

# SCIENTIFIC REPORTS



OPEN

## Reproducibility of structural brain connectivity and network metrics using probabilistic diffusion tractography

Shang-Yueh Tsai <sup>1,2</sup>

The structural connectivity network constructed using probabilistic diffusion tractography can be characterized by the network metrics. In this study, short-term test-retest reproducibility of structural networks and network metrics were evaluated on 30 subjects in terms of within- and between-subject coefficient of variance ( $CV_{ws}$ ,  $CV_{bs}$ ), and intra class coefficient (ICC) using various connectivity thresholds. The short-term reproducibility under various connectivity thresholds were also investigated when subject groups have same or different sparsity. In summary, connectivity threshold of 0.01 can exclude around 80% of the edges with  $CV_{ws} = 73.2 \pm 37.7\%$ ,  $CV_{bs} = 119.3 \pm 44.0\%$  and  $ICC = 0.62 \pm 0.19$ . The rest 20% edges have  $CV_{ws} < 45\%$ ,  $CV_{bs} < 90\%$ ,  $ICC = 0.75 \pm 0.12$ . The presence of 1% difference in the sparsity can cause additional within-subject variations on network metrics. In conclusion, applying connectivity thresholds on structural network to exclude spurious connections for the network analysis should be considered as necessities. Our findings suggest that a connectivity threshold over 0.01 can be applied without significant effect on the short-term when network metrics are evaluated at the same sparsity in subject group. When the sparsity is not the same, the procedure of integration over various connectivity thresholds can provide reliable estimation of network metrics.

Diffusion Tensor Imaging (DTI) and the associate tractography methods can be used to reconstruct fiber bundles based on voxel based directional information<sup>1-6</sup>. Based on the tractography, the connectivity matrix also known as structural network can be constructed from structural connectivity calculated by counting the number of fibers between cortical regions in the brain<sup>7-10</sup>. Further, graph theory-based network analysis has been applied on the connectivity matrix to investigate the topology properties of entire network instead of individual analysis of large number of tracts<sup>11-14</sup>. The network metrics such as global efficiency, cluster coefficients have been found to be relevant in age and gender of healthy human brain<sup>15,16</sup>. The change in network metrics have already been found in several neurological and psychological disorders<sup>17-21</sup>.

The network metrics can be estimated in various ways with several intermediate steps. The reproducibility of network metrics can be therefore affected by many factors. For example: the number of directions of DTI acquisition, structural connectivity estimated from deterministic or probabilistic tractography, the structural network in binary or weighted, the definition of weighting of weighted network, the number of nodes in the structural network, different kind of thresholds applied on the connectivity. Because potential sources of bias and errors in each step can induce the variability of the outcome measures, a full characterization of the reproducibility of structural network and of network metrics is therefore essential for the applications<sup>22-25</sup>. In summary, Vaessen and associates<sup>25</sup> reported the within subject coefficient of variation (CV) of network metrics for three kinds of DTI scheme (6, 15, 32 directions) for binary structural network. Owen and colleagues<sup>24</sup> reported the CV and intraclass correlation coefficient (ICC) for intra- and inter-site to address the issue on the number of nodes for both binary and weighted network. Buchanan and colleagues<sup>26</sup> compared the repeatability of network metrics from weighted network using two tractography algorithms (deterministic and probabilistic), two seeding approaches (white matter and gray matter), and three definitions of network weightings. They showed that probabilistic tractography has higher ICC on network metrics. Bonilha and colleagues<sup>27</sup> also showed that probabilistic tractography has

<sup>1</sup>Graduate Institute of Applied Physics, National Chengchi University, Taipei, Taiwan. <sup>2</sup>Research Center for Mind, Brain and Learning, National Chengchi University, Taipei, Taiwan. Correspondence and requests for materials should be addressed to S.-Y.T. (email: [sytsai@nccu.edu.tw](mailto:sytsai@nccu.edu.tw))

higher ICC on network metrics. Andreotti and colleagues<sup>23</sup> reported the CV and ICC for global and local network metrics and addressed the issues on the effect of density thresholds. In their report, the weighted network is constructed using a connectivity threshold to exclude edges having low connectivity and using a density threshold to maintain the total number of edges (sparsity) across subjects, which is considered important in the comparison of network metrics between subjects<sup>28,29</sup>. They showed that applying the density and connectivity thresholds has significant effect on the network metrics and improves the reliability of network metrics (ICC and CV). Although there is argument that connectivity thresholds should not be applied or it will eliminate the real network property of individuals<sup>24,26</sup>, the strategy of connectivity thresholds have been adopted in studying age and gender differences of network metrics<sup>15</sup>, and in studying patients with attention-deficit/hyperactivity disorder<sup>21</sup>. In these studies, connectivity thresholds were applied on whole subject groups to exclude edges. In this way, the structural network has same sparsity and identical position of the connections for all subjects.

The probabilistic approach is considered more effective in the calculation of connectivity between cortical regions compared to the deterministic approach<sup>27,30,31</sup>. However, little is known about the reproducibility of the connectivity matrix constructed from probabilistic tractography and how the connectivity thresholds affect the reproducibility of network metrics. We think the issues about what connectivity thresholds should be applied can be addressed by investigating the reproducibility of structural connectivity at different connectivity levels. Then, the reproducibility of network metrics at various connectivity thresholds can be reported. Because the sparsity may differ when connectivity thresholds are applied, it is important understand how the reproducibility of the network metrics alter when the sparsity is not the same. In addition, the strategy of cost normalization is used in the calculation of network metrics to maintain the sum of all weighting at the same level<sup>15,16,21</sup>. It is interested to investigate the reproducibility of the network metrics before and after cost normalization.

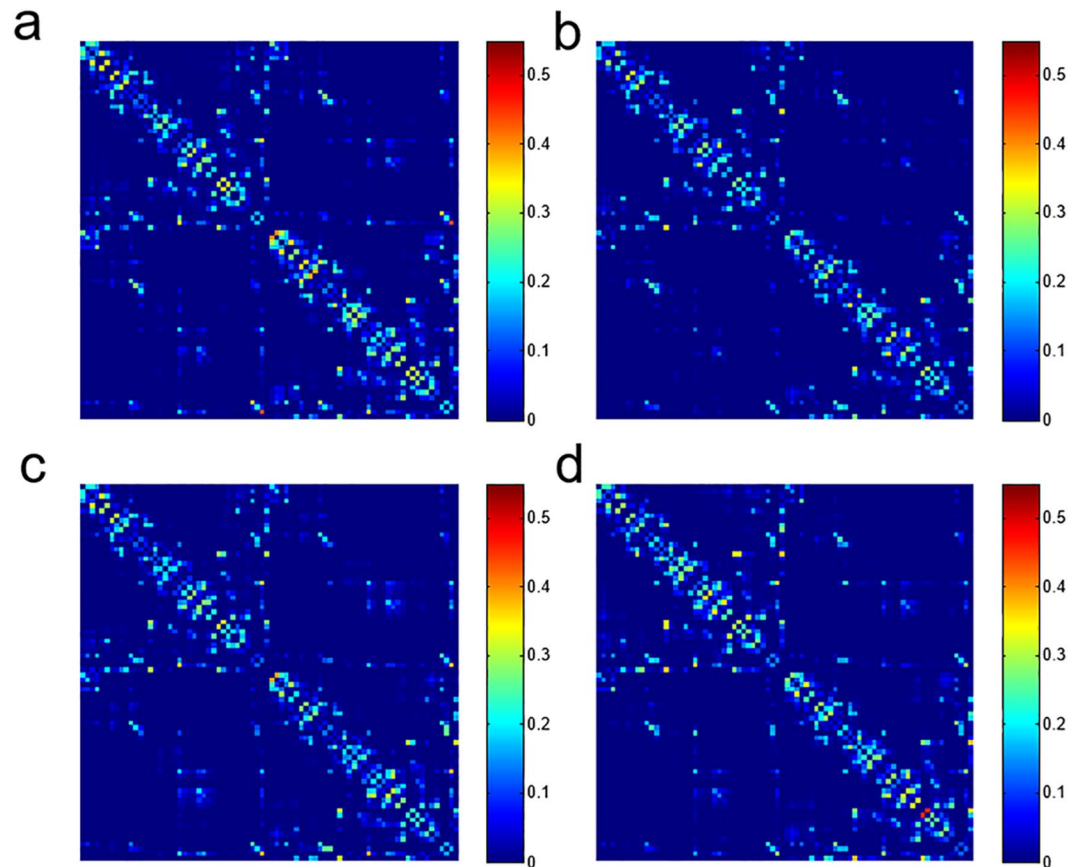
In this study, we constructed the structural connectivity matrix among cortical regions using probabilistic tractography from repeated scans on healthy subjects. We calculated the network metrics based on weighted structural network. Three specific issues are investigated. (1) We investigate the short-term reproducibility of structural connectivity matrix at various levels of connectivity. (2) We investigate the short-term reproducibility of network metrics at various connectivity thresholds. (3) We investigate the effects of three processing strategies on the short-term reproducibility of network metrics including integration over all connectivity thresholds, maintenance of sparsity and cost normalization.

## Results

**Structural connectivity matrix.** Figure 1 shows the structural connectivity matrixes of 78 cortical regions from 4 subjects. Edges with stronger connectivity are mostly along the diagonal line and aside in the left-upper and right-lower part of the matrix. These edges belong to link with shorter inter-region distance and ROI sizes are larger in these regions. The within subject similarity of the connectivity of all subjects are 0.98 for and between subject similarity is 0.95. The distribution of structural connectivity matrix is similar to those in the previous reports using probabilistic tractography<sup>15,21,27</sup>. The averaged structural connectivity for 30 subjects with corresponding matrix of  $CV_{bs}$ ,  $CV_{ws}$  and ICC are shown in Fig. 2. The edges with higher connectivity have lower  $CV_{bs}$  and  $CV_{ws}$ . Further,  $CV_{ws}$  are in general lower than  $CV_{bs}$ . The histogram of the number of edges versus connectivity (connectivity > 0.01) and corresponding  $CV_{bs}$ ,  $CV_{ws}$  and ICC are shown in Fig. 3. Overall, both  $CV_{bs}$  and  $CV_{ws}$  decrease as increasing connectivity. The  $CV_{bs}$  are all higher than  $CV_{ws}$ . The ICC of the edges with connectivity over 0.01 is  $0.75 \pm 0.12$  and the mean of ICC in each group range from 0.66 to 0.83. No observable trend in ICC versus connectivity is found but the variations of  $CV_{ws}$  and  $CV_{bs}$  decrease for edges with connectivity over 0.1. For all 3003 edges, there are 2428 (80.85%) of the edges having connectivity less than 0.01. A significant increase of the  $CV_{bs}$  ( $118 \pm 44.5\%$ ) and  $CV_{ws}$  ( $73.2 \pm 37.7\%$ ) is found for edges with connectivity less than 0.01 and the ICC of this group is  $0.62 \pm 0.19$ .

**Network metrics.** The sparsity for connectivity thresholds applied together and separate are shown in Fig. 4. The sparsity decrease from 0.32 to 0.08 as connectivity thresholds increasing from 0.01 to 0.1. The sparsity is the same for  $N1_{com}$  and  $N2_{com}$ . A slight difference is found in the sparsity between  $N1_{sep}$  and  $N2_{sep}$  (Fig. 4b). The sparsity of  $N1_{sep}$  and  $N2_{sep}$  also differ from  $N1_{com}$  and  $N2_{com}$  in the range of  $-4.7 \times 10^{-3}$  to  $2.3 \times 10^{-3}$  (Fig. 4c). Whether the thresholds are applied together and separate, the spatial distribution of edges on the connectivity matrix is similar at the same threshold (Fig. 4d).

The network metrics ( $E_{glob}$ ,  $E_{loc}$ ,  $C_p$ ,  $L_p$ ) estimated with and without cost normalization are shown in Fig. 5. Quantitative values are different for the network metrics estimated with and without cost normalization. But the difference among  $N1_{com}$ ,  $N2_{com}$ ,  $N1_{sep}$ ,  $N2_{sep}$  for each network metric is small relative to the difference caused by connectivity thresholds by visual inspection. The  $CV_{bs}$  and  $CV_{ws}$  of network metrics are shown in Fig. 6. When the sparsity is the same ( $N1_{com}$  and  $N2_{com}$ ), the  $CV_{bs-com}$  range from 6% to 11% and the  $CV_{ws-com}$  range from 4% to 7% for network metrics without normalization (Fig. 6a). For network metrics with cost normalization, the  $CV_{bs-com}$  are in the range of 1% to 3.5% and the  $CV_{ws-com}$  are in the range of 0.4% to 1.8% (Fig. 6b). In general, both the  $CV_{ws-com}$  and  $CV_{bs-com}$  decreases smoothly as increasing connectivity thresholds (lower sparsity) and the  $CV_{bs-com}$  are larger than the  $CV_{ws-com}$ . When the sparsity is not the same ( $N1_{sep}$  and  $N2_{sep}$ ), there is less than 0.1% difference between  $CV_{bs-sep}$  and  $CV_{bs-com}$  for all connectivity thresholds and for all network metrics estimated with or without cost normalization. Large variation is found between  $CV_{ws-sep}$  to  $CV_{ws-com}$  in  $C_w$  and  $E_{loc}$  with cost normalization at several connectivity thresholds (up to 3%) (Fig. 6b). The ICC of network metrics are shown in Fig. 7. The ICC of network metrics without cost normalization from  $N_{sep}$  and  $N_{com}$  are consistent at the range of 0.5 to 0.6 for all connectivity thresholds (Fig. 7a,b). The ICC of network metrics with cost normalization are at range of 0.67 to 0.85 for  $N_{com}$  (Fig. 7c). However, the ICC vary in the range of 0.1 to 0.8 through the connectivity thresholds in  $E_{loc}$  and  $C_w$  with cost normalization for  $N_{sep}$  (Fig. 7d).



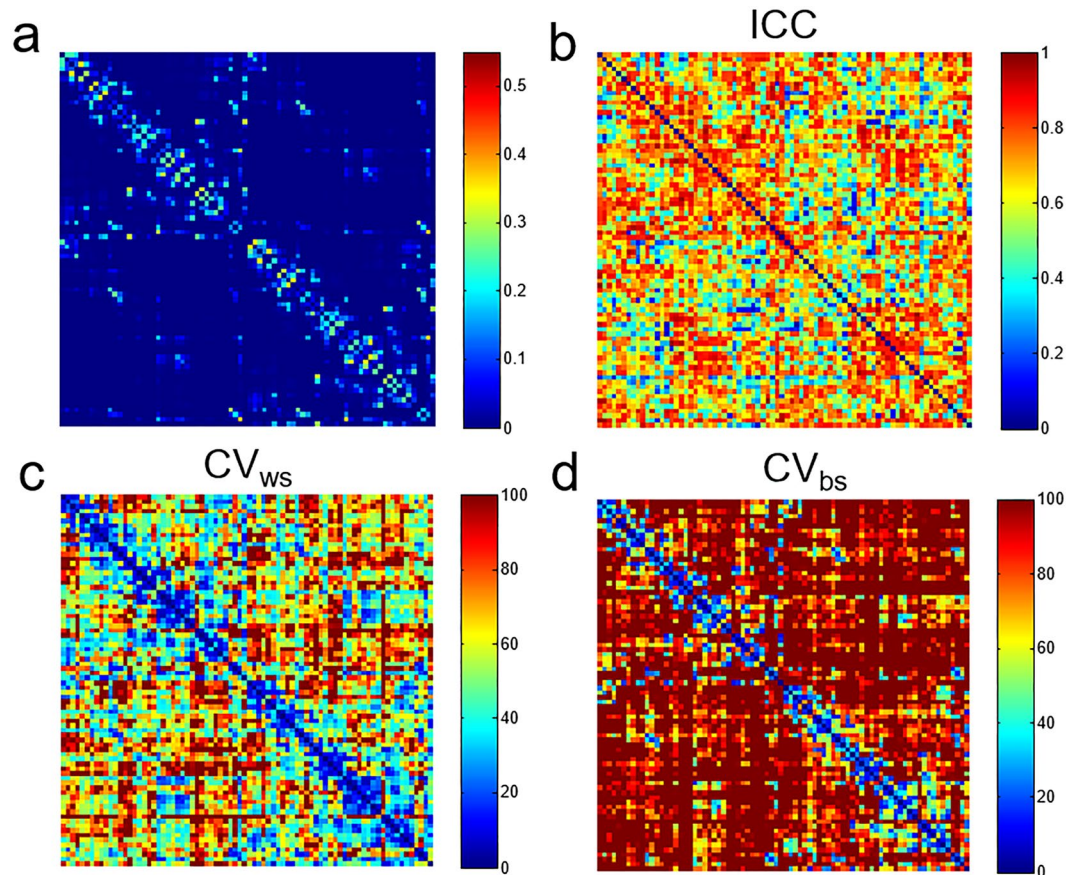
**Figure 1.** Structural connectivity matrices of four subjects. Each matrix element (edge) represents the weighted connectivity (0 to 1) from probabilistic tractography between cortical regions by AAL template. The 78 cortical regions are in left to right hemisphere order and names of cortical regions are listed in Supplementary Table 1.

The integrated network metrics are summarized in Table 1. For the network metrics without cost normalization, the  $CV_{bs}$  are 6% to 8% and the  $CV_{ws}$  are 4% to 5.5%. The ICCs are in the level of fair to good (0.53 to 0.60) for  $N_{sep}$  and  $N_{com}$ . There is only minor difference of  $CV_{ws}$  between  $N_{sep}$  and  $N_{com}$  found on  $C_w$  (0.13%) and  $E_{loc}$  (0.07%). The network metrics with cost normalization show lower  $CV_{bs}$  (0.9% to 1.64%) and  $CV_{ws}$  (0.7% to 1.05%) compared to those without cost normalization. The ICCs increase but are still in the fair to good level (0.69 to 0.76). A considerable difference of  $CV_{ws}$  between  $N_{sep}$  and  $N_{com}$  is found in  $C_w$  (0.3%) and  $E_{loc}$  (0.15%). That lowers the ICC to the level of 0.5 and 0.63 for  $C_w$  and  $E_{loc}$  in  $N_{sep}$ .

## Discussion

In this study, the short-term reproducibility of structural connectivity matrix among 78 cortical regions estimated using probabilistic tractography was reported based on the connectivity. The edges with higher connectivity have less between-subject variations (lower  $CV_{bs}$ ) and better reproducibility (lower  $CV_{ws}$ ) (Figs 2 and 3). The connectivity threshold of 0.01 excluded around 80% of the edges and these edges show poor to moderate reproducibility ( $CV_{ws} = 73.2 \pm 37.7\%$ ;  $CV_{bs} = 119.3 \pm 44.0\%$ ;  $ICC = 0.62 \pm 0.19$ ) compared to the other 20% edges (connectivity  $> 0.01$ ) showing moderate to good reproducibility ( $CV_{ws} < 45\%$ ;  $CV_{bs} < 90\%$ ;  $ICC = 0.75 \pm 0.12$ ). A connectivity threshold at 0.01 can be considered as good choice to enroll edges for network analysis. The short-term reproducibility of network metrics ( $C_w$ ,  $L_w$ ,  $E_{loc}$ ,  $E_{glo}$ ) show smooth change for the connectivity thresholds between 0.01 to 0.1 at the same sparsity (Fig. 6). We found that the network metrics estimated without cost normalization have moderate reproducibility ( $CV_{ws} < 5.5\%$ ,  $CV_{ws} < 8.0\%$ ,  $ICC = 0.5\text{--}0.6$ ) (Table 1) and is less sensitive to the difference in the sparsity. The procedure of cost normalization can improve the reproducibility ( $ICC = 0.69$  to  $0.76$ ) in associate with reduced within and between subject variations ( $CV_{ws} < 1.1\%$  and  $CV_{bs} < 1.7\%$ ). The presence of 1% difference in the sparsity (Fig. 4b,c) can cause additional within subject variations on the estimation of  $C_w$  and  $E_{loc}$  with cost normalization (Fig. 6b), which leads to lower ICC (Fig. 7d). The ICC is still in the moderate level when the networks metrics are estimated over various connectivity thresholds (Table 1).

**Short-term reproducibility of connectivity matrix.** For connectivity matrixes, the edges with the higher connectivity mostly coming from the links with shorter tracking distance have better reproducibility (lower  $CV_{ws}$ ) and less inter-subject variation (lower  $CV_{bs}$ ). One reason is that tracking for longer distance is more likely to be interrupted by the successive tracking process, which yield lower connectivity and larger variations. The findings are

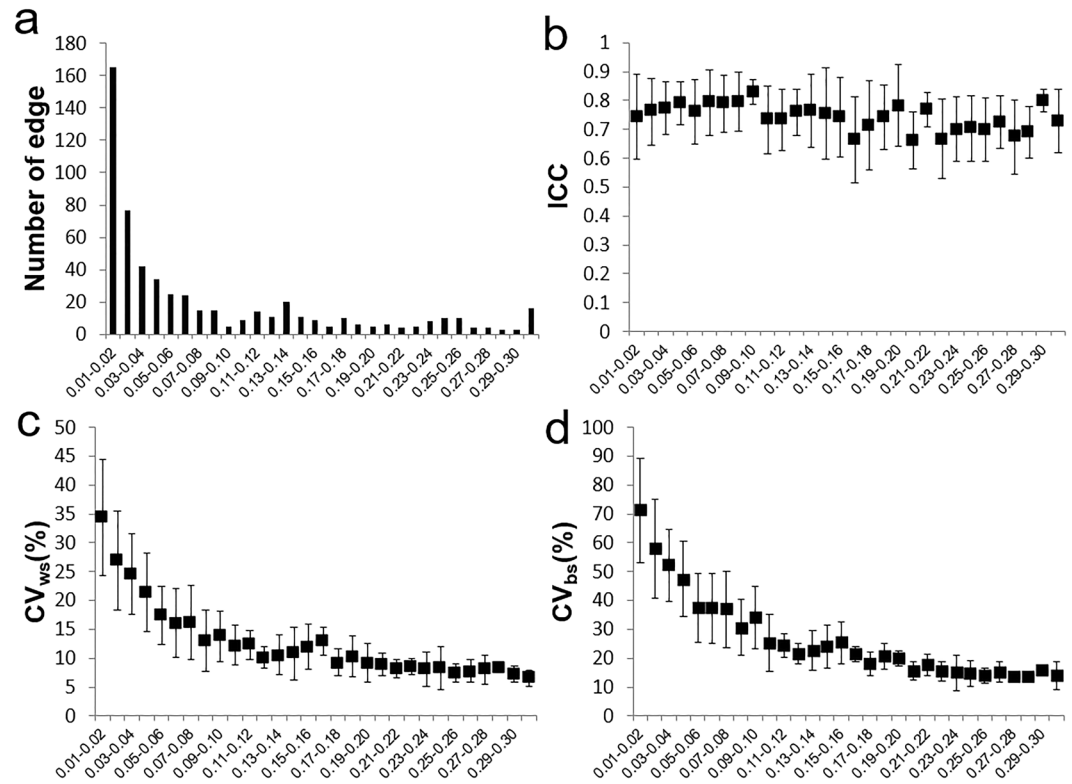


**Figure 2.** (a) Averaged connectivity matrix from 30 subjects and matrices of (b) ICC, (c)  $CV_{ws}$  (%), (d)  $CV_{bs}$  (%). Note edges with higher connectivity have lower  $CV_{ws}$  and  $CV_{bs}$ .

in agreement with previous reports<sup>27</sup>. The outcome also indicates that tracking using probabilistic model between cortical regions are still subject to the intricate fiber pathways in the brain especially when the seeding points are mostly belonging to gray matter where the fiber has uncertain direction. Connectivity thresholds is therefore necessary to exclude the edges with low connectivity, which can possibly lead to unreliable parameters calculated based on the connectivity matrix. The information of the reproducibility based on the level of connectivity provided in this study can be useful in setting the connectivity thresholds to exclude edges depending on the sensitivity needed for successive analysis. For example, we found that the variations of  $CV_{ws}$  and  $CV_{bs}$  drop at the connectivity threshold of 0.1. For the study directly using inter-region connectivity for statistical analysis, we suggest to use the connectivity threshold at 0.1, where  $CV_{ws} < 15\%$  and  $CV_{bs} < 32\%$  can be reached for 173 edges (5.76%).

**Short-term reproducibility of network metrics.** For network metrics estimated at the same sparsity ( $N1_{com}$  and  $N2_{com}$ ), both  $CV_{ws-com}$  and  $CV_{bs-com}$  decrease smoothly in comparable trend as increasing connectivity thresholds (Fig. 6), so the ICC are at same level for all connectivity thresholds and for all network metrics (Fig. 7a,c). The finding implies that the choice of connectivity thresholds in the range of 0.01 to 0.1 may not be crucial for the short-term reproducibility of the estimated network metrics at the same sparsity. Around 1% difference in the sparsity is found between  $N1_{sep}$  and  $N2_{sep}$  (Fig. 4b,c). The difference in sparsity here is from the difference in connectivity for edges in the shorter-term repetition. Our findings show that the difference in the sparsity leads to relatively larger variation on the  $CV_{ws-sep}$  and ICC on  $E_{loc}$  and  $C_w$  with cost normalization (Figs 6b and 7d). Even though, the  $CV_{ws-sep}$  of the integrated  $E_{loc}$  and  $C_w$  are still less than  $CV_{bs-sep}$  and give moderate reproducibility (ICC > 0.5). Therefore, integration over various connectivity thresholds for the estimation of  $E_{loc}$  and  $C_w$  with cost normalization is suggested when the connectivity thresholds are applied without maintaining sparsity.

**Comparison with the previous literature.** Among all studies reporting the reproducibility of network metrics, we compared our results to those using probabilistic tractography and weighted network. In Buchanan *et al.*<sup>26</sup>, they have reported the test-retest reliability of network metrics using three kinds of definition in network weighting. The  $CV_{ws}$  is 3.6% to 4.8% for  $L_w$  and 5.3% to 7.2% for  $C_w$ . The  $CV_{bs}$  is 4.5% to 7.1% for  $L_w$  and is 8.5% to 9.6% for  $C_w$ . The ICC of  $L_w$  and  $C_w$  are in the moderate-to-good range (0.59 ~ 0.76). Andreotti *et al.*<sup>23</sup> have reported the  $CV_{ws}$  for network metrics calculated by averaging over a range of density thresholds. The definition of network weighting is similar to our study but the range of sparsity differs from our study. According to their results, the  $CV_{ws}$  of  $L_w$ ,  $E_{loc}$ ,  $E_{glo}$  is at the level of 4% and  $CV_{ws}$  of  $C_w$  is around 8%. The ICC is the moderate-to-good



**Figure 3.** (a) Histogram of number of edges versus connectivity groups (0.01 to 0.30 with 0.01 steps) and corresponding (b) ICC, (c)  $CV_{ws}$  (%), (d)  $CV_{bs}$  (%) of 30 groups. For connectivity less than 0.01, there are 2428 (80.85%) out of total 3003 edges with the  $CV_{bs}$  ( $118 \pm 44.5\%$ ),  $CV_{ws}$  ( $73.2 \pm 37.7\%$ ) and ICC ( $0.62 \pm 0.19$ ).

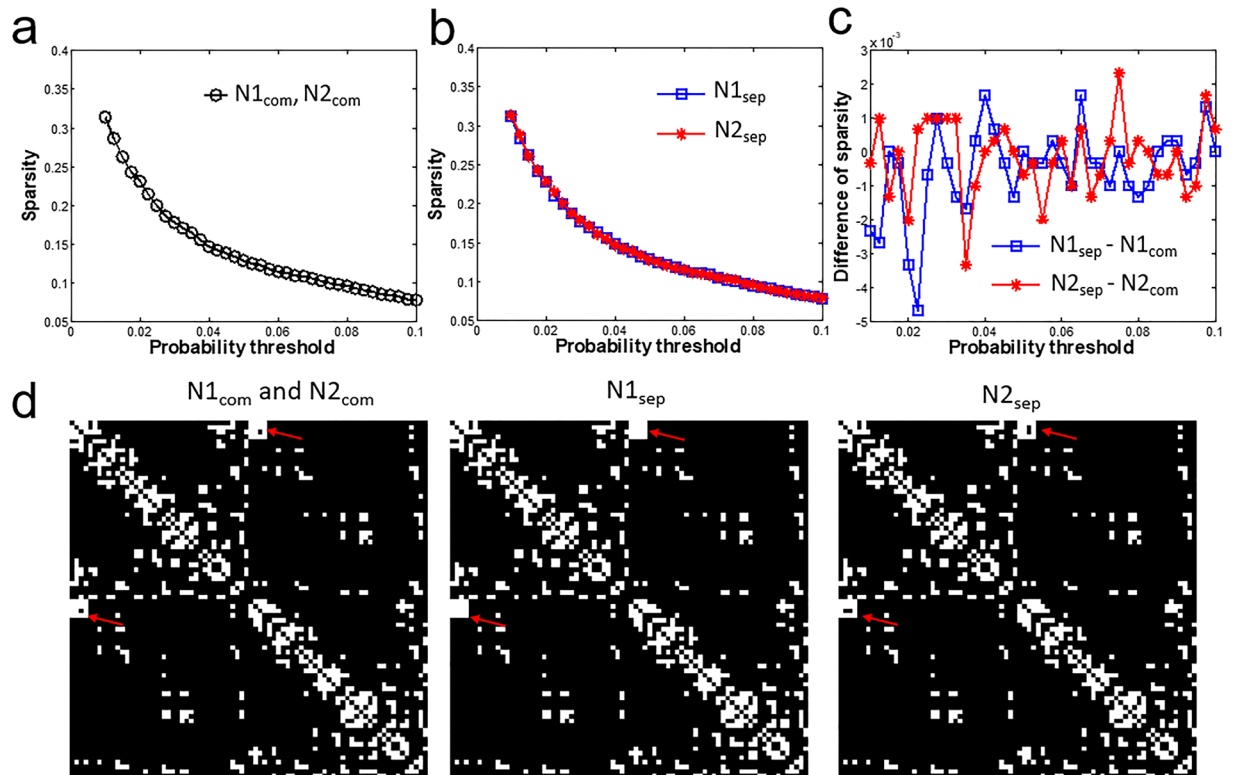
range (0.66 ~ 0.89). Despite there is difference in network processing procedures between our study and these two previous reports, the  $CV_{ws}$  of network metrics without cost normalization are close to their reports (Table 1). Both of these two previous reports found higher  $CV_{ws}$  for  $C_w$  than other network metrics. Same trend is found in our study. The ICC reported in Andreotti *et al.*<sup>23</sup>, Buchanan *et al.*<sup>26</sup> and our study are in different levels where our study show lowest ICC among the three studies. The possible explanation is the difference in the subject groups. This leads to difference in the between subject variations. One should notice that the between subject variation ( $CV_{bs}$ ) in the subject sample of this study is expectedly much lower than the general healthy groups because the study population is very homogeneous (healthy young subjects at comparable age and educational level)<sup>23,25,26</sup>. The smaller within and between subject variation helps in inferring the methodological difference in the construction of structural network and in the estimation of network metrics.

**Limitation.** In this study, probabilistic tractography is preferred based on previous reports<sup>26,27,31</sup>. Even our analyses were not able to directly compare the deterministic and probabilistic tractography, there are still other works reporting acceptable repeatability for network metrics from deterministic tractography<sup>28,32,33</sup>. Here, the weighting of the network is defined according the most commonly used weighting scheme in the literatures<sup>21,24,33</sup>. However, the weighting may include effect of both relevant and irrelevant diffusion characteristics of WM tracks. It is therefore difficult to correctly interpret the weight of network. Question on how the definition of the weighting influence the reliability and which weight is more suitable to represent the connectivity are not addressed in this study. The test-retest repeatability of network metrics was investigated in a relative short time interval (~ 30 min.). The subjects stay in the scanner for consecutive measurement makes the  $CV_{ws}$  smaller than other studies where repeated scans were made in different days<sup>24,27</sup>, it is also important to understand the long-term or cross site repeatability<sup>24,27</sup>.

In conclusion, we suggest to apply the connectivity thresholds to exclude spurious connections for network analysis. When the sparsity is the same, a connectivity threshold over 0.01 can serve as an acceptable choice without significant effect on the short-term reproducibility of network metrics. When the sparsity is not the same for subject group, the procedure of integration over various connectivity thresholds can be considered to give reliable estimation of network metrics.

## Methods

**Data acquisition.** Experiments were conducted in 30 healthy volunteers (15 females/15 males; mean age  $\pm$  standard deviation:  $22.03 \pm 1.82$  years; age range: 20–26 years). Before being included in the study, all participants gave their informed consent to the protocol, which was approved by the Research Ethic Committee of National Chengchi University. All experiments were performed in accordance with the approved guidelines. Data were collected on a 3T MR system (Skyra, SIEMENS Medical Solutions, Erlangen, Germany) with a 32-channel

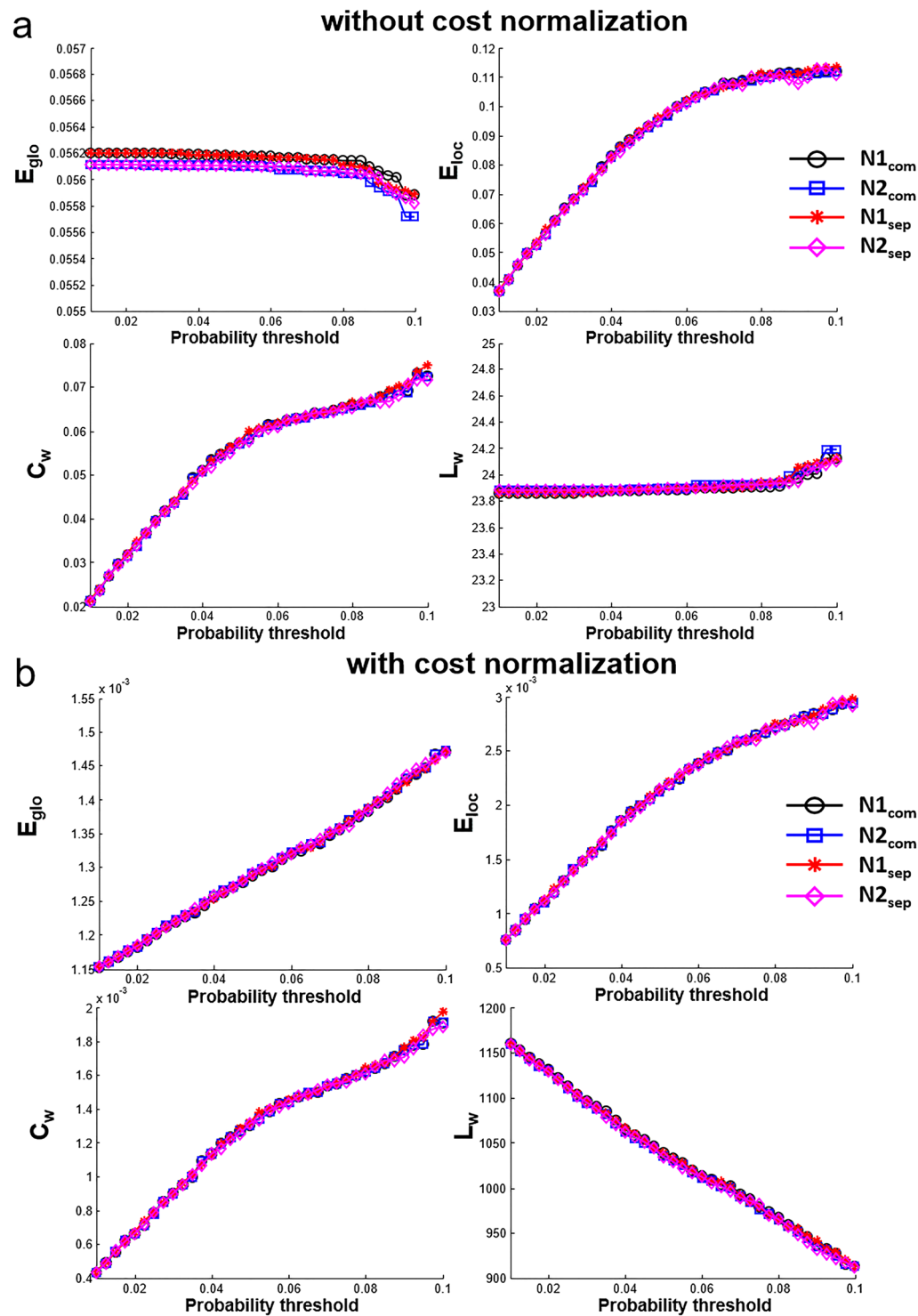


**Figure 4.** The sparsity at 37 connectivity thresholds from 0.01 to 0.10 with a step of 0.0025. The connectivity thresholds are applied on (a) N1 and N2 together (N1<sub>com</sub> and N2<sub>com</sub>) and (b) N1 and N2 separately (N1<sub>sep</sub> and N2<sub>sep</sub>). (c) Difference of N1<sub>sep</sub> to N1<sub>com</sub> and of N2<sub>sep</sub> and N2<sub>com</sub>. Note the sparsity decrease from 0.32 to 0.08 as increasing connectivity thresholds and the difference in sparsity is less than  $5 \times 10^{-3}$ . (d) masks of connectivity matrix at connectivity thresholds of 0.0575 where the sparsity is 0.1171 for N1<sub>com</sub>, N2<sub>com</sub>, 0.1182 for N1<sub>sep</sub>, and 0.1175 for N2<sub>sep</sub>. The differences on the edges are indicated by the red arrow.

head coil array. We acquired a high-resolution 3D MPRAGE (Magnetization Prepared Rapid Acquisition Gradient Echo) anatomical scan (TR/TE/flip angle: 2530 ms/3.03 ms/7 degrees; FOV:  $224 \times 256 \times 192 \text{ mm}^3$ ; voxel size:  $1 \times 1 \times 1 \text{ mm}^3$ ). DTI datasets were obtained using single shot spin echo EPI sequence. We used 30 diffusion directions with b-value  $1000 \text{ s/mm}^2$  and 5 additional images with b-value =  $0 \text{ s/mm}^2$ . Experiment parameters were TR = 8800 ms, TE = 90 ms, FOV =  $256 \times 256 \text{ mm}^2$ , MAT =  $128 \times 128$ , slice thickness = 2 mm, slice = 61, NEX = 4, acceleration factor = 3. The total acquisition time were 30 minutes. DTI protocols were repeated twice on each subject for the assessment of short-term test-retest reproducibility (N1 and N2). The subjects were asked to stay in the scanner and rest for 30 minutes between N1 and N2.

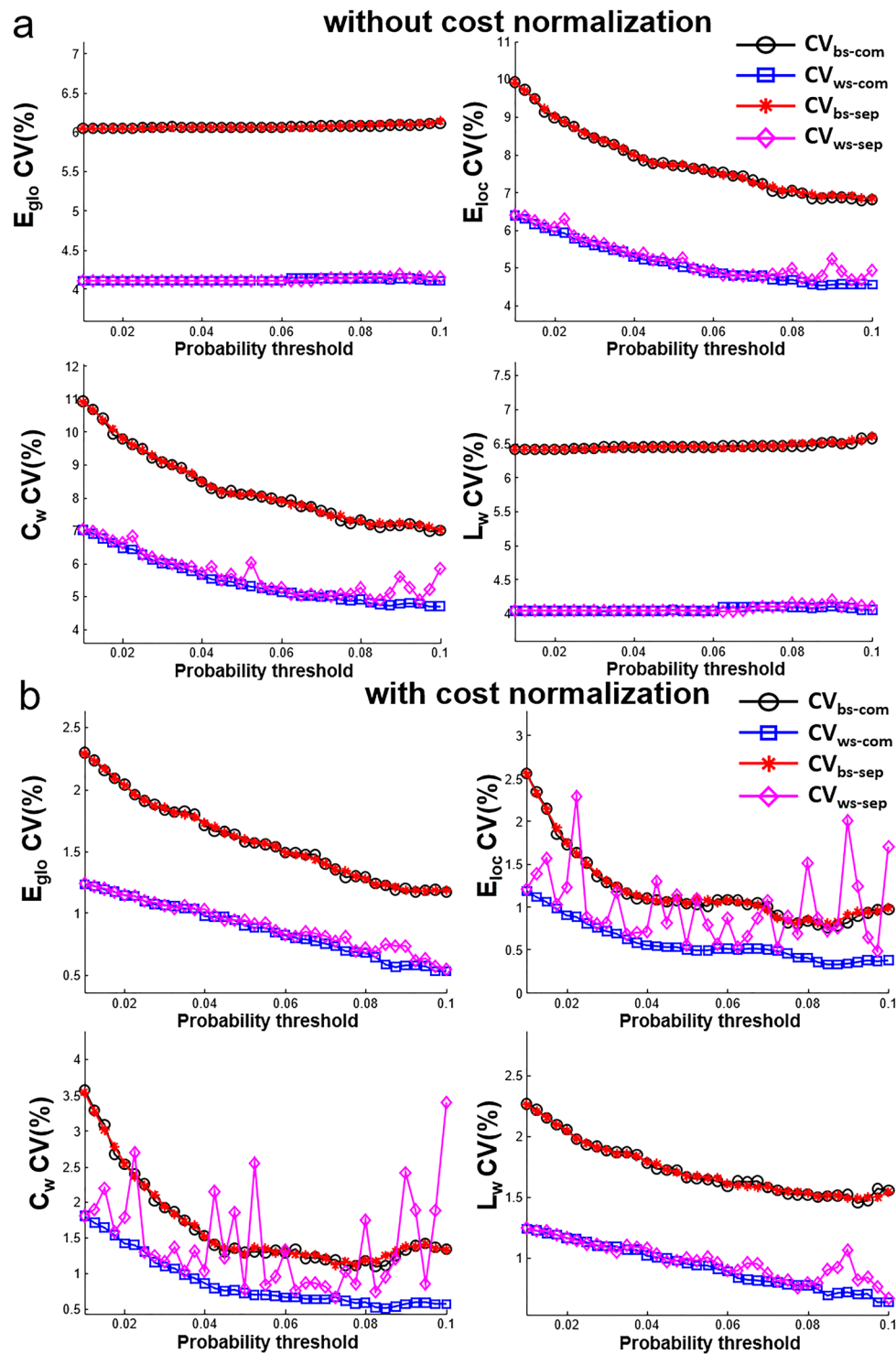
**DTI analysis and tractography.** Before DTI analysis, 78 cortical regions (39 for each hemisphere) were extracted using the automated anatomical labeling (AAL) template in standard MNI space<sup>34</sup>. Labels and names of cortical regions can be found in Supplementary Table S1. Each cortical region represents a network node. These AAL masks were then transformed to DTI native space for each individual using following procedures. First, T1 images were coregistered to non-diffusion weighting image of DTI data sets. The coregistered T1 image was transformed to standard MNI template using nonlinear transformation. The transformation matrix was then applied to warp the defined AAL mask to DTI native space for each subject. The procedures were carried out using FLIRT and FNIRT tool (FSL, version 4.1; <http://www.fmrib.ox.ac.uk/fsl>). The AAL masks in each subject were further refined by removing WM voxels that are not neighbor to GM voxels<sup>15</sup>.

For the analysis of DTI data sets, the procedures were performed using the FMRIB Software Library (FSL, version 4.1; FMRIB's Diffusion Toolbox [FDT]<sup>5</sup>; Oxford Centre for Functional MRI of the Brain [FMRIB], UK; <http://www.fmrib.ox.ac.uk/fsl>) as described in previous studies<sup>15,21,26</sup>. DICOM images were converted to Neuroimaging Informatics Technology Initiative (NIFTI) format using the MRICron tool. Images were visually checked for observable artifacts, and no volume was discarded. Eddy current correction was applied with the eddy\_correct tool, using the default settings. In summary, the first non-diffusion-weighted image was set as the target image, into which the remaining images (120 diffusion weighted image and 3 non-diffusion weighted images) were registered using an affine transformation to adjust for distortions caused by eddy currents and head motion. Then bet tool was used for skull stripping and bedpostx tool was used to build up a two-fibre per voxel model for fiber tracking. Probabilistic tractography was applied to estimate the connectivity probability among 78 cortical regions using PROBTRACKS tool. For the seed region, 5000 fibers streamlines grow from each voxel with tracking parameters of 0.5 mm step size, 500 mm maximum trace length,  $\pm 80^\circ$  curvature threshold. This



**Figure 5.** Network metrics ( $E_{\text{glo}}$ ,  $E_{\text{loc}}$ ,  $C_w$ ,  $L_w$ ) of  $N1_{\text{com}}$ ,  $N2_{\text{com}}$ ,  $N1_{\text{sep}}$ ,  $N2_{\text{sep}}$  at 37 connectivity thresholds estimated (a) without cost normalization and (b) with cost normalization. Note that all network metrics have consistent trend o as increasing connectivity thresholds and minor difference can be observed on the network metrics among  $N1_{\text{com}}$ ,  $N2_{\text{com}}$ ,  $N1_{\text{sep}}$ ,  $N2_{\text{sep}}$ .

yields  $5000 \times n$  streamlines from the seed region where  $n$  is the number of voxel in seed region<sup>5</sup>. The connectivity probability between seed region  $i$  and target region  $j$  was then calculated as the ratio of number of fibers passing through target region  $j$  to the total number of fibers from seed region  $i$ . The probability from  $i$  to  $j$  is not equal to that from  $j$  to  $i$  but they are correlated (all Pearson  $> 0.92$ ,  $p < 10^{-15}$ ). We then calculated the unidirectional connectivity probability between region  $i$  to region  $j$  ( $P_{ij}$ ) as the average of these two probabilities<sup>15,21</sup>. For each subject a  $78 \times 78$  symmetric connectivity matrix was estimated.



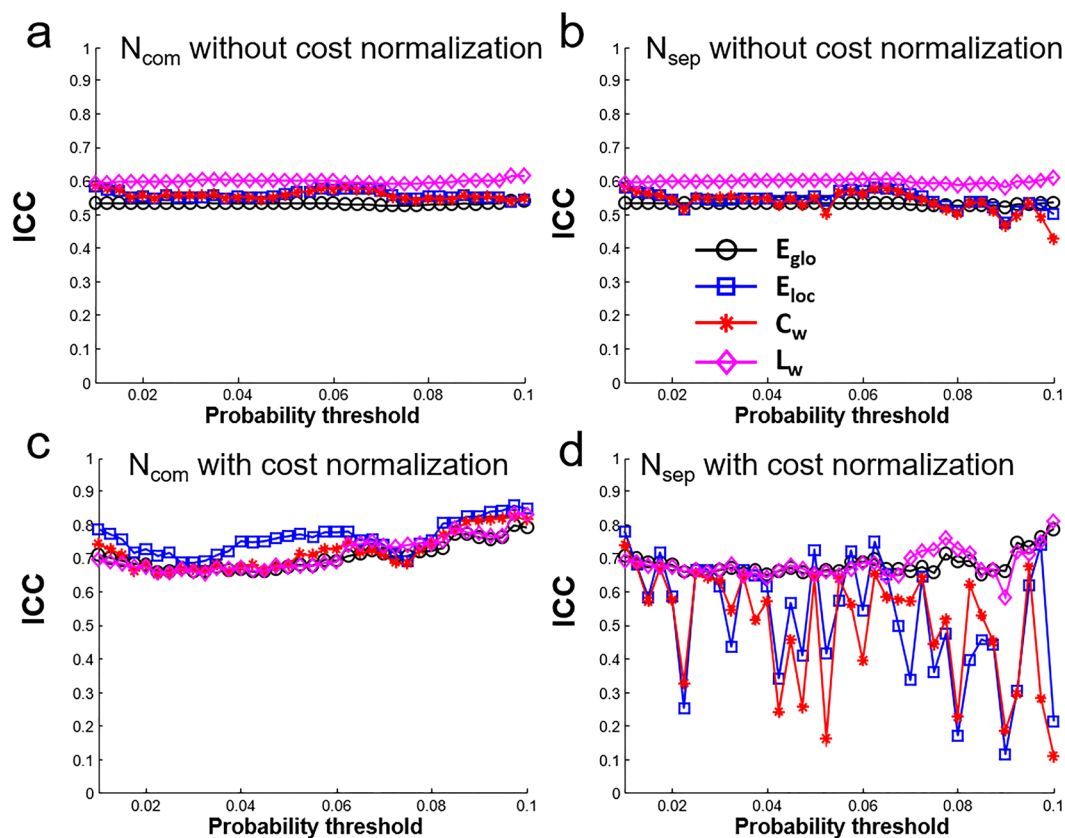
**Figure 6.**  $CV_{ws}$  and  $CV_{bs}$  of network metrics ( $E_{glo}$ ,  $E_{loc}$ ,  $C_w$ ,  $L_w$ ) at 37 connectivity thresholds estimated (a) without cost normalization and (b) with cost normalization. Both  $CV_{bs}$  and  $CV_{ws}$  vary smoothly as increasing connectivity thresholds. When sparsity is not the same ( $N1_{sep}$  and  $N2_{sep}$ ), dramatic variations of  $CV_{ws}$  are found in  $E_{loc}$ ,  $C_w$  estimated with cost normalization.

**Construction of weighted network.** The connectivity matrix has non-zero probabilities for all 3003 edges. Thresholds on connectivity were used to remove spurious connections that have small connectivity probabilities. Two cortical regions were considered unconnected and set to zero in the connectivity matrix, if the mean connectivity probability across subjects plus two times of standard deviation is less than the connectivity



without cost normalization		N1	N2	CV <sub>bs</sub> (%)	CV <sub>ws</sub> (%)	ICC
com	E <sub>glo</sub>	2.08 ± 0.13	2.07 ± 0.13	6.06	4.12	0.53
	E <sub>loc</sub>	3.28 ± 0.25	3.27 ± 0.25	7.54	5.00	0.55
	C <sub>w</sub>	2.00 ± 0.16	1.99 ± 0.16	7.92	5.26	0.55
	L <sub>w</sub>	884 ± 55.9	885.3 ± 58.2	6.45	4.06	0.60
sep	E <sub>glo</sub>	2.08 ± 0.13	2.07 ± 0.13	6.07	4.12	0.53
	E <sub>loc</sub>	3.29 ± 0.25	3.27 ± 0.25	7.56	5.07	0.54
	C <sub>w</sub>	2.01 ± 0.16	1.99 ± 0.16	7.93	5.39	0.54
	L <sub>w</sub>	884.8 ± 55.9	884.9 ± 58.2	6.45	4.06	0.60
com	E <sub>glo</sub>	$(4.82 \pm 0.08) \times 10^{-2}$	$(4.83 \pm 0.07) \times 10^{-2}$	1.49	0.82	0.69
	E <sub>loc</sub>	$(7.75 \pm 0.07) \times 10^{-2}$	$(7.75 \pm 0.07) \times 10^{-2}$	0.90	0.44	0.76
	C <sub>w</sub>	$(4.73 \pm 0.06) \times 10^{-2}$	$(4.72 \pm 0.06) \times 10^{-2}$	1.27	0.70	0.69
	L <sub>w</sub>	38159 ± 656	38076 ± 593	1.64	0.89	0.70
sep	E <sub>glo</sub>	$(4.82 \pm 0.08) \times 10^{-2}$	$(4.83 \pm 0.07) \times 10^{-2}$	1.49	0.83	0.68
	E <sub>loc</sub>	$(7.78 \pm 0.07) \times 10^{-2}$	$(7.74 \pm 0.07) \times 10^{-2}$	0.91	0.59	0.63
	C <sub>w</sub>	$(4.76 \pm 0.06) \times 10^{-2}$	$(4.71 \pm 0.06) \times 10^{-2}$	1.28	1.05	0.50
	L <sub>w</sub>	38152 ± 659	38052 ± 590	1.64	0.92	0.69

**Table 1.** Integrated network metrics (E<sub>glo</sub>, E<sub>loc</sub>, C<sub>w</sub>, L<sub>w</sub>) from N1<sub>com</sub>, N2<sub>com</sub>, N1<sub>sep</sub>, N2<sub>sep</sub> and the associate CV<sub>bs</sub>, CV<sub>ws</sub>, ICC.



**Figure 7.** ICC of network metrics (E<sub>glo</sub>, E<sub>loc</sub>, C<sub>w</sub>, L<sub>w</sub>) at 37 connectivity thresholds estimated (a,b) without cost normalization and (c,d) with cost normalization. Note The ICC is at level of 0.5 to 0.6 for network metrics without cost normalization and the ICC is at 0.67 to 0.85 for network metrics with cost normalization. However, dramatic variations of ICC are found for E<sub>loc</sub>, C<sub>w</sub> estimated with cost normalization when sparsity is not the same (N1<sub>sep</sub> and N2<sub>sep</sub>).

threshold. The sparsity defined as the number of non-zero edges divided by the total number of edges in the diffusion connectivity matrix can be calculated. To ensure same sparsity for N1 and N2, connectivity thresholds were applied on N1 and N2 group together denoted as N1<sub>com</sub> and N2<sub>com</sub>:

$$\text{average}(P_{ij}^{Total}) + 2\text{std}(P_{ij}^{Total}) < \text{threshold}. \quad (1)$$

where  $P_{ij}^{Total}$  is the connectivity probability from N1 and N2 (60 DTI scans from 30 subjects). In this way, all subjects have the same number of edges and the edges are located at the same position of the connectivity matrix. Further, we applied the connectivity thresholds on N1 and N2 separately, denote as  $N1_{sep}$  and  $N2_{sep}$ :

$$\text{average}(P_{ij}^{N1}) + 2 * \text{std}(P_{ij}^{N1}) < \text{threshold} \quad (2)$$

$$\text{average}(P_{ij}^{N2}) + 2 * \text{std}(P_{ij}^{N2}) < \text{threshold} \quad (3)$$

where  $P_{ij}^{N1}$  and  $P_{ij}^{N2}$  is the connectivity from N1 (30 DTI scans from 30 subjects) and N2 (30 DTI scans from 30 subjects) respectively. The resultant connectivity matrixes may have different sparsity for the repeated scans under the same connectivity threshold. This is to examine the potential variation of the network metrics when sparsity cannot be adjusted at same level by the connectivity thresholds. A series of connectivity thresholds at the range of 0.01 to 0.1 with 0.0025 steps were chosen based on Gong *et al.*<sup>15</sup>. The range of sparsity in this study is 8% to 32%, which is similar to previous studies<sup>15,35</sup>. To construct the weighted network from connectivity matrix, the weights of each edge is computed as  $w_{ij} = P_{ij}$ . For each subject, 37 weighted networks were constructed corresponding to 37 connectivity thresholds.

**Network measures.** For each network, four network metrics were calculated. They are network cluster coefficient ( $C_w$ ), characteristic path length ( $L_w$ ), global efficiency ( $E_{glob}$ ) and local efficiency ( $E_{loc}$ )<sup>36,37</sup>. The definition of these network metrics is given in the supplementary methods. Because these network metrics were computed for a range of sparsity under a series of connectivity thresholds, the summary network metrics were integrals of each metrics over the range of the sparsity. The network metrics were additionally calculated using the weights scaled by the sum of all weights for all edges to control each subject's cost at same level, known as cost normalization<sup>15,21</sup>. All network analysis was performed in Matlab (The MathWorks, Natick, USA) using Brain Connectivity Toolbox<sup>9</sup>.

**Quantification of reproducibility.** The similarity was calculated by the Pearson's correlation coefficients of connectivity from structural connectivity matrixes. Within subject similarity was computed by averaging 30 correlation coefficients of connectivity calculated between 2 repeated measurements. Between subject similarity was computed by averaging correlation coefficients of connectivity calculated among 30 subjects. The CV and ICC were used as indices of reproducibility<sup>23,25</sup>. The CV is defined as the standard deviation divided by the overall measurement mean. For  $CV_{ws}$ , the standard deviation was calculated between 2 repeated measurements (N1 and N2) for each subject. The  $CV_{ws}$  was given by mean of within subject standard deviation divided by the overall mean. For  $CV_{bs}$ , the standard deviation was calculated among 30 subjects in N1 and N2, respectively. The  $CV_{bs}$  was given by mean of between subject standard deviation divided by the overall mean. The ICC was computed using two-way mixed single measures using the absolute agreement within the repeated measurement. The ICC was computed using MATLAB toolbox created by Arash Salarian ([www.mathworks.com/matlabcentral/fileexchange/22099](http://www.mathworks.com/matlabcentral/fileexchange/22099)). The ICC values are classified as: poor reproducibility (<0.5), moderate reproducibility (0.5–0.75), good reproducibility (0.75–0.9), excellent reproducibility (>0.9)<sup>23,38,39</sup>. To summarize the structural connectivity matrix, edges were separated into 31 groups according to the connectivity probability from 0 to 0.3 in 0.01 step. The indexes of the reproducibility were calculated for each group.

## References

1. Lebel, C., Benner, T. & Beaulieu, C. Six is enough? Comparison of diffusion parameters measured using six or more diffusion-encoding gradient directions with deterministic tractography. *Magn Reson Med* **68**, 474–483, <https://doi.org/10.1002/mrm.23254> (2012).
2. Van Hecke, W. *et al.* On the construction of an inter-subject diffusion tensor magnetic resonance atlas of the healthy human brain. *NeuroImage* **43**, 69–80, S1053-8119(08)00797-0, <https://doi.org/10.1016/j.neuroimage.2008.07.006> (2008).
3. Glasser, M. F. & Rilling, J. K. DTI tractography of the human brain's language pathways. *Cereb Cortex* **18**, 2471–2482, <https://doi.org/10.1093/cercor/bhn011> (2008).
4. Huang, H. *et al.* DTI tractography based parcellation of white matter: application to the mid-sagittal morphology of corpus callosum. *NeuroImage* **26**, 195–205, <https://doi.org/10.1016/j.neuroimage.2005.01.019> (2005).
5. Behrens, T. E., Berg, H. J., Jbabdi, S., Rushworth, M. F. & Woolrich, M. W. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage* **34**, 144–155, <https://doi.org/10.1016/j.neuroimage.2006.09.018> (2007).
6. Ciccarelli, O. *et al.* Probabilistic diffusion tractography: a potential tool to assess the rate of disease progression in amyotrophic lateral sclerosis. *Brain* **129**, 1859–1871, <https://doi.org/10.1093/brain/awl100> (2006).
7. Sporns, O. The human connectome: origins and challenges. *NeuroImage* **80**, 53–61, <https://doi.org/10.1016/j.neuroimage.2013.03.023> (2013).
8. Sporns, O. The human connectome: a complex network. *Ann N Y Acad Sci* **1224**, 109–125, <https://doi.org/10.1111/j.1749-6632.2010.05888.x> (2011).
9. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage* **52**, 1059–1069, <https://doi.org/10.1016/j.neuroimage.2009.10.003> (2010).
10. Hagmann, P. *et al.* MR connectomics: Principles and challenges. *Journal of neuroscience methods* **194**, 34–45, <https://doi.org/10.1016/j.jneumeth.2010.01.014> (2010).
11. Zalesky, A. *et al.* Whole-brain anatomical networks: does the choice of nodes matter? *NeuroImage* **50**, 970–983, <https://doi.org/10.1016/j.neuroimage.2009.12.027> (2010).

12. Zalesky, A., Fornito, A. & Bullmore, E. T. Network-based statistic: identifying differences in brain networks. *NeuroImage* **53**, 1197–1207, <https://doi.org/10.1016/j.neuroimage.2010.06.041> (2010).
13. Gong, G. *et al.* Mapping anatomical connectivity patterns of human cerebral cortex using *in vivo* diffusion tensor imaging tractography. *Cereb Cortex* **19**, 524–536, <https://doi.org/10.1093/cercor/bhn102> (2009).
14. Iturria-Molina, Y. *et al.* Characterizing brain anatomical connections using diffusion weighted MRI and graph theory. *NeuroImage* **36**, 645–660, <https://doi.org/10.1016/j.neuroimage.2007.02.012> (2007).
15. Gong, G. *et al.* Age- and gender-related differences in the cortical anatomical network. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **29**, 15684–15693, <https://doi.org/10.1523/JNEUROSCI.2308-09.2009> (2009).
16. Yan, C. *et al.* Sex- and brain size-related small-world structural cortical networks in young adults: a DTI tractography study. *Cereb Cortex* **21**, 449–458, <https://doi.org/10.1093/cercor/bhq111> (2011).
17. Filippi, M., Preziosa, P. & Rocca, M. A. Brain mapping in multiple sclerosis: Lessons learned about the human brain. *NeuroImage* <https://doi.org/10.1016/j.neuroimage.2017.09.021> (2017).
18. Collin, G. *et al.* Brain network analysis reveals affected connectome structure in bipolar I disorder. *Human brain mapping* **37**, 122–134, <https://doi.org/10.1002/hbm.23017> (2016).
19. Munsell, B. C. *et al.* Evaluation of machine learning algorithms for treatment outcome prediction in patients with epilepsy based on structural connectome data. *NeuroImage* **118**, 219–230, <https://doi.org/10.1016/j.neuroimage.2015.06.008> (2015).
20. Prescott, J. W. *et al.* The Alzheimer structural connectome: changes in cortical network topology with increased amyloid plaque burden. *Radiology* **273**, 175–184, <https://doi.org/10.1148/radiol.14132593> (2014).
21. Cao, Q. *et al.* Probabilistic diffusion tractography and graph theory analysis reveal abnormal white matter structural connectivity networks in drug-naïve boys with attention deficit/hyperactivity disorder. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **33**, 10676–10687, <https://doi.org/10.1523/JNEUROSCI.4793-12.2013> (2013).
22. Welton, T., Kent, D. A., Auer, D. P. & Dineen, R. A. Reproducibility of graph-theoretic brain network metrics: a systematic review. *Brain Connect* **5**, 193–202, <https://doi.org/10.1089/brain.2014.0313> (2015).
23. Andreotti, J. *et al.* Repeatability analysis of global and local metrics of brain structural networks. *Brain Connect* **4**, 203–220, <https://doi.org/10.1089/brain.2013.0202> (2014).
24. Owen, J. P. *et al.* Test-retest reliability of computational network measurements derived from the structural connectome of the human brain. *Brain Connect* **3**, 160–176, <https://doi.org/10.1089/brain.2012.0121> (2013).
25. Vaessen, M. J. *et al.* The effect and reproducibility of different clinical DTI gradient sets on small world brain connectivity measures. *NeuroImage* **51**, 1106–1116, <https://doi.org/10.1016/j.neuroimage.2010.03.011> (2010).
26. Buchanan, C. R., Pernet, C. R., Gorgolewski, K. J., Storkey, A. J. & Bastin, M. E. Test-retest reliability of structural brain networks from diffusion MRI. *NeuroImage* **86**, 231–243, <https://doi.org/10.1016/j.neuroimage.2013.09.054> (2014).
27. Bonilha, L. *et al.* Reproducibility of the Structural Brain Connectome Derived from Diffusion Tensor Imaging. *PLoS one* **10**, e0135247, <https://doi.org/10.1371/journal.pone.0135247> (2015).
28. Bassett, D. S., Brown, J. A., Deshpande, V., Carlson, J. M. & Grafton, S. T. Conserved and variable architecture of human white matter connectivity. *NeuroImage* **54**, 1262–1279, <https://doi.org/10.1016/j.neuroimage.2010.09.006> (2011).
29. van Wijk, B. C., Stam, C. J. & Daffertshofer, A. Comparing brain networks of different size and connectivity density using graph theory. *PLoS one* **5**, e13701, <https://doi.org/10.1371/journal.pone.0013701> (2010).
30. Khalsa, S., Mayhew, S. D., Chechlac, M., Bagary, M. & Bagshaw, A. P. The structural and functional connectivity of the posterior cingulate cortex: comparison between deterministic and probabilistic tractography for the investigation of structure-function relationships. *NeuroImage* **102**(Pt 1), 118–127, <https://doi.org/10.1016/j.neuroimage.2013.12.022> (2014).
31. Moldrich, R. X. *et al.* Comparative mouse brain tractography of diffusion magnetic resonance imaging. *NeuroImage* **51**, 1027–1036, <https://doi.org/10.1016/j.neuroimage.2010.03.035> (2010).
32. Smith, R. E., Tournier, J. D., Calamante, F. & Connelly, A. The effects of SIFT on the reproducibility and biological accuracy of the structural connectome. *NeuroImage* **104**, 253–265, <https://doi.org/10.1016/j.neuroimage.2014.10.004> (2015).
33. Cheng, H. *et al.* Characteristics and variability of structural networks derived from diffusion tensor imaging. *NeuroImage* **61**, 1153–1164, <https://doi.org/10.1016/j.neuroimage.2012.03.036> (2012).
34. Tzourio-Mazoyer, N. *et al.* Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* **15**, 273–289, <https://doi.org/10.1006/nimg.2001.0978> (2002).
35. Achard, S. & Bullmore, E. Efficiency and cost of economical brain functional networks. *PLoS Comput Biol* **3**, e17, <https://doi.org/10.1371/journal.pcbi.0030017> (2007).
36. Latora, V. & Marchiori, M. Efficient behavior of small-world networks. *Phys Rev Lett* **87**, 198701, <https://doi.org/10.1103/PhysRevLett.87.198701> (2001).
37. Watts, D. J. & Strogatz, S. H. Collective dynamics of ‘small-world’ networks. *Nature* **393**, 440–442, <https://doi.org/10.1038/30918> (1998).
38. Koch, G. G. *et al.* A general methodology for the analysis of experiments with repeated measurement of categorical data. *Biometrics* **33**, 133–158 (1977).
39. Koo, T. K. & Li, M. Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* **15**, 155–163, <https://doi.org/10.1016/j.jcm.2016.02.012> (2016).

## Acknowledgements

The author thanks the Taiwan Mind & Brain Imaging Center (TMBIC) and National Chengchi University for the instrumental availability. The TMBIC is supported by the Ministry of Science and Technology, Taiwan. This work was supported in part by grants from the Ministry of Science and Technology, Taiwan (MOST 106-2221-E-004-001, 105-2221-E-004-003, 105-2420-H-004-003-MY2).

## Author Contributions

S.-Y.T. conceived and designed the experiments, performed the experiments, analyzed the data, interpreted the data, and drafted the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-29943-0>.

**Competing Interests:** The author declares no competing interests.

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018