

# Graph-based network analysis of resting-state fMRI: test-retest reliability of binarized and weighted networks

Jie Xiang<sup>1</sup> · Jiayue Xue<sup>1</sup> · Hao Guo<sup>1</sup> · Dandan Li<sup>1</sup> · Xiaohong Cui<sup>1</sup> · Yan Niu<sup>1</sup> · Ting Yan<sup>2</sup> · Rui Cao<sup>1</sup> · Yao Ma<sup>1</sup> · Yanli Yang<sup>1</sup> · Bin Wang<sup>1</sup> 

© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

In the past decade, resting-state functional magnetic resonance imaging (rs-fMRI) and graph-based measures have been widely used to quantitatively characterize the architectures of brain functional networks in healthy individuals and in patients with abnormalities related to psychopathic and neurological disorders. To accurately evaluate the topological organization of brain functional networks, the definition of the nodes and edges for the construction of functional networks is critical. Furthermore, both types of brain functional networks (binarized networks and weighted networks) are widely used to analyze topological organization. However, how to best select the network type is still debated. Consequently, we investigated the test-retest reliability of brain functional networks with binarized and weighted edges using two independent datasets and four strategies for defining nodes. We revealed fair to good reliability for a majority of network topological attributes and overall higher reliabilities for weighted networks than for binarized networks. For regional nodal efficiency, weighted networks also showed higher reliability across nodes. Thus, our findings imply that weighted networks contain more information, leading to more stable results. In addition, we found that the reliability of brain functional networks was influenced by the node definition strategy and that more precise of nodal definition were associated with higher reliability. The time effect of reliability was restricted, as no differences between long-term and short-term reliability were observed. In conclusion, our results suggest that weighted networks have better reliability because they reflect more topological information, implying broader applications of weighted networks related to normal and disordered function of the human brain.

**Keywords** Resting-state fMRI · Graph-based measures · Binarized and weighted edges · Test-retest reliability

## Introduction

The human brain is considered an extremely complex system characterized as a functional connectome that links related functional areas via a structural connectome consisting of

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11682-019-00042-6>) contains supplementary material, which is available to authorized users.

✉ Dandan Li  
lidandan@tyut.edu.cn

✉ Bin Wang  
wangbin01@tyut.edu.cn

<sup>1</sup> College of Information and Computer, Taiyuan University of Technology, Taiyuan, China

<sup>2</sup> Translational Medicine Research Center, Shanxi Medical University, Taiyuan, China

neural elements and neural connections (Biswal et al. 2009; Finn et al. 2015; Sporns et al. 2005). Resting-state functional magnetic resonance imaging (rs-fMRI) and graph-based measures (Heuvel and Sporns 2013; Smith et al. 2013; Zuo and Xing 2014) have emerged as powerful tools for studying the human brain architecture. Many studies have suggested that the functional architectures of human brain networks can be modulated by psychopathic and neurological disorders (He et al. 2009; Wang et al. 2010c) and are affected by age (Achard and Bullmore 2007; Meunier et al. 2009; Wang et al. 2010b), sex (Tian et al. 2011), intelligence (Van Den Heuvel et al. 2009), and genetic variations (Gilman et al. 2011). Exploring the underlying mechanism and architecture of the connectome can provide valuable insights into the configuration and efficiency of brain functional networks, which is important for our understanding of human brain architectures and function from a systematic perspective (Horn et al. 2014; Zhou et al. 2012; Zuo et al. 2012).

Based on graph theory, the human brain can be modeled as a complex network represented graphically by a collection of nodes and edges (Wang et al. 2010a). According to studies of brain networks, the strategy used to define edges has a great impact on the topological organization of the resulting brain network (Rubinov and Sporns 2010; Sporns 2013). In general, networks can be defined as binarized edges if the edges have equal weights, or edges can be assigned weighted edges. Small-worldness (Watts and Strogatz 1998) is an important parameter describing the organizational principles governing complex networks of biology and is key in exploring the topological organization of the brain. For brain networks, the property of small-worldness is a basic metric of binarized networks (Van Wijk et al. 2010) and weighted networks (Bassett et al. 2011b, 2012; Charles 2014). Assortativity, modularity and synchronization are also reflected in brain networks. Therefore, both binarized networks and weighted networks are desirable for exploring brain topologies. Furthermore, based on these metrics, binarized and weighted network analyses have been widely used to reveal the mechanisms of normal brain topologies (Rubinov and Sporns 2011; Zuo et al. 2012) and to explore the changes in brain topologies caused by diseases (Tijms et al. 2013; Zhao et al. 2012) and during tasks (Kim et al. 2015; Weber et al. 2013).

Strategies for defining edges (binarized and weighted) have recently received increased attention in this field. In earlier studies, binarized network analysis had historically been the first choice in neural imaging due to the limited signal-to-noise ratio (SNR) of brain data (Achard et al. 2006). However, while binarized network analysis can effectively measure some high density connections, it is not the best method for understanding complex topologies (Bassett et al. 2011a; Klimm et al. 2014). Moreover, Farine (2014) found that weighted edges significantly reduce the impact of random associations on a perfectly assortative network compared with binarized networks. Furthermore, Bassett and Bullmore (2016) proposed differences in the small-worldness between binarized and weighted networks and explored the influence of each network type on the topological metrics of brain networks. These previous studies showed that weighted networks may be more appropriate than binarized networks for analyzing the topological architectures of brain networks. To compare the reliability of binarized networks and weighted networks in describing real brain networks, some progress made by test-retest reliability studies based on binarized and weighted methods to construct brain functional networks has been reported. Guo et al. (2012) found that weighted networks outperformed binarized networks in test-retest reliability of clustering coefficients at the node level. Inconsistently, another study found that overall fair to good reliabilities were detected for global and local connectivity metrics based on binarized network analysis (Cao et al. 2014). These

inconsistent findings suggest that the reliabilities of binarized and weighted methods in depicting brain function networks remain unclear.

Here, using two independent test-retest datasets from Institute of Psychology, Chinese Academy of Sciences (IPCAS\_1) (Zhao et al. 2013) and New York University, Child Study Center (NYU-CSC) (Wang et al. 2011), we examined the reliability of functional network topological architectures measured using binarized and weighted edges. To demonstrate the difference in reliability, we also adopted four strategies for defining nodes and evaluated long-term and short-term reliability. We compared the reliabilities of network topological attributes at the global and nodal levels between binarized and weighted networks. Our findings provide evidence of how to choose the edge type in brain graphs and encourage the use of multiple atlases to enhance human functional connectome studies.

## Materials and methods

### Subjects

Data for 30 right-handed healthy college students (mean age:  $20.93 \pm 1.72$ ; 9 males) are publicly available from IPCAS\_1 ([https://doi.org/10.15387/fcp\\_indi.corr.ipcas1](https://doi.org/10.15387/fcp_indi.corr.ipcas1)). Because one of the subjects lacked information for three rs-fMRI scans, we selected 29 subjects for our test-retest study (8 males).

### Data acquisition

During the rest scan, a fixation cross was presented to the first group of 29 subjects, and the subjects were instructed to rest while focusing on the fixation cross. Three resting-state scans were obtained for each subject using a Siemens 3 T scanner. Researchers acquired the echo-planar imaging (EPI) functional volumes of each scan (time repetition (TR) = 2500 ms; time echo (TE) = 30 ms; flip angle (FA) =  $90^\circ$ , number of slices = 32, matrix =  $64 \times 64$ ; field of view (FOV) = 256 mm, acquisition voxel size =  $3 \times 3 \times 3 \text{ mm}^3$ ) and structural MRI data using sagittal T1-weighted magnetization prepared rapid gradient echo (MPRAGE) sequences (TR = 2530 ms; TE = 2.51 ms; inversion time = 1100 ms; FA =  $7^\circ$ ; number of slices = 128; FOV = 256 mm). The mean interval between scan 1 and 2 was 29 min, and the mean interval to scanning 3 was 5–24 days.

### Data preprocessing

Image preprocessing followed standard routines implemented in the Statistical Parametric Mapping software (SPM8, <http://www.fil.ion.ucl.ac.uk/spm/software/spm8/>). In brief, all fMRI

images were removed from the first 10 time points and slice-timing corrected, realigned for head motion, and spatially normalized to the Montreal Neurological Institute (MNI) template (voxel size  $3 \times 3 \times 3 \text{ mm}^3$ ). Finally, the normalized images were further smoothed with a 6 mm full-width at half-maximum (MFWH) Gaussian kernel and temporally bandpass filtered (0.01–0.1 Hz) to reduce the effects of low-frequency drift and high-frequency physiological noise.

## Node definition

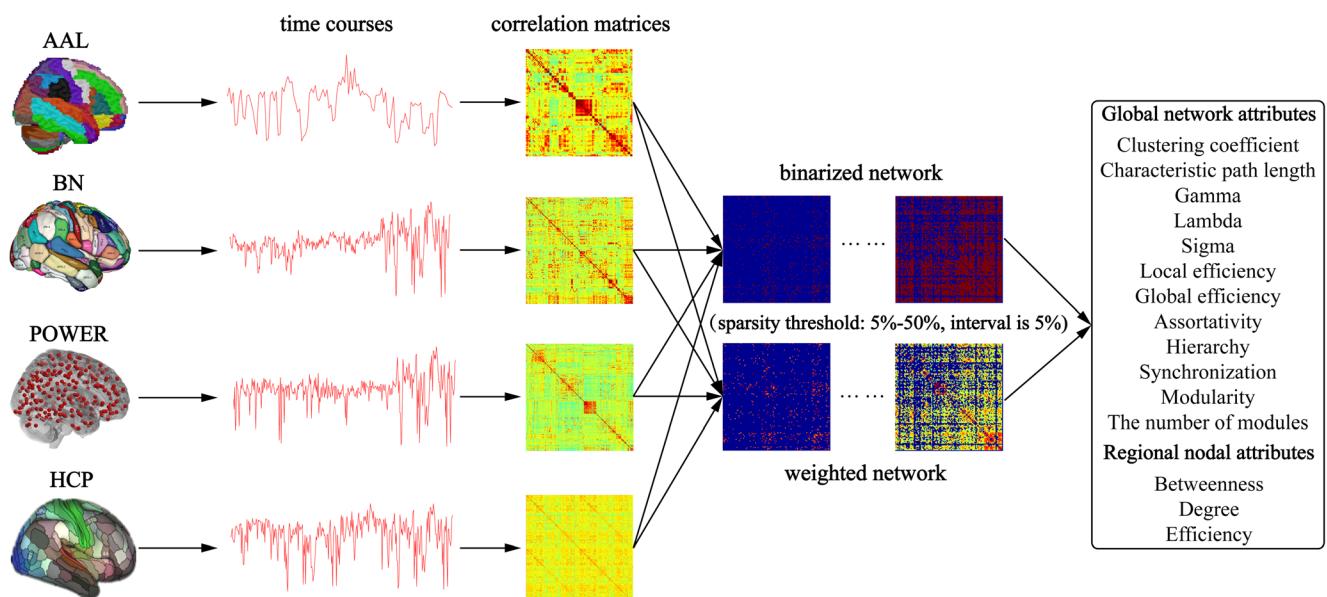
A network comprises nodes and edges connecting the nodes. To examine the potential influence of different node definition strategies on the reliability of network estimates, we adopted four widely used parcellation atlases (Fig. 1). One of the structural atlases included 90 anatomical brain regions defined by the structural Automated Anatomical Labeling atlas (AAL) (Tzouriomazoyer et al. 2002). Another structural atlas, the Brainnetome atlas (BN atlas), provided a fine-grained, cross-validated atlas and contained information on both anatomical and functional connections, with 210 cortical and 36 subcortical subregions (Fan et al. 2016). Additionally, a functional parcellation (POWER atlas) proposed by a prior study combines meta-analytic data from active fMRI experiments and functional connectivity mapping (Power et al. 2011). The functional template consists of a set of 264 hypothetical functional nodes that are modeled as 10-mm diameter spheres that span the cerebral cortex and subcortical structures. Another cortical areal

parcellation (HCP atlas) contains 180 distinct areas per hemisphere, symmetrically arranged across the two hemispheres (Glasser et al. 2016).

## Edge definition and network thresholding

After defining the nodes, the mean time series for each node in both atlases were extracted from the preprocessed images, and Pearson correlation was applied to the mean time series. Overall, brain networks were constructed with a defined sparsity. The sparsity threshold ensures that all resultant networks have comparable topological structures with the same number of edges (Wang et al. 2011). Hence, we selected the equal-interval sparsity threshold range (ranging from 0.05 to 0.5 with a partition interval of 0.05), and individual brain networks were constructed at the same sparsity level across subjects and scans.

Then, positive edges in networks were subjected to brain network analysis, and different types of networks were divided into binarized networks and weighted networks according to the definition of edges. In the surviving connections under a given threshold, the edges of the binarized networks were assigned a value of 1, and the edges of the weighted networks were assigned the original correlation coefficient. Additionally, a value of 0 was assigned to the nonexistent connections both in binarized and weighted networks. In total, eight adjacency matrices were generated for each scan. All network topological attributes based on graph theoretical methods were subsequently calculated from these derived matrices (Fig. 1).



**Fig. 1 Diagram of the construction and analysis of networks in this study.** The time series were extracted from preprocessed fMRI data using four different brain atlases, and binarized and weighted brain networks were later constructed from the extracted time series based on different

sparsity thresholds. Long-term and short-term test-retest reliability values were calculated for all network topological attributes. See the methods for details

## Network metrics

We explored two sets of network topological attributes: 1) global network metrics: small-world parameters (clustering coefficient  $C_p$ , characteristic path length  $L_p$ , normalized clustering coefficient  $\gamma$ , normalized characteristic path length  $\lambda$  and small-worldness  $\sigma$ ), network efficiency (local efficiency  $E_{loc}$  and global efficiency  $E_g$ ), assortativity  $r$ , hierarchy  $\beta$ , synchronization  $S$ , modularity  $Q$  and the number of modules  $N_M$ ; 2) regional nodal metrics: betweenness  $b_i$ , degree  $k_i$ , and efficiency  $e_i$ . All the computations of network topological attributes were performed using MATLAB codes termed GRETNAs (<http://www.nitrc.org/projects/gretna/>). Table 1 provides detailed descriptions of the above attributes.

## Test-retest reliability

Consistent with prior works (Braun et al. 2012; Cao et al. 2014; Plichta et al. 2012), we used a common index, namely, the intraclass correlation (ICC) (Shrout and Fleiss 2015), to investigate the test-retest reliability of all the network topological attributes mentioned above. The ICC values were calculated according to the following equation:

$$ICC = \frac{MS_\lambda - MS_\varepsilon}{MS_\lambda + (k-1)MS_\varepsilon + \frac{k}{n}(MS_\pi - MS_\varepsilon)} \quad (1)$$

$MS_\lambda$  is the between-subject mean square,  $MS_\varepsilon$  is the residual mean square,  $MS_\pi$  is the between-scan mean square,  $n$  is the number of subjects, and  $k$  is the number of repeated

observations per subject. Here, ICC values close to 1 indicate perfect agreement between two scans, while ICC values close to 0 (or negative) indicate poor or no agreement. In this study, we used the standard record of reliability (Sampat et al. 2006; Winer 1962), with an ICC value from 0 to 0.25 indicating poor reliability; 0.25 to 0.4 indicating low reliability; 0.4 to 0.6 indicating fair reliability; 0.6 to 0.75 indicating good reliability and 0.75 to 1.0 indicating excellent reliability.

## Statistical analysis

Because the network construction was performed over the sparsity threshold, the ICC was a function of the threshold. In addition, to provide a threshold-independent reliability assessment, we also calculated the area under the curve (AUC, i.e., the integral) for small-world parameters and regional nodal metrics that were used to compute a single ICC scalar for each network measure. As the ICC is already a statistical indicator, our further statistical analysis could only be performed based on the AUC. Other statistical analyses were performed using the statistic toolbox SPSS 19. To further explore significant differences, we calculated 1) the paired sample t-test in reliability among binarized network analysis and weighted network analysis; 2) the independent sample t-test among long-term retest analysis and short-term retest analysis; 3) the differences in reliability among the AAL, BN, POWER and HCP networks (one-way repeated-measure ANOVA).

**Table 1** Brief descriptions of the complex network metrics examined in this study

Attribute	Character	Description
Global network attributes		
Clustering coefficient	$C_p$	The extent of local interconnectivity or cliquishness of a network
Characteristic path length	$L_p$	The extent of overall communication efficiency of a network
Gamma	$\gamma$	The deviation of $C_p$ of a network from those of surrogate random networks
Lambda	$\lambda$	The deviation of $L_p$ of a network from those of surrogate random networks
Sigma	$\sigma$	The small-worldness indicating the extent of a network between randomness and order
Local efficiency	$E_{loc}$	The ability of a network to transmit information at the global level
Global efficiency	$E_g$	The ability of a network to transmit information at the local level
Assortativity	$r$	The tendency of nodes to link those nodes with the same or similar number of edges
Hierarchy	$\beta$	How likely it is that all nodes oscillate with the same wave pattern
Synchronization	$S$	How likely that all nodes fluctuate in the same wave pattern
Modularity	$Q$	The degree to which a network is organized into a modular or community structure
The number of modules	$N_M$	Nodes can be divided into several densely connected connections, with connections being sparse
Regional nodal attributes		
Betweenness	$b_i$	The influence that one node has over the flow of information between all other nodes in the network
Degree	$k_i$	The number of edges linked to a node
Efficiency	$e_i$	The ability of a node to propagate information with the other nodes in a network

## Results

### Comparison of reliabilities between binarized and weighted networks

Overall, among the networks with four different node definitions (AAL, BN, POWER, HCP) and two types of scanning interval (long-term and short-term), we observed that the test-retest reliability of global network metrics was higher based on weighted network analysis than based on binarized network analysis. Among all the attributes,  $C_p$ ,  $L_p$ ,  $\gamma$ ,  $\lambda$ ,  $\sigma$ ,  $E_{loc}$  and  $E_g$  had good to excellent reliability (mean ICC > 0.6) in weighted networks but fair to good reliability in binarized networks (mean ICC < 0.6). For example, based on the BN atlas,  $C_p$  showed good reliability (mean ICC = 0.7) in weighted networks (Fig. 2c) but fair reliability (mean ICC = 0.42) in binarized networks (Fig. 2a). In addition,  $\beta$  and  $S$  showed fair to good reliability (mean ICC < 0.6), but  $Q$ ,  $N_M$  and  $r$  showed poor to fair reliability (mean ICC < 0.4) in binarized networks, while these attributes were also more reliable in weighted networks (Fig. 2).

The reliabilities of the network were slightly different across sparsities, making it difficult to compare attributes among the network constructed method. Thus, we further analyzed the reliabilities of AUC across the sparsity and performed paired sample t-tests based on the AUC of each node to test the differences in the reliability of network attributes. The reliability of attributes including  $C_p$ ,  $L_p$ ,  $E_{loc}$  and  $E_g$  exhibited significantly higher ICCs in weighted networks than in binarized networks (all:  $t > 3.425$ ,  $p < 0.001$ , Fig. 3), except for the reliability of  $L_p$ , based on AAL networks. Notably, both long- and short-term retest assessments had consistent differences in reliability between the weighted networks and binarized networks.

### Nodal reliability in binarized and weighted networks

For regional nodal metrics, efficiency showed relatively high reliability and differences between binarized networks and weighted networks (shown in Figs. 4 and 5). For instance, based on the AAL atlas, the BN atlas, the POWER atlas and the HCP atlas, more reliable nodes were found in weighted networks than in binarized networks (Fig. 4a, b, 5a, b). Furthermore, the short-term reliability of  $e_i$  based on weighted networks (mean ICC = 0.48) was higher than that based on binarized networks (mean ICC = 0.32) using the POWER atlas (Fig. 5a, b). In terms of node distribution, nodes with good reliability (ICC > 0.6) were mainly distributed in the visual network (e.g., calcarine and lingual), the somatosensory and motor network (e.g., Heschl's gyrus) and the default-mode network (e.g., posterior cingulum and middle temporal gyrus) both in binarized networks and weighted networks. In addition, other brain areas that belong to the somatosensory and motor network (e.g., rolandic operculum, supplementary

motor area and insula) were found to be more reliable in weighted networks than in binarized networks.

Overall, we found that weighted networks had considerably more reliable nodes, irrespective of the factors of atlas and scanning interval. In contrast, binarized networks in general showed fewer reliable nodes. To more clearly reflect the effects of binarized and weighted measures on the network topology throughout the brain, we calculated the percentages of the ranges of nodal reliabilities using the four brain atlases (shown in Figs. 4c and 5c). More than half of the nodes had fair to excellent test-retest reliability in weighted networks, whereas less than half of the nodes were reliable in binarized networks (Fig. 5c).

### Comparison of reliabilities among three node definition strategies

As shown in Figs. 2 and 3, in the weighted network, the global network metrics were influenced by different strategies of node definition. For instance,  $\gamma$ ,  $\lambda$  and  $\sigma$  showed fair reliability based on the AAL atlas but good reliability based on the BN, POWER and HCP atlases (Fig. 2). In addition, the reliabilities based on the BN, POWER and HCP atlases were significantly higher than those based on the AAL atlas for the  $C_p$  ( $F = 4.667$ ,  $p < 0.01$ ),  $L_p$  ( $F = 4.094$ ,  $p < 0.01$ ) and  $E_{loc}$  ( $F = 9.434$ ,  $p < 0.001$ ) of long-term networks and for the  $L_p$  ( $F = 8.908$ ,  $p < 0.001$ ),  $E_{loc}$  ( $F = 2.589$ ,  $p < 0.05$ ) and  $E_g$  ( $F = 9.385$ ,  $p < 0.001$ ) of short-term networks (Fig. 3). For regional nodal metrics, the fair to excellent reliability of  $e_i$  using weighted networks based on the BN atlas, the POWER atlas and the HCP atlas was also higher than that based on the AAL atlas (Fig. 4c and 5c). Notably, the differences in strategies of node definition appeared to be absent in binarized networks.

### Independent data set

A publicly available retest data was included to replicate and validate the significant findings. Consistent with the principal data set, the reliabilities of topological attributes of weighted networks were higher than those of binarized networks. In addition, the rest-reliability based on the BN, POWER and HCP atlases was higher than that based on the AAL atlas (Fig. S1). For nodal reliability, weighted networks had considerably more reliable nodes than binarized networks (Fig. S2, S3, S4, S5). The detailed data and experimental results of the independent data set are included in the supplementary material.

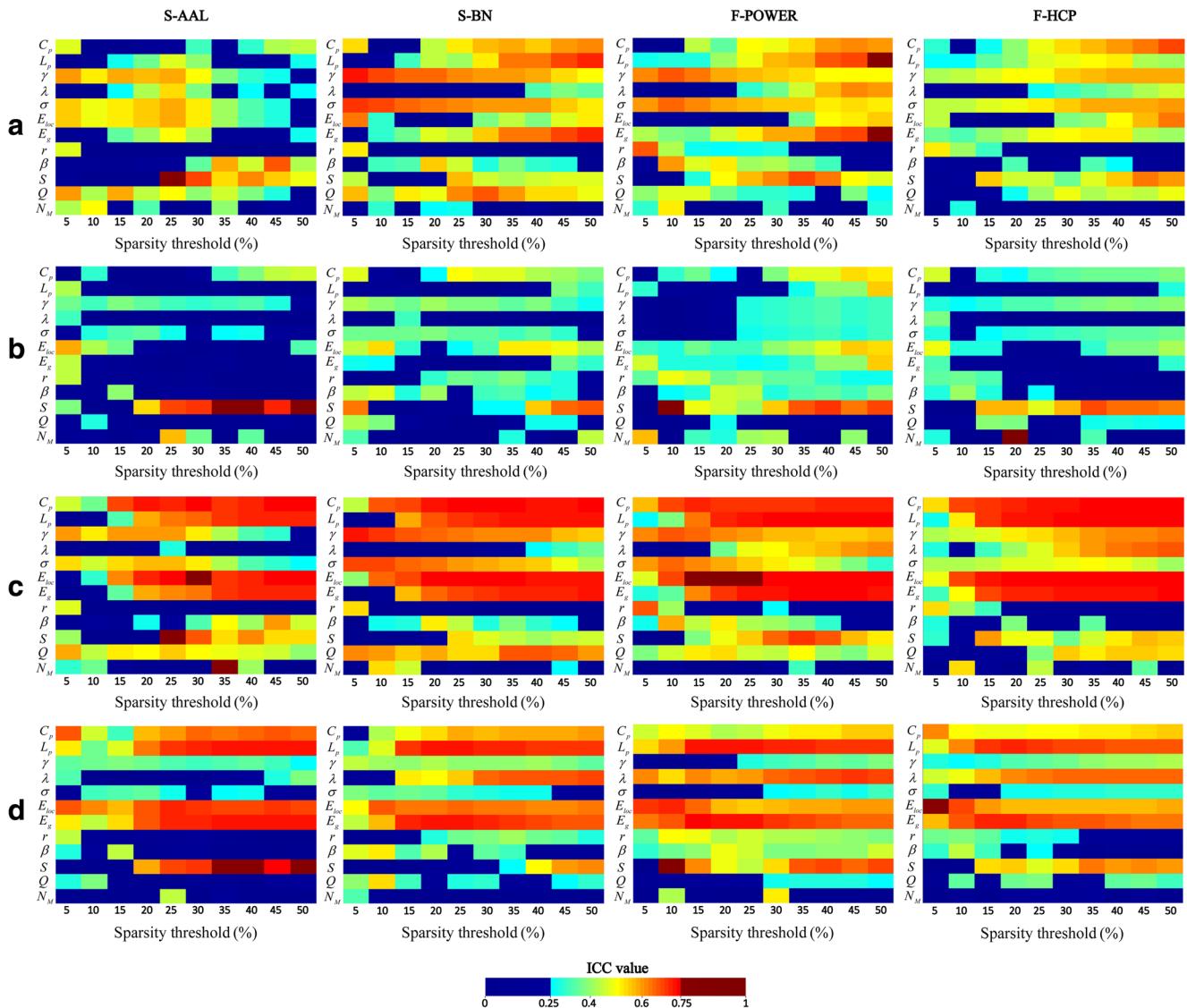
## Discussion

In this study, both the principal and independent data supported that weighted networks have better test-retest

reliability in global network metrics than binarized networks. The nodal efficiency also showed higher reliability in weighted networks than in binarized networks. In addition, we found that the reliability of brain functional networks was influenced by the node definition strategy; more precise parcellation schemes resulted in more realistic network connections and were associated with higher reliability. The time effect on the reliability was restricted, as no differences between long-term and short-term reliability were observed. Our results suggest that weighted networks have better reliability due to their ability to reflect more topological information.

## Comparison of reliabilities between binarized and weighted networks

Overall, we found that global network metrics generally showed fair test-retest reliability, including  $C_p$ ,  $L_p$ ,  $E_{loc}$  and  $E_g$ . The small-worldness (Watts and Strogatz 1998) is an important parameter for describing the organizational principles governing complex networks of biology. Specifically, small-world networks can be described by high local clustering and minimum path length, which are reflected by the clustering coefficient and the characteristic path length, respectively (Wang et al. 2010a). Moreover, global efficiency and local



**Fig. 2 Reliabilities of global network metrics as a function of sparsity threshold.** ICC values from 0 to 1 are mapped in dark blue to dark red color. **a, b** represent long-term and short-term reliabilities, respectively, using four brain atlases based on binarized network analysis; **c, d** represent long-term and short-term reliability, respectively, using four brain

atlases based on weighted network analysis. Long-term: mean 5–24 days; short-term: mean 29 min. Four brain atlases were described in detail. The 12 global network metrics of the ordinate are described in detail in Table 1

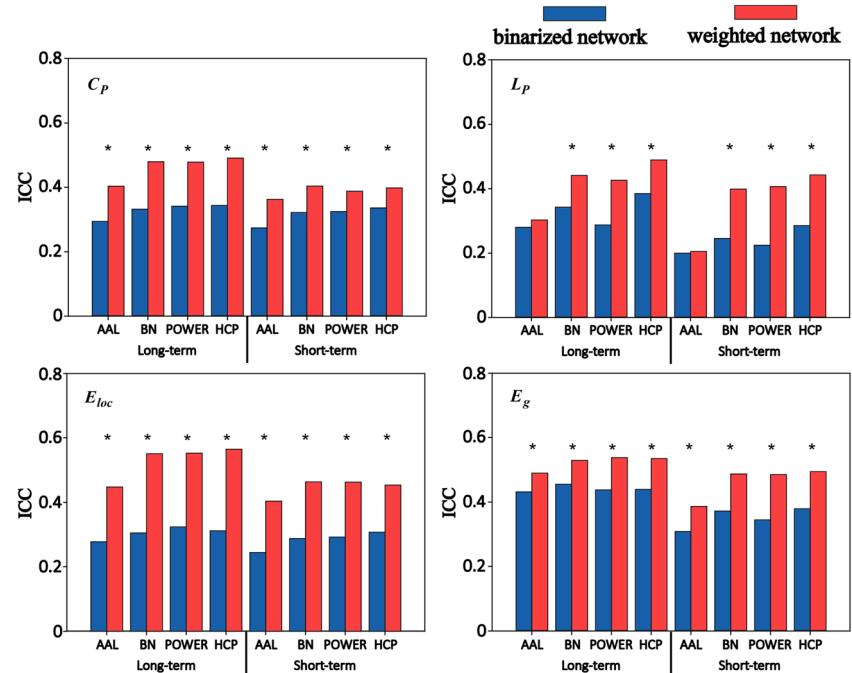
efficiency measure the ability of a network to transmit information at the global and local level (Latora and Marchiori 2001), respectively, and can reflect the topological connection efficiency of small-world networks. Our findings suggested that networks with small-world attributes achieved good to excellent levels of internal consistency in describing the organization of brain functional networks.

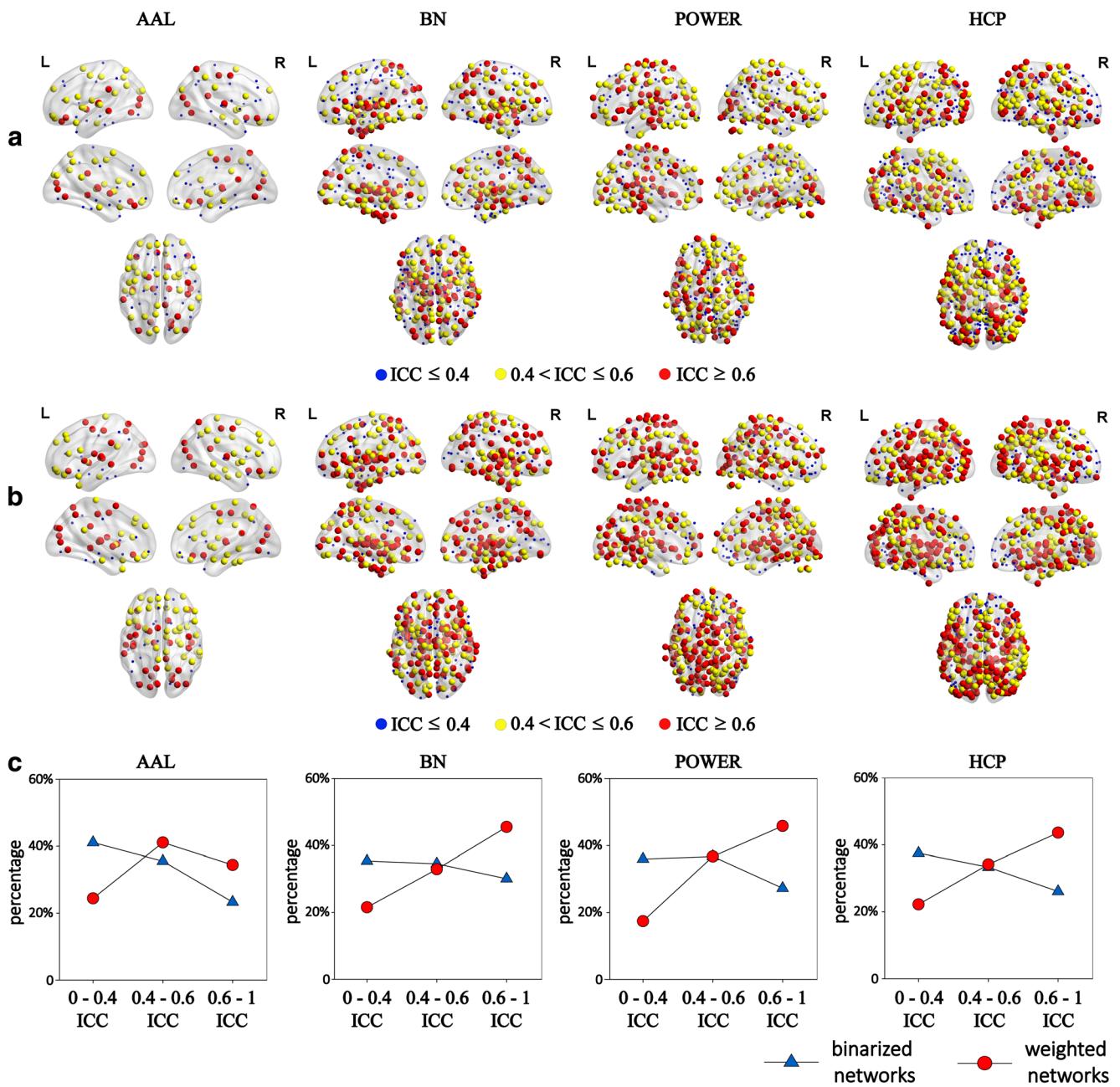
Moreover, for  $C_p$ ,  $L_p$ ,  $E_{loc}$  and  $E_g$ , weighted networks had higher reliabilities than binarized networks. Consistently, Guo et al. (2012) also observed that weighted networks outperformed binarized networks in the test-retest reliability of clustering coefficients. Another study found overall fair to good reliabilities for global and local connectivity metrics based on binarized network analysis (Cao et al. 2014). It is well known that the fundamental difference between binarized networks and weighted networks is the amount of information contained in the edges. Because connection strength is taken into account, weighted networks are thought to more accurately characterize the network topology and have higher reliability in network topology (Cole et al. 2010; Guo et al. 2012). In the human brain, networks with complex types also exhibit a large heterogeneity in terms of capacity and connection strength (Latora and Marchiori 2001; Sporns 2002), suggesting that weighted networks with more information may be more suitable for describing the topology of brain networks. Conversely, the construction of binarized networks represents an extreme simplification of brain networks, and binarized networks of the nodes of the same structure are as simple as networks obtained in graph theory (Bassett et al. 2012). Furthermore, while binarized networks were very common in early applications of graph theory to neural data (Wijk

et al. 2010), they are inaccurate for architectures that may be defined by weighted edges (Rubinov and Sporns 2011). Constructing binarized networks by simply determining the connections may cause a large amount of information to be lost and result in binarized networks that are less stable than the topology of weighted networks. Thus, weighted networks may be more suitable in functional network studies as they can embrace more novel findings that may extend our knowledge of brain function networks.

In addition to small-world parameters, weighted networks also had higher reliabilities for hierarchy, synchronization, modularity, and assortativity than binarized networks. Hierarchy favors top-down relationships among nodes and reveals the close connections between other nodes and hubs (Ravasz and Barabási 2003). Furthermore, synchronization represents the degree of coupling between nodes (Motter et al. 2005). All these parameters indicate that weighted networks contain more information that can better depict the related activities among nodes, leading to higher reliabilities of hierarchy and synchronization. Moreover, modularity refers to a set of nodes with denser links among themselves but sparser links with the rest of the network (Newman 2006). Likewise, assortativity indicates that the nodes in the network that have many connections tend to be connected to other nodes with many connections (Newman 2002). These attributes all indicate a clustering phenomenon among nodes and may result in lower reproducibility of constructing brain networks due to differences between scans. The weak reliability of these parameters was more obvious in binarized networks, which might be due to the loss of information in binarized networks.

**Fig. 3 Paired sample t-test between binarized and weighted network analysis.** Apart from the reliability of the characteristic path length based on AAL networks, other attributes exhibited significant differences under all conditions (\* $p < 0.001$ ). Long-term: mean 5–24 days; short-term: mean 29 min. Four brain atlases were described in detail. The clustering coefficient, characteristic path length, local efficiency and global efficiency are described in detail in Table 1





**Fig. 4** Long-term reliabilities of efficiency over averaged densities. Nodes with fair to excellent reliabilities are mapped in yellow ( $ICC = 0.40\text{--}0.6$ ) or red ( $ICC \geq 0.6$ ); those with low reliabilities are mapped in blue. **a**, **b** represent the reliabilities using four brain atlases based on binarized

network analysis and weighted network analysis, respectively; **c** represents the percentages of the ranges of nodal reliabilities using four brain atlases based on binarized network analysis and weighted network analysis. Long-term: mean 5–24 days

## Reliable nodes in binarized and weighted networks

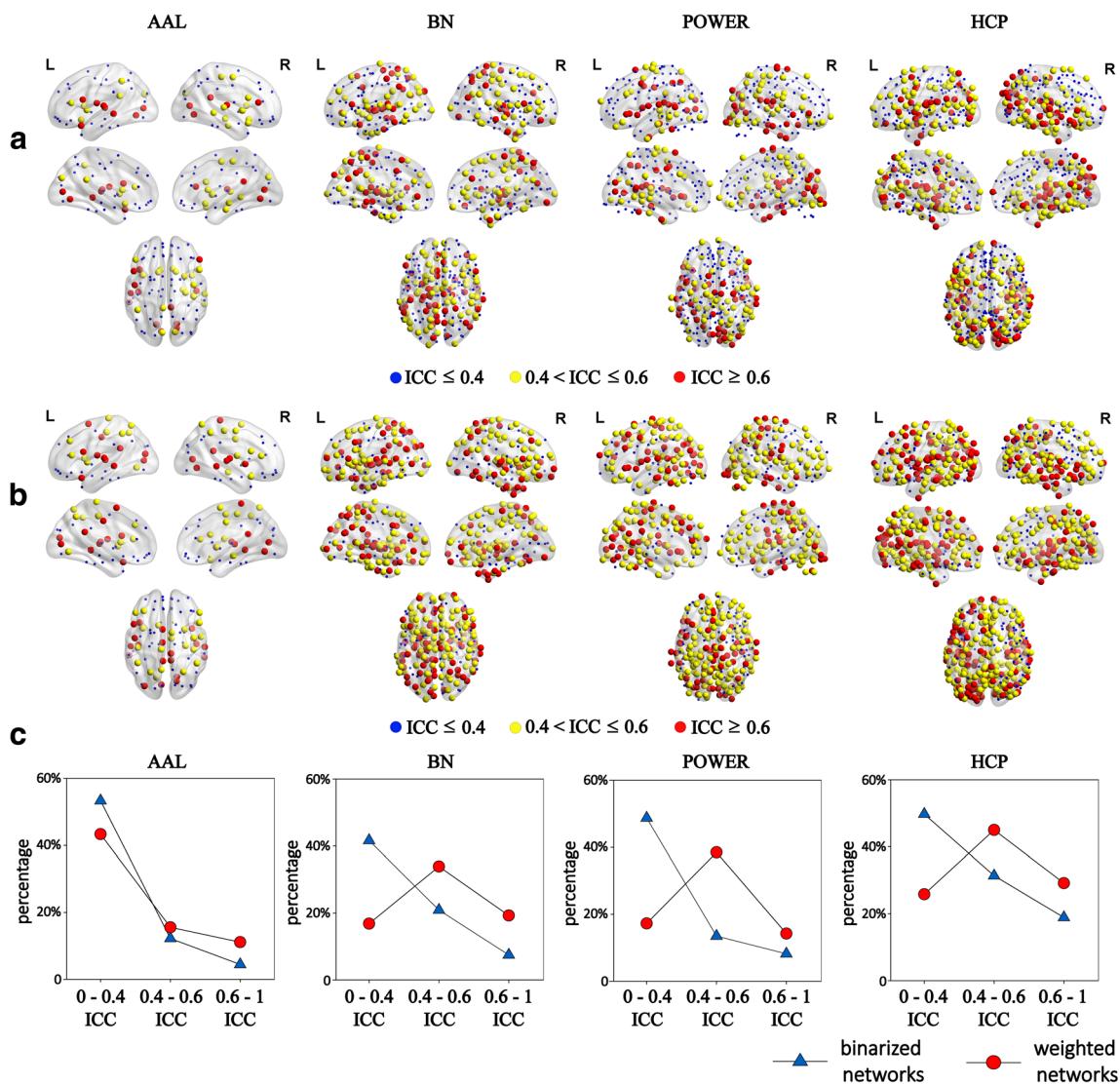
Across both binarized and weighted networks, the nodes with excellent reliability were predominantly distributed in the default-mode network. Interestingly, the default-mode network is the main functional system on which cognitive function of the human brain depends (Cocito et al. 2012; Doria et al. 2010; Koch et al. 2010). A default-mode network is engaged when individuals are left to think to themselves undisturbed (Mazoyer et al. 2001; Raichle et al. 2001; Shulman

et al. 1997) and is preferentially active when individuals are not focused on the external environment (Buckner et al. 2008). Furthermore, maintaining a resting state is considered to be a specific task and can activate the default-mode network (Raichle et al. 2001). Hence, activation of the default-mode network by a resting-state environment may not correspond to a large amount of information transmission, and the connection information that already existed in the networks was relatively simple. Namely, this could result in the presence of a default-mode network with high test-retest reliability.

The nodes in the visual, somatosensory and motor subnetworks also showed higher test-retest reliability. Of note, the most important factor in the visual network is sensing the physical metrics of the spatial layout (Wandell et al. 2007), while the somatosensory and motor networks are primarily active when external stimuli are received. Nevertheless, in the resting state, the subjects were not disturbed by the outside world, and therefore, few stimuli were present during the resting state in regard to the construction of these networks. Hence, these networks may not be active and correspond to less transmission of information. In other words, due to less activity, these networks were relatively stable and robust in the resting state regardless of the method of topological construction used.

## Effects of strategies of node definition in binarized and weighted networks on reliability

We observed comparable test-retest reliabilities for the examined strategies of node definition for the attributes and significantly higher reliabilities for the BN, POWER and HCP networks relative to the AAL networks. There are two possible explanations for the superior performance of the strategies of node definition. First, structural ROIs of the AAL atlas mainly depend on the anatomical features of the sulcal pattern by comparing roughly divided methods (Tzouriomazoyer et al. 2002), whereas structural ROIs of the BN atlas were derived from diffusion tensor imaging (DTI) connections, such that they contain anatomical and functional connection



**Fig. 5 Short-term reliabilities of efficiency over averaged densities.** Nodes with fair to excellent reliabilities are mapped in yellow ( $ICC\ 0.40-0.6$ ) or red ( $ICC \geq 0.6$ ), those with low reliabilities are mapped in blue. **a, b** represent the reliabilities using four brain atlases based on binarized

network analysis and weighted network analysis, respectively; **c** represents the percentages of the ranges of nodal reliabilities using four brain atlases based on binarized network analysis and weighted network analysis. Short-term: mean 29 min

information (Fan et al. 2016). Additionally, functional ROIs of the POWER and HCP atlas were derived from a prior study combining meta-analytic data from active rs-fMRI experiments, which hold specific functional information (Glasser et al. 2016; Power et al. 2011). Thus, it is possible that functional connection information could affect the test-retest reliability evaluation of brain functional networks. Second, the number of nodes in the BN atlas, the POWER atlas and the HCP atlas was approximately three to four times that of the AAL atlas, and the size of the node set has a considerable influence on the theoretical analysis of graphs (Butts 2009; Zalesky et al. 2010). In addition, previous studies demonstrated that network analyses are sensitive to the accurate identification and separation of biological functional units in the brain (Dosenbach et al. 2010; Power et al. 2011; Smith et al. 2011; Wig et al. 2011), and network characteristics are sensitive to the strategy used to define nodes (Hayasaka and Laurienti 2010; Sanabria Diaz et al. 2010).

Of note, a higher number of nodes in the brain network corresponded to a greater probability of short-distance connections between nodes. Furthermore, short-distance connections had not only strong weighting but also were topologically aggregated (Bassett and Bullmore 2016). Consequently, a relatively large number of brain nodes may result in weighted networks that are more accurate than binarized networks. Thus, weighted networks with a large number of nodes are thought to more closely map real brain functional networks and show higher reliability. In addition, the BN, POWER and HCP atlases may be more similar to real brain networks in weighted networks that contain more information. Therefore, different strategies used to define nodes may have an impact on the structure and information transfer in brain function networks.

## Limitations

There are some limitations to our research. First, we considered various commonly used network topological attributes to evaluate the reliability of brain functional networks, although other measures may differently reflect the topological stability of binarized and weighted networks, such as rich club (Aso et al. 2011), strength and transitivity (Cao et al. 2014). Second, while we adopted different methods to build the networks, the reliability assessment in this study still depended on some common processing methods, such as the preprocessing steps (Braun et al. 2012) and the selected thresholds (Achard and Bullmore 2007; Achard et al. 2006; Gong et al. 2009; He et al. 2008). Third, while we provided data on the reliability of graph attributes derived from three established strategies of node definition, the conclusions and recommendations may not necessarily be generalizable to other parcellation strategies (Cao et al. 2014; Wang et al. 2011).

## Conclusion

In summary, we studied the test-retest reliability of graph-based network attributes derived from rs-fMRI data and the influence of several factors on the stability of network topology. Specifically, this study revealed overall fair to good reliability for a majority of graph metrics based on functional networks in both binarized and weighted networks. Importantly, higher reliabilities were observed in weighted networks than in binarized networks. For regional nodal efficiency, the weighted networks also showed higher reliability across nodes. In addition, the detected reliabilities were influenced by network type and the strategy used to define nodes, with higher reliabilities detected for graphs generated from the BN, POWER and HCP atlases. Our results demonstrated that weighted networks are more reliable and stable considering the reproducibility of brain networks, providing methodological advice on exploring the brain functional networks by rs-fMRI.

**Acknowledgements** This project is supported by the National Natural Science Foundation of China (61503272, 61873178 and 61876124), the Natural Science Foundation of Shanxi (201801D121135), the International Science and Technology Cooperation Project of Shanxi (201803D421047), and the Youth Science and Technology Research Fund (201701D221119). Also, we would like to thank PhD Lynne Hyman for the professional language editing services.

**Funding** This project is supported by the National Natural Science Foundation of China (61503272, 61873178 and 61876124), the Natural Science Foundation of Shanxi (201801D121135), the International Science and Technology Cooperation Project of Shanxi (201803D421047), and the Youth Science and Technology Research Fund (201701D221119).

## Compliance with ethical standards

**Conflict of interest** Jie Xiang, Jiayue Xue, Hao Guo, Dandan Li, Xiaohong Cui, Yan Niu, Ting Yan, Rui Cao, Yao Ma, Yanli Yang and Bin Wang declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Ethics approval** The dataset of IPCAS\_1 and NYU CSC are used in the study. The IPCAS\_1 dataset was approved by the Institute of Psychology, Chinese Academy of Sciences, the NYU CSC dataset was approved by the New York University, Child Study Center.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Achard, S., & Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Computational Biology*, 3(2), e17.
- Achard, S., Salvador, R., Whitcher, B., Suckling, J., & Bullmore, E. (2006). A resilient, low-frequency, small-world human brain

- functional network with highly connected association cortical hubs. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 26(1), 63–72.
- Aso, T., Okamura, S., Matsuguchi, T., Sakamoto, N., Sata, T., & Niho, Y. (2011). Rich-Club Organization of the Human Connectome. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 31(44), 15775–15786.
- Bassett, D. S., & Bullmore, E. T. (2016). Small-world brain networks revisited. *Neuroscientist A Review Journal Bringing Neurobiology Neurology & Psychiatry*, 23(5).
- Bassett, D. S., Brown, J. A., Deshpande, V., Carlson, J. M., & Grafton, S. T. (2011a). Conserved and variable architecture of human white matter connectivity. *Neuroimage*, 54(2), 1262–1279.
- Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., & Grafton, S. T. (2011b). Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences of the United States of America*, 108(18), 7641–7646.
- Bassett, D. S., Nelson, B. G., Mueller, B. A., Camchong, J., & Lim, K. O. (2012). Altered resting state complexity in schizophrenia. *Neuroimage*, 59(3), 2196–2207.
- Biswal, B. B., Mennes, M., Zuo, X. N., Gohel, S., Kelly, C., Smith, S. M., Beckmann, C. F., Adelstein, J. S., Buckner, R. L., & Colcombe, S. (2009). Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 71(10), 4734–4739.
- Braun, U., Plichta, M. M., Esslinger, C., Sauer, C., Haddad, L., Grimm, O., Mier, D., Mohnke, S., Heinz, A., Erk, S., Walter, H., Seiferth, N., Kirsch, P., & Meyer-Lindenberg, A. (2012). Test-retest reliability of resting-state connectivity network characteristics using fMRI and graph theoretical measures. *NeuroImage*, 59(2), 1404–1412.
- Buckner, R. L., Andrewshanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, 1124(1), 1–38.
- Butts, C. T. (2009). Revisiting the foundations of network analysis. *Science*, 325(5939), 414–416.
- Cao, H., Plichta, M. M., Schäfer, A., Haddad, L., Grimm, O., Schneider, M., Esslinger, C., Kirsch, P., Meyer Lindenberg, A., & Tost, H. (2014). Test-retest reliability of fMRI-based graph theoretical properties during working memory, emotion processing, and resting state. *Neuroimage*, 84(1), 888–900.
- Charles, L. (2014). Resolving structure in human brain organization: Identifying mesoscale Organization in Weighted Network Representations. *PLoS Computational Biology*, 10(10), e1003712.
- Cocito, C., Vanlinden, F., & Branlant, C. (2012). Exposure to subliminal arousing stimuli induces robust activation in the amygdala, hippocampus, anterior cingulate, insular cortex and primary visual cortex: A systematic meta-analysis of fMRI studies. *Neuroimage*, 59(3), 2962–2973.
- Cole, M. W., Pathak, S., & Schneider, W. (2010). Identifying the brain's most globally connected regions. *Neuroimage*, 49(4), 3132–3148.
- Doria, V., Beckmann, C. F., Arichi, T., Merchant, N., Groppo, M., Turkheimer, F. E., Counsell, S. J., Murgasova, M., Aljabar, P., & Nunes, R. G. (2010). Emergence of resting state networks in the preterm human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 107(46), 20015–20020.
- Dosenbach, N. U. F., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., Nelson, S. M., Wig, G. S., Vogel, A. C., & Lessovschlaggar, C. N. (2010). Prediction of individual brain maturity using fMRI. *Science*, 329(5997), 1358–1361.
- Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., & Laird, A. R. (2016). The human Brainnetome atlas: A new brain atlas based on connectional architecture. *Cerebral Cortex*, 26(8), 3508–3526.
- Farine, D. R. (2014). Measuring phenotypic assortment in animal social networks: Weighted associations are more robust than binary edges. *Animal Behaviour*, 89(3), 141–153.
- Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., Papademetris, X., & Constable, R. T. (2015). Functional connectome fingerprinting: Identifying individuals based on patterns of brain connectivity. *Nature Neuroscience*, 18(11), 1664–1671.
- Gilman, S. R., Iossifov, I., Levy, D., Ronemus, M., Wigler, M., & Vitkup, D. (2011). Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron*, 70(5), 898–907.
- Glasser, M. F., Smith, S. M., Marcus, D. S., Andersson, J. L. R., Auerbach, E. J., Behrens, T. E. J., Coalson, T. S., Harms, M. P., Jenkinson, M., & Moeller, S. (2016). The human connectome Project's neuroimaging approach. *Nature Neuroscience*, 19(9), 1175–1187.
- Gong, G., He, Y., Concha, L., Lebel, C., Gross, D. W., Evans, A. C., & Beaulieu, C. (2009). Mapping anatomical connectivity patterns of human cerebral cortex using *in vivo* diffusion tensor imaging Tractography. *Cerebral Cortex*, 19(3), 524–536.
- Guo, C. C., Kurth, F., Zhou, J., Mayer, E. A., Eickhoff, S. B., Kramer, J. H., & Seeley, W. W. (2012). One-year test-retest reliability of intrinsic connectivity network fMRI in older adults. *Neuroimage*, 61(4), 1471–1483.
- Hayasaka, S., & Laurienti, P. J. (2010). Comparison of characteristics between region-and voxel-based network analyses in resting-state fMRI data. *Neuroimage*, 50(2), 499–508.
- He, Y., Chen, Z., & Evans, A. (2008). Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer's disease. *Journal of Neuroscience*, 4(4), T284–T285.
- He, Y., Dagher, A., Chen, Z., Charil, A., Zijdenbos, A., Worsley, K., & Evans, A. (2009). Impaired small-world efficiency in structural cortical networks in multiple sclerosis associated with white matter lesion load. *Brain: A Journal of Neurology*, 132(Pt 12), 3366–3379.
- Heuvel, M. P. V. D., & Sporns, O. (2013). Network hubs in the human brain. *Trends in Cognitive Sciences*, 17(12), 683–696.
- Horn, A., Ostwald, D., Reisert, M., & Blankenburg, F. (2014). The structural-functional connectome and the default mode network of the human brain. *Neuroimage*, 102142–102151.
- Kim, J., Chey, J., Kim, S. E., & Kim, H. (2015). The effect of education on regional brain metabolism and its functional connectivity in an aged population utilizing positron emission tomography. *Neuroscience Research*, 94(3), 50–61.
- Klimm, F., Bassett, D. S., Carlson, J. M., Mucha, P. J., (2014). Resolving Structural Variability in Network Models and the Brain. *PLoS Computational Biology*, 10.3(2014-3-27) 10(3), e1003491.
- Koch, W., Teipel, S., Mueller, S., Buerger, K., Bokde, A. L. W., Hampel, H., Coates, U., Reiser, M., & Meindl, T. (2010). Effects of aging on default mode network activity in resting state fMRI: Does the method of analysis matter? *Neuroimage*, 51(1), 280–287.
- Latora, V., & Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical Review Letters*, 87(19), 198701.
- Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houdé, O., Crivello, F., Joliot, M., Petit, L., & Tzourio-Mazoyer, N. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Research Bulletin*, 54(3), 287–298.
- Meunier, D., Achard, S., Morcom, A., & Bullmore, E. (2009). Age-related changes in modular organization of human brain functional networks. *Neuroimage*, 44(3), 715–723.
- Motter, A. E., Changsong, Z., & Jürgen, K. (2005). Network synchronization, diffusion, and the paradox of heterogeneity. *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, 71(2), 016116.
- Newman, M. E. (2002). Assortative mixing in networks. *Physical Review Letters*, 89(20), 208701.
- Newman, M. E. (2006). Modularity and community structure in networks. *APS March Meeting pp.*, 8577–8582.

- Plichta, M. M., Schwarz, A. J., Grimm, O., Morgen, K., Mier, D., Haddad, L., Gerdes, A. B., Sauer, C., Tost, H., & Esslinger, C. (2012). Test-retest reliability of evoked BOLD signals from a cognitive-emotive fMRI test battery. *Neuroimage*, 60(3), 1746–1758.
- Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., & Schlaggar, B. L. (2011). Functional network Organization of the Human Brain. *Neuron*, 72(4), 665–678.
- Raichle, M. E., Macleod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2), 676–682.
- Ravasz, E., & Barabási, A. L. (2003). Hierarchical organization in complex networks. *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, 67(2), 026112.
- Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage*, 52(3), 1059–1069.
- Rubinov, M., & Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. *Neuroimage*, 56(4), 2068–2079.
- Sampat, M. P., Whitman, G. J., Stephens, T. W., Broemeling, L. D., Heger, N. A., Bovik, A. C., & Markey, M. K. (2006). The reliability of measuring physical characteristics of spiculated masses on mammography. *British Journal of Radiology* 79 Spec No, 2-(special\_issue\_2), S134.
- Sanabriadiaz, G., Meliagarcía, L., Iturriamedina, Y., Alemángómez, Y., Hernándezgonzález, G., Valdésurrutia, L., Galán, L., & Valdéssosa, P. (2010). Surface area and cortical thickness descriptors reveal different attributes of the structural human brain networks. *Neuroimage*, 50(4), 1497–1510.
- Shrout, P. E., & Fleiss, J. L. (2015). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420.
- Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., & Petersen, S. E. (1997). Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *Journal of Cognitive Neuroscience*, 9(5), 648–663.
- Smith, S. M., Miller, K. L., Salimikhoshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., Ramsey, J. D., & Woolrich, M. W. (2011). Network modelling methods for fMRI. *Neuroimage*, 54(2), 875–891.
- Smith, S. M., Vidaurre, D., Beckmann, C. F., Glasser, M. F., Jenkinson, M., Miller, K. L., Nichols, T. E., Robinson, E., Salimikhoshidi, G., & Woolrich, M. W. (2013). Functional connectomics from resting-state fMRI. *Trends in Cognitive Sciences*, 17(12), 666–682.
- Sporns, O. (2002). Network analysis, complexity, and brain function. *Complexity*, 8(1), 56–60.
- Sporns, O. (2013). Making sense of brain network data. *Nature Methods*, 10(6), 491–493.
- Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Computational Biology*, 1(4), e42.
- Tian, L., Wang, J., Yan, C., & He, Y. (2011). Hemisphere- and gender-related differences in small-world brain networks: A resting-state functional MRI study. *Neuroimage*, 54(1), 191–202.
- Tijms, B. M., Wink, A. M., De, H. W., Wm, V. D. F., Stam, C. J., Scheltens, P., & Barkhof, F. (2013). Alzheimer's disease: Connecting findings from graph theoretical studies of brain networks. *Neurobiology of Aging*, 34(8), 2023–2036.
- Tzouriomazoyer, 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(1), 273–289.
- Van Den Heuvel, M., Stam, C. J., Kahn, R. S., & Hulshoff Pol, H. E. (2009). Efficiency of functional brain networks and intellectual performance. *Journal of Neuroscience*, 29(23), 7619–7624.
- Van Wijk, B. C., Stam, C. J., & Daffertshofer, A. (2010). Comparing brain networks of different size and connectivity density using graph theory. *PLoS One*, 5(10), e13701.
- Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, 56(2), 366–383.
- Wang, J., Zuo, X., & He, Y. (2010a). Graph-based network analysis of resting-state functional MRI. *Frontiers in Systems Neuroscience*, 4(16), 16.
- Wang, L., Li, Y., Metzak, P., He, Y., & Woodward, T. S. (2010b). Age-related changes in topological patterns of large-scale brain functional networks during memory encoding and recognition. *Neuroimage*, 50(3), 862–872.
- Wang, L., Zhu, C., He, Y., Zang, Y., Cao, Q., Zhang, H., Zhong, Q., & Wang, Y. (2010c). Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder. *Human Brain Mapping*, 30(2), 638–649.
- Wang, J. H., Zuo, X. N., Gohel, S., Milham, M. P., Biswal, B. B., & He, Y. (2011). Graph theoretical analysis of functional brain networks: Test-retest evaluation on short- and long-term resting-state functional MRI data. *PLoS One*, 6(7), e21976.
- Watts, D.J., Strogatz, S.H., (1998). Collective dynamics of 'small-world' networks, Collective dynamics of 'small-world' networks.
- Weber, M. J., Detre, J. A., Thompsonschild, S. L., & Avants, B. B. (2013). Reproducibility of functional network metrics and network structure: A comparison of task-related BOLD, resting ASL with BOLD contrast, and resting cerebral blood flow. *Cognitive, Affective, & Behavioral Neuroscience*, 13(3), 627–640.
- Wig, G. S., Schlaggar, B. L., & Petersen, S. E. (2011). Concepts and principles in the analysis of brain networks. *Annals of the New York Academy of Sciences*, 1224(1), 126–146.
- Wijk, B. C. M. V., Stam, C. J., & Daffertshofer, A. (2010). Comparing brain networks of different size and connectivity density using graph theory. *PLoS One*, 5(10), e13701.
- Winer, B. J. (1962). Statistical principles in experimental design. *International Student Edition*, 29, 7304–7309.
- Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Yücel, M., Pantelis, C., & Bullmore, E. T. (2010). Whole-brain anatomical networks: Does the choice of nodes matter? *Neuroimage*, 50(3), 970–983.
- 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(3), e33540.
- Zhao, K., Yan, W. J., Chen, Y. H., Zuo, X. N., & Fu, X. (2013). Amygdala volume predicts inter-individual differences in fearful face recognition. *PLoS One*, 8(8), e74096.
- Zhou, J., Gennatas, E. D., Kramer, J. H., Miller, B. L., & Seeley, W. W. (2012). Predicting regional neurodegeneration from the healthy brain functional connectome. *Neuron*, 73(6), 1216–1227.
- Zuo, X. N., & Xing, X. X. (2014). Test-retest reliabilities of resting-state fMRI measurements in human brain functional connectomics: A systems neuroscience perspective. *Neuroscience and Biobehavioral Reviews*, 45, 100–118.
- Zuo, X. N., Ehmke, R., Mennes, M., Imperati, D., Castellanos, F. X., Sporns, O., & Milham, M. P. (2012). Network centrality in the human functional connectome. *Cerebral Cortex*, 22(8), 1862–1875.