

# Graph Theoretical Analysis of Functional Brain Networks: Test-Retest Evaluation on Short- and Long-Term Resting-State Functional MRI Data

Jin-Hui Wang<sup>1</sup>, Xi-Nian Zuo<sup>2,3</sup>, Suril Gohel<sup>4</sup>, Michael P. Milham<sup>3</sup>, Bharat B. Biswal<sup>4</sup>, Yong He<sup>1\*</sup>

**1** State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China, **2** Laboratory for Functional Connectome and Development, Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China, **3** Phyllis Green and Randolph Cöwen Institute for Pediatric Neuroscience, New York University Langone Medical Center, New York, New York, United States of America, **4** Department of Radiology, University of Medicine and Dentistry of New Jersey, Newark, New Jersey, United States of America

## Abstract

Graph-based computational network analysis has proven a powerful tool to quantitatively characterize functional architectures of the brain. However, the test-retest (TRT) reliability of graph metrics of functional networks has not been systematically examined. Here, we investigated TRT reliability of topological metrics of functional brain networks derived from resting-state functional magnetic resonance imaging data. Specifically, we evaluated both short-term (<1 hour apart) and long-term (>5 months apart) TRT reliability for 12 global and 6 local nodal network metrics. We found that reliability of global network metrics was overall low, threshold-sensitive and dependent on several factors of scanning time interval (TI, long-term>short-term), network membership (NM, networks excluding negative correlations>networks including negative correlations) and network type (NT, binarized networks>weighted networks). The dependence was modulated by another factor of node definition (ND) strategy. The local nodal reliability exhibited large variability across nodal metrics and a spatially heterogeneous distribution. Nodal degree was the most reliable metric and varied the least across the factors above. Hub regions in association and limbic/paralimbic cortices showed moderate TRT reliability. Importantly, nodal reliability was robust to above-mentioned four factors. Simulation analysis revealed that global network metrics were extremely sensitive (but varying degrees) to noise in functional connectivity and weighted networks generated numerically more reliable results in compared with binarized networks. For nodal network metrics, they showed high resistance to noise in functional connectivity and no NT related differences were found in the resistance. These findings provide important implications on how to choose reliable analytical schemes and network metrics of interest.

**Citation:** Wang J-H, Zuo X-N, Gohel S, Milham MP, Biswal BB, et al. (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. doi:10.1371/journal.pone.0021976

**Editor:** Marcus Kaiser, Newcastle University, United Kingdom

**Received:** February 25, 2011; **Accepted:** June 14, 2011; **Published:** July 19, 2011

**Copyright:** © 2011 Wang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the Natural Science Foundation of China (Grant Nos. 81030028 and 30870667), Beijing Natural Science Foundation (Grant No. 7102090) and the Scientific Research Foundation for the Returned Overseas Chinese Scholars (State Education Ministry). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: yong.he@bnu.edu.cn

## Introduction

The human brain is a highly complex system represented as a structurally interconnected network by a dense of cortico-cortical axonal pathways (i.e., structural connectome, [1]) and a functionally synchronized network by external or intrinsic coherent neural activity (i.e., functional connectome, [2]). Mapping the brain connectome and highlighting the underlying organizational principles are fundamental for our understanding of the brain architecture. Recent studies have manifested that human brain connectome networks can be constructed using neuroimaging (e.g., functional MRI (fMRI) and diffusion tensor imaging (DTI)) or electrophysiological (e.g., electroencephalography (EEG) and magnetoencephalography (MEG)) data and further investigated by graph theoretical approaches. These brain networks have consistently demonstrated many non-trivial topological properties, such as small-worldness, modularity and highly connected hubs (for reviews, see [3,4,5,6,7]), and exhibited distinct alterations associated with different neurocognitive disorders (for reviews, see [8,9]).

While graph theoretical approaches provide valuable insights into normal brain architecture and pathological mechanism for brain disorders, the test-retest reliability has not been systematically investigated. Reliable measures are fundamental to infer trustworthy conclusions and to serve as potential clinical biomarkers. In response to the demand, several groups examined the TRT reliability/reproducibility of graph network metrics. In anatomical world, Vaessen et al. [10] assessed the reproducibility of anatomical brain networks derived from DTI data and reported high inter-scan reproducibility of network metrics across sampling schemes (e.g., number of gradient directions and gradient amplitude). Bassett et al. [11] demonstrated high reproducibility and low variability of graph metrics for both DTI and diffusion spectrum imaging data derived networks. As for functional imaging arena, Deuker et al. [12] investigated the TRT reliability of functional brain networks using MEG data and reported high reliability during a working memory task but relatively low under resting condition for network metrics. More recently, Telesford et al. [13] constructed functional brain networks using baseline

**Table 1.** Brief descriptions of several specific acronyms used in the present study.

Abbreviation	Full name	Explanation
TI	Scanning time interval	Short-term: <1 hour apart Long-term: >5 months apart
NM	Network membership	Network (+): only positive correlations
		Network (+/-): both positive and negative correlations
NT	Network type	Binarized: binarized network
		Weighted: weighted network
ND	Node definition	S-: Structural ROIs
		F-: Functional ROIs
S-AAL	Structural Anatomical Automatic Labeling atlas	This atlas includes 90 regions
S-HOA	Structural Harvard-Oxford atlas	This atlas includes 112 regions
F-DOS	Functional ROIs from ref (40)	This set of ROIs includes 160 regions

doi:10.1371/journal.pone.0021976.t001

fMRI data during an executive task and demonstrated excellent reproducibility for both small-world properties and network efficiency metrics. Despite these progresses, however, the TRT reliability of network metrics derived from resting-state fMRI (R-fMRI) dataset has not been well documented so far.

R-fMRI is a promising tool to non-invasively map intrinsic functional connectivity patterns of the human brain in vivo [2,14,15,16] and has been extensively used to investigate inherent brain network topological organization (for a review, see [17]). Of note, several previous R-fMRI studies suggest that the strength of

interregional functional connectivity is dynamic in time (from seconds to minutes) and frequency domains [18] and can be modulated by the levels of current conscious awareness [19,20,21,22], emotional state [23] and cognitive demand prior to resting-state scanning [24,25,26]. Using R-fMRI, Shehzad et al. [27] have demonstrated modest to good TRT reliability for some specific functional connections. However, these states or experiments related alterations in functional connectivity may further interact with the global network topology [28]. To our best knowledge, there are no studies to systematically examine the

**Table 2.** Brief descriptions of complex network metrics examined in the present study.

Parameter	Character	Descriptions
Regional nodal parameters		
<sup>1</sup> Degree	$k_i$	The number of connections linked directly to a node
<sup>1</sup> Efficiency	$e_i$	How efficient an index node communicates with the other nodes
<sup>2</sup> Betweenness	$b_i$	The influences of an index node over information flow between other nodes
<sup>1</sup> Clustering coefficient	$c_i$	The extent of interconnectivity among the neighbors of an index node
<sup>2</sup> Participant coefficient	$p_i$	The ability of an index node in keeping the communication between its own module and the other modules
<sup>2</sup> Normalized $p_i$	$p_i^{nor}$	The normalized $p_i$ after correcting for the effects of number of modules
Global network parameters		
<sup>1</sup> Clustering coefficient	$C_p$	The extent of local clustering or cliquishness of a network
<sup>1</sup> Characteristic path length	$L_p$	The extent of overall routing efficiency of a network
<sup>2</sup> Gamma	$\gamma$	The deviation of $C_p$ of a network from those of surrogate random networks
<sup>2</sup> Lambda	$\lambda$	The deviation of $L_p$ of a network from those of surrogate random networks
<sup>2</sup> Sigma	$\sigma$	The small-worldness indicating the extent of a network between randomness and order
<sup>1</sup> Local efficiency	$E_{loc}$	How efficient of information propagation over a node's direct neighbors
<sup>1</sup> Global efficiency	$E_{glob}$	How efficient of information propagation through the whole network
<sup>2</sup> Assortativity	$r$	The tendency of nodes to link those nodes with similar number of edges
<sup>2</sup> Hierarchy	$\beta$	How likely it is that all nodes oscillate with the same wave pattern
<sup>1</sup> Synchronization	$S$	How likely that all nodes fluctuate in the same wave pattern
<sup>2</sup> Modularity	$Q$	The extent that nodes can be divided several subsets with dense connections within them but sparse between them

Note that the formulas listed here are only for binarized networks. For details for weighted networks, see Text S1.

<sup>1</sup>First-order network metrics which are dependent on only one graph property.

<sup>2</sup>Second-order network metrics which are dependent on more than one property or are defined as ratios of one property.

doi:10.1371/journal.pone.0021976.t002

TRT reliability of network topological metrics derived from R-fMRI data. Accordingly, systematic and direct work is clearly warranted.

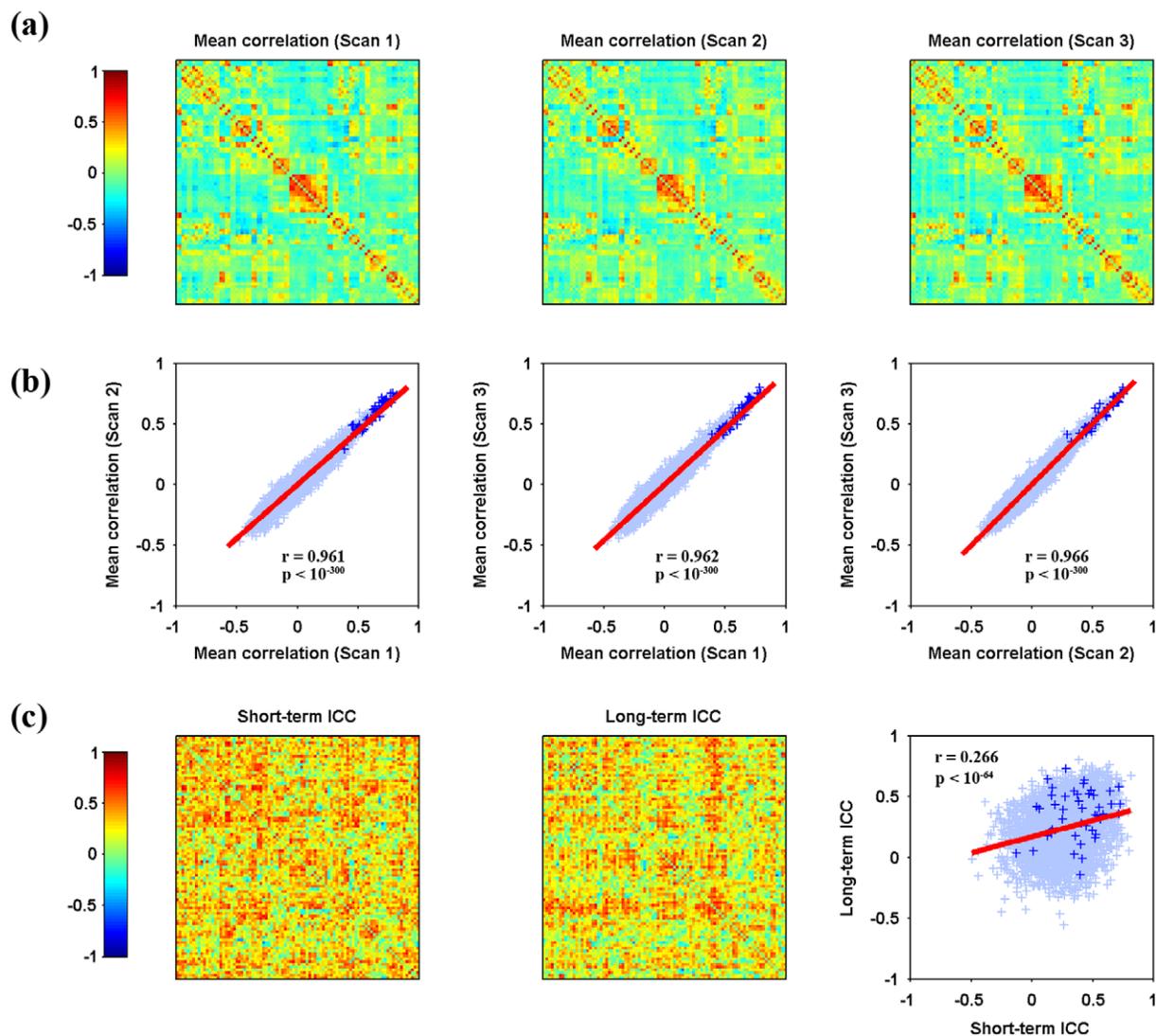
In the current study, we implemented a comprehensive estimation of TRT reliability for both global network properties and regional nodal characteristics of intrinsic functional brain networks constructed using a public TRT R-fMRI dataset (<http://www.nitrc.org/projects/trt>). This dataset allows us to examine both short-term (<1 hour apart) and long-term (>5 months apart) network reliability. Moreover, given numerous discrepancies in the analytical strategies of existing brain network studies (e.g., how to define network nodes or how to deal with negative correlations/connections), we further evaluated the effects of three factors on network reliability. They are: (1) network node definition (ND, i.e., structural regions of interest (ROIs) based node definition or

functional ROIs based node definition); (2) network membership (NM, i.e., inclusion or exclusion negative correlations); and (3) network type (NT, i.e., binarized or weighted networks). Table 1 lists those acronyms specific to the current study.

## Methods

### Subjects

We used a TRT R-fMRI dataset of 25 participants (mean age  $30.7 \pm 8.8$ , 9 males) that is publicly available at NITRC (<http://www.nitrc.org/projects/trt>). The dataset has been used to examine TRT reliability of seed-based resting-state functional connectivity (RSFC) [27], independent component analysis and dual regression [29], amplitude of low-frequency fluctuations [30] and functional homotopy [31].



**Figure 1. Spatial similarity and TRT reliability patterns of S-AAL-based RSFC.** Mean Pearson correlation matrices (a), consistency of overall patterns between mean matrices (b) and TRT reliability of individual connections as well as the relationship between short-term and long-term reliability (c) are illustrated. The mean correlation matrices exhibited high similarity from both visual inspection (a) and quantitative spatial correlation analyses (b). Further TRT reliability analyses revealed a portion of connections exhibiting fair to excellent reliability (c, also see Fig. 2). Moreover, short-term reliability was significantly ( $p < 0.05$ ) correlated with long-term reliability among connections (c). Functional connections linking inter-hemisphere homotopic regions, as highlighted by plus signs (+), showed high connectivity strength and many of them exhibited high reliability. TRT, test-retest; RSFC, resting-state functional connectivity; S-AAL, structural ROIs from Anatomical Automatic Labeling atlas. Of note, the structural ROIs were listed as in Table S1.

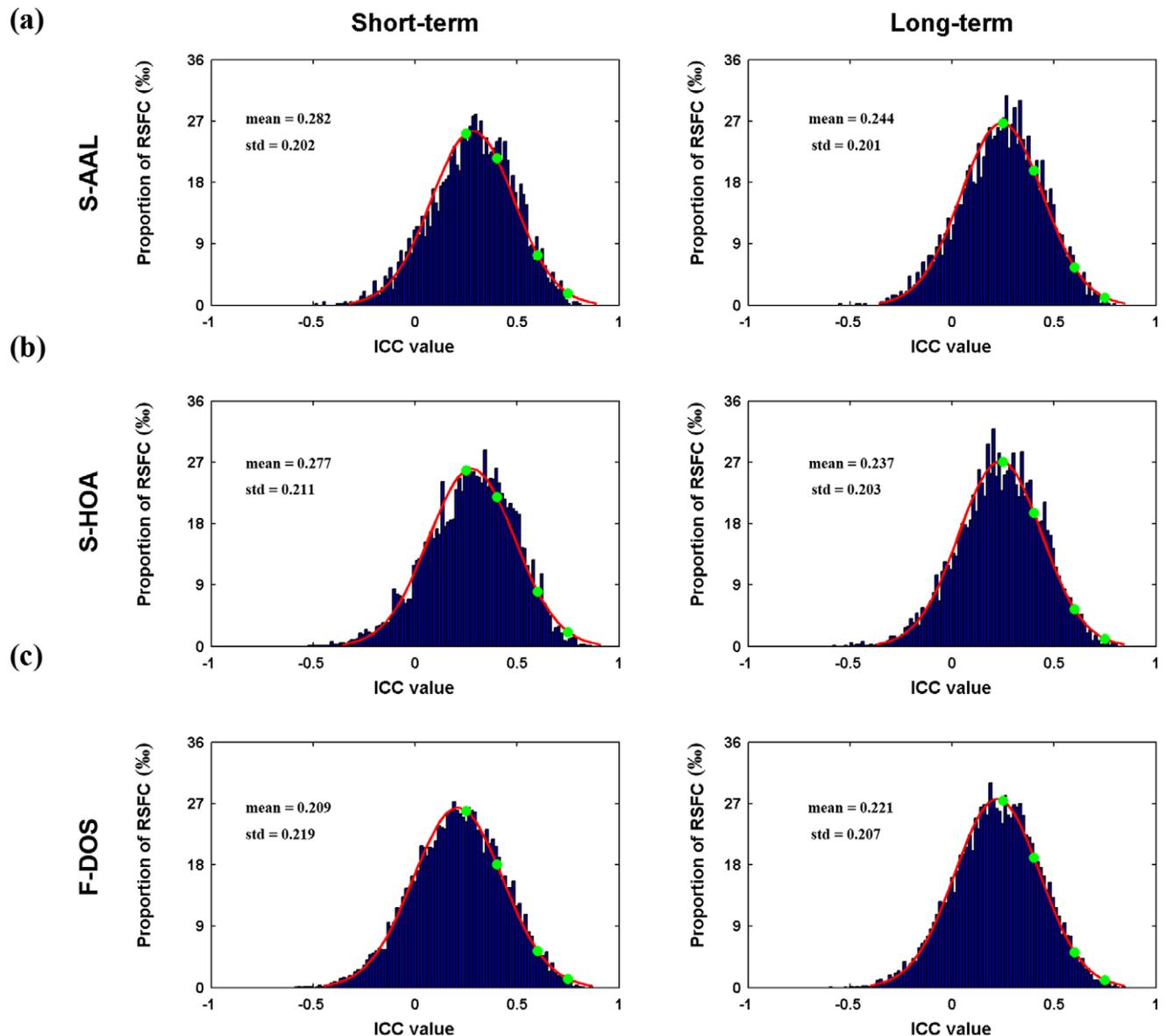
doi:10.1371/journal.pone.0021976.g001

## Data acquisition

Three resting-state scans were obtained for each participant using a Siemens Allegra 3.0-Tesla scanner. Each scan consisted of 197 contiguous EPI functional volumes (time repetition (TR)=2000 ms; time echo (TE)=25 ms; flip angle (FA)=90°, number of slices = 39, matrix = 64×64; field of view (FOV) = 192 mm; acquisition voxel size = 3×3×3 mm<sup>3</sup>). Scans 2 and 3 were conducted in a single-scan session, 45 minutes apart, and were 5–16 months (mean 11±4) after scan 1. All individuals were asked to relax and remain still with eyes open during the scan. Additionally, a high-resolution T1-weighted magnetization prepared gradient echo sequence was also obtained (MPRAGE, TR = 2500 ms; TE = 4.35 ms; inversion time = 900 ms; FA = 8°; number of slices = 176; FOV = 256 mm).

## Data preprocessing

Data preprocessing was performed using SPM5 package (<http://www.fil.ion.ucl.ac.uk/spm>). First, all images were corrected for intra-volume acquisition time offsets between slices using the Sinc interpolation and inter-volume geometrical displacement due to head movement using six-parameter (rigid body) transformation. Then all functional images were normalized into the Montreal Neurological Institute space using an optimum 12-parameter affine transformation and nonlinear deformations, and then resampled to 3-mm isotropic voxels. Finally, the normalized images were further temporally band-pass filtered (0.01–0.1 Hz) to reduce the effects of low-frequency drift and high-frequency physiological noise. Notably, for the extraction of mean nodal time courses of functional defined ROIs, spatial smoothing with 6-mm



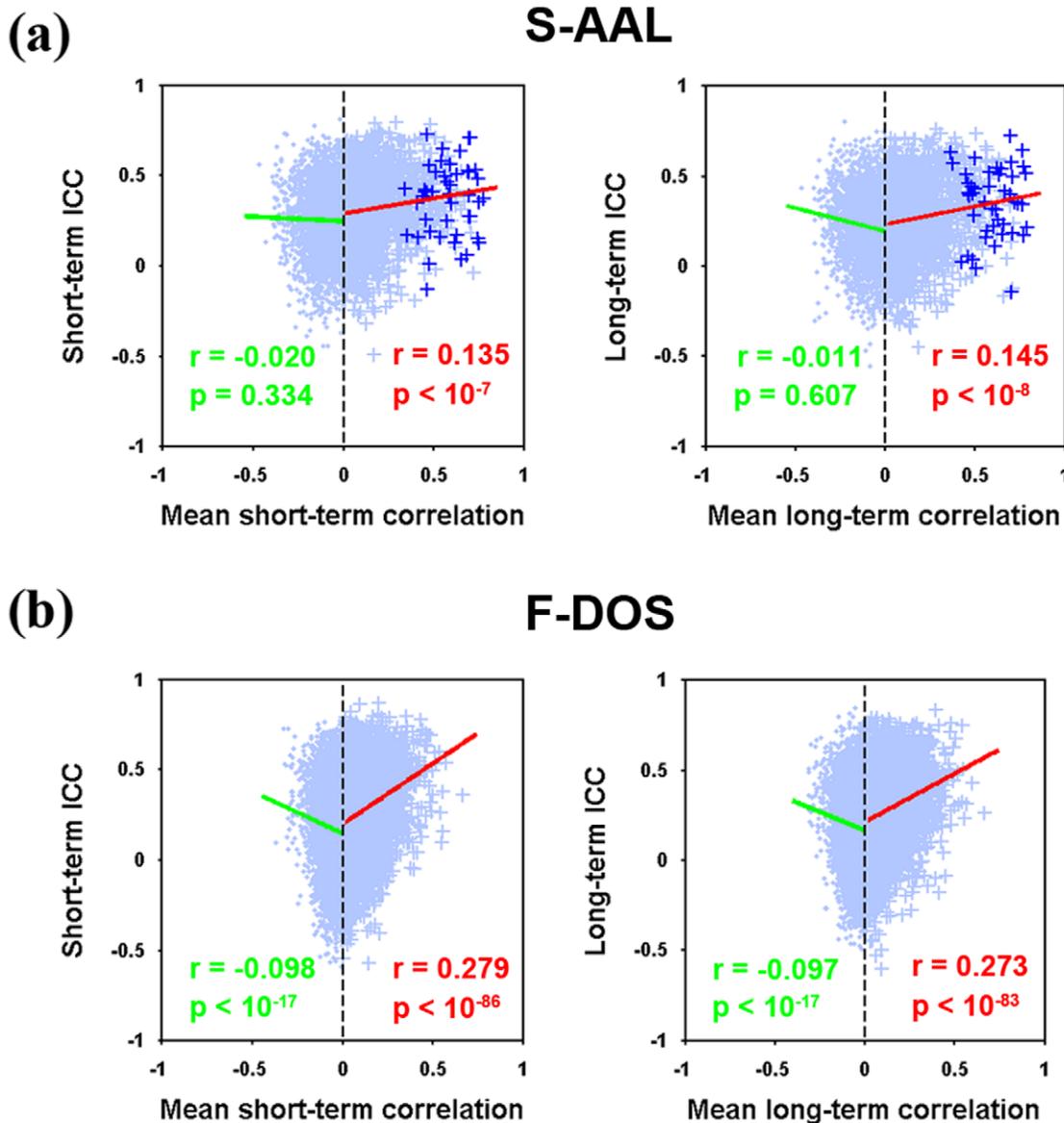
**Figure 2. TRT reliability distribution of RSFC.** Both short-term and long-term TRT reliability exhibit approximately normal distribution for all ROI sets. The mean reliability was about 0.28 (short-term) and 0.24 (long-term) for both structural ROIs-based RSFC while relatively low values were observed for functional ROIs-based RSFC. Green dots indicate the critical values used in the present study to grade reliability. RSFC, resting-state functional connectivity; TRT, test-retest. doi:10.1371/journal.pone.0021976.g002

full width at half maximum (FWHM) Gaussian kernel was performed before band-pass filtering (see below for node definition).

### Functional connectivity matrix and network construction

**Node definition (ND).** A network (i.e., graph) is comprised of nodes and edges connecting nodes. In the current study, nodes represent ROIs and edges represent RSFC between ROIs. Given the accumulating evidence of effects of node definition on network topology [32,33,34,35,36], two strategies of defining ROIs (i.e., anatomical and functional ROIs) were employed to provide a comprehensive assessment of TRT reliability of brain networks

across different node definitions. Specifically, to obtain structurally defined ROIs, a prior Anatomical Automatic Labeling atlas (AAL) [37] and Harvard-Oxford atlas (HOA) [38,39] were separately used to divide the whole brain into different number of regions. These two structural atlases parcellated the whole brain into 45 and 56 regions in each hemisphere and were termed as S-AAL and S-HOA, respectively. To obtain functionally defined ROIs, 160 spheres (radius = 5 mm) were generated around the peak coordinates previously identified from meta-analytic studies of multiple brain functions [40,41] and were termed as F-DOS. These ROIs are comprised of discrete spherical ROIs and not completely cover the cerebral cortex and cerebellum (Fig. S1). All



**Figure 3. Relationship between RSFC and TRT reliability.** Scatter plots of mean connectivity strength against corresponding ICC values are depicted to show the relationship for both S-AAL (a) and F-DOS (b) based correlation matrices. The trend lines were obtained by linear least-square fit. Significant ( $p < 0.05$ ) positive correlations were found between positive RSFC and their corresponding ICC values for both ROIs sets and for both short-term and long-term scanning. In addition, significant negative correlations were also found for negative RSFC with their corresponding ICC values but only for F-DOS-based correlation matrices. These findings suggest higher reliability for stronger RSFC. Functional connections linking inter-hemisphere homotopic regions are highlighted by plus signs (+) for S-AAL but not for F-DOS because of the absence of direct correspondence. RSFC, resting-state functional connectivity; TRT, test-retest; S-AAL, structural ROIs from Anatomical Automatic Labeling atlas; F-DOS, functional ROIs from Dosenbach et al. (2006, 2010).

doi:10.1371/journal.pone.0021976.g003

the ROIs are associated with five different kinds of functions of error-processing, default-mode, memory, language and sensorimotor. There were no any overlaps between ROIs. Names of the three sets of ROIs and their corresponding abbreviations are listed in Table S1, S2 and S3.

**Edge definition.** To measure inter-ROI RSFC, for each of the three sets of ROIs, a mean time series for each ROI was calculated by averaging the time series of all voxels within that ROI. Several potential nuisance signals associated with physiological processes were further removed. Specifically, we regressed out estimated head-motion profiles and global signal from each ROI's mean time series [32,42]. The residuals were then used to estimate inter-ROI RSFC that were quantified by Pearson correlation coefficient. For each subject at each scan, three correlation matrices (corresponding to three sets of ROIs)

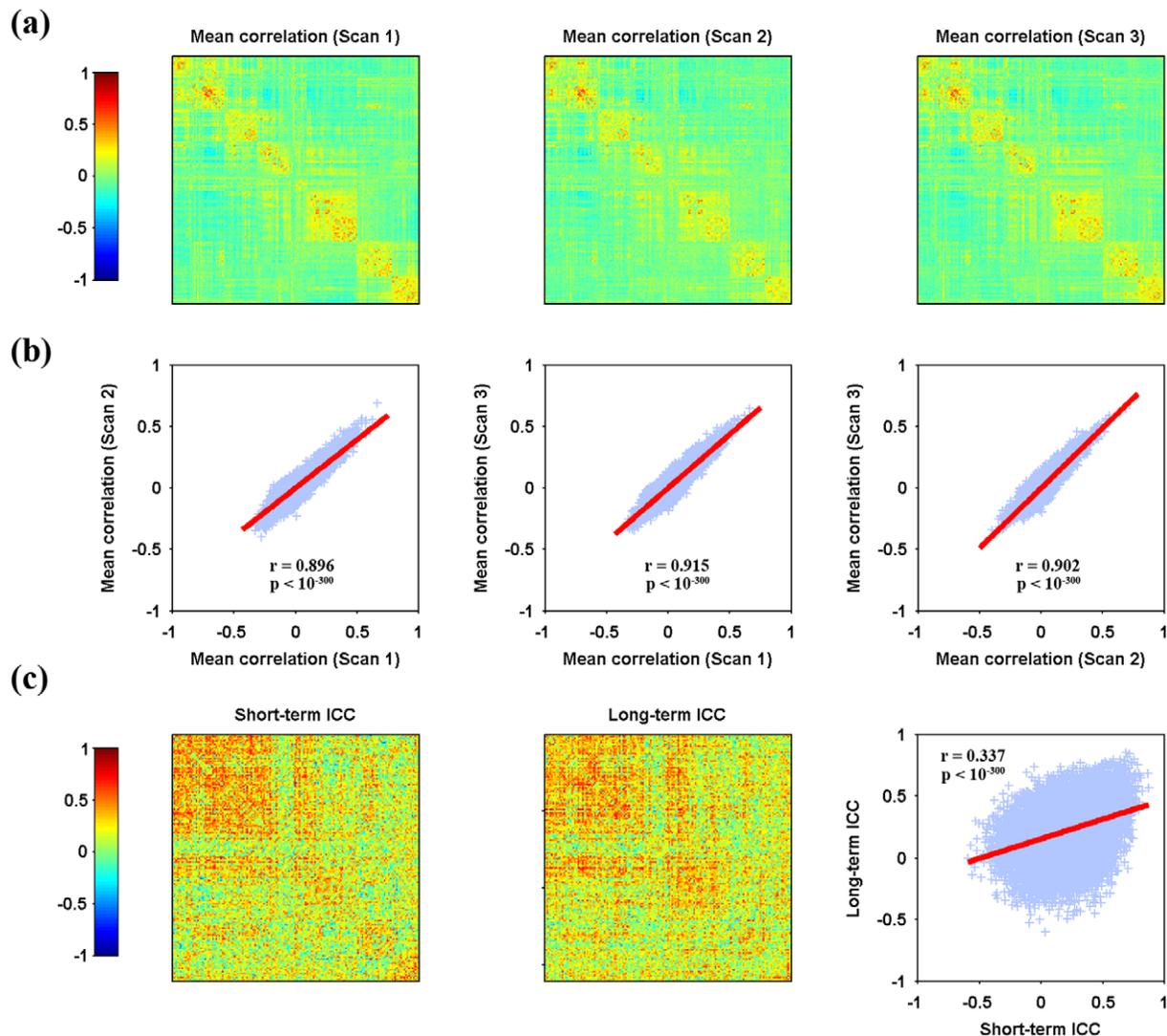
consisting of Pearson correlation coefficients between each pair of ROIs were therefore generated.

**Network type (NT).** Individual correlation matrices  $C_{ij} = [c_{ij}]$  derived above was converted into both a binarized network

$$A_{ij} = [a_{ij}] = \begin{cases} 1, & \text{if } |c_{ij}| > r_{thr}; \\ 0, & \text{others} \end{cases} \quad (1)$$

and a weighted network

$$W_{ij} = [w_{ij}] = \begin{cases} |c_{ij}|, & \text{if } |c_{ij}| > r_{thr}; \\ 0, & \text{others} \end{cases} \quad (2)$$



**Figure 4. Spatial similarity and TRT reliability patterns of F-DOS-based RSFC.** Mean Pearson correlation matrices (a), consistency of overall patterns between mean matrices (b) and TRT reliability of individual connections as well as the relationship between short-term and long-term reliability (c) are illustrated. The mean correlation matrices exhibited high similarity from both visual inspection (a) and quantitative spatial correlation analyses (b). Further TRT reliability analyses revealed many connections exhibiting fair to excellent reliability (c, also see Fig. 2). Moreover, a significant ( $p < 0.05$ ) correlation was found in the ICC matrices between short-term and long-term scans (c). No inter-hemisphere homotopic functional connections were highlighted because of the absence of direct inter-hemisphere correspondence for these ROIs. TRT, test-retest; RSFC, resting-state functional connectivity; F-DOS, functional ROIs from Dosenbach et al. (2006, 2010). Of note, the functional ROIs were listed as in Table S3. doi:10.1371/journal.pone.0021976.g004

where  $r_{thr}$  is a pre-defined correlation threshold. To determine  $r_{thr}$ , sparsity measure,  $S_{thr}$  (defined as the ratio of the number of actual edges divided by the maximum possible number of edges in a network) was applied to each correlation matrix. Using the sparsity threshold, a subject-specific  $r_{thr}$  was determined to threshold each correlation matrix such that the resulting networks have the same sparsity level (i.e., the same number of edges) across subjects and scans. Currently, there is no a definitive way to accurately determine threshold and previous studies construct brain networks either under a single threshold (e.g., [43,44]) or over a continuous threshold range (e.g., [45,46]) in terms of specific constraint conditions. Here, brain networks were constructed over the full range of sparsity, i.e.,  $0 < S_{thr} < 1$  for the whole correlation matrices of both positive and negative correlations and  $0 < S_{thr} < \min [S_{thr-\max}(i,j)]$  for only positive correlation matrices.  $S_{thr-\max}(i,j)$  is a data-specific maximum of  $S_{thr}$  for  $i$ th subject at  $j$ th scans (note that scan2 and scan3 were used to determine the threshold range for short-term reliability estimation and scan1 and the average of scan2 and scan3 were used for long-term reliability estimation). Characterization of network topology over continuous sparsity levels allows us to trace the trajectory of TRT reliability of network properties over varying network structures and to identify specific threshold range of high reliability.

**Network membership (NM).** Given the disagreements in treating negative correlations in R-fMRI network studies (e.g., [46,47]), the thresholding procedure was performed on both

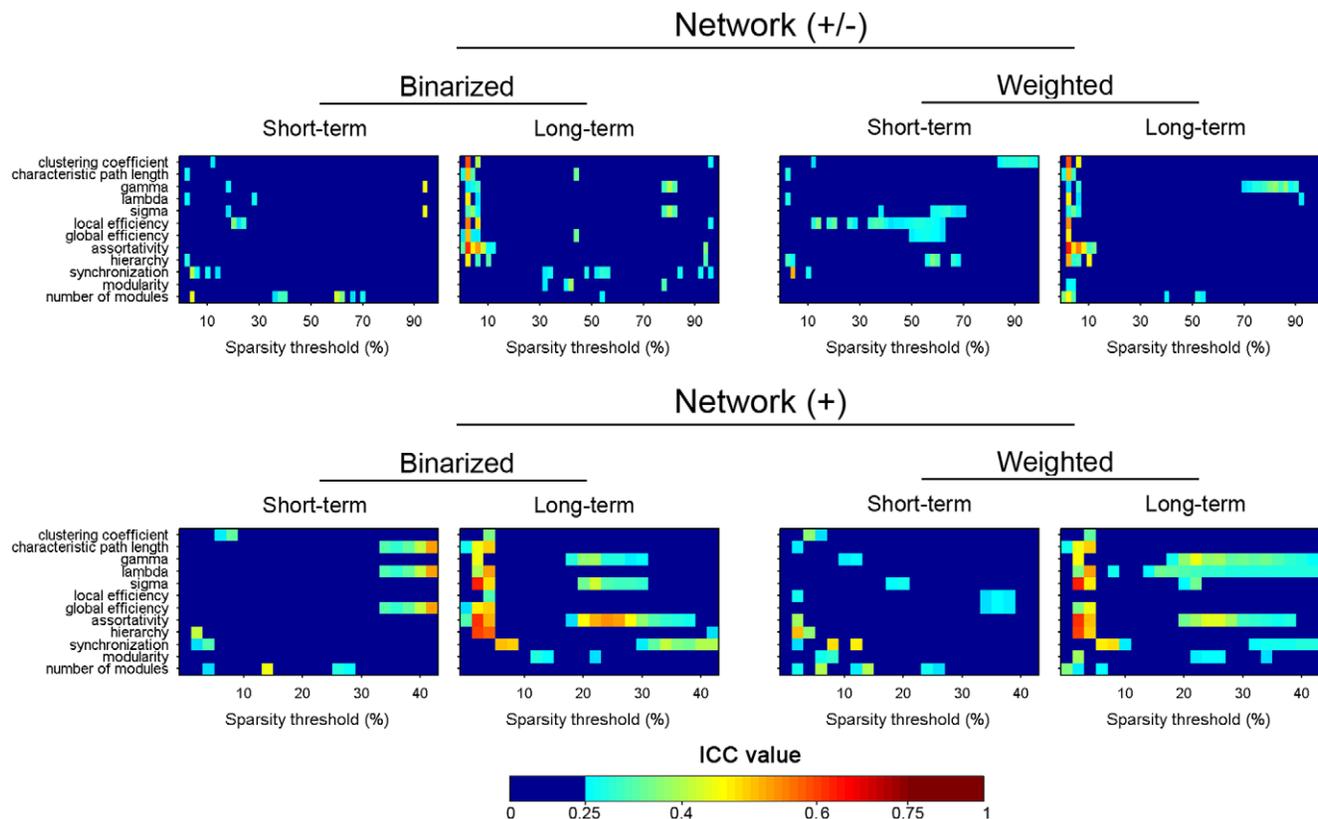
the whole correlation matrices consisting of positive and negative connections and positive correlation matrices consisting of only positive connections (i.e., negative correlations were set to 0).

### Network metrics

We explored two sets of network topological attributes: 1) regional nodal characteristics: degree  $k_i$ , efficiency  $e_i$ , betweenness  $b_i$ , cluster coefficient  $c_i$ , participant coefficient  $p_i$ , and normalized participant coefficient  $p_i^{nor}$ ; 2) 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  $E_{glob}$ ), assortativity  $r$ , hierarchy  $\beta$ , synchronization  $S$ , modularity  $Q$  and the number of modules  $N_M$ . All the computations of network metrics were performed using in-house custom MATLAB codes termed as GRETNA. Text S1 and Table 2 give detailed descriptions for above metrics.

### Test-retest reliability

To investigate the TRT reliability of all graph metrics mentioned above, we used a common index of intraclass correlation [48]. For each global and nodal network measure derived under each combination of the three factors mentioned above, individual values were first merged into two  $25 \times 2$  matrices (rows corresponding to subjects and column corresponding to



**Figure 5. TRT reliability of global network metrics as a function of sparsity threshold for S-AAL-based networks.** ICC values less than 0.25 were mapped to a single color of dark blue as well dark red color for ICC values greater than 0.75, respectively. Network (+/-), networks constructed using absolute both positive and negative correlations; Network (+), networks constructed using only positive correlations; Binarized, binarized network analysis; Weighted, weighted network analysis; TRT, test-retest; S-AAL, structural ROIs from Anatomical Automatic Labeling atlas. doi:10.1371/journal.pone.0021976.g005

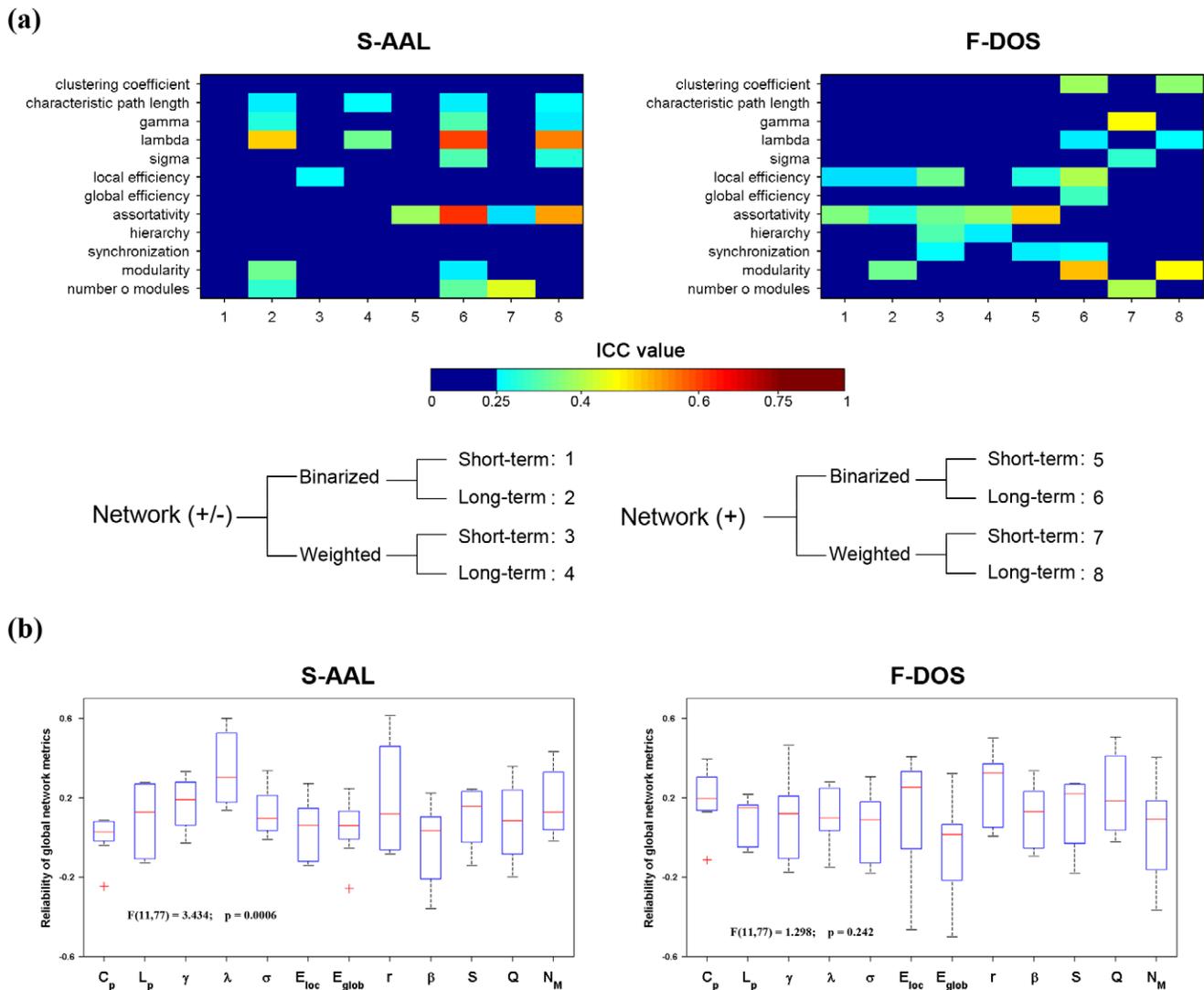
scans), with one representing short-term intra-session across scans 2 and 3 and the other long-term inter-session between scan 1 and the average of scans 2 and 3. Additional long-term reliability estimation using scan 1 and scan 3 alone outputted similar results (Table S4). Therefore, long-term reliability results were reported based on scan 1 and the average of scan 2 and scan 3. Of note, the average was done on individual functional connectivity matrices rather than graph metrics between scan 2 and scan 3, followed by graph metric calculation. Using a one-way ANOVA on each of the two matrices, with random subject effects, we split the total sum of the squares into between-subject ( $MS_b$ ) and within-subject ( $MS_w$ , i.e., residual error) sum of squares. Finally, ICC values were calculated according to the following equation where  $k$  is the number of repeated observations per subject [48]:

$$ICC = \frac{MS_b - MS_w}{MS_b + (k - 1)MS_w} \quad (3)$$

Of note, the ICC derived from (3) has a relationship with the F-value derived from the one-way ANOVA as follows [49]:

$$ICC = \frac{F - 1}{F + (k - 1)} \quad (4)$$

ICC is close to 1 for reliable measures that show low within-subject variance relative to between-subject variance and 0 (negative) otherwise. In the current study, reliability was recorded in terms of



**Figure 6. TRT reliability of summarized global network metrics (a) and metric-related differences in reliability (b).** Areas under curves (AUCs) of each metrics were used to provide threshold-independent reliability estimation. Different metrics showed variable levels of reliability. Several of them were moderately reliable (e.g., lambda for S-AAL-based networks). Subsequent statistical analysis revealed significant differences in TRT reliability among the 12 global network metrics for S-AAL- but not for F-DOS-based networks. ICC values less than 0.25 were mapped to a single color of dark blue as well dark red color for ICC values greater than 0.75, respectively in (a). Network (+/-), networks constructed using absolute both positive and negative correlations; Network (+), networks constructed using only positive correlations; Binarized, binarized network analysis; Weighted, weighted network analysis; TRT, test-retest; S-AAL, structural ROIs from Anatomical Automatic Labeling atlas; F-DOS, functional ROIs from Dosenbach et al. (2006, 2010). doi:10.1371/journal.pone.0021976.g006

the criteria of [50,51], with an ICC value from 0 to 0.25 indicating poor; 0.25 to 0.4 indicating low; 0.4 to 0.6 indicating fair; 0.6 to 0.75 indicating good and 0.75 to 1.0 indicating excellent reliability. Since the network construction was done over a continuous range of sparsity threshold, ICC is a function of the threshold. To provide a threshold-independent reliability assessment, we also calculated the area under curve (AUC, i.e., the integral) for each network metric [52] that was used to compute a single ICC scalar for each network measure. Finally, we compared the consistency between ICC-based reliability and Pearson correlation coefficient-based similarity analysis (across subjects) for network metrics between scans, which was restricted to S-AAL-based networks.

### Simulation analysis

To investigate the effects of numerical changes in RSFC on network metric reliability, we performed simulation analyses as follows: individual functional connectivity matrices based on S-AAL were calculated using dataset of scan1 and their corresponding network metrics (both global and nodal metrics) were used as reference values. Then, for each correlation matrix, different levels of independent Gaussian noise were added and all network metrics were recomputed. The added Gaussian noise were zero mean and the variances varied across six equally spaced levels corresponding to 8.3%, 16.7%, 25.0%, 33.3%, 41.7% and 50.0% of actual functional connectivity variances for each subject. This procedure assures the same proportion of noise added to each correlation matrix. Of note, the procedure of noise addition was performed 5 times. Therefore, 25 (subjects)  $\times$  6 (noise levels)  $\times$  5 (random times) = 750 functional connectivity matrices were simulated in total. Finally, the TRT reliability of each metric was calculated between the reference values and those obtained from simulated functional networks under each level of noise and then averaged across 5 rounds of noise addition.

## Results

### TRT reliability of RSFC: S-AAL

**Consistency of overall RSFC patterns.** The mean RSFC matrices across subjects were calculated (after Fisher's r-to-z

transform) for each of the three TRT scans. Initial visual inspection suggested that mean RSFC matrices showed highly similar spatial patterns between different time points (Fig. 1a). Further quantitative spatial correlation analysis (Pearson correlation) confirmed the visual inspection, as revealed by high correlations in the mean correlation values between each pair among the three scans (Fig. 1b, scan1 vs. scan2:  $r=0.961$ ,  $p<10^{-300}$ ; scan1 vs. scan3:  $r=0.962$ ,  $p<10^{-300}$ ; scan2 vs. scan3:  $r=0.966$ ,  $p<10^{-300}$ ).

**Reliability of RSFC.** ICC-based TRT reliability analysis on individual functional connections demonstrated an approximate normal distribution of the ICC values for all 4005 (i.e.,  $90 \times 89/2$ ) connections with a mean around 0.25 for both short-term and long-term scans (Fig. 2a). In terms of the category used in the present study, 1203 (~30.0%) functional connections exhibited fair to excellent reliability for short-term scans (fair: 1006, ~25.1%; good: 191, ~4.8%; excellent: 6, ~0.2%) and 914 connections (~22.8%) for long-term scans (fair: 796, ~19.9%; good: 114, ~2.9%; excellent: 4, ~0.1%). The majority of RSFC (up to 70% for both short- and long-term scans) showed low or poor reliability. Additionally, a significantly positive correlation (Pearson correlation,  $r=0.266$ ,  $p<10^{-64}$ ) was found in the ICC values across connections between short-term and long-term scans (Fig. 1c).

**Relationship between connectivity and reliability.** To explore the relationship between connectivity strength and reliability, linearly fitted lines were obtained separately for positive connections and negative connections with their corresponding ICC values. We found significantly positive correlations (Pearson correlation) between positive connections and their ICC values for both short-term ( $r=0.135$ ,  $p<10^{-7}$ ) and long-term ( $r=0.145$ ,  $p<10^{-8}$ ) scans (Fig. 3a). No significant correlations were found between negative connections and their ICC values ( $p>0.3$  for both short-term and long-term scans) (Fig. 3a). These findings indicate that reliability of functional connectivity was partly determined by their strength, whereas functional connectivity strength had limited predictive ability to their reliability since the small amount of variance in the functional connectivity reliability explained by their strength ( $R^2<3\%$ ).

**Table 3.** Summary of the main findings in the present study.

	Factors effecting network metric reliability		
	Time interval	Network type	Network membership
S-AAL			
Global metrics	Long>Short (Binarized)	N.S.	Network (+)>Network (+/-)
Nodal metrics	N.S.	N.S.	N.S.
S-HOA			
Global metrics	N.S.	Binarized>Weighted	N.S.
Nodal metrics	N.S.	N.S.	N.S.
F-DOS			
Global metrics <sup>a</sup>	N.S.	N.S.	N.S.
Nodal metrics	N.S.	N.S.	N.S.

S-AAL, structural ROIs from Anatomical Automatic Labeling atlas; S-HOA, structural ROIs from Harvard-Oxford atlas; F-DOS, functional ROIs from ref (40); Network (+/-), networks constructed using absolute both positive and negative correlations; Network (+), networks constructed using only positive correlations; Binarized, binarized network analysis; Weighted, weighted network analysis; N.S., non-significant.

<sup>a</sup>There were no significant differences in reliability among global network metrics for F-DOS based networks while nodal network metrics for F-DOS and both global and nodal metrics for S-AAL and S-HOA based networks were significantly different.

doi:10.1371/journal.pone.0021976.t003



Network (+), networks constructed using only positive correlations; Binarized, binarized network analysis; Weighted, weighted network analysis; TRT, test-retest; S-AAL, structural ROIs from Anatomical Automatic Labeling atlas. doi:10.1371/journal.pone.0021976.g007

### TRT reliability of RSFC: S-HOA

**Consistency of overall RSFC patterns.** The mean S-HOA-based RSFC matrices across subjects also showed highly similar spatial patterns revealed by visual inspection (Fig. S2a) and quantitative spatial correlation analyses (Fig. S2b,  $r > 0.95$  between any two scans).

**Reliability of RSFC.** Similar to S-AAL, approximate normal distributions were also found for the reliability of S-HOA-based RSFC which had comparable mean ( $\sim 0.25$ ) (Fig. 2b). Also consistent with S-AAL, although there were quite a few connections showing fair to good to excellent reliability (1874,  $\sim 30.1\%$  for short-term and 1356,  $\sim 21.8\%$  for long-term scans), most connections were poorly reliable (4342,  $\sim 69.9\%$  for short-term and 4860,  $\sim 78.2\%$  for long-term scans). Finally, short-term reliability was found to positively correlate with long-term reliability across connections (Fig. S2c, Pearson correlation,  $r = 0.307$ ,  $p < 10^{-135}$ ).

**Relationship between connectivity and reliability.** Consistent with S-AAL, positive correlations were found between positive RSFC and their reliability for the S-HOA-based correlations (short-term:  $r = 0.166$ ,  $p < 10^{-17}$ ; long-term:  $r = 0.148$ ,  $p < 10^{-13}$ ), indicating limited determination of functional connectivity strength on their reliability ( $R^2 < 3\%$ ). Additionally, a negative correlation was demonstrated between negative correlations and their reliability for long-term scans ( $r = -0.033$ ,  $p = 0.048$ ) (Fig. S3).

### TRT reliability of RSFC: F-DOS

**Consistency of overall RSFC patterns.** Relative to structural ROIs-based RSFC matrices (both S-AAL and S-HOA), the similarity in the spatial patterns across scans decreased for the mean RSFC matrices derived on the basis of 160 functional ROIs but still remained high (Scan1 vs. Scan2:  $r = 0.896$ ,  $p < 10^{-300}$ ; Scan1 vs. Scan3:  $r = 0.915$ ,  $p < 10^{-300}$ ; Scan2 vs. Scan3:  $r = 0.902$ ,  $p < 10^{-300}$ ) (Fig. 4a and b).

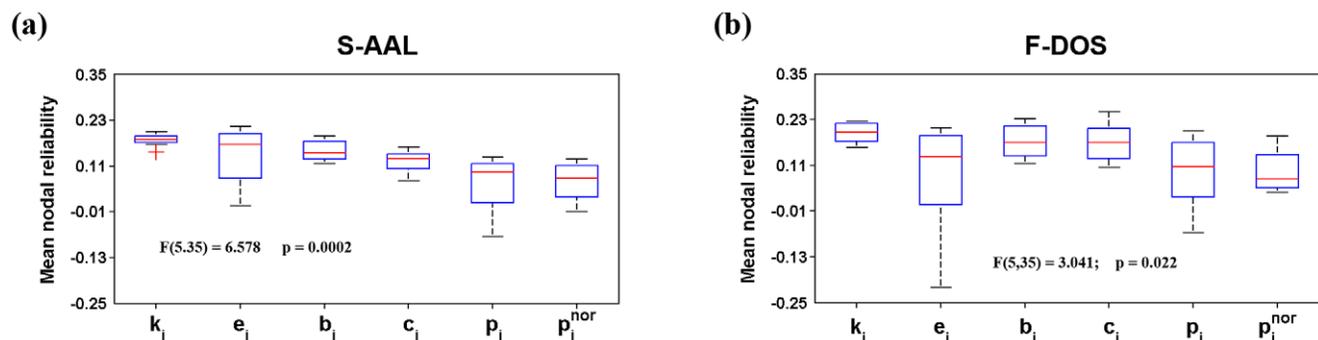
**Reliability of RSFC.** Normal distributions were also found for TRT reliability of functional ROIs-based RSFC, however lower mean ICC values ( $\sim 0.20$ ) were obtained in this case in comparison with structural ROIs-based RSFC ( $\sim 0.25$ ) (Fig. 2c). Moreover, higher percentage of connections (up to  $\sim 80.0\%$ ) showed poor and low reliability for both short-term and long-term scanning procedure, with  $\sim 20.0\%$  showing fair to good to

excellent reliability. Of note, those reliable connections were mainly related with ROIs designated as default mode network according to previous study [40]. In addition, short-term reliability was found to positively correlate with long-term reliability across all connections (Fig. 4c, Pearson correlation,  $r = 0.337$ ,  $p < 10^{-300}$ ).

**Relationship between connectivity and reliability.** In the case of functional ROIs based RSFC matrices, functional connectivity strength explained relatively more in comparison with structural ROIs based matrices but still low variance ( $R^2 < 8\%$ ) in connectivity reliability (positive correlations and their reliability:  $r = 0.279$ ,  $p < 10^{-86}$  for short-term and  $r = 0.273$ ,  $p < 10^{-83}$  for long-term scans; negative correlations and their reliability:  $r = -0.098$ ,  $p < 10^{-17}$  for short-term and  $r = -0.097$ ,  $p < 10^{-17}$  for long-term scans) (Fig. 3b).

### TRT reliability of network metrics: S-AAL

**Reliability of global network metrics.** In the present study, individual networks were constructed at the same sparsity level by applying subject-specific correlation thresholds to individual correlation matrices (see Fig. S4 for the corresponding correlation thresholds under each sparsity level). Sparsity threshold ensures all resultant networks to have comparable topological structures of the same number of edges. Figure 5 shows the TRT reliability of 12 global network metrics over the whole sparsity range. Generally, most global network metrics exhibited poor to low reliability, irrespective of the factors of TI, NT and NM. For example, clustering coefficient  $C_p$  was found to uniformly exhibit poor reliability ( $ICC < 0.25$ ) under all conditions. Nonetheless, we noted that some global metrics (e.g.,  $\lambda$  and assortativity  $r$ ) exhibited modest long-term reliability when the networks were sparsely connected (sparsity  $< 10\%$ ). Interestingly, we found that global network reliability appeared to depend on the factors of TI and NM but relatively insensitive to NT by qualitative visual inspection. Specifically, long-term scans seemed to be associated with better reliability in compared with short-term scans and the exclusion of negative correlations enhanced network reliability (Fig. 5). These were reflected in both increased ICC values and the broadened threshold range of high ICC. Finally, a threshold-independent reliability scalar was obtained for each global network metric by using the AUC. Again, several specific global metrics (e.g.,  $\lambda$ ) demonstrated moderate long-term reliability under certain analytical schemes (Fig. 6a, left). Of note, we found that

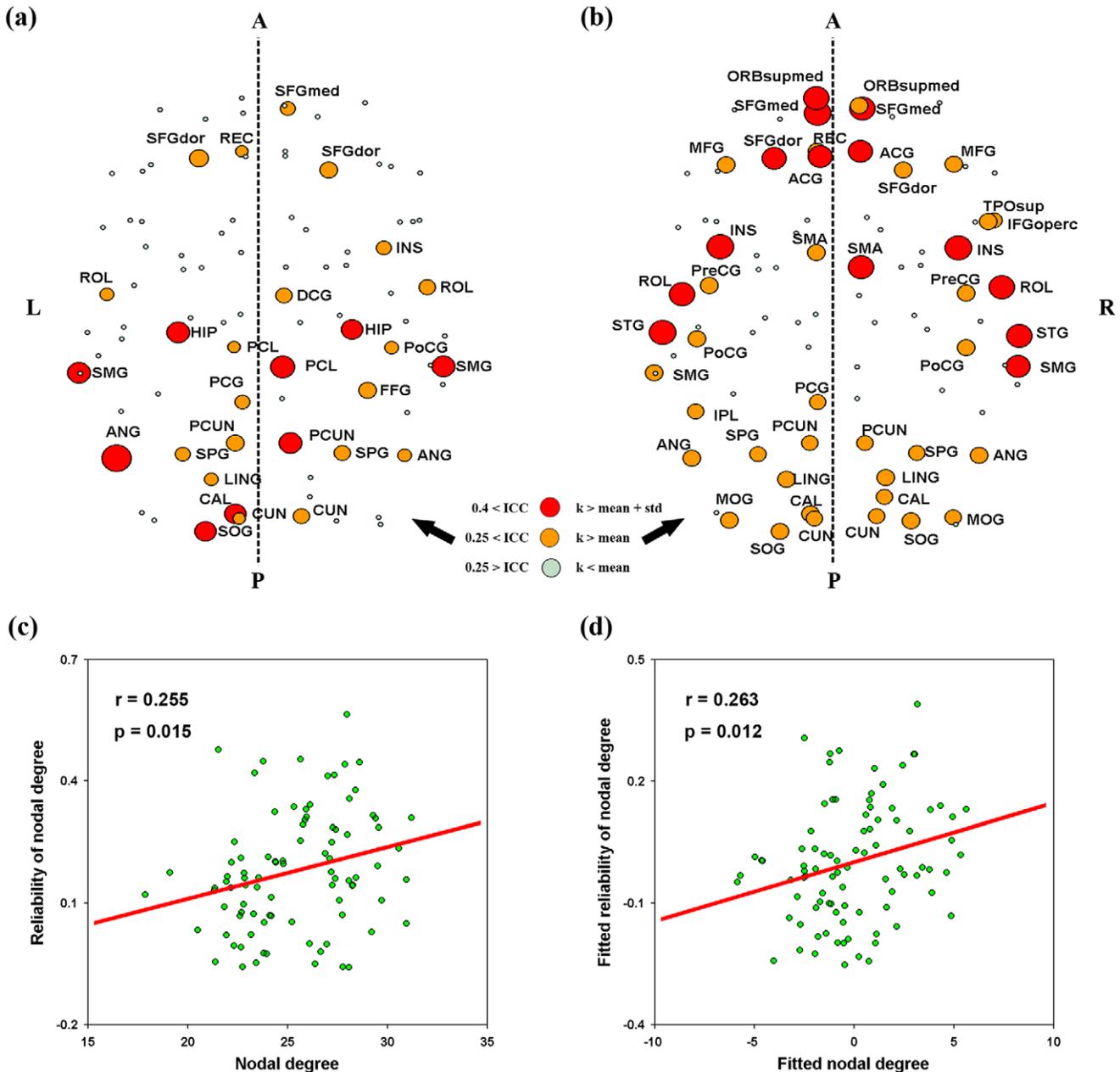


**Figure 8. Boxplot of mean nodal TRT reliability for S-AAL- (a) and F-DOS- (b) based networks.** Significant differences were found in the mean nodal reliability among the six nodal metrics examined with nodal degree showing the highest ICC values and least variances for both ROIs sets. TRT, test-retest; S-AAL, structural ROIs from Anatomical Automatic Labeling atlas; F-DOS, functional ROIs from Dosenbach et al. (2006, 2010). doi:10.1371/journal.pone.0021976.g008

assortativity  $r$  showed moderated both short-term and long-term reliability for networks of positive correlations.

Beyond the descriptive results mentioned above, we further performed statistical analyses to test the differences in reliability among 12 global metrics (one-way repeated-measure ANOVA) and the effects of TI, NM and NT on the reliability of global network metrics (three-factor repeated-measure ANOVA). Those AUC-based ICC values were used for the statistical analyses. The results showed that TRT reliability differed significantly

( $F(11,77) = 3.434, p = 0.001$ ) among 12 global network metrics with lambda  $\lambda$  showing the highest reliability (Fig. 6b, left). Furthermore, TI ( $F(1,11) = 8.176, p = 0.016$ ) and NM ( $F(1,11) = 4.492, p = 0.058$ ) showed significant or marginally significant main effects on global network reliability, respectively. In addition, a significant interaction was observed between TI and NT ( $F(1,11) = 5.317, p = 0.042$ ). NT and other interactions were not significant ( $p > 0.05$ ). Further post-hoc comparisons (paired t-tests) revealed that long-term scans outperformed short-term scans only



**Figure 9. Nodal TRT reliability of degree and its relationship with nodal degree centrality for S-AAL-based networks.** (a) Nodal TRT reliability was mapped in anatomical space after average across scanning time interval, network type and network membership because of no effects of these factors on nodal reliability. (b) Nodal degree centrality (AUCs) was also mapped in anatomical space which was averaged across subjects and factors of scanning time interval, network type and network membership. Trend lines were further obtained by linear least-square fit to reveal the relationship between nodal degree centrality and their corresponding reliability after with (d) and without (c) correcting for the effects of regional size. Of note, the full names of region's abbreviations were listed as in Table S1. TRT, test-retest; S-AAL, structural ROIs from Anatomical Automatic Labeling atlas;  $k$ , nodal degree; A, anterior; P, posterior; L, left; R, right. doi:10.1371/journal.pone.0021976.g009

for binarized networks ( $t(23)=5.100, p<10^{-4}$ ) but not for weighted networks ( $t(23)=1.333, p=0.196$ ) and excluding negative correlations increased the reliability ( $t(47)=3.228, p=0.002$ ) of global network metrics. See Table 3 for the summary of all statistical results.

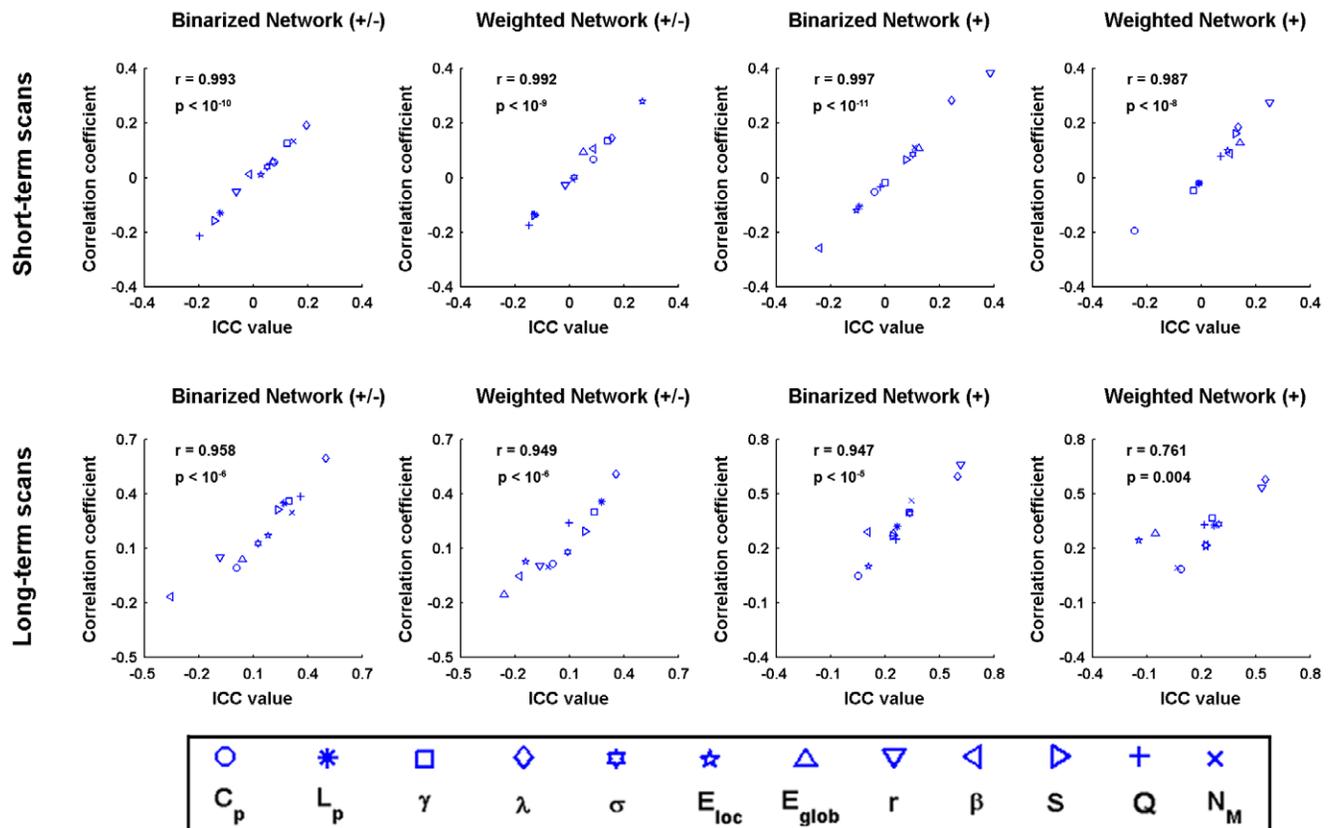
**Reliability of local nodal metrics.** Nodal reliability was estimated based on AUCs. We found that nodal reliability showed: (1) unobvious differences associated with factors of TI, NM and NT; (2) different patterns across nodal metrics; and (3) a spatially heterogeneous distribution over the whole brain (Fig. 7).

First, nodal reliability patterns did not show remarkable differences associated with the factors of TI, NM and NT by visual inspection. To test whether or not there exist differences in the TRT reliability associated with these factors, three-factor repeated-measure ANOVA was further performed on the mean ICC values over all nodes. Results revealed that none of these three factors had significant main effects or interactions on the mean nodal reliability ( $p>0.05$ ) (Table 3).

Second, nodal reliability exhibited variable patterns across nodal attributes under each combination of the three factors. Further one-factor repeated-measure ANOVA on the mean nodal reliability over regions supported this finding that there was significant ( $F(5,35)=6.578, p=0.0002$ ) differences among the six nodal metrics examined, with the highest ICC values and least variance for nodal degree (Fig. 8a).

Finally, nodal reliability distributed non-uniformly over the brain, an observation irrespective of nodal metrics and factors of TI, NM and NT. To highlight those reliable regions, we selectively

mapped nodal reliability of degree of all regions after averaging over factors of TI, NM and NT (Fig. 9a). This was because nodal degree showed higher reliability and less variance as compared to other nodal metrics and was robust to TI and NM as well as NT. As shown in Figure 9a, some association and limbic/paralimbic cortex regions [53] exhibited fair reliability that were predominantly located in bilateral parietal and occipital lobes, such as association cortex regions of the left angular gyrus (ANG), right paracentral lobule (PCL), right precuneus (PCUN), bilateral supramarginal gyrus (SMG), bilateral dorsolateral superior frontal gyrus (SFGdor), right medial superior frontal gyrus (SFGmed) and left superior occipital gyrus (SOG), and limbic/paralimbic regions of the bilateral hippocampus (HIP) and the left posterior cingulate gyrus. In addition, one primary cortex region of the left calcarine fissure (CAL) was also found to be fairly reliable. To test whether or not nodal reliability was related with nodal centrality, we also mapped the mean nodal degree over TI, NM and NT (Fig. 9b) and found visually different patterns between nodal degree and nodal reliability. The most reliable regions located on the posterior while the most connected regions on the anterior portions of the brain. Further quantitative correlation analysis revealed that only tiny variance ( $R^2<7\%$ ) in nodal reliability could be explained by nodal degree centrality in both cases of with ( $r=0.255, p=0.015$ , Fig. 9c) and without ( $r=0.263, p=0.012$ , Fig. 9d) correction for regional nodal size. To test whether there exist a relationship between spatial location and nodal reliability, we compared nodal degree reliability between anterior ( $y>0$ ) and posterior ( $y<0$ ) regions. The results revealed that posterior regions were more



**Figure 10. The similarity between inter-scan ICC-based reliability and inter-scan Pearson correlation coefficients for S-AAL-based networks.** The reliability and correlation analyses revealed highly consistent results ( $r>0.9$  under most conditions), ruling out the possibility of linear scaling biases of network metrics across test and retest scans that will lead to low TRT reliability. doi:10.1371/journal.pone.0021976.g010

reliable than anterior regions even if nodal mean functional connectivity differences were corrected ( $t(87) = 2.801, p = 0.006$ ). In addition, we also found dramatically different patterns across nodal metrics even for those most reliable regions except for the right PCL (Fig. S5a).

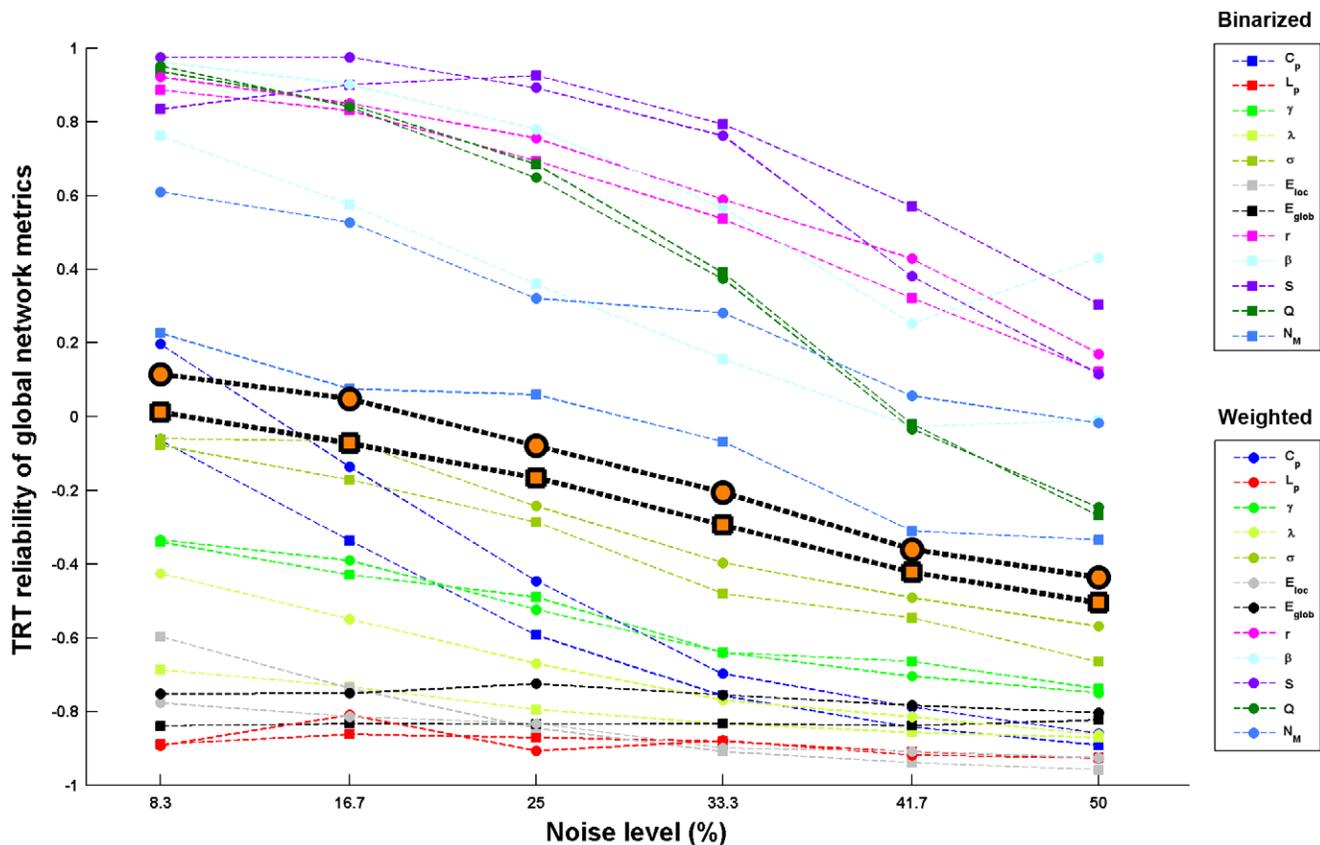
**Consistency between ICC-based reliability analysis and Pearson correlation analysis.** To test the possibility of linear scaling biases across test and retest scans which may result in low TRT reliability, we calculated the inter-scan Pearson correlation coefficient for each global network metric (AUC) across subjects for both short-term and long-term scans. Further scatter plots between ICC values and Pearson correlation coefficients revealed highly correlated patterns ( $r > 0.9$  under most conditions) (Fig. 10), suggesting consistent results revealed by the two measures.

**Simulation results.** By simulating functional connectivity matrices with different levels of noise, we found that: 1) for global network metrics, the TRT reliability was sensitive ( $F(5,55) = 23.303, p < 10^{-11}$ , repeated two-way ANOVA) to disturbances in functional connectivity values and weighted network analysis generated numerically more ( $F(1,11) = 5.183, p = 0.044$ , repeated two-way ANOVA) reliable results than binarized network analysis (Fig. 11); 2) for nodal network metrics, although sensitive to the levels of noise ( $F(5,25) = 7.762, p < 10^{-3}$ , repeated two-way ANOVA), they were highly resistant to numerical changes in functional connectivity and there were no differences ( $F(1,5) =$

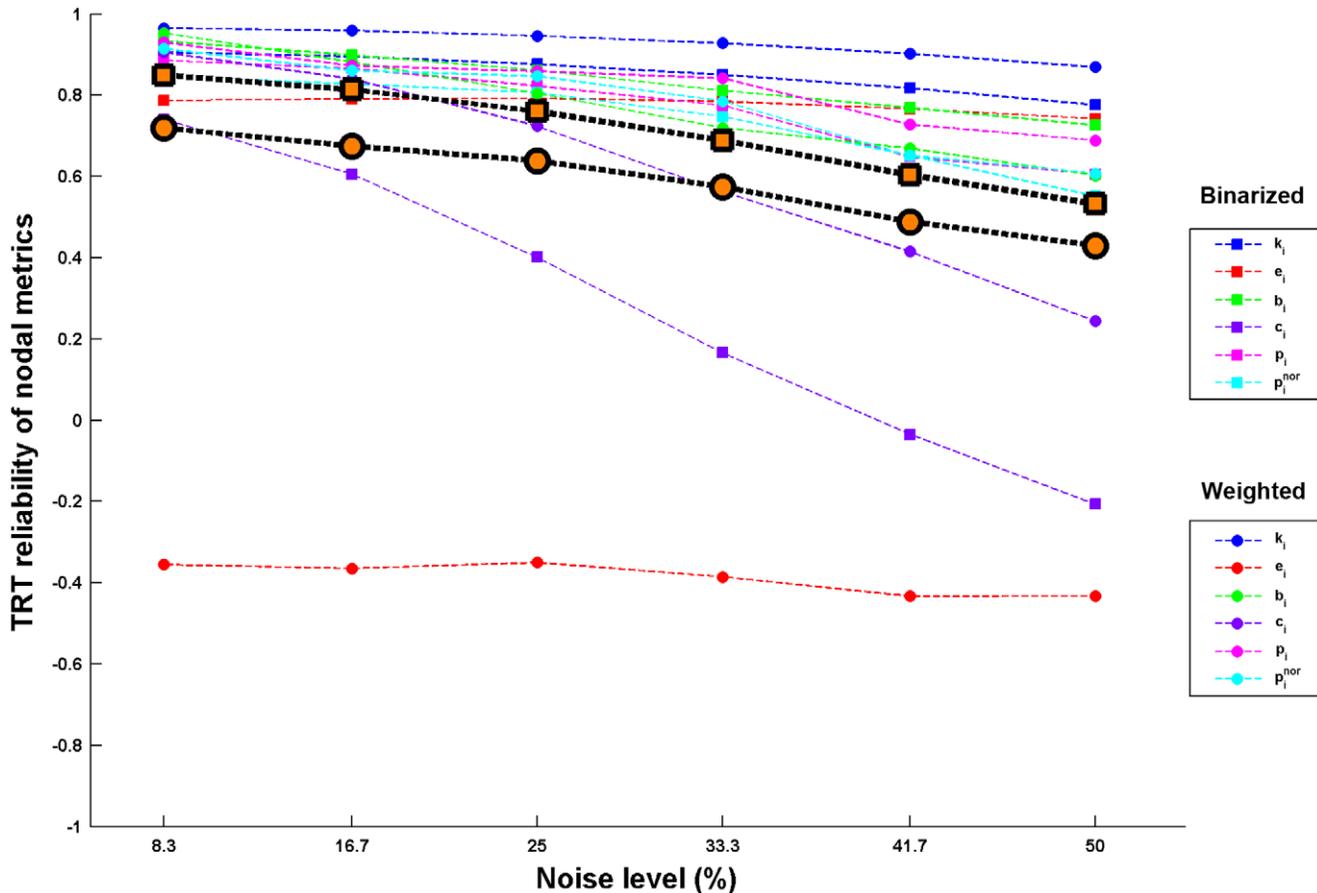
$0.312, p = 0.601$ , repeated two-way ANOVA) in the resistance to noise between binarized and weighted network analyses (Fig. 12); 3) there were no differences in numerical stability against noise in functional connectivity ( $p > 0.05$  under each noise level) between the first-order and second-order network metrics (Table 2); 4) nodal network metrics were more numerically reliable than global network against noise in functional connectivity ( $p < 10^{-3}$  under each noise level). Of note, although sensitive to functional connectivity noise, the degree varied dramatically among global metrics. For instance, small-world parameters and network efficiency were extremely sensitive to even little noise in functional connectivity while assortativity, hierarchy, synchronization and modularity were relatively resistant to noise (Fig. 11).

**TRT reliability of network metrics: S-HOA**

**Reliability of global network metrics.** Analogous to results from S-AAL-based networks, S-HOA-based networks also showed overall low (Fig. S6a) but metric- (Fig. S6b) and threshold- (Fig. S7) sensitive reliability. However, unlike the finding of modest long-term reliability of multiple global metrics for S-AAL-based networks (Fig. 5 and Fig. 6a, left), S-HOA-based networks were mainly related with moderate short-term reliability in multiple global metrics (except for lambda) (Fig. S6a and Fig. S7). Of note, synchronization  $S$  was found to repeatedly show overall moderate reliability (Fig. S6a). Subsequent statistical comparisons revealed



**Figure 11. TRT reliability of global network metrics as a function of noise in RSFC for S-AAL-based networks.** Global network metrics were sensitive to disturbances of RSFC and weighted network analysis generated numerically more stable results in comparison with binarized network analysis. The highlighted black border marks are the average reliability across metrics for binarized (square) and weighted (circle) network analysis, respectively. Of note, the sensitivity varied dramatically among metrics. Small-world parameters and network efficiency were extremely sensitive to even little noise in functional connectivity while assortativity, hierarchy, synchronization and modularity were relatively resistant to noise. TRT, test-retest; RSFC, resting-state functional connectivity. doi:10.1371/journal.pone.0021976.g011



**Figure 12. TRT reliability of nodal network metrics as a function of noise in RSFC for S-AAL-based networks.** Nodal network metrics were sensitive to disturbances of RSFC and no differences were observed in the resistance to noise in functional connectivity between binarized and weighted network analysis. The highlighted black border marks are the average reliability across metrics for binarized (square) and weighted (circle) network analysis, respectively. Of note, although sensitive, nodal network metrics showed strong tolerance of disturbances in RSFC. TRT, test-retest; RSFC, resting-state functional connectivity.  
doi:10.1371/journal.pone.0021976.g012

that TRT reliability of global network metrics were modulated by NT factor ( $F(1,11) = 6.819$ ,  $p = 0.024$ ) with higher reliability observed for binarized networks ( $t(47) = 2.248$ ,  $p = 0.029$ , paired t-test) (Table 3).

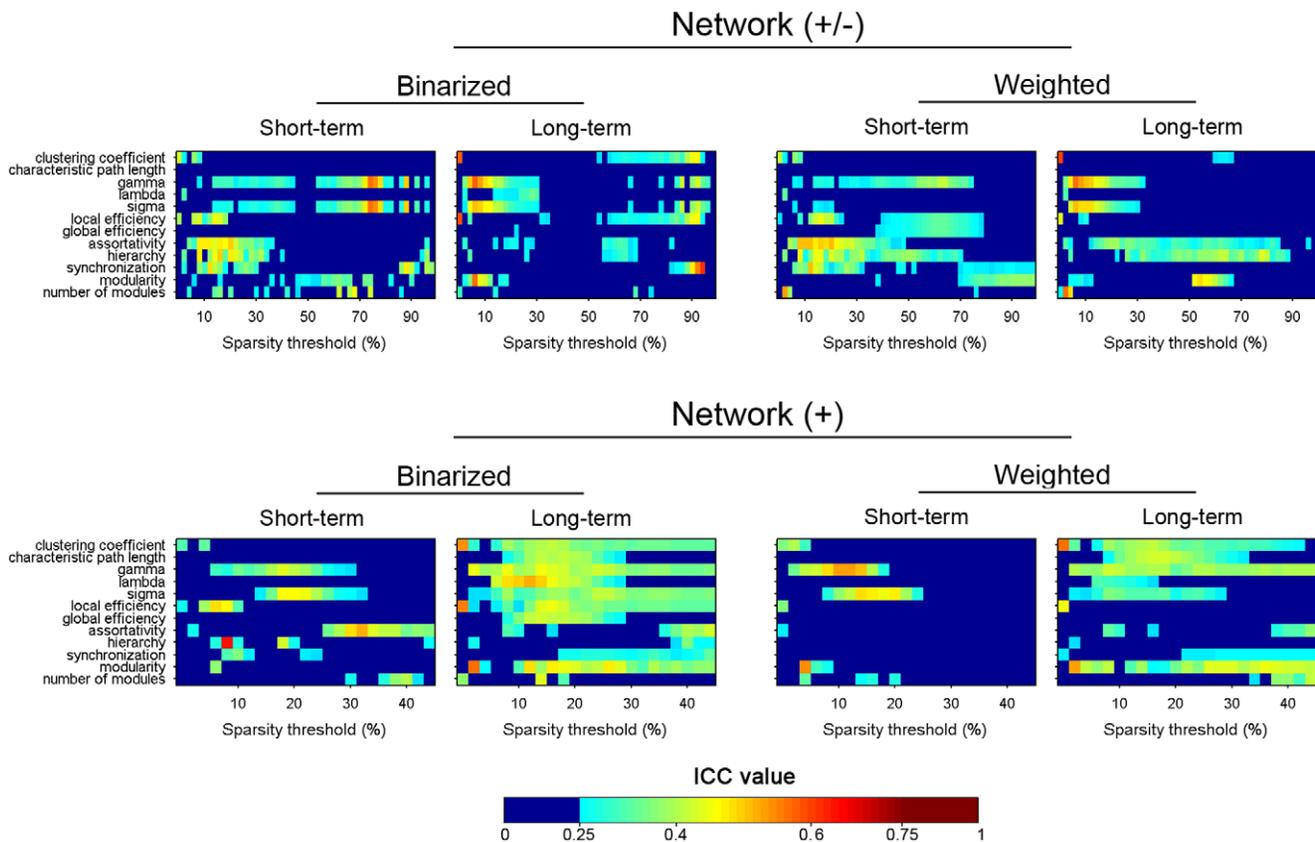
**Reliability of local nodal metrics.** Nodal reliability of S-HOA-based networks (Fig. S8) exhibited the same patterns as those for S-AAL-based networks of 1) factors independent ( $p > 0.05$  for all the factors of TI, NM and NT as well as all possible interactions) (Table 3), 2) metric-sensitive ( $F(5,35) = 12.098$ ,  $p < 10^{-6}$ , degree was the most reliable and least variable) (Fig. S9), and 3) spatial heterogeneous distribution over the brain (Fig. S10a). The most reliable regions were also mainly unimodal and heteromodal association cortex regions and limbic/paralimbic regions of temporal and parietal lobes that were not replicated by other nodal metrics (Fig. S5b). Also, nodal centrality (Fig. S10b) showed no significant relationship ( $R^2 < 2\%$ ) with nodal reliability (Fig. S10c and d).

#### TRT reliability of network metrics: F-DOS

**Reliability of global network metrics.** In compared with structural ROIs-based networks, functional ROIs-based networks showed fair reliability in more global metrics over wider threshold range, especially for networks of positive correlations (Fig. 13). For example, small-world parameters (clustering coefficient  $C_p$ ,

characteristic path length  $L_p$ , normalized clustering coefficient  $\gamma$ , normalized characteristic path length  $\lambda$  and small-worldness  $\sigma$ ) were fairly reliable (predominantly for long-term reliability) for positive networks. The threshold-independent reliability was presented in the right panel of Figure 6a. Subsequent statistical analyses revealed that, in contrast with the measure-related differences in global network reliability observed for structural ROIs based-networks (Fig. 6b, left and Fig. S6b), there was no significant differences ( $F(11,77) = 1.298$ ,  $p = 0.242$ ) among global metrics (Fig. 6b, right) for functional ROIs-based networks. Furthermore, unlike the sensitivity of global network reliability to experimental factor of TI and graph-based analytical strategies of NM and NT for structural ROIs-based networks, reliability of functional ROIs-based networks was robust against these factors ( $p > 0.05$ ) (Table 3).

**Reliability of local nodal metrics.** Figure 14 delineated the nodal reliability for functional ROIs-based networks. No significant ( $p > 0.05$ ) effects were observed for TI, NM and NT on mean nodal reliability (Table 3), consistent with findings from structural ROIs-based networks (both S-AAL and S-HOA). Also analogous to findings of structural ROIs-based networks, nodal degree was found to show the highest reliability and least variance in compared with others ( $F(5,35) = 3.041$ ,  $p = 0.022$ ) (Fig. 8b). After averaged over factors of TI, NM and NT, mean nodal



**Figure 13. TRT reliability of global network metrics as a function of sparsity threshold for F-DOS-based networks.** ICC values less than 0.25 were mapped to a single color of dark blue as well dark red color for ICC values greater than 0.75, respectively. Multiple network metrics showed modest reliability in certain threshold range. Network (+/-), networks constructed using absolute both positive and negative correlations; Network (+), networks constructed using only positive correlations; Binarized, binarized network analysis; Weighted, weighted network analysis; TRT: test-retest; F-DOS, functional ROIs from Dosenbach et al. (2006, 2010). doi:10.1371/journal.pone.0021976.g013

degree reliability showed that there were quite a few reliable regions distributed in bilateral temporal, parietal and the right frontal lobes (Fig. 15a). The nodal centrality pattern (Fig. 15b) can only explain a small fraction ( $R^2 < 6\%$ ) of nodal reliability pattern (Fig. 15c and d). We also noted that the most reliable regions were predominantly located in the right hemisphere (Fig. 15a) and varied across nodal metrics (Fig. S5c).

## Discussion

In the present study, we examined the test-retest reliability of topological metrics of intrinsic connectivity networks derived from human brain R-fMRI data. First, we replicated previous findings that RSFC exhibited modest to high test-retest reliability [27]. Further reliability analyses of network metrics highlighted several main findings: 1) that global network metrics showed overall poor to low but threshold-sensitive reliability; 2) that local nodal metrics were fairly reliable for association and limbic/paralimbic cortex regions; 3) that reliability of network metrics (both global and local) differed significantly among the measures examined; 4) that reliability of global network metrics depended on multiple experiment and analytical factors while nodal reliability was robust to these factors; and 5) that weighted networks (compared to binarized networks) and nodal (compared to global) network metrics were numerically more reliable in the face of noise in functional connectivity. Taken together, we provided a systematically quantitative TRT reliability evaluation of topological metrics

of R-fMRI based brain networks. Our findings suggested continued usage of graph theoretical approaches to explore brain networks and had potential relevance for guiding graph analytical schemes for R-fMRI to achieve reliable results.

For global network metrics, we observed overall low TRT reliability. This observation was consistent with results reported in a previous MEG study [12]. Indeed, compared with task engagement, Deuker et al. (2009) found that resting state was related with significantly lower reliability of global network metrics. This may be related to variable mental states of participants across scans which induce variations in RSFC [19,20,21,22,23,24,25]. Such discrepancies in RSFC especially in shortcuts or inter-module/component connections [54] may further affect the topological organization of the overall connectivity network [28]. To test the possibility, we examined the differences in RSFC strength between scans by paired t-tests. The results revealed that no connections showed significant differences ( $p < 0.05$ , corrected) across scans, implying the temporal stability of RSFC [27,29,31]. Despite of non-significant differences, our simulation analyses indicated that global network metrics were extremely sensitive to numerical changes in RSFC, especially for small-world parameters and network efficiency. Another possible origin of low TRT reliability is due to low between-subject variance or low ability of global network metrics to differentiate subjects. With that said, the low TRT reliability of global metrics may suggest high consistency of global properties of intrinsic brain networks across subjects. Finally, the noise resulting from MRI



**Figure 14. TRT reliability of nodal metrics for F-DOS-based networks.** Nodal reliability varied across nodal attributes and spatial locations. The full names of region's abbreviations were listed as in Table S3. ICC values less than 0.25 were mapped to a single color of dark blue as well dark red color for ICC values greater than 0.75, respectively. Network (+/-), networks constructed using absolute both positive and negative correlations; Network (+), networks constructed using only positive correlations; Binarized, binarized network analysis; Weighted, weighted network analysis; TRT, test-retest; F-DOS, functional ROIs from Dosenbach et al. (2006, 2010). doi:10.1371/journal.pone.0021976.g014

data acquisition and coregistration inaccuracy may also influence network reliability, which should be elucidated in the future work.

Despite of the overall low TRT reliability, some global metrics showed relatively high reliability. For example, lambda showed moderate long-term reliability for structural both S-AAL and S-HOA-based networks. This may be due to the correction of absolute characteristic path length to referenced random networks which compensates for underlying differences of baseline networks. Further statistical analyses revealed significant differences in TRT reliability among global metrics, suggesting an obvious heterogeneity among different global network metrics in reliably capturing intrinsic brain architecture. Moreover, the profiles of global network reliability presented threshold sensitive patterns indicating the importance of threshold selection for reliable results. These findings raise the question of how to determine threshold for brain network studies. A compromise strategy is to investigate brain networks over a continuum threshold range under the circumstance that no sufficient knowledge exists for prior threshold selection.

Several factors were found to significantly affect the TRT reliability of global network metrics. First, inclusion of negative functional connectivity in brain networks tended to decrease TRT reliability of global network properties. Previous evidence has manifested that negative correlations showed greater population and state related variance in the spatial maps [55] and lower TRT reliability [27] relative to positive correlations. Consistent with these findings, our results suggest that negative connectivity should be treated with cautions for resting-state brain network studies, which may reduce the TRT reliability. It should be noted that the emergence of negative connectivity is related with the global signal regression, a currently controversial step in preprocessing R-fMRI data [56,57,58,59].

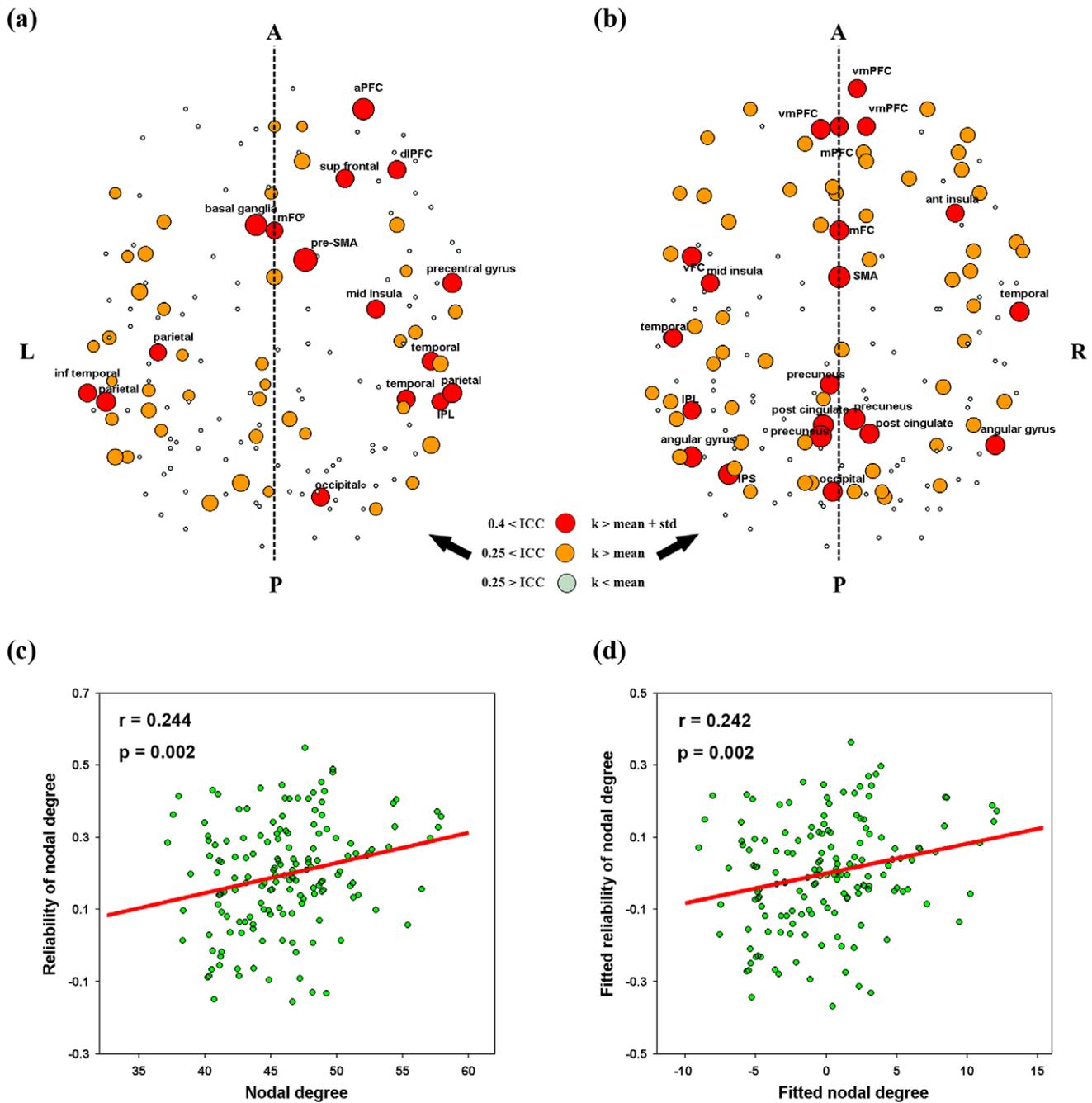
Second, binarized networks outperformed weighted networks in TRT reliability of global network metrics. This finding seemed counterintuitive. Indeed, weighted networks could characterize network topology more precisely and detect more subtle network topological changes than binarized networks due to the consideration of connectivity strength [60]. However, this is not necessary to mean better reliability for weighted networks since the possibility that weighted networks may introduce simultaneously extra noise or overly model individual specific details. All these may lead to more within-subject variance (i.e., variance across scans) and thus lower reliability. Of note, our simulation results showed that weighted networks generated numerically more stable results against noise in functional connectivity in comparison to binarized networks. This suggests that the observed reliability derived from actual R-fMRI data were affected by various factors, not a single factor of numerical changes in functional connectivity.

Third, long-term scans showed higher TRT reliability of global network metrics than short-term scans. This finding was contrast to previous findings that RSFC exhibited higher TRT reliability for short-term interval scans [27]. It may reflect the fact that the average of scan2 and scan3 in the current study can potentially improve the estimation of long-term reliability, i.e., reduce within-session noise [29,30]. To test this interpretation, we further calculated the long-term TRT reliability by using scan1 and scan3 and again found a long-term-larger-than-short-term pattern,

indicating a robust finding. Nevertheless, further work is needed to verify this finding and aid in our understanding of how network topology interacts with the scanning procedure of time interval.

Finally, TRT reliability of global network metrics was modulated by strategies of network node definition. Specifically, reliability of only structural (S-AAL and S-HOA) rather than functional (F-DOS) ROIs-based networks depended on the factors of TI, NM and NT. The discrepancy may reflect different approaches of generating ROIs. Structural ROIs were obtained mainly in terms of anatomical features of sulcal pattern (S-AAL) [37] or standard anatomical boundaries (S-HOA) [38,39] whereas functional ROIs were derived from previous meta-analyses of fMRI activation studies which carried specific functional information [40,41]. Furthermore, even for structural ROIs based networks, the modulations of TI, NM and NT differed across parcellations. Previous studies have demonstrated that network properties were sensitive to nodal definition based on parcellation strategies [32,36] and spatial scales [33,34,35]. Nevertheless, it's hard to conclude which approach or which parcellation is better since all of them are valid and important approaches to uncover brain connectivity architecture from different perspectives [40,43,61,62]. Here, our results provide references for studying intrinsic brain networks, for example, binarized networks should be preferred for S-HOA-based intrinsic brain networks according to our results.

For local nodal metrics, nodal degree showed the highest reliability and least variance across factors of TI, NM and NT among the six nodal metrics. Using this metric, we found that some association cortex and limbic/paralimbic regions exhibited fair to good TRT reliability for S-AAL and S-HOA derived networks, such as precuneus, angular gyrus, superior frontal gyrus, paracentral lobule, supramarginal gyrus, anterior cingulate gyrus, hippocampus and parahippocampal gyrus. Most of these regions have been identified to serve as structural or functional hubs/connectors in human brain networks [43,44,54,60,63,64,65,66]. For F-DOS derived networks, more regions were modestly reliable, predominately located in the right frontal lobe and bilateral parietal and temporal lobes. Hubs are essential in supporting the performance of high cognitive functions of the human brain by integrating specialized brain regions into coordinated networks. Buckner and colleagues [64] demonstrated that the topography of human brain cortical hubs is highly similar across populations and robust against task states, therefore reflecting a stable property of brain functional architecture. Here, our results indicate that those reliable regions qualitatively tend to serve as hubs in intrinsic functional brain networks. Nonetheless, our quantitative analysis found that nodal reliability showed low correlations (although significant) with nodal centrality ( $R^2 < 10\%$ ), suggesting limited predictive ability of nodal centrality on reliability. These findings imply that there may exist other factors affecting nodal reliability, such as the spatial locations of nodes or regions. Indeed, we found that posterior regions were more reliable than anterior regions even after correcting for the differences in functional connectivity across regions. This may reflect the nature of the brain in which the neural dynamics of spatially different brain regions are differently constrained in the resting-state. It would be an



**Figure 15. Nodal TRT reliability of degree and its relationship with nodal degree centrality for F-DOS-based networks.** (a) Nodal TRT reliability was mapped in anatomical space after average across scanning time interval, network type and network membership because of no effects of these factors on nodal reliability. (b) Nodal degree centrality (AUCs) was also mapped in anatomical space which was averaged across subjects and factors of scanning time interval, network type and network membership. Trend lines were further obtained by linear least-square fit to reveal the relationship between nodal degree centrality and their corresponding reliability after with (d) and without (c) correcting for the effects of regional size. Of note, the full names of region's abbreviations were listed as in Table S1. TRT, test-retest; F-DOS, functional ROIs from Dosenbach et al. (2006, 2010); k, nodal degree; A, anterior; P, posterior; L, left; R, right. doi:10.1371/journal.pone.0021976.g015

interesting question for future studies. Additionally, we noted that the most reliable regions appeared to predominately locate in the right hemisphere for F-DOS and S-HOA based networks. Hence, exploring brain functional asymmetry from the perspective of reliability may provide more insights into the brain's functional architecture.

Nodal reliability was found to be robust against the factors of TI, NM and NT, regardless of different node definition strategies. Simulation analyses revealed that nodal metrics were highly tolerant of fluctuations in functional connectivity values and were numerically more stable than global network metrics in the face of connectivity noise. The reliable and robust features of nodal

metrics propose local nodal metrics as reliable candidates to reveal topological organization of intrinsic functional brain networks.

There are several issues that remained to be addressed in future. First, the reliability analyses of graph-based network metrics were conducted after several R-fMRI preprocessing steps. To date, how different preprocessing strategies affect the TRT reliability of network metrics is rarely investigated. Specifically, in the current study, RSFC were obtained based on band-pass filtered data (0.01–0.1 Hz). Previous R-fMRI studies have demonstrated frequency specific features for RSFC [67,68,69] and network topology [43,46,70,71]. Accordingly, exploring the impacts of different preprocessing steps on TRT reliability, especially the filtering frequency bands, is an important topic to determine specific processing schemes for consistent, reliable results. Second, we limited our examination of TRT reliability to 12 global network properties and 6 nodal characteristics, which were widely used to characterize brain network architectures. However, there are still a lot of other network metrics, such as motif [72] and vulnerability [73] (for reviews, see [74,75]) whose TRT reliability need to be evaluated in future. Finally, using R-fMRI, we examined the TRT reliability of intrinsic functional brain networks. Previous studies have performed similar analyses of structural or functional brain networks using DTI, MEG, or fMRI data during resting state or cognitive task engagement [10,11,12,13]. Despite these advances, a systematic reliability evaluation using multimodal data from the same cohort of population is warranted to gain more insights into human brain's structural and functional architectures.

In conclusion, we studied the TRT reliability of graph-based network metrics derived from resting-state fMRI data and the effects of several factors on the reliability. Based on our findings, we provide some methodological recommendations for resting-state fMRI community in dealing with brain connectome studies. First, negative correlations need to be excluded or considered with cautions for S-AAL-based brain network studies. Second, binarized networks should be preferred for S-HOA-based brain network studies as compared to weighted networks. Third, reliability of functional ROIs-based networks was robust against the three factors of scanning time, network membership and network type. Finally, nodal metrics (especially nodal degree) could produce more reliable results and are more resilient to functional connectivity disturbances, which should be popularized in future brain network studies. Nonetheless, we pointed out that further work is necessary to standardize the methodological framework on this burgeoning field.

## Supporting Information

**Figure S1** Spatial locations of functionally defined ROIs. These ROIs broadly but not completely cover the cerebral cortex and cerebellum without any overlap between ROIs and were associated with five functions of error-processing, default-mode, memory, language and sensorimotor. A, anterior; P, posterior; L, left; R, right.  
(DOC)

**Figure S2** Spatial similarity and TRT reliability patterns of S-HOA-based RSFC. Mean Pearson correlation matrices (a), consistency of overall patterns between mean matrices (b) and TRT reliability of individual connections as well as the relationship between short-term and long-term reliability (c) are illustrated. The mean correlation matrices exhibited high similarity from both visual inspection (a) and quantitative spatial correlation analyses (b). Further TRT reliability analyses revealed many connections exhibiting fair to excellent reliability (c, also see

Fig. 2). Moreover, a significant ( $p < 0.05$ ) correlation was found in the ICC matrices between short-term and long-term scans (c). Functional connections linking inter-hemisphere homotopic regions, as highlighted by plus signs (+), showed high connectivity strength and many of them exhibited high reliability. TRT, test-retest; RSFC, resting-state functional connectivity; S-HOA, structural ROIs from Harvard-Oxford atlas. Of note, the structural ROIs were listed in the order as in Table S2.  
(DOC)

**Figure S3** Relationship between RSFC and TRT reliability for S-HOA-based correlation matrices. Scatter plots of mean connectivity strength against corresponding ICC values are depicted to show the relationship. The trend lines were obtained by linear least-square fit. Significant ( $p < 0.05$ ) positive correlations were found between positive RSFC and their corresponding ICC values for both short-term and long-term scanning. In addition, significant negative correlations were also found for negative RSFC with their corresponding ICC values but only for long-term scanning. These findings suggest higher reliability for stronger RSFC. Functional connections linking inter-hemisphere homotopic regions are highlighted by plus signs (+). RSFC, resting-state functional connectivity; TRT, test-retest; S-HOA, structural ROIs from Harvard-Oxford atlas.  
(DOC)

**Figure S4** The absolute correlation thresholds under each sparsity level for all the three sets of ROIs based networks. The correlation thresholds decrease with the increase of sparsity and are comparable across scans and across subjects for each set of ROIs-based networks. Of note, negative correlations were included.  
(DOC)

**Figure S5** Ranks of reliable regions revealed by nodal degree over other nodal metrics. (a) S-AAL-based networks; (b), S-HOA-based networks; (c) F-DOS-based networks. The ranks of those most reliable regions in terms of nodal degree (regions with  $ICC > 0.4$  in Fig. 9a, Fig. S10a and Fig. 15a) changed dramatically over nodal metrics for all ROIs sets, indicating inconsistency for most reliable regions. The full names of region's abbreviations were listed as in Table S1, S2 and S3.  
(DOC)

**Figure S6** TRT reliability of summarized global network metrics (a) and metric-related differences in reliability (b). The area under curve (AUC) of each metric was used to provide threshold-independent reliability estimation. Different metrics showed variable levels of reliability. Several of them were moderately reliable (e.g., lambda and synchronization). Subsequent statistical analysis revealed significant differences in TRT reliability among the 12 global network metrics, with lambda showing relatively high reliability and low variance. ICC values less than 0.25 were mapped to a single color of dark blue as well dark red color for ICC values greater than 0.75, respectively in (a). Network (+/-), networks constructed using absolute both positive and negative correlations; Network (+), networks constructed using only positive correlations; Binarized, binarized network analysis; Weighted, weighted network analysis; TRT, test-retest; S-HOA, structural ROIs from Harvard-Oxford atlas.  
(DOC)

**Figure S7** TRT reliability of global network metrics as a function of sparsity threshold for S-HOA-based networks. ICC values less than 0.25 were mapped to a single color of dark blue as well dark red color for ICC values greater than 0.75, respectively. Network (+/-), networks constructed using absolute both positive

and negative correlations; Network (+), networks constructed using only positive correlations; Binarized, binarized network analysis; Weighted, weighted network analysis; TRT: test-retest; S-HOA, structural ROIs from Harvard-Oxford atlas. (DOC)

**Figure S8** TRT reliability of nodal metrics for S-HOA-based networks. Nodal reliability varied across nodal attributes and spatial locations. Moreover, removing negative correlations seemed to result in more regions showing higher reliability in more nodal attributes (predominantly for binarized networks). The full names of region's abbreviations were listed as in Table S2. ICC values less than 0.25 were mapped to a single color of dark blue as well dark red color for ICC values greater than 0.75, respectively. Network (+/-), networks constructed using absolute both positive and negative correlations; Network (+), networks constructed using only positive correlations; Binarized, binarized network analysis; Weighted, weighted network analysis; TRT, test-retest; S-HOA, structural ROIs from Harvard-Oxford atlas. (DOC)

**Figure S9** Boxplot of mean nodal TRT reliability for S-HOA-based networks. Significant differences were found in the mean nodal reliability among the six nodal metrics examined with nodal degree showing the highest ICC values and least variances. TRT, test-retest; S-HOA, structural ROIs from Harvard-Oxford atlas. (DOC)

**Figure S10** Nodal TRT reliability of degree and its relationship with nodal degree centrality for S-HOA-based networks. (a) Nodal TRT reliability was mapped in anatomical space after average across scanning time interval, network type and network membership because of no effects of these factors on nodal reliability. (b) Nodal degree centrality (AUCs) was also mapped in anatomical space which was averaged across subjects and factors of scanning

time interval, network type and network membership. Trend lines were further obtained by linear least-square fit to reveal the relationship between nodal degree centrality and their corresponding reliability after with (d) and without (c) correcting for the effects of regional size. Of note, the full names of region's abbreviations were listed as in Table S2. TRT, test-retest; S-HOA, structural ROIs from Harvard-Oxford atlas; k, nodal degree; A, anterior; P, posterior; L, left; R, right. (DOC)

(DOC)

**Table S1** Regions of interest from S-AAL.

(DOC)

**Table S2** Regions of interest from S-HOA.

(DOC)

**Table S3** Regions of interest from F-DOS.

(DOC)

**Table S4** Correlation coefficients between long-term reliability estimated by scan1 and the average of scan 2 and scan 3 and those estimated by scan 1 and scan 3 alone. (DOC)

(DOC)

**Text S1** Mathematical definitions of network metrics.

(DOC)

## Acknowledgments

We thank Mr. Jonathan Adelstein for his assistance in language editing.

## Author Contributions

Conceived and designed the experiments: YH. Performed the experiments: MPM. Analyzed the data: JHW XNZ. Contributed reagents/materials/analysis tools: JHW XNZ SG MPM BBB YH. Wrote the paper: JHW XNZ YH.

## References

- Sporns O, Tononi G, Kotter R (2005) The human connectome: A structural description of the human brain. *PLoS Comput Biol* 1: e42.
- Biswal BB, Mennes M, Zuo XN, Gohel S, Kelly C, et al. (2010) Toward discovery science of human brain function. *Proc Natl Acad Sci U S A* 107: 4734–4739.
- Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10: 186–198.
- He Y, Evans A (2010) Graph theoretical modeling of brain connectivity. *Curr Opin Neurol* 23: 341–350.
- Guye M, Bettus G, Bartolomei F, Cozzone PJ (2010) Graph theoretical analysis of structural and functional connectivity MRI in normal and pathological brain networks. *MAGMA*.
- Stam CJ, Reijneveld JC (2007) Graph theoretical analysis of complex networks in the brain. *Nonlinear Biomed Phys* 1: 3.
- Bullmore ET, Bassett DS (2010) Brain Graphs: Graphical Models of the Human Brain Connectome. *Annu Rev Clin Psychol*.
- He Y, Chen Z, Gong G, Evans A (2009) Neuronal networks in Alzheimer's disease. *Neuroscientist* 15: 333–350.
- Bassett DS, Bullmore ET (2009) Human brain networks in health and disease. *Curr Opin Neurol* 22: 340–347.
- Vaessen MJ, Hofman PA, Tijssen HN, Aldenkamp AP, Jansen JF, et al. (2010) The effect and reproducibility of different clinical DTI gradient sets on small world brain connectivity measures. *Neuroimage* 51: 1106–1116.
- Bassett DS, Brown JA, Deshpande V, Carlson JM, Grafton ST (2011) Conserved and variable architecture of human white matter connectivity. *Neuroimage* 54: 1262–1279.
- Deuker L, Bullmore ET, Smith M, Christensen S, Nathan PJ, et al. (2009) Reproducibility of graph metrics of human brain functional networks. *Neuroimage* 47: 1460–1468.
- Telesford QK, Morgan AR, Hayasaka S, Simpson SL, Barret W, et al. (2010) Reproducibility of graph metrics in fMRI networks. *Front Neuroinformatics* 4: 12.
- Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med* 34: 537–541.
- Fox MD, Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 8: 700–711.
- Zhang D, Raichle ME (2010) Disease and the brain's dark energy. *Nat Rev Neurol* 6: 15–28.
- Wang J, Zuo XN, He Y (2010) Graph-based network analysis of resting-state functional MRI. *Front Syst Neurosci* 4: 16.
- Chang C, Glover GH (2010) Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *Neuroimage* 50: 81–98.
- Greicius MD, Kiviniemi V, Tervonen O, Vainionpaa V, Alahuhta S, et al. (2008) Persistent default-mode network connectivity during light sedation. *Hum Brain Mapp* 29: 839–847.
- Horovitz SG, Braun AR, Carr WS, Picchioni D, Balkin TJ, et al. (2009) Decoupling of the brain's default mode network during deep sleep. *Proc Natl Acad Sci U S A* 106: 11376–11381.
- Horovitz SG, Fukunaga M, de Zwart JA, van Gelderen P, Fulton SC, et al. (2008) Low frequency BOLD fluctuations during resting wakefulness and light sleep: a simultaneous EEG-fMRI study. *Hum Brain Mapp* 29: 671–682.
- Martuzzi R, Ramani R, Qiu M, Rajeevan N, Constable RT (2010) Functional connectivity and alterations in baseline brain state in humans. *Neuroimage* 49: 823–834.
- Harrison BJ, Pujol J, Ortiz H, Fornito A, Pantelis C, et al. (2008) Modulation of brain resting-state networks by sad mood induction. *PLoS ONE* 3: e1794.
- Waites AB, Stanislavsky A, Abbott DF, Jackson GD (2005) Effect of prior cognitive state on resting state networks measured with functional connectivity. *Hum Brain Mapp* 24: 59–68.
- Albert NB, Robertson EM, Miall RC (2009) The resting human brain and motor learning. *Curr Biol* 19: 1023–1027.
- Lewis CM, Baldassarre A, Committeri G, Romani GL, Corbetta M (2009) Learning sculpts the spontaneous activity of the resting human brain. *Proc Natl Acad Sci U S A* 106: 17558–17563.
- Shehzad Z, Kelly AM, Reiss PT, Gee DG, Gotimer K, et al. (2009) The resting brain: unconstrained yet reliable. *Cereb Cortex* 19: 2209–2229.
- Spoormaker VI, Schroter MS, Gleiser PM, Andrade KC, Dresler M, et al. (2010) Development of a Large-Scale Functional Brain Network during Human Non-Rapid Eye Movement Sleep. *J Neurosci* 30: 11379–11387.

29. Zuo XN, Kelly C, Adelstein JS, Klein DF, Castellanos FX, et al. (2010) Reliable intrinsic connectivity networks: Test-retest evaluation using ICA and dual regression approach. *NeuroImage* 49: 2163–2177.
30. Zuo XN, Di Martino A, Kelly C, Shehzad ZE, Gee DG, et al. (2010) The oscillating brain: complex and reliable. *Neuroimage* 49: 1432–1445.
31. Zuo XN, Kelly C, Di Martino A, Mennes M, Margulies DS, et al. (2010) Growing together and growing apart: regional and sex differences in the lifespan developmental trajectories of functional homotopy. *J Neurosci* 30: 15034–15043.
32. Wang J, Wang L, Zang Y, Yang H, Tang H, et al. (2009) Parcellation-dependent small-world brain functional networks: a resting-state fMRI study. *Hum Brain Mapp* 30: 1511–1523.
33. Hayasaka S, Laurienti PJ (2010) Comparison of characteristics between region- and voxel-based network analyses in resting-state fMRI data. *Neuroimage* 50: 499–508.
34. Zalesky A, Fornito A, Harding IH, Cocchi L, Yucel M, et al. (2010) Whole-brain anatomical networks: does the choice of nodes matter? *Neuroimage* 50: 970–983.
35. Fornito A, Zalesky A, Bullmore ET (2010) Network scaling effects in graph analytic studies of human resting-state fMRI data. *Front Syst Neurosci* 4: 22.
36. Sanabria-Diaz G, Melic-Garcia L, Iturria-Molina Y, Aleman-Gomez Y, Hernandez-Gonzalez G, et al. (2010) Surface area and cortical thickness descriptors reveal different attributes of the structural human brain networks. *Neuroimage* 50: 1497–1510.
37. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, et al. (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15: 273–289.
38. Kennedy DN, Lange N, Makris N, Bates J, Meyer J, et al. (1998) Gyri of the human neocortex: an MRI-based analysis of volume and variance. *Cereb Cortex* 8: 372–384.
39. Makris N, Meyer JW, Bates JF, Yeterian EH, Kennedy DN, et al. (1999) MRI-Based topographic parcellation of human cerebral white matter and nuclei II. Rationale and applications with systematics of cerebral connectivity. *Neuroimage* 9: 18–45.
40. Dosenbach NU, Nardos B, Cohen AL, Fair DA, Power JD, et al. (2010) Prediction of individual brain maturity using fMRI. *Science* 329: 1358–1361.
41. Dosenbach NU, Visscher KM, Palmer ED, Miezin FM, Wenger KK, et al. (2006) A core system for the implementation of task sets. *Neuron* 50: 799–812.
42. Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, et al. (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America* 102: 9673–9678.
43. 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. *J Neurosci* 26: 63–72.
44. Gong G, He Y, Concha L, Lebel C, Gross DW, et al. (2009) Mapping Anatomical Connectivity Patterns of Human Cerebral Cortex Using In Vivo Diffusion Tensor Imaging Tractography. *Cereb Cortex* 19: 524–536.
45. He Y, Chen Z, Evans A (2008) Structural Insights into Aberrant Topological Patterns of Large-Scale Cortical Networks in Alzheimer's Disease. *J Neurosci* 28: 4756–4766.
46. Achard S, Bullmore E (2007) Efficiency and cost of economical brain functional networks. *PLoS Comput Biol* 3: e17.
47. Fair DA, Cohen AL, Power JD, Dosenbach NU, Church JA, et al. (2009) Functional brain networks develop from a “local to distributed” organization. *PLoS Comput Biol* 5: e1000381.
48. Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 86: 420–428.
49. Stratford PW (1989) Confidence limits for your ICC. *Phys Ther* 69: 237–238.
50. Winer BJ (1971) *Statistical principles in experimental design*. New York: McGraw-Hill press. pp 283–293.
51. Sampat MP, Whitman GJ, Stephens TW, Broemeling LD, Heger NA, et al. (2006) The reliability of measuring physical characteristics of spiculated masses on mammography. *Br J Radiol* 79 Spec No 2: S134–140.
52. He Y, Dagher A, Chen Z, Charil A, Zijdenbos A, et al. (2009) Impaired small-world efficiency in structural cortical networks in multiple sclerosis associated with white matter lesion load. *Brain* 132: 3366–3379.
53. Mesulam MM (2000) *Principles of behavioural and cognitive neurology*. New York: Oxford University Press. xviii+540 p.
54. He Y, Wang J, Wang L, Chen ZJ, Yan C, et al. (2009) Uncovering intrinsic modular organization of spontaneous brain activity in humans. *PLoS One* 4: e5226.
55. Tian L, Jiang T, Liang M, Li X, He Y, et al. (2007) Stabilities of negative correlations between blood oxygen level-dependent signals associated with sensory and motor cortices. *Hum Brain Mapp* 28: 681–690.
56. Fox MD, Zhang D, Snyder AZ, Raichle ME (2009) The global signal and observed anticorrelated resting state brain networks. *J Neurophysiol* 101: 3270–3283.
57. Murphy K, Birn RM, Handwerker DA, Jones TB, Bandettini PA (2009) The impact of global signal regression on resting state correlations: are anticorrelated networks introduced? *Neuroimage* 44: 893–905.
58. Weissenbacher A, Kasess C, Gerstl F, Lanzenberger R, Moser E, et al. (2009) Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies. *Neuroimage* 47: 1408–1416.
59. Birn RM, Diamond JB, Smith MA, Bandettini PA (2006) Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. *Neuroimage* 31: 1536–1548.
60. Cole MW, Pathak S, Schneider W (2010) Identifying the brain's most globally connected regions. *Neuroimage* 49: 3132–3148.
61. Salvador R, Suckling J, Coleman MR, Pickard JD, Menon D, et al. (2005) Neurophysiological architecture of functional magnetic resonance images of human brain. *Cereb Cortex* 15: 1332–1342.
62. Alexander-Bloch AF, Gogtay N, Meunier D, Birn R, Clasen L, et al. (2010) Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Front Syst Neurosci* 4: 147.
63. Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, et al. (2008) Mapping the Structural Core of Human Cerebral Cortex. *PLoS Biol* 6: e159.
64. Buckner RL, Sepulcre J, Talukdar T, Krienen FM, Liu H, et al. (2009) Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *J Neurosci* 29: 1860–1873.
65. Yan C, Gong G, Wang J, Wang D, Liu D, et al. (2010) Sex- and Brain Size-Related Small-World Structural Cortical Networks in Young Adults: A DTI Tractography Study. *Cereb Cortex*.
66. Chen ZJ, He Y, Rosa-Neto P, Germann J, Evans AC (2008) Revealing Modular Architecture of Human Brain Structural Networks by Using Cortical Thickness from MRI. *Cereb Cortex* 18: 2374–2381.
67. Salvador R, John Suckling, Schwarzbauer C, Bullmore E (2005) Undirected graphs of frequency-dependent functional connectivity in whole brain networks. *Phil Trans R Soc B* 360: 937–946.
68. Salvador R, Martinez A, Pomarol-Clotet E, Gomar J, Vila F, et al. (2008) A simple view of the brain through a frequency-specific functional connectivity measure. *Neuroimage* 39: 279–289.
69. Salvador R, Suckling J, Schwarzbauer C, Bullmore E (2005) Undirected graphs of frequency-dependent functional connectivity in whole brain networks. *Philos Trans R Soc Lond B Biol Sci* 360: 937–946.
70. Supekar K, Musen M, Menon V (2009) Development of large-scale functional brain networks in children. *PLoS Biol* 7: e1000157.
71. Supekar K, Menon V, Rubin D, Musen M, Greicius MD (2008) Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLoS Comput Biol* 4: e1000100.
72. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, et al. (2002) Network motifs: simple building blocks of complex networks. *Science* 298: 824–827.
73. Costa LD, Rodrigues FA, Travenço G, Boas PRV (2007) Characterization of complex networks: A survey of measurements. *Advances in Physics* 56: 167–242.
74. Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52: 1059–1069.
75. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU (2006) Complex networks: Structure and dynamics. *Physics Reports* 424: 175–308.