



Published in final edited form as:

Neuroimage. 2012 July 16; 61(4): 1153–1164. doi:10.1016/j.neuroimage.2012.03.036.

Characteristics and variability of structural networks derived from diffusion tensor imaging

Hu Cheng^{a,*}, Yang Wang^b, Jinhua Sheng^b, William G. Kronenberger^c, Vincent P. Mathews^b, Tom A. Hummer^c, and Andrew J. Saykin^{a,b,c}

¹Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405, USA

²Center for Neuroimaging, Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN 46202, USA

³Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN 46202, USA

Abstract

Structural brain networks were constructed based on diffusion tensor imaging (DTI) data of 59 young healthy male adults. The networks had 68 nodes, derived from FreeSurfer parcellation of the cortical surface. By means of streamline tractography, the edge weight was defined as the number of streamlines between two nodes normalized by their mean volume. Specifically, two weighting schemes were adopted by considering various biases from fiber tracking. The weighting schemes were tested for possible bias toward the physical size of the nodes. A novel thresholding method was proposed using the variance of number of streamlines in fiber tracking. The backbone networks were extracted and various network analyses were applied to investigate the features of the binary and weighted backbone networks. For weighted networks, a high correlation was observed between nodal strength and betweenness centrality. Despite similar small-worldness features, binary networks and weighted networks are distinctive in many aspects, such as modularity and nodal betweenness centrality. Inter-subject variability was examined for the weighted networks, along with the test–retest reliability from two repeated scans on 44 of the 59 subjects. The inter-/intra-subject variability of weighted networks was discussed in three levels — edge weights, local metrics, and global metrics. The variance of edge weights can be very large. Although local metrics show less variability than the edge weights, they still have considerable amounts of variability. Weighting scheme one, which scales the number of streamlines by their lengths, demonstrates stable intra-class correlation coefficients against thresholding for global efficiency, clustering coefficient and diversity. The intra-class correlation analysis suggests the current approach of constructing weighted network has a reasonably high reproducibility for most global metrics.

Keywords

Structural network; Variability; DTI

Introduction

Recently, interest in understanding the connectivity of the brain from a structural perspective has grown with the development of important technological and methodological

neuroimaging tools. For instance, the advance of diffusion magnetic resonance imaging (MRI) has made it possible to construct structural networks of the human brain with the help of fiber tracking techniques (Gong et al., 2009; Hagmann et al., 2007, 2008). Likewise, improvements in the speed, accuracy and reliability of gray matter parcellation techniques have enabled better modeling of specific brain regions. Thanks to these advances, structural networks can reveal the topological architecture of neural connections and macroscopic pathways of signal transmission within the brain, making structural network analysis a valuable tool to characterize brain anatomy and functions (Honey et al., 2009; van den Heuvel et al., 2009, 2010).

Diffusion tensor imaging (DTI) has been a widely used technique to quantify the white matter connections through tractography (Mori and van Zijl, 2002). Advanced diffusion techniques using high angular resolution reconstruction schemes such as diffusion spectrum imaging (DSI) (Wedeen et al., 2005, 2008) and HARDI (Frank, 2001) or Q-ball imaging (Tuch, 2004) are desirable to resolve fiber crossing and fiber kissing within a voxel. However, the acquisition time for DSI and Q-ball imaging can be very long and technically challenging. Even though DSI can generate more fibers than DTI and has an advantage of mapping fiber bundles over long distances (Gigandet et al., 2010), DTI remains an attractive imaging tool for structural networks when scan time is a crucial factor.

The topological representation of a network is a collection of nodes and edges between pairs of these nodes (Sporns, 2010). This network can be classified as weighted or binary, and directed or undirected. To construct a weighted, undirected brain network, the gray matter is first parcellated into small regions serving as nodes of the network. The links between pairs of nodes are characterized by the white matter tracts derived with diffusion tensor imaging methods.

Many different parcellation methods have been adopted to construct the structural network. Besides anatomical based parcellation (Desikan et al., 2006; Hagmann et al., 2010) used in this work, there are also template-based parcellations, such as automated anatomical labeling (Gong et al., 2009; Lo et al., 2010; Tzourio-Mazoyer et al., 2002; van den Heuvel et al., 2010) and random parcellation (Hagmann et al., 2007; Zalesky et al., 2010). Different weighting schemes and parcellation schemes can result in different networks. Even with the same weighting scheme, it has been found that the metrics of a structural network are largely dependent on the choice of nodes (Zalesky et al., 2010). FreeSurfer parcellation is based on the anatomical features of individual subjects (Fischl et al., 2004). It mitigates the errors introduced in image warping between the individual subjects and the template while it retains the anatomical features of the parcellated regions. This is an advantage when comparing different subjects.

Many studies have shown distinctive alteration of structural and functional networks of human brain associated with development/aging or in neurodegenerative diseases, making various network metrics potential biomarkers (Hagmann et al., 2010; Lo et al., 2010; van den Heuvel et al., 2010). To better understand possible alterations of brain structural networks in various conditions, it is important to investigate the common features and variability of a normal structural network across healthy subjects. While the binary backbone network from 80 young adults has been studied using the DTI data (Gong et al., 2009), neither the backbone network nor its inter-subject variability has been characterized for the weighted network.

On the other hand, construction of the network involves several procedures that may introduce errors and variances. For example, the fiber tracking can introduce errors because of algorithm limitations or low spatial resolution and signal-to-noise ratio (SNR) of DTI.

Recently the test–retest reliability of functional network measures has been examined by several groups (Braun et al., 2011; Wang et al., 2011; Zuo et al., 2011). A similar work is needed for structural networks. The test–retest reliability of the network measures is very critical for the practical application of network analysis and may shed light on experimental design, data processing, group comparison, and longitudinal evaluation. The inter- and intra-subject variability has been discussed on networks from DTI data (Bassett et al., 2010). However, that report was focused on the comparison between DSI and DTI with a very limited sample ($n=7$), and the analysis was on template-based or random parcellation. A larger sample is preferred for the analysis of variability, and ideally the network is constructed using an individual-based parcellation to better extract the variability from true subject difference.

In the current study, we have applied the network analysis on relatively homogeneous samples, with the networks constructed from DTI data and FreeSurfer parcellation. Two weighting schemes (WS) were adopted and tested for any bias from node size. The backbone networks for the specific parcellation and weighting schemes were obtained. Some features of the backbone networks were investigated and compared with the binary features. The inter- and intra-subject variability were examined for different network metrics and characterized in various ways.

Methods

MRI acquisition and tractography

Fifty-nine male young adults with an average age of 24.0 ± 3.2 years were included in this study. They were all healthy volunteers with no history of a neurological or psychiatric disorder. A written informed consent form was obtained for each of the participants. The study was approved by the local institutional review board of the Indiana University School of Medicine.

In order to characterize anatomical networks, participants underwent structural MRI scans, acquired on a 3.0 T TIM Trio scanner using a 12-channel head coil. For white matter tracking, an SE-EPI DTI sequence was performed using the following parameters: matrix= 128×128 ; FOV= 256×256 mm; TE/TR= $77/8300$ ms; 68 transversal slices with 2 mm thickness; 48 diffusion directions with gradients $b=1000$ s/mm², and 8 samplings at $b=0$. In addition, a high resolution T1-weighted MP-RAGE imaging (160 contiguous 1 mm slices, TR: 2300 ms, TE: 2.91 ms, TI: 900 ms, FA: 9, NEX: 1, BW/Pixel: 240, FOV: 256 mm, matrix 256×256) was taken for subsequent brain parcellation. To study intra-subject variability, forty-four of the fifty-nine subjects received another DTI scan using the same protocol. The mean time between two DTI scans was one week.

The DTI data was first preprocessed with FDT toolbox of FSL (<http://www.fmrib.ox.ac.uk/fsl/>) to correct artifacts induced by head motion and eddy currents. To perform these corrections, all the image volumes were registered to the first b_0 image via an affine transformation. The processed DTI images were then exported to Diffusion Toolkit (<http://trackvis.org/>) for fiber tracking, using the streamline tractography algorithm called FACT (fiber assignment by continuous tracking) (Mori et al., 1999). The FACT algorithm initializes streamlines from seed points and propagates these streamlines along the vector of the largest principle axis of the diffusion tensor within each voxel until certain termination criteria are met. The stop angle threshold was set to 35° (default setting in Diffusion Toolkit), which means that if the angle change between two voxels is greater than 35° , the tracking process stops. The outcome of tractography using FACT is affected by the initial position of the seed points. To reduce biases from initial seed positioning, 30 seeds were randomly selected within each voxel. This process significantly reduces the variance of

network due to fiber tracking, based on our comparison using different numbers of seeds (Cheng et al., 2011). In the end, a spline filtering was applied to smooth the streamlines.

Parcellation

Anatomical parcellation was performed using FreeSurfer 4.5 (<http://surfer.nmr.mgh.harvard.edu/>) on the high resolution MP-RAGE image. Parcellation was an automated operation on each subject to obtain 68 gyral-based regions of interest (ROIs), with 34 in each hemisphere (Table 1). These ROIs were on the FreeSurfer space, coregistered with the skull-stripped brain (brain.mgz). The brain image in FreeSurfer space was registered to the low resolution b0 image of DTI data using FLIRT toolbox in FSL, and the warping parameters were applied to the FreeSurfer ROIs, so that these ROIs were mapped to the DTI image. These ROIs were then used for constructing the structural network.

The parcellated regions vary in volume across subjects. For each ROI, the volume across all the subjects can be represented by a vector. A coefficient of variation (CV) for the ROI volume was computed to characterize its variation across subjects. The CV of a quantity is defined as the standard deviation normalized by its mean value. In addition, the Pearson's correlation coefficient of the 68 ROI volumes was calculated for each pair of subjects. A correlation matrix was obtained for all 59 subjects to show the level of similarity of the parcellation.

Network construction

The DTI fiber tracking results were used to calculate three quantities using in-house Matlab (The Mathworks, Inc, Natick, MA, USA) scripts: fiber length for each fiber, number of fibers between any two ROIs obtained in parcellation, and mean fiber length between two ROIs. Those quantities were further employed for constructing the structural network.

In this work, the nodes were chosen to be the 68 ROIs in DTI image space. The weight of edges was defined as the density of the fibers connecting a pair of nodes, which is the number of fibers between two ROIs divided by the mean volume of the two ROIs. The number of fibers is proportional to the fiber length from a deterministic tracking algorithm if initial seeds are selected in all voxels. The number of tracts between two ROIs should be scaled by the fiber length to account for this effect. Therefore, the weights were computed as the ratio between the sum of inverse of the fiber length and the mean volume of two ROIs (weighting scheme 1), as described in Eq. (1):

$$W_{ij} = \frac{2}{n_i + n_j} \sum_m \frac{1}{L_{ij}^m}, \quad (1)$$

where n_i denotes the number of voxels in ROI $_i$, L_{ij}^m denotes the length of the m th fiber between ROI $_i$ and ROI $_j$. A fiber is considered to connect two ROIs if, and only if, its end points fall in the two ROIs. However, scaling with the fiber length may not be necessary in Eq. (1) considering the fact that the success rate of tracking is lower for longer fibers. A second weighting scheme (scheme 2) was used to appreciate that effect (Eq. (2)):

$$w_{ij} = \frac{2NF_{ij}}{n_i + n_j}, \quad (2)$$

where NF_{ij} denotes the number of fibers between ROI $_i$ and ROI $_j$.

The weighted network was constructed for each subject. The networks using weighting scheme 1 and those using weighting scheme 2 were named network 1 and network 2, respectively. The weight, by definition, is the number of fibers scaled by volume of the nodes. To check if there is a systematic bias relating to ROI size, we examined the correlation between the variance of node strength and the node size. If ROI size is related to node strength, variance would be associated with ROI size, indicated by a significant correlation between the variance of node strength and the node size.

Thresholding

Because of noise in the DTI data and oversimplification of the tensor modeling, the fiber tracking may generate many spurious connections that result in false edges in the network. An appropriate thresholding is needed prior to any network analysis. Thresholding is more appropriate for fiber tracking data rather than the constructed network data which involves size of the nodes. Thresholding is a challenging problem due to lack of a proper null hypothesis (Jbabdi and Johansen-Berg, 2011). The natural null hypothesis of zero fibers seems to be too optimistic because it is unlikely that no streamlines interconnect a pair of regions in the absence of a genuine anatomical connection. On the other hand, knowing the variance of the fiber counts can be informative to thresholding in the context of z-test. Using a similar method proposed by National Electrical Manufacturers Association for noise quantification of MRI images (NEMA, 2008), the variance of number of fibers can be estimated from two trials of fiber tracking. For each fiber track data, a complementary network M is constructed using the number of fibers:

$$M_{ij} = NF_{ij}. \quad (3)$$

The difference of the two complementary matrices M^1 and M^2 is nonzero and related to the variance of the number of fibers. Assuming the variance of fiber counts is equal for all the edges (with non-zero value of number of fibers), then the standard deviation of the number of fibers of each edge is proportional to the standard deviation of the edge differences:

$$\sigma_1 = std\left(\left(M^1 - M^2\right)\Big|_{M^1 > 0 \& M^2 > 0}\right) / \sqrt{2}. \quad (4)$$

For a true edge with weak connection, the mean value of stream-lines in tractography is a positive number but we expect to see a truncated normal distribution for those non-zero values if many trials of fiber tracking are performed. The distribution for the values greater than the mean value of streamlines is indistinguishable from the right half of a normal distribution; hence we borrow the z-score as a measure of confidence level but restrict the z-score analysis on the positive side (one-tail z-score).

For a non-existing edge, the expectation value of the normal distribution is zero. A threshold of the fiber counts can be determined from the standard deviation σ_1 , along with the one-tail z-score Z_p for a specific confidence level. Then the process is repeated to obtain a more accurate standard deviation σ_2 :

$$\sigma_2 = std\left(\left(M^1 - M^2\right)\Big|_{M^1 > Z_p \sigma_1 \& M^2 > Z_p \sigma_1}\right) / \sqrt{2}. \quad (5)$$

The new standard deviation can be used to create a mask for the weighted network. For instance, for a 95% confidence interval of the fibers, the weights meet the following condition:

$$w_{ij}=0 \text{ if } M_{ij} < 1.645\sigma_2. \quad (6)$$

Based on the 59 subjects, the standard deviation of the number of fibers is 23.8 ± 2.6 . For simplicity, we used 24.0 as the universal variance of the number of fibers and created masks for the networks of all subjects. The mask is independent of the weighting schemes and can be treated as the binary network.

Backbone networks

The network of an individual subject can be viewed as a small variation from a backbone network, from which the common feature of the structural connectivity of human brain can be extracted (Gong et al., 2009). To obtain the backbone network, the weighted network was first obtained using 95% confidence for z-test, which corresponds to $Z_p=1.65$. Then we ran a one-tailed *t*-test on each weight of the weighted network (after thresholding) for all subjects against the null hypothesis that the weight is zero for a specific edge. A Bonferroni correction was imposed in this approach to account for multiple comparisons (i.e., $68 \times 67 / 2 = 2278$ pairs of regions) at $P < 0.05$. All the edges surviving the test constitute a binary backbone network. Once the binary backbone network is obtained, the corresponding weighted backbone network can be constructed by assigning each edge a weight computed as the mean value of the weights on that edge from all subjects. The procedure was applied to both network 1 and network 2. The corresponding binary backbone networks are named as BBN1 and BBN2; the weighted backbone networks corresponding to BBN1 and BBN2 are named as WBN1 and WBN2, respectively.

Network analysis

Various network metrics were computed for the network of each subject and the backbone networks. The local quantities of the network were computed for nodal degree, nodal strength, betweenness centrality, path length, and clustering coefficient. Nodal degree is the number of links connected to a node. Nodal strength is the sum of neighboring link weights of a node. Assuming the network is written in a matrix form M_{ij} , the nodal strength can be expressed as:

$$str(i) = \sum_j M_{ij}. \quad (7)$$

The mean strength for a node is defined as the nodal strength divided by the nodal degree. Betweenness centrality is the fraction of all shortest paths in the network that pass through a given node. The path length between a node and its neighbor is defined as the inverse of the edge weight. The nodal path length is the average shortest path length between a node and all other nodes. The clustering coefficient characterizes the local segregation, which can be measured by the fraction of triangles around an individual node. The global metrics were computed for mean clustering coefficient (γ), characteristic path length (λ), global efficiency (γ_g), small-worldness index (σ), maximized modularity (Q), optimal number of modules (m), assortativity coefficient (a), and diversity (d). Mean clustering coefficient, γ , is the global mean of clustering coefficient of each node. The characteristic path length, λ , is the average shortest path length between all pairs of nodes. Global efficiency is the average inverse shortest path length. Small-worldness index makes

reference to values of the two key metrics (clustering and path length) in a population of random (randomized) graphs, which is defined as:

$$swi = \frac{gam/gam_{rand}}{lam/lam_{rand}}, \quad (8)$$

where gam_{rand} and lam_{rand} are the mean clustering coefficient and characteristic path length for a random network. Maximized modularity evaluates the density of communities relative to a random model. Optimal module partitions divide the network into modules such that the modularity metric is maximized. The assortativity coefficient is a correlation coefficient between the degrees of all nodes on two opposite ends of a link. Those metrics have been well-defined previously and have been used to characterize the structural network of human brains in a number of studies (Bassett et al., 2010; Rubinov and Sporns, 2010). All the computations were performed in Matlab using the brain Connectivity Tool Box (Rubinov and Sporns, 2010). Forty randomized networks were generated to compute the small-worldness index.

Efficiency, path length, and clustering coefficient are scalable with the weights; their absolute values do not make much sense without being compared to those from a random network that preserves the basic characteristics of the original network. As we are comparing a relatively homogeneous population that shares the common network structure, these metrics were normalized with the total weights according to Eqs. (9) and (10) throughout this paper so that the comparison of these metrics with those from random networks is not necessary:

$$lam_{normalized} = \frac{lam}{(\text{number of all possible edges})} (\text{total weight}), \quad (9)$$

$$X_{normalized} = \frac{X}{(\text{total weight})} (\text{number of all possible edges}), \quad X = gam, \text{gef}, \quad (10)$$

where the number of possible edges is $67 \times 34 = 2278$ in our case. The above normalization is equivalent to the normalization of total weights for all subjects, which can effectively reduce variability due to difference in brain sizes or SNR of DTI data.

Measures of variability

Both intra-subject variability and inter-subject variability were examined in this work. A slightly different threshold of fiber counts was used to construct the weighted networks for the variability analysis. The 95% confidence for z-test was used in conjunction with the Bonferroni correction, resulting $Z_p = 3.30$. The intra-subject variability was characterized by the variance of the network between two scan sessions. While inter-subject variability mainly characterizes subject difference of brain connectivity, it can be affected by the intra-subject variability.

Intra-subject variability

The intra-subject variability was first characterized by the Pearson's correlation coefficient between all the edge weights of two different networks from different scanning sessions. The correlation analysis was performed for both weighting schemes. The intra-subject variability on the global network metrics was also examined using the Pearson's correlation from measurements of 44 subjects.

Inter-subject variability

The Pearson's correlation analysis was used to quantify inter-subject variability as well and a 59×59 correlation matrix was obtained. The mean value and standard deviation for various local and global network metrics were also calculated.

Intra-class correlation coefficient (ICC)

Variances were computed across all the subjects to explore the variations of various network quantities (e.g., lam, gam, etc.). However, the absolute value of the variability of network properties from subject to subject is meaningless without being compared to a reference. We used the ICC and the coefficient of variation to characterize inter-subject variability of various network properties in order to specifically test the hypothesis that variation within an individual between scanning sessions is larger than the variation across individuals (Bassett et al., 2010). The ICC was introduced to measure how much between-subject variation contributed to the total variance. The ICC is defined as:

$$ICC = \frac{\sigma_{bs}^2}{\sigma_{ws}^2 + \sigma_{bs}^2}, \quad (11)$$

where σ_{ws} denotes the within-subject variance and σ_{bs} denotes the between-subject variance. ICC was computed for all the network metrics to quantify how much inter-subject variability is caused by subject differences for those metrics. The coefficient of variation (CV), defined earlier, was also calculated as a measure of inter-subject variability.

Some of the metrics of the weighted network may be largely dependent on the thresholding. To explore the effect of threshold on ICC values, we varied the threshold of fiber counts from 5 to 200 in the network construction and computed the network metrics and ICC for each threshold.

Results

FreeSurfer parcellation

The cerebral cortex was parcellated into 68 ROIs, equally distributed into the left and right hemispheres (Table 1). In the 59 subjects, total ROI volume varied with subject (CV=8.13%).

This variation may simply reflect subject differences in brain size. Fig. 1a shows the mean size and standard deviation of each ROI across all the subjects. The average CV for all ROIs was 17.2%. Although each individual ROI has relatively large fluctuation across subjects, the correlation of the ROI sizes across subjects is very high, as shown in Fig. 1b. The mean correlation coefficient is 0.97. By normalizing each ROI size with the total ROI size, the average coefficient of variation for all ROIs decreased to 14.9%. This indicates that despite the correction of brain volume, a significant variability in anatomy still exists between subjects. The correlation between node ROI size and standard deviation of node strength was evaluated to see if the edge weight is modulated by the ROI size. Figs. 1c and d are scatter plots of the normalized standard deviation of node strength vs. node ROI size for both weighting schemes. The Pearson's correlation coefficient is -0.06 for WS 1 and -0.14 for WS 2, indicating little ROI-size related bias of the weighting schemes.

Backbone network

The binary backbone network was clustered into five modules for BBN1 and four modules for BBN2, while the weighted networks were classified into seven and six modules

respectively (Fig. 2). The modules have large overlaps between BBN1 (Fig. 2e), WBN1 (Fig. 2g) and WBN2 (Fig. 2h), but distinctive between BBN2 (Fig. 2f) and others. Good left–right symmetry can be seen for the network connectivity. The cross coefficient of edge weights between left and right hemisphere is 0.974 and 0.967, for WBN1 and WBN2, respectively.

Global metrics of the backbone networks are summarized in Table 2. The normalized efficiency is very close for all networks. The structural network exhibits a small-world topology. Smallworldness indices are similar for binary and weighted networks but varying with number of edges: 4.38 and 4.44 for BBN1 and WBN1; 3.86 and 3.93 for BBN2 and WBN2. There were moderate differences between most of the global metrics from binary networks and corresponding weighted ones. For instance, the optimal modularity was 0.537 and 0.666 for BBN1 and WBN1; 0.552 and 0.645 for BBN2 and WBN2. The weighting scheme seems to have little impact on the global metrics except for small-worldness index. There is slight difference in the degree distribution of the two backbone networks BBN1/WBN1 and BBN2/WBN2 (Fig. 3a). They share nine nodes that have the highest degrees: left and right superiorfrontal (SFL-L and SFL-R), left and right precuneus (PCN-L and PCN-R), left and right precentral (PRC-L and PRC-R), right lingual (LGL-R), right superior temporal (STL-R), and right insula (ISL-R). There is a high correlation of the degree distribution between left and right hemisphere for both BBN1 ($r=0.94$) and BBN2 ($r=0.95$). Likewise, the total number of degrees was similar between hemispheres for both BBN1 (left: 205; right: 203) and BBN2 (left: 220; right: 220). Although the total degree is about the same for both hemispheres, the ranks of degree are not the same. The total rank of degree in BBN1 is 1365 (left) vs. 981 (right) and 1257 (left) vs. 1089 (right) in BBN2.

The histogram of the network weights is displayed in Fig. 3b (WS 1) and Fig. 3c (WS 2). It is clear that the weights are not normally distributed. The asymmetric long tail in the histogram suggests the distribution is more likely a gamma distribution.

The betweenness centrality of the nodes can change dramatically from binary network to weighted network (Fig. 4). Correlation coefficient of the betweenness coefficient between WBN1 and BBN1 is 0.28; between WBN2 and BBN2 it is 0.29. The correlation is 0.90 between WBN1 and WBN2 and 0.80 between BBN1 and BBN2. The nodes with large betweenness centrality are considered hubs of the network. The top six nodes with high betweenness centrality in BBN1 are left and right superiorfrontal (SFL-L, SFL-R), left and right precentral (PRC-L, PRC-R), left precuneus (PCN-L), and left parahippocampal (PHL-L). This result agrees in part with previous research (Gong et al., 2009). Without overlapping, the top six nodes with a high betweenness centrality in the corresponding weighted network are left and right fusiform (FSF-L, FSF-R), left and right lingual (LGL-L, LGL-R), and left and right insula (ISL-L, ISL-R). Among the six top hubs in WBN1, FSF-L, FSF-R, ISL-L and LGL-R are also hubs in WBN2.

As weights are measures of the fiber density, it is expected to see higher fiber density from a node with more connections. However, a close look of the mean nodal strength shows little correlation between nodal degree and the mean fiber density of a node, defined as the average strength per edge. Instead, a higher correlation between mean nodal strength and nodal betweenness centrality is observed (Fig. 5), indicating that the hubs have more efferent connections than other nodes.

Different metrics behave differently for the weighted networks under thresholding. Fig. 6 shows the relationship between total degree after thresholding and three global metrics: clustering coefficient, maximized modularity, and efficiency. The clustering coefficient goes up and down very mildly in a wide range of thresholding but drops rapidly lately for

both WS 1 and WS 2. The turning point is around 200 for the total degrees. The maximized modularity goes up slowly with a little fluctuation with fewer edges. Conversely, the efficiency does not change at all until reaching a critical threshold, after which it goes down steadily. The total degree corresponding to the critical threshold is 257 and 295 for WS 1 and WS 2, respectively.

Intra-subject variability

The intra-subject correlation coefficients for the 44 repeated subjects are plotted in Figs. 7a and b for WS 1 (0.73–0.95, \bar{x} 0.89±0.046) and WS 2 (0.60–0.93, \bar{x} 0.84±0.070), respectively. The intra-subject variation with weighting scheme 2 is slightly higher than that with weighting scheme 1. The scattered plots of global metrics of the two scans are displayed in Fig. 7c (WS 1) and Fig. 7d (WS 2). The correlation coefficients of these metrics across the 44 subjects were not very high, and varying in a wide range for different metrics and weighting schemes.

Inter-subject variability

There was a high degree of inter-subject correlation between the edge weights of the networks for all 59 subjects. The correlation matrix is shown for WS 1 (Fig. 8a). The average correlation coefficients, 0.72 for WS 1 and 0.66 for WS2, are lower than the corresponding intra-subject correlation. The CV values for each edge weight in the network are high (Fig. 8b). Many edges have CVs higher than 1, indicating very large variation of the edge weights across subjects.

The mean values and coefficients of variation for some of the global measures are listed in Table 3. The two weighting schemes show little difference (<5%) for the mean values of most of the global metrics (deg, lam, msc, qsc, swi, gef) in the group level. The group averages of these metrics are also very close to the values of the backbone networks (Table 2). The coefficient of variation changes dramatically across different metrics. Some of the global measures have very small variability. For instance, the coefficient of variation is around 3% for maximized modularity and around 4% for global efficiency. Some metrics have moderate variability, such as the path length and clustering coefficient. The assortativity coefficient has the largest variation.

The mean value of local metrics was computed for nodal path length, nodal clustering coefficient, nodal strength and nodal betweenness centrality (Fig. 9). Although the variability of the corresponding global metrics is relatively small, the variation is very large for those of the local properties. The mean CV is 0.371 for strength, 0.851 for betweenness centrality, 0.216 for path length, and 0.656 for clustering coefficient. The largest fluctuations of path length occur at the corpus callosum and frontal pole ROIs, which have very small volumes as parcellated by FreeSurfer. Excluding those four nodes, the mean CV value for path length is 0.136.

The results of ICC analysis of the inter-subject variability and intrasubject variability are shown in Table 4. We only computed ICC for six metrics that show small to moderate variations across subjects (Table 3). Despite similar mean values and coefficients of variation, the ICC coefficients are quite different for the two weighting schemes. The ICC coefficient varies from 0.54 to 0.67 for WS 1 but varies from 0.3 to 0.64 for WS 2. For WS 1, the lowest ICC occurs for the clustering coefficient, which has the highest ICC value for WS 2. The top three metrics with highest ICC coefficients for WS 1 are the three with lowest ICC values for WS 2.

The relation between ICC and thresholding is shown in Fig. 10, where the threshold of the number of fibers changes from 5 to 200. For WS 1, ICCs of path length, clustering

coefficient, and smallworldness index change wildly with thresholding, whereas ICCs of global efficiency, maximized modularity, and diversity show less variation with thresholding. The ICC for path length is very small for WS 1 at the beginning of thresholding but gets larger and more stable after the threshold reaches 110, after which both weighting schemes have an ICC value of around 0.5. The ICC for clustering coefficient is fairly stable after the threshold is greater than 75, fluctuating around 0.5 for WS 1 and around 0.6 for WS 2. The ICC of maximized modularity (qsc) lies mainly between 0.5 and 0.7 for WS 1 and between 0.4 and 0.6 for WS 2. The ICC of small-world index (swi) varies significantly with thresholding, even when the threshold is greater than 110. The ICCs of global efficiency (gef) and diversity (div) are much larger than those for WS 2. The results in Table 4 correspond to the points at threshold=80 in Fig. 10.

Discussion

Structural networks of the brain were constructed for young healthy male adults using DTI data and FreeSurfer parcellation. Basic features of the structural network were investigated on a backbone network, with a focus on the differences between binary and weighted networks. Some hubs, such as precuneus and superior frontal cortex, agree with other literature despite a different parcellation scheme for binary networks (Gong et al., 2009). The binary networks and weighted networks share similar small-worldness. However, the weighted network and the binary counterpart can be very distinct in many aspects. For instance, the difference in clustering coefficient and path length indicates that they have different segregation and integration properties (Table 2). The hubs are dramatically different, as shown in Fig. 4. The global-local properties are also different in terms of modularity (Fig. 2 and Table 4). In weighted backbone networks, some weak connections are not essential for global efficiency (Fig. 6), which might be a feature for human brain in terms of robustness. Therefore, it is more appropriate to characterize the brain connectivity with the weighted network.

Several different weighting schemes have been proposed to make the weights of network links meaningful for deterministic tracking data. For instance, Bassett et al. (2010) have defined the weights as the number of fibers between any two regions but also studied an alternative weighting scheme: the number of fibers normalized by the mean volume of the two regions. Other studies have used the weighting scheme proposed by Hagmann et al. (Hagmann et al., 2007; Lo et al., 2010; van den Heuvel et al., 2010), which performs further correction on the distance bias in tractography algorithms. The differing definitions of the network weight may affect the network properties dramatically. Hence, there is a need to justify the choice of weighting scheme. The weighting schemes, which are rather arbitrary, were tested by checking its bias toward node sizes. The weighting schemes adopted in our study showed no bias with node size. The two WSs showed little difference in values of the global metrics, even under different thresholding. The main difference in the two WSs is on the ICC coefficients.

The global metrics of the network were compared between different scans and different subjects to address the issue of test-retest reliability and between-subject variability for DTI-based network construction. As expected, intra-subject variability was lower than inter-subject differences, although there was remarkable consistency throughout the sample. For both weighting schemes, the coefficients of variation were all below 10% for most global metrics. It is important to minimize the intra-subject variability so that measures of intersubject variability are more accurate. Despite a number of differences in the network construction, such as parcellation scheme and fiber tracking seed density, the ICC values with WS 2 are comparable to the results from Bassett et al. (2010). However, the ICCs seem to be affected by the weighting schemes. The ICC analysis demonstrated that most ICCs are

higher than 0.5 for WS 1, indicating that the variance of the networks from two different scans is smaller than that from subject difference. The ICCs with WS 2 are mostly below 0.5, except for gam and swi. The different results between the two weighting schemes may be related to different weightings on shorter fibers that may differ significantly between subjects.

The local metrics show larger fluctuation across subjects. Analysis of nodal strength, betweenness centrality, path length, clustering coefficients shows that the mean CV values are much higher than those of global metrics. These data suggest that the specific aspects of the network structure may have a large regional variation from subject to subject.

In practice, the weighted network needs to be thresholded before network analysis, though thresholding may have an adverse effect on the measured global metrics and their variability. The global efficiency characterizes the integrity of the network. Because a connection with a very small weight (corresponding to a long path) has no advantage over shorter paths across several nodes, the global efficiency is very stable against thresholding (Fig. 6c) and so is its ICC (Fig. 10). Other metrics, such as clustering coefficient and maximized modularity, which are more or less related to segregation, can be influenced by thresholding. Our results show that, besides global efficiency, the intra-class correlation coefficients under thresholding showed small variation for clustering coefficient, maximized modularity and diversity. The ICCs for small-worldness index and path length are more sensitive to thresholding.

There are several sources that contribute to the variation of the edge weights. Noise in the DTI data, spatial resolution and partial volume effects may affect the quality of fiber quantification. The tractography algorithm, including number of random seeds in fiber tracking, can also have some effect on the variance of network. Specifically, fewer random seeds will lead to larger variance in the number of fibers from fiber tracking, although the effect in this study was diminished by choosing 30 seeds per voxel in fiber tracking (Cheng et al., 2011). In addition, the reliability of network construction relies on the accuracy of parcellation and the mapping to DTI image space. The parcellation can have errors due to SNR limitations of the T1-weighted image or the algorithm itself. The registration between T1-weighted image and DTI can also have some errors due to image distortion and partial volume effects. All these factors affect the test–retest reliability of structural networks. The ICC analysis can be useful for experiment design and data processing. For instance, low ICC values indicate a need for more subjects in practical studies to obtain sufficient statistical power, more sophisticated tractography methods, or a need to improve the quality of data acquisition. Despite many findings of alteration of some network metrics in diseases and neuropsychological patients, the moderate values of ICC indicate that structural network constructed from DTI is still facing some technical challenge before being used as a reliable biomarker.

In summary, weighted networks were constructed for healthy adults based on DTI streamline fiber tracking and FreeSurfer parcellation. The backbone network was extracted and its variability across subjects was investigated. Our results show that weighted network is very distinctive from its binary version in a variety of topological properties. For weighted networks, considerable intra-subject and inter-subject variance was observed on the edge weights, causing variances in network metrics at different levels. However, the intraclass correlation analysis indicates a reasonably high reproducibility for several global metrics.

Acknowledgments

The authors would like to thank Dr. Olaf Sporns for his support in data analysis and insightful discussions. This study is funded in part by Indiana CTSI grant. We also thank Siemens Healthcare for providing work-in-progress diffusion sequence for our DTI data acquisition.

References

- Bassett DS, Brown JA, Deshpande V, Carlson JM, Grafton ST. Conserved and variable architecture of human white matter connectivity. *NeuroImage*. 2010
- Braun U, Plichta 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. Test-retest reliability of resting-state connectivity network characteristics using fMRI and graph theoretical measures. *NeuroImage*. 2011
- Cheng H, Wang Y, Sheng J, Sporns O, Kronenberger WG, Mathews VP, Hummar TA, Saykin AJ. Optimization of seed density in DTI tractography for structural networks. *J. Neurosci. Methods*. 2011; 203:264–272. [PubMed: 21978486]
- Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*. 2006; 31:968–980. [PubMed: 16530430]
- Fischl B, van der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM. Automatically parcellating the human cerebral cortex. *Cereb. Cortex*. 2004; 14:11–22. [PubMed: 14654453]
- Frank LR. Anisotropy in high angular resolution diffusion-weighted MRI. *Magn. Reson. Med*. 2001; 45:935–939. [PubMed: 11378869]
- Gigandet, X.; Kober, T.; Hagmann, P.; Cammoun, L.; Meuli, R.; Thiran, J-P.; Krueger, G. ISMRM 18th Annual Meeting. Stockholm; Sweden: 2010. A connectome-based comparison of diffusion MR acquisition schemes; p. P1067
- Gong G, He Y, Concha L, Lebel C, Gross DW, Evans AC, Beaulieu C. Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cereb. Cortex*. 2009; 19:524–536. [PubMed: 18567609]
- Hagmann P, Kuran M, Gigandet X, Thiran P, Wedeen VJ, Meuli R, Thiran J-P. Mapping human whole-brain structural networks with diffusion MRI. *PLoS Biol*. 2007:e597.
- Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O. Mapping the structural core of human cerebral cortex. *PLoS Biol*. 2008; 6:e164. [PubMed: 20076716]
- Hagmann P, Sporns O, Madan N, Cammoun L, Pienaar R, Wedeen VJ, Meuli R, Thiran JP, Grant PE. White matter maturation reshapes structural connectivity in the late developing human brain. *Proc. Natl. Acad. SciUSA*. 2010; 107:19067–19072.
- Honey CJ, Sporns O, Cammoun L, Gigandet X, Thiran JP, Meuli R, Hagmann P. Predicting human resting-state functional connectivity from structural connectivity. *Proc. Natl. Acad. SciUSA*. 2009; 106:2035–2040.
- Jbabdi S, Johansen-Berg H. Tractography: where do we go from here? *Brain Connect*. 2011; 1:169–182. [PubMed: 22433046]
- Lo CY, Wang PN, Chou KH, Wang J, He Y, Lin CP. Diffusion tensor tractography reveals abnormal topological organization in structural cortical networks in Alzheimer's disease. *J. Neurosci*. 2010; 30:16876–16885. [PubMed: 21159959]
- Mori S, van Zijl PCM. Fiber tracking: principles and strategies — a technical review. *NMR Biomed*. 2002; 15:465–480.
- Mori S, Crain B, Chacko V, van Zijl P. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann. Neurol*. 1999; 45:265–269. [PubMed: 9989633]
- NEMA. MS 6-2008. Determination of signal-to-noise ratio and image uniformity for single-channel non-volume coils in diagnostic MR imaging. 2008
- Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*. 2010; 52:1059–1069. [PubMed: 19819337]

- Sporns, O. *Network of the Brain*. Cambridge, Massachusetts: 2010.
- Tuch DS. Q-Ball imaging. *Magn. Reson. Med.* 2004; 52:1358–1372. [PubMed: 15562495]
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage.* 2002; 15:273–289. [PubMed: 11771995]
- van den Heuvel MP, Mandl RC, Kahn RS, Hulshoff Pol HE. Functionally linked resting state networks reflect the underlying structural connectivity architecture of the human brain. *Hum. Brain Mapp.* 2009; 30:3127–3141. [PubMed: 19235882]
- van den Heuvel MP, Mandl RC, Stam CJ, Kahn RS, Hulshoff Pol HE. Aberrant frontal and temporal complex network structure in schizophrenia: a graph theoretical analysis. *J. Neurosci.* 2010; 30:15915–15926. [PubMed: 21106830]
- Wang J-H, Zuo X-N, Gohel S, Milham MP, Biswal BB, He Y. Graph theoretical analysis of functional brain networks: test-retest evaluation on short- and long-term resting-state functional MRI data. *PLoS One.* 2011; 6:e21976. [PubMed: 21818285]
- Wedeen VJ, Hagmann P, Tseng WY, Reese TG, Weisskoff RM. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magn. Reson. Med.* 2005; 54:1377–1386. [PubMed: 16247738]
- Wedeen VJ, Wang RP, Schmahmann JD, Benner T, Tseng WY, Dai G, Pandya DN, Hagmann P, D'Arceuil H, de Crespigny AJ. Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *NeuroImage.* 2008; 41:1267–1277. [PubMed: 18495497]
- Zalesky A, Fornito A, Harding IH, Cocchi L, Yonelinas M, Pantelis C, Bullmore ET. Whole-brain anatomical networks: does the choice of nodes matter? *Neuro-Image.* 2010; 50:970–983. [PubMed: 20035887]
- Zuo XN, Ehmke R, Mennes M, Imperati D, Castellanos FX, Sporns O, Milham MP. Network centrality in the human functional connectome. *Cereb. Cortex.* 2011

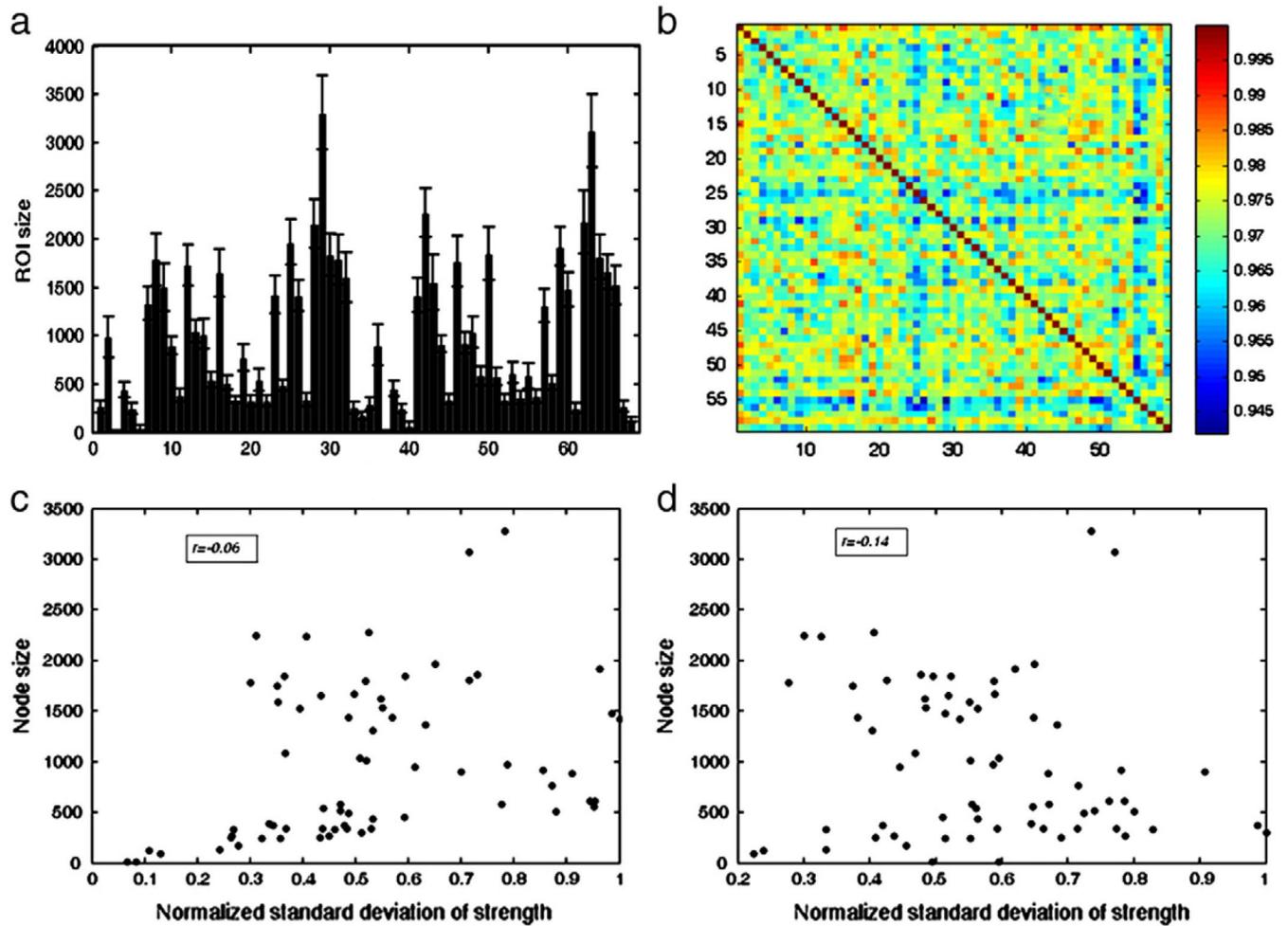


Fig. 1.

Top panel is the inter-subject variability for 59 subjects showing mean and standard deviation of the ROI size of the nodes (a) and correlation matrix of the ROI sizes between subjects (b); Bottom panel is the bias test of the weighting scheme showing the scatter plots of standard deviation of nodal strength and nodal ROI size, with weighting scheme 1 (c) and weighting scheme 2 (d). The cross correlation coefficient between nodal strength and node size is -0.06 in (c) and -0.14 in (d).

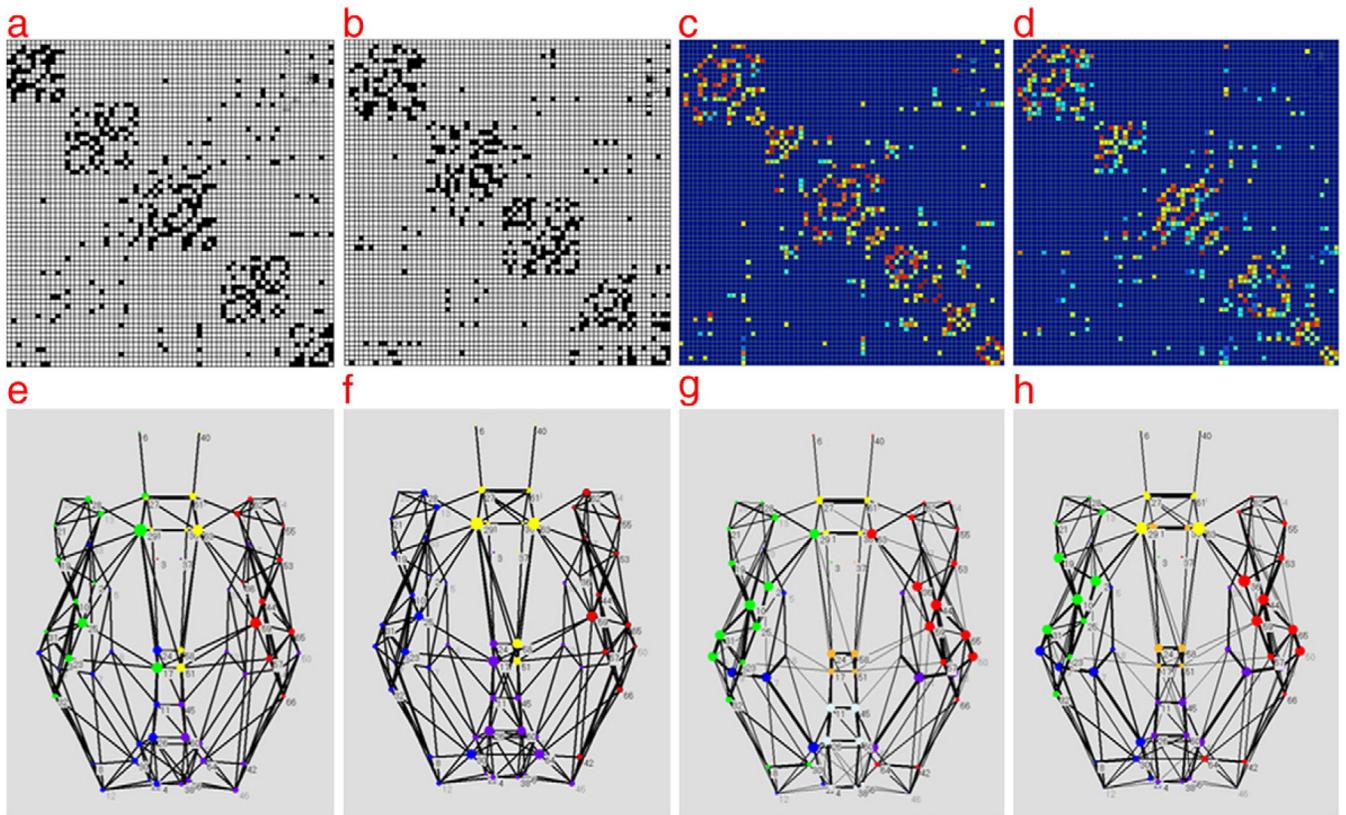


Fig. 2. Matrix and graph representations of the backbone networks. Top row: network in matrix forms for a) BBN1; b) BBN2; c) WBN1; d) WBN2; corresponding graph forms are shown in the bottom. In matrix representations, the nodes and edges are clustered into different modules, which are distinguished by different colors in graph representations. The node positions in the graph is an approximate 2D mapping of the physical location of the ROIs in the brain, left corresponds to the left hemisphere, right corresponds to the right hemisphere, up is front and bottom is posterior.

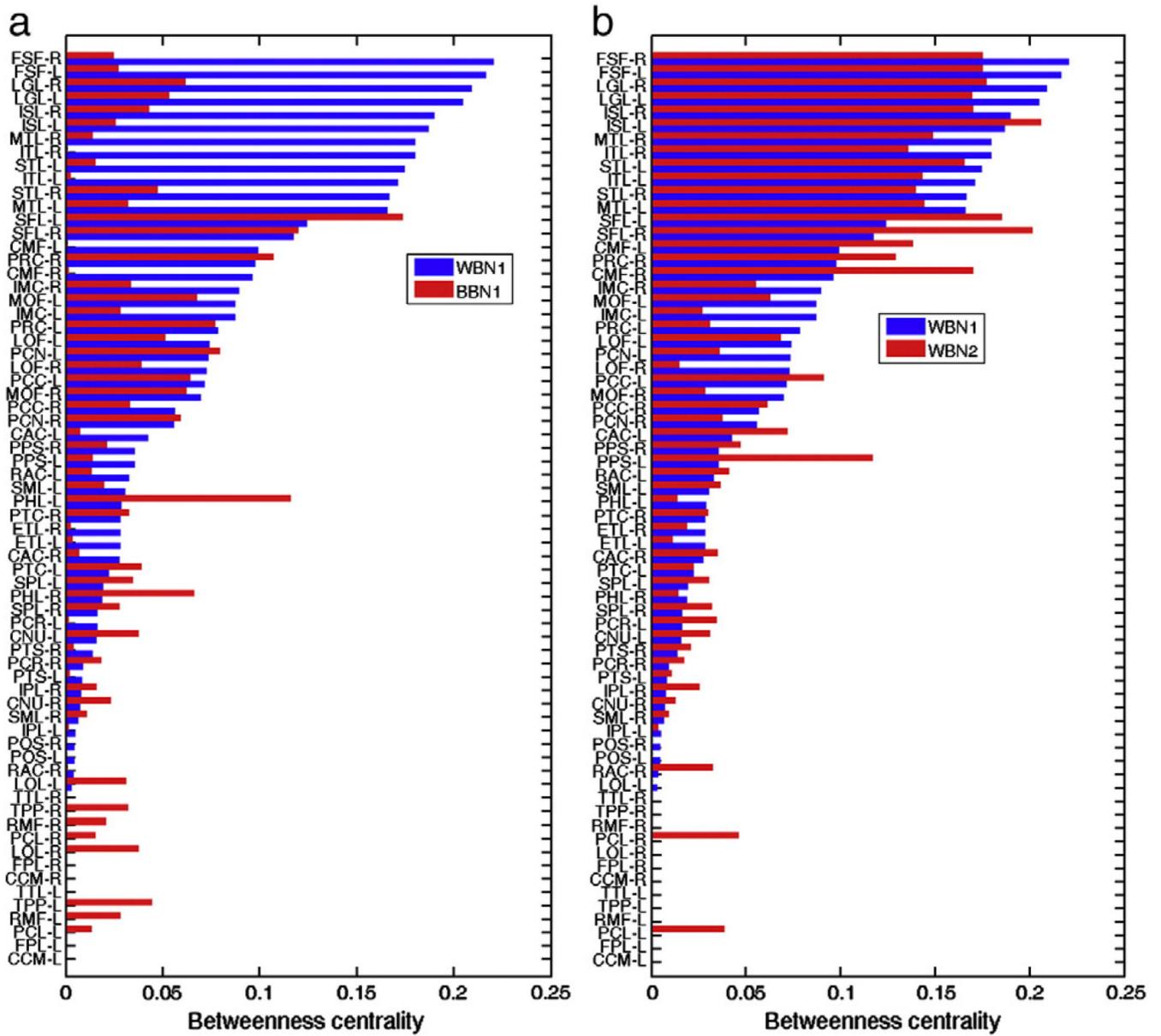


Fig. 4. Comparison of node betweenness centrality between WBN1 and BBN1 (a) and WBN1 and WBN2 (b). The order of nodes is based on the order of betweenness centrality in WBN1.

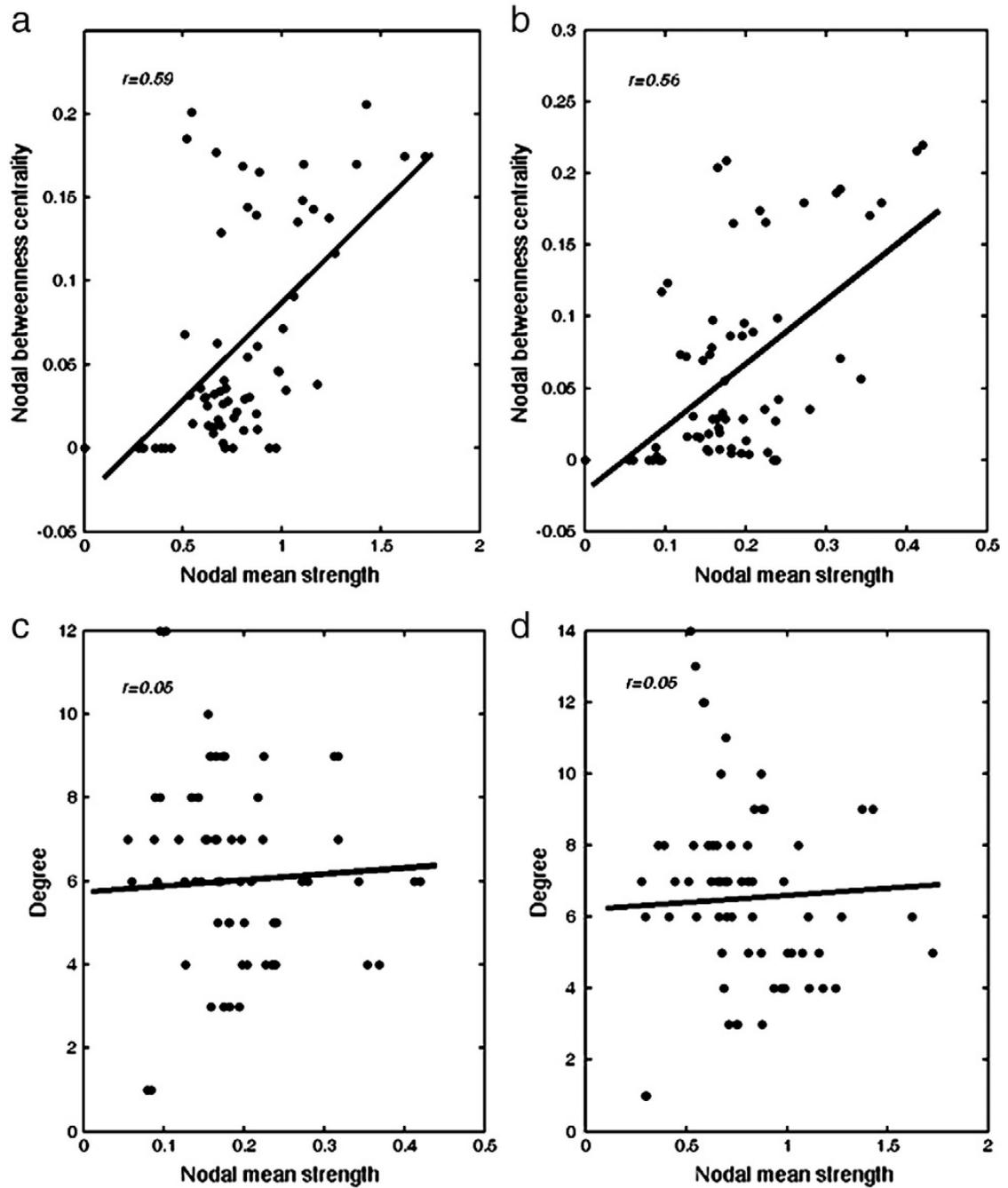


Fig. 5. Scatter plots of the nodal mean strength (nodal strength divided by nodal degree) and nodal betweenness centrality (a, b); nodal mean strengths and nodal degree (c, d) for WBN1 and WBN2.

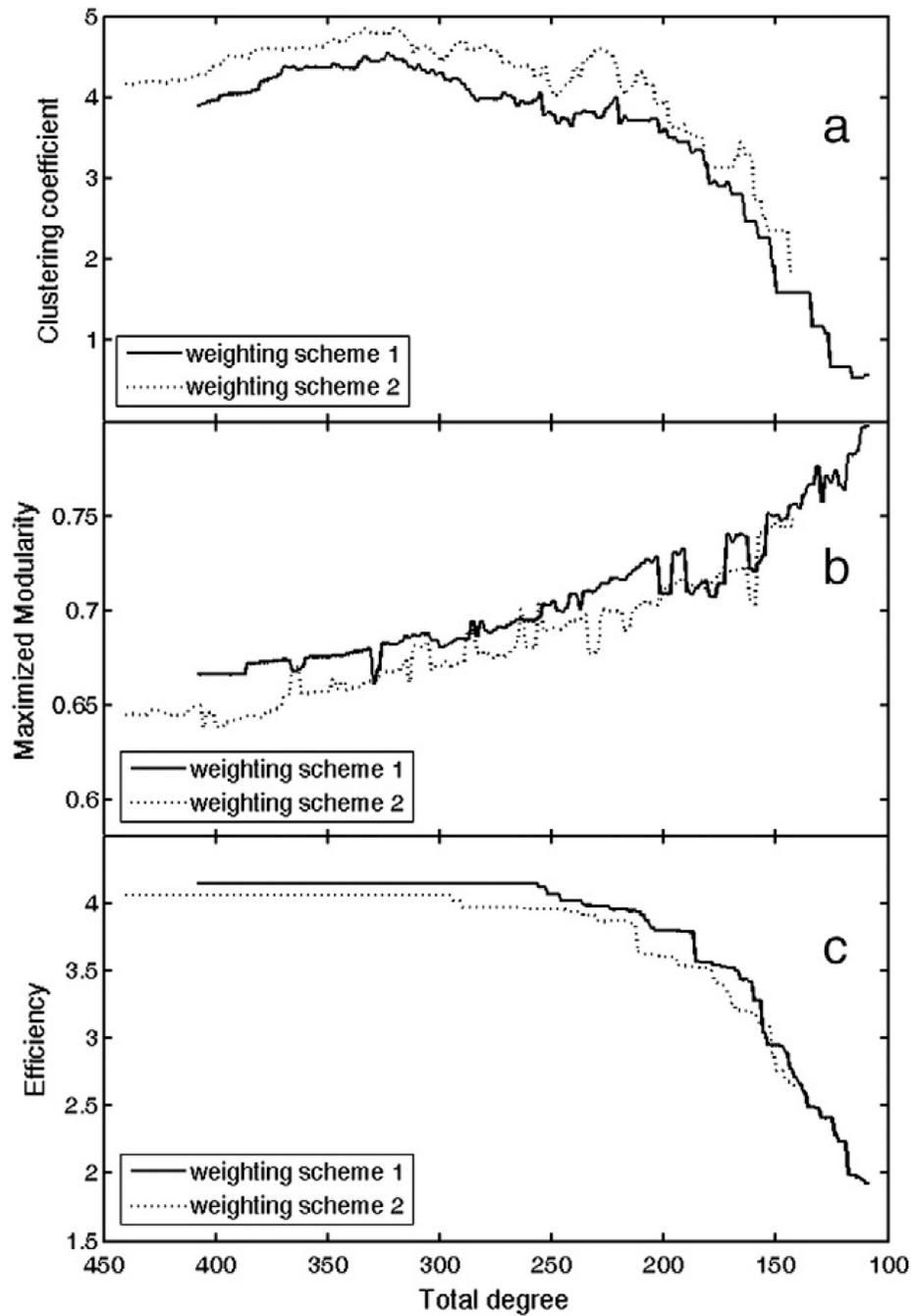


Fig. 6. Thresholding effects for three global metrics: clustering coefficient (a), maximized modularity (b), and efficiency (c), plotted as functions of total degree after thresholding.

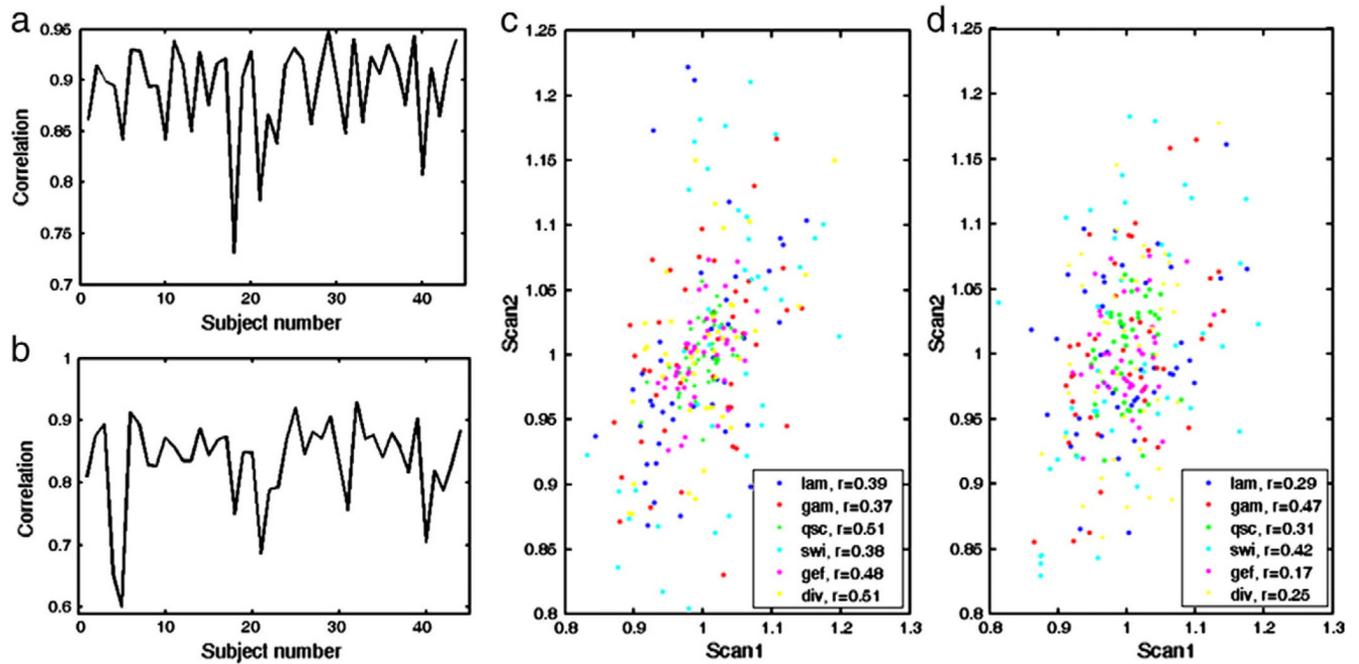


Fig. 7. Intra-subject variability characterized by correlation of network edge weights between two scans (a, b), and scatter plots of global metrics between two scans (c, d) for both weighting schemes. The correlation coefficients of the global metrics across subjects are listed in (c) and (d).

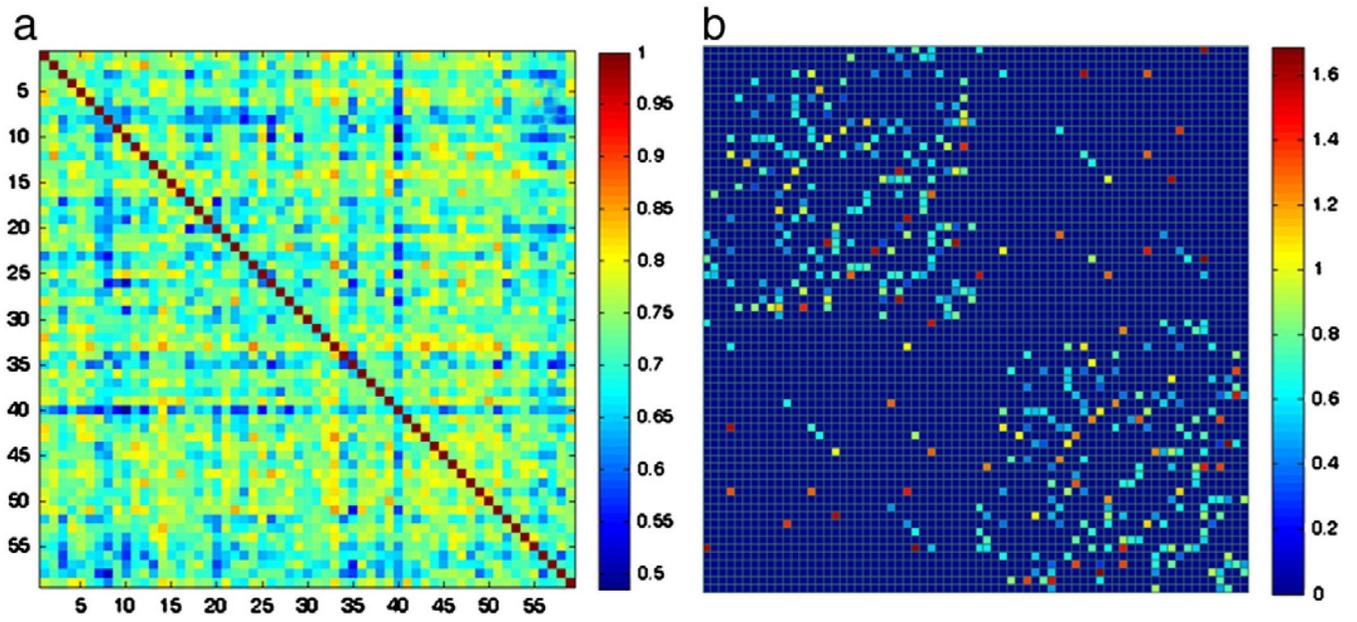


Fig. 8. Inter-subject variability for WS 1 characterized by correlation of network edge weights between subjects (a); and coefficient of variation of the edge weights (b).

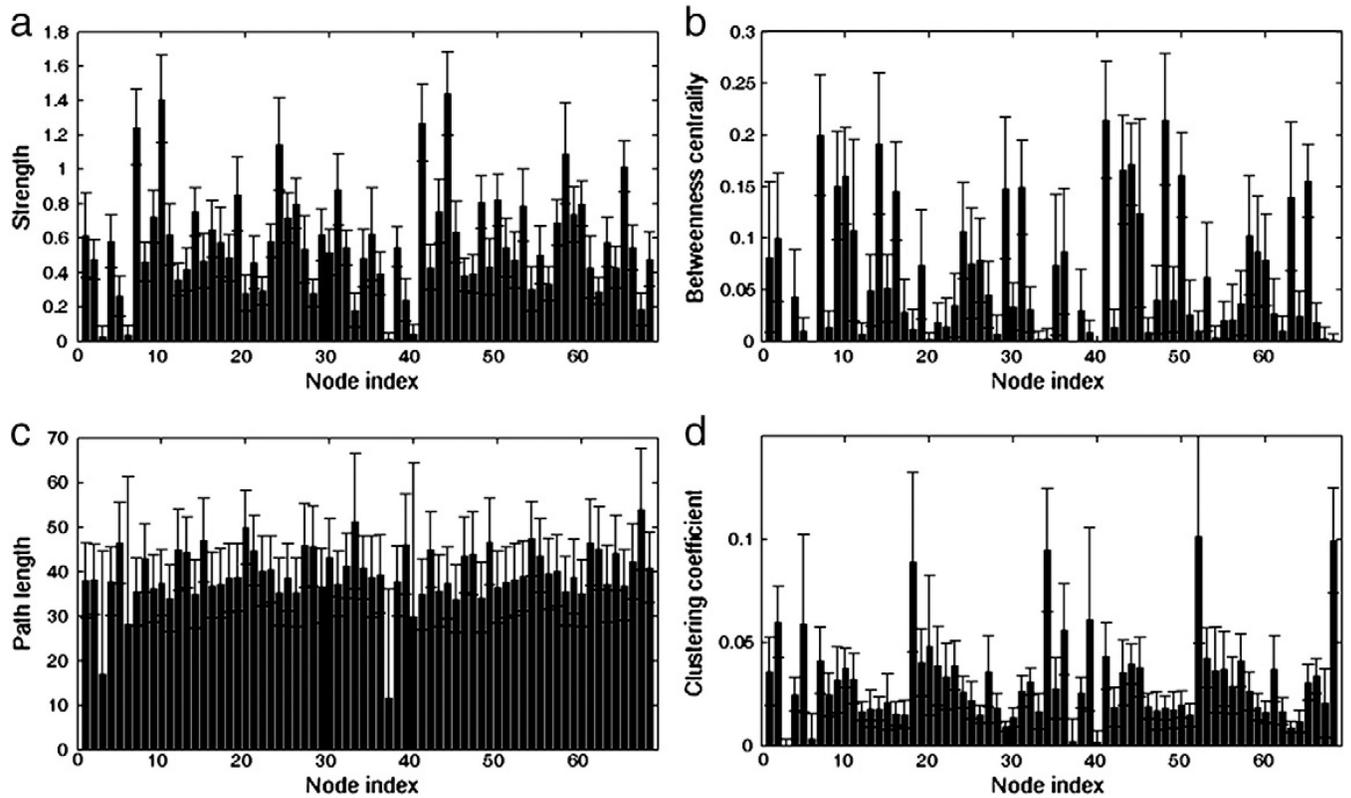


Fig. 9. Mean values and standard deviation of local network metrics for each node across 59 subjects with WS 1: (a) strength; (b) betweenness centrality; (c) path length; (d) clustering coefficient.

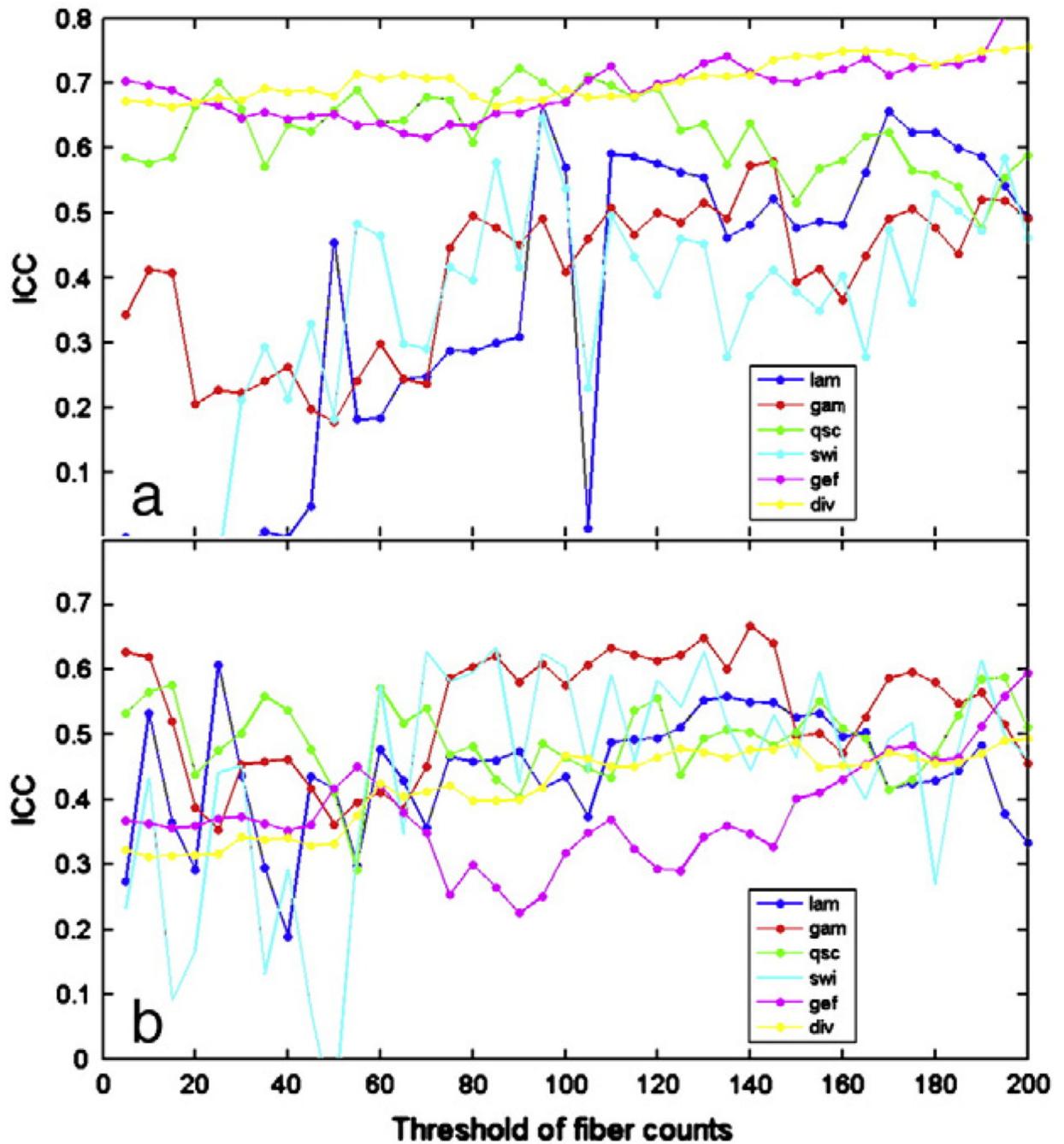


Fig. 10. Plot of the Intra-class correlation coefficient for different network metrics as a function of threshold of fiber counts in constructing the weighted network with weighting scheme 1 (A) and weighting scheme 2 (B).

Table 1

Names of the parcellated cortex regions from FreeSurfer and their abbreviations. Nodes 1–34 correspond to those regions in left hemisphere, nodes 35–68 (in parenthesis) correspond to the counterpart in right hemisphere.

Node	Abbreviation	Parcellated cortical labels
1	(35)	CAC Caudalanteriorcingulate
2	(36)	CMF Caudalmiddlefrontal
3	(37)	CCM Corpuscallosum
4	(38)	CNU Cuneus
5	(39)	ETL Entorhinal
6	(40)	FPL Frontalpole
7	(41)	FSF Fusiform
8	(42)	IPL Inferiorparietal
9	(43)	ITL Inferiortemporal
10	(44)	ISL Insula
11	(45)	IMC Isthmuscingulate
12	(46)	LOL Lateraloccipital
13	(47)	LOF Lateralorbitofrontal
14	(48)	LGL Lingual
15	(49)	MOF Medialorbitofrontal
16	(50)	MTL Middletemporal
17	(51)	PHL Parahippocampal
18	(52)	PCL Paracentral
19	(53)	PPS Parsopercularis
20	(54)	POS Parsorbitalis
21	(55)	PTS Parstriangularis
22	(56)	PCR Pericalcarine
23	(57)	PTC Postcentral
24	(58)	PCC Posteriorcingulate
25	(59)	PRC Precentral
26	(60)	PCN Precuneus
27	(61)	RAC Rostralanteriorcingulate
28	(62)	RMF Rostralmiddlefrontal
29	(63)	SFL Superiorfrontal
30	(64)	SPL Superiorparietal
31	(65)	STL Superiortemporal
32	(66)	SML Supramarginal
33	(67)	TPP Temporalpole
34	(68)	TTL Transversetemporal

Table 2

Various global network metrics for backbone networks. Weighted networks with weighting schemes 1 and 2 are denoted as WBN1 and WBN2 respectively; their binary counterparts are denoted as BBN1 and BBN2. The global metrics are: deg (total degree), lam (characteristic path length), gam (mean clustering coefficient), msc (optimal number of modules), qsc (maximized modularity), swi (small-worldness index), ass (assortativity coefficient), gef (global efficiency), and div (diversity).

	deg	lam	gam	msc	qsc	swi	ass	gef	div
BBN1	408	0.287	5.14	5	0.537	4.38	-.0043	4.09	0.404
BBN2	440	0.296	5.11	4	0.552	3.86	.0418	3.93	0.427
WBN1	408	0.335	3.91	7	0.666	4.44	-0.0053	4.15	0.537
WBN2	440	0.341	4.15	6	0.645	3.93	.131	4.05	0.501

Table 3

Various global network metrics for 59 subjects from scan 1 with both weighting schemes. The global metrics are: deg (total degree), lam (characteristic path length), gam (mean clustering coefficient), msc (optimal number of modules), qsc (maximized modularity), swi (small-worldness index), ass (assortativity coefficient), gef (global efficiency), and div (diversity).

	WS 1		WS 2	
	Mean	CV	Mean	CV
deg	414	4.36%	414	4.35%
lam	0.334	8.25%	0.336	6.69%
gam	3.58	6.89%	3.89	6.61%
msc	7.54	21.9%	7.25	19.3%
qsc	0.672	2.85%	0.656	3.13%
swi	3.52	13.55%	3.37	9.04%
ass	-0.0143	-5.56	0.0807	1.02
gef	4.06	3.78%	4.06	4.15%
div	0.599	6.28%	0.556	5.45%

Table 4

ICC coefficients for various network metrics with both weighting schemes. The global metrics are: lam (characteristic path length), gam (mean clustering coefficient), qsc (maximized modularity), swi (small-worldness index), gef (global efficiency), and div (diversity).

Metrics	WS 1		WS 2	
	ICC coefficient	P-value	ICC coefficient	P-value
lam	0.281	0.14473	0.445	0.0298
gam	0.544	0.00569	0.641	0.000529
qsc	0.673	0.000191	0.474	0.0190
swi	0.550	0.00506	0.591	0.0207
gef	0.644	0.000485	0.296	0.127
div	0.675	0.000176	0.390	0.0545