γ = 3.5). These results indicated that LNBM reproduced synthetic networks capturing all of the key topological characteristics of functional brain networks. Specially, LNBM minimized the mismatch between an observed brain network and synthetic networks with regard to the property of global efficiency (P = 0.7845). Therefore, LNBM was the best model for simulating the dynamic mechanism that caused changes of connections in brain networks of AD patients. In contrast, the performance of common neighbors and Adamic–Adar was primarily limited by mismatches in local efficiency and clustering coefficient. In addition, we found that although resource allocation achieved higher P-values for characteristic path length and betweenness, it failed to match real AD networks well for topological properties in terms of clustering coefficient and local efficiency. Moreover, the results shown in **Table 3** confirmed that no minimum P-values were found for PA for any network properties. However, P-values for all of these properties were not as large as observed for LNBM, which resulted in a smaller *Energy* of PA. To prepare a detailed illustration of the results described above, average values of different properties were also provided in synthetic networks (**Figure 4**).

Discussion

This study presented several simple models for brain network simulation to explore mechanisms underlying alterations of brain network connections of AD patients. Our work makes several contributions. First, we showed that there are apparent changes in topological properties between brain networks in NC and AD patients, and confirmed that a one-parameter model considering only anatomical distance could not simultaneously capture all the complex topological properties of real brain networks. Moreover, this study verified that applying additional topological factors to the one-parameter model could dramatically improve the simulation of realistic brain networks features. Notably, our proposed model LNBM was the best-fitting model. Synthetic networks reproduced by LNBM provided a good account of not only characteristic path length and betweenness properties, but also

clustering coefficient, modularity, and efficiency. Therefore, our results demonstrated that the calculation of connection probability for the formation of links between nodes in LNBM corresponded to mechanisms causing the changes of topological properties during the pathogenesis of AD.

It has been demonstrated that cognitive functions such as learning, attention, and working memory over neurodevelopment are remarkably associated with networks topologies (Guo et al., 2014). Moreover, neurodevelopmental disorders can lead to cognitive deficits and cause diseases such as AD, and this cognitive decline can be reflected by changes in network topologies. Here, our findings suggest that alterations of topological properties, *i.e.* modularity and clustering coefficient, were observed between functional network topologies of NC and AD groups, which further verified the notion described above. Recent studies have indicated that these alterations are caused by both decreases and increases of connections between brain regions (Tijms et al., 2013; Toussaint et al., 2014). However, in previous studies, models were established by either adding or deleting edges to predict developmental changes between brain regions. Although network topologies generated by these models can successfully match several properties of real target networks, these efforts were not realistic without considering these two operations together. Thus, these models failed to identify potential mechanisms that caused topological alternations between NC and AD groups. Therefore, for more complete understanding of mechanisms underlying these changes, we focused on an explicit simulation model of the brain that considered both the addition and deletion of edges.

Considerable progress for simulation of brain networks has been made in previous models emphasizing the importance of penalizing anatomical distance (Samu et al., 2014; Roberts et al., 2016). This work is motivated by one outstanding idea that the metabolic cost of establishing a connection between brain regions is proportional to anatomical distance. The more considerable the anatomical distance between two brain regions, the more substantial the metabolic cost. Consequently, the brain prefers to build connections between anatomically neighboring nodes to minimize global costs. Many aspects of the current study support this expectation. For example, Alexander-Bloch et al. (2013) reported that the average distance of connected edges in human brain networks was short, while Meunier et al. (2010) confirmed that topological modules in brain networks were usually composed of anatomically neighboring brain regions. Therefore, connections in the module were often over short distances. As a result, most previous models adopted anatomical distance as the parameter in brain network simulations to achieve the lowest wiring cost, *i.e.*, Dis.

However, models considering only the distance parameter have generated much debate in recent years. One obvious limitation of previous models is that they ignore the fact that brain networks also exhibit high topological efficiency (Bullmore and Sporns, 2012), which has mainly been attributed to the existence of long-distance connections between anatomically localized brain regions. Brain network efficiency is a critical property of complex networks for evaluating the capacity of exchanging information. Previous research has demonstrated that the topological property of efficiency is associated with the cognitive functions of human brain networks (Bassett and Sporns, 2017). This finding was contrary to the cost conservation principle in prior studies. Moreover, extensive work has suggested that distance penalization cannot reproduce network topologies with the same properties as real brain networks (Vértes et al., 2012; Betzel et al., 2016). These findings are in accordance with our results shown in **Table 1**. Therefore, the formation of brain networks cannot be modeled by minimization cost alone. Both distance penalization and topological structures should be simultaneously considered to generate more realistic simulation networks.

This work investigated several competitive models that tradeoff between topological structures and distance penalization, and further discussed differential influences of these particular topological structures on the reorganization of brain networks. Our common neighbors model considers neighbors overlap when calculating the connection probability between two nodes. Jaccard, Adamic–Adar, and resource allocation are variations of common neighbors. Jaccard defines similarity for link establishment with consideration of both the size of neighbor overlap and union of neighbor sets. Adamic–Adar and resource allocation consider both neighbor overlap and the degree of each common neighbor. In the PA model, one node gets a higher preferential attachment probability for a larger degree of connections to other nodes, *i.e.*, the hub node. As shown in **Table 3**, our results indicate that with the best fitting parameters estimated by simulated annealing, models considering both network topologies and anatomical distances reproduce better likeness of brain networks than those generated by distance penalization alone. Comparing *P*-values of different models in **Table 3**, we also observed that by designing a naïve Bayesian model for the classification of connection strength after deleting and adding edges continuously, LNBM is absolutely the best model for generating synthetic networks that best fit the real target network of AD. Importantly, our results indicate that alterations in the connections between topologies of NC and AD are not random, and follow a propriety mechanism. Our findings draw attention to the critical need for understanding the relationship between structural mechanisms and decline of cognitive function.

Our study has some limitations. First, by setting different thresholds for θ, we can obtain real brain networks with a different number of connections. Thus, it might be insufficient to verify our findings with only one θ, and future study performed on a larger range of θ is necessary to further confirm our findings. However, because the number of connections is proportional to θ, we assume that the number of connections in real AD brain networks is still less than that of NC with a different θ. Thus, variations of θ cannot significantly influence our proposed evolution mechanism. Second, the evaluation of fitness between synthetic networks and real target networks was performed according to several selective properties. Whether these properties are the most appropriate measures for network comparison is not yet known. Third, the link prediction index used in this study is based on local topologies for reducing the computational complexity; as such, some global topological structure-based methods should also be considered. We aim to overcome these limitations in future studies.

In conclusion, we have presented how network modeling in graph theory can be used for functional connectivity simulation of AD brain networks. We focused on the evolution

of brain networks with regard to how connections are established or deleted between nodes. Our efforts aim to reveal the mechanisms underlying alterations in brain function during the progression from NC to AD, and contribute to the understanding of neural regeneration patterns in brain networks. We have demonstrated that both topological structures and anatomical distance significantly influence modeling of this development. As such, flexible reconfiguration of the complex mechanism of interactions between brain regions enables further understanding of how network structure and function influence human cognition and behavior.

Acknowledgments: *We would like to thank Cai Wei (from Neusoft) and He Xuan (from Northeastern University) for useful comments on the final version of this manuscript. We also thank the Alzheimer's Disease Neuroimaging Initiative (ADNI) for allowing us to access the database. ADNI is a longitudinal multicenter study, data collection and sharing for this project was funded by the National Institutes of Health Grant U19 AG024904. ADNI aims at achieving the goal of early detection, tracking, and treatment of Alzheimer's disease (AD).*

Author contributions: *Study conception and design: SZS and HZ; literature retrieval: SZS and JFW; experiment implementation and data analysis: SZS and BW; paper preparation: SZS, BW and JFW; paper editing and review: XL. All authors approved the final version of this paper for publication.*

Conflicts of interest: *The authors declare that they have no competing interests.*

Financial support: *This work was supported in part by Fundamental Research Funds for the Central Universities in China, No. N161608001 and No. N171903002. The funding sources had no role in study design, conception, analysis or interpretation of data, writing and deciding to submit this paper for publication.*

Institutional review board statement: *The ADNI was performed in accordance with the Good Clinical Practice guidelines, US 21CFR Part 50–Protection of Human Subjects, and Part 56–Institutional Review Boards (IRBs)/Research Good Clinical Practice guidelines Institutional Review Boards (IRBs)/ Research Ethics Boards (REBs).*

Declaration of patient consent: *The authors state that they have obtained approval from the ADNI Data Sharing and Publications Committee for use of the data. ADNI certify that they have obtained all appropriate legal guardian consent forms. In the forms, the legal guardians have given their consent for the patients' images and other clinical information to be reported in the journal. The patients' legal guardians understand that the patients' names and initials will not be published and due efforts will be made to conceal their identity.*

Reporting statement: *The writing and editing of the article was performed in accordance with the STrengthening the Reporting of OBservational Studies in Epidemiology Statement.*

Biostatistics statement: *The statistical methods of this study were reviewed by the biostatistician of Northeastern University, China.*

Copyright license agreement: *The Copyright License Agreement has been signed by all authors before publication.*

Data sharing statement: *Datasets analyzed during the current study are available from the corresponding author on reasonable request.*

Plagiarism check: *Checked twice by iThenticate.*

Peer review: *Externally peer reviewed.*

Open peer reviewer: *Aysegul Yildiz-Unal, Mugla Sitki Kocman University, Turkey.*

Additional file: *Open peer review report 1.*

References

Alexander-Bloch AF, Vertes PE, Stidd R, Lalonde F, Clasen L, Rapoport J, Giedd J, Bullmore ET, Gogtay N (2013) The anatomical distance of functional connections predicts brain network topology in health and schizophrenia. Cereb Cortex 23:127-138.

Alzheimer's Association (2018) 2018 Alzheimer's disease facts and figures. Alzheimers Dement 14:367-429.

Avena-Koenigsberger A, Misic B, Sporns O (2017) Communication dynamics in complex brain networks. Nat Rev Neurosci 19:17-33.

Barabasi AL, Albert R (1999) Emergence of scaling in random networks. Science 286:509-512.

Bassett DS, Sporns O (2017) Network neuroscience. Nat Neurosci 20:353-364.

Bassett DS, Zurn P, Gold JI (2018) On the nature and use of models in network neuroscience. Nat Rev Neurosci 19:566-578.

Betzel RF, Avena-Koenigsberger A, Goñi J, He Y, de Reus MA, Griffa A, Vértes PE, Mišic B, Thiran JP, Hagmann P, van den Heuvel M, Zuo XN, Bullmore ET, Sporns O (2016) Generative models of the human connectome. Neuroimage 124:1054-1064.

Brier MR, Thomas JB, Fagan AM, Hassenstab J, Holtzman DM, Benzinger TL, Morris JC, Ances BM (2014) Functional connectivity and graph theory in preclinical Alzheimer's disease. Neurobiol Aging 35:757-768.

Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10:186-198.

Bullmore E, Sporns O (2012) The economy of brain network organization. Nat Rev Neurosci 13:336-349.

Cheng C, Chen J, Cao X, Guo H (2016) Comparison of local information indices applied in resting state functional brain network connectivity prediction. Front Neurosci 10:585.

Dimitriadis SI, Liparas D, Alzheimer's Disease Neuroimaging Initiative (2018) How random is the random forest? Random forest algorithm on the service of structural imaging biomarkers for Alzheimer's disease: from Alzheimer's disease neuroimaging initiative (ADNI) database. Neural Regen Res 13:962-970.

Frere S, Slutsky I (2018) Alzheimer's disease: from firing instability to homeostasis network collapse. Neuron 97:32-58.

Guimerà R, Sales-Pardo M (2009) Missing and spurious interactions and the reconstruction of complex networks. Proc Natl Acad Sci U S A 106:22073-22078.

Guo H, Cheng C, Cao X, Xiang J, Chen J, Zhang K (2014) Resting-state functional connectivity abnormalities in first-onset unmedicated depression. Neural Regen Res 9:153-163.

Harrison TM, Burggren AC, Small GW, Bookheimer SY (2016) Altered memory-related functional connectivity of the anterior and posterior hippocampus in older adults at increased genetic risk for Alzheimer's disease. Hum Brain Mapp 37:366-380.

Jalili M (2017) Graph theoretical analysis of Alzheimer's disease: Discrimination of AD patients from healthy subjects. Inform Sci 384:145-156.

Kaiser M, Hilgetag CC (2004) Modelling the development of cortical systems networks. Neurocomputing 58-60:297-302.

Liben-Nowell D, Kleinberg J (2007) The link-prediction problem for social networks. J Am Soc Inf Sci 58:1019-1031.

Liu Z, Zhang QM, Lü L, Zhou T (2011) Link prediction in complex networks: A local naïve Bayes model. Europhys Lett 96:48007.

Meunier D, Lambiotte R, Bullmore ET (2010) Modular and hierarchically modular organization of brain networks. Front Neurosci 4:200.

Morris JC (1993) The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 43:2412-2414.

Muldoon SF (2018) Multilayer network modeling creates opportunities for novel network statistics: Comment on "Network science of biological systems at different scales: A review" by Gosak et al. Phys Life Rev 24:143-145.

Roberts JA, Perry A, Lord AR, Roberts G, Mitchell PB, Smith RE, Calamante F, Breakspear M (2016) The contribution of geometry to the human connectome. Neuroimage 124:379-393.

Samu D, Seth AK, Nowotny T (2014) Influence of wiring cost on the large-scale architecture of human cortical connectivity. PLoS Comput Biol 10:e1003557.

Sperling R, Mormino E, Johnson K (2014) The evolution of preclinical Alzheimer's disease: implications for prevention trials. Neuron 84:608-622.

Sporns O (2011) Networks of the Brain. Cambridge, USA: The MIT Press.

Stam CJ (2014) Modern network science of neurological disorders. Nat Rev Neurosci 15:683-695.

Tan TT, Wang D, Huang JK, Zhou XM, Yuan X, Liang JP, Yin L, Xie HL, Jia XY, Shi J, Wang F, Yang HB, Chen SJ (2017) Modulatory effects of acupuncture on brain networks in mild cognitive impairment patients. Neural Regen Res 12:250-258.

Tijms BM, Wink AM, de Haan W, van der Flier WM, Stam CJ, Scheltens P, Barkhof F (2013) Alzheimer's disease: connecting findings from graph theoretical studies of brain networks. Neurobiol Aging 34:2023-2036.

Tombaugh TN, McIntyre NJ (1992) The mini-mental state examination: a comprehensive review. J Am Geriatr Soc 40:922-935.

Toussaint PJ, Maiz S, Coynel D, Doyon J, Messe A, de Souza LC, Sarazin M, Perlbarg V, Habert MO, Benali H (2014) Characteristics of the default mode functional connectivity in normal ageing and Alzheimer's disease using resting state fMRI with a combined approach of entropy-based and graph theoretical measurements. Neuroimage 101:778-786.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15:273-289.

Vértes PE, Alexander-Bloch AF, Gogtay N, Giedd JN, Rapoport JL, Bullmore ET (2012) Simple models of human brain functional networks. Proc Natl Acad Sci U S A 109:5868-5873.

Wang WW, Lu YC, Tang WJ, Zhang JH, Sun HP, Feng XY, Liu HQ (2018) Small-worldness of brain networks after brachial plexus injury: A resting-state functional magnetic resonance imaging study. Neural Regen Res 13:1061-1065.

Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393:440-442.

World Health Organization (2017) Dementia. http://www.who.int/en/news-room/fact-sheets/detail/dementia.

Yan CG, Zang YF (2010) DPARSF: A MATLAB toolbox for "Pipeline" data analysis of resting-state fMRI. Front Syst Neurosci 4:13.

Zhang X, Moore C, Newman MEJ (2017) Random graph models for dynamic networks. Eur Phys J B 90:200.

Zhao QB, Feng HB, Tang YY (2007) Modelling human cortical network in real brain space. Chin Phys Lett 24:3582-3585.

Zhao X, Liu Y, Wang X, Liu B, Xi Q, Guo Q, Jiang H, Jiang T, Wang P (2012) Disrupted small-world brain networks in moderate Alzheimer's disease: a resting-state fMRI study. PLoS One 7:e33540.

P-Reviewer: Yildiz-Unal A; C-Editor: Zhao M; S-Editors: Yu J, Li CH; L-Editors: Deusen AV, Stow A, Qiu Y, Song LP; T-Editor: Liu XL