Surface-based Single-subject Morphological Brain Networks: Effects of Morphological Index, Brain Parcellation and Similarity Measure, Sample Size-varying Stability and Test-retest Reliability

Yinzhi Li[1,5#], Ningkai Wang[1,5#], Hao Wang[2], Yating Lv[3], Qihong Zou[4] and Jinhui Wang[1,5*]

[1]Institute for Brain Research and Rehabilitation, Guangdong Key Laboratory of Mental Health and Cognitive Science, Center for Studies of Psychological Application, South China Normal University, Guangzhou, China

[2]Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China

[3]Institute of Psychological Sciences, Hangzhou Normal University, Hangzhou, Zhejiang, China

[4]Center for MRI Research, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China

[5]Key Laboratory of Brain, Cognition and Education Sciences (South China Normal University), Ministry of Education

[#]These authors contribute equally to this work.

[*]**Correspondence author**

23 Jinhui Wang, Institute for Brain Research and Rehabilitation, Guangdong Key

24 Laboratory of Mental Health and Cognitive Science, Center for Studies of

25 Psychological Application, South China Normal University, Guangzhou 510631,

26 China. Email: jinhui.wang.1982@m.scnu.edu.cn

Abstract

Morphological brain networks, in particular those at the individual level, have become an important approach for studying the human brain connectome; however, relevant methodology is far from being well-established in their formation, description and reproducibility. Here, we extended our previous study by constructing and characterizing single-subject morphological similarity networks from brain volume to surface space and systematically evaluated their reproducibility with respect to effects of different choices of morphological index, brain parcellation atlas and similarity measure, sample size-varying stability and test-retest reliability. Using the Human Connectome Project dataset, we found that surface-based single-subject morphological similarity networks shared common small-world organization, high parallel efficiency, modular architecture and bilaterally distributed hubs regardless of different analytical strategies. Nevertheless, quantitative values of all interregional similarities, global network measures and nodal centralities were significantly affected by choices of morphological index, brain parcellation atlas and similarity measure. Moreover, the morphological similarity networks varied along with the number of participants and approached stability until the sample size exceeded ~70. Using an independent test-retest dataset, we found fair to good, even excellent, reliability for most interregional similarities and network measures, which were also modulated by different analytical strategies, in particular choices of morphological index. Specifically, fractal dimension and sulcal depth outperformed gyrification index and cortical thickness, higher-resolution atlases outperformed lower-resolution

atlases, and Jensen-Shannon divergence-based similarity outperformed Kullback-Leibler divergence-based similarity. Altogether, our findings propose surface-based single-subject morphological similarity networks as a reliable method to characterize the human brain connectome and provide methodological recommendations and guidance for future research.

Keywords: morphological brain network, cortical surface, structural MRI, sample size, test-retest reliability

Introduction

Morphological brain networks depict patterns of interregional relations in regional brain morphology on the basis of structural magnetic resonance imaging. Historically, morphological brain networks are mainly derived via population-based morphological covariance network methods by estimating interregional covariance across a cohort of participants in a certain morphological index, such as gray matter volume, cortical thickness and surface area (Bassett et al., 2008; He et al., 2007; Sanabria-Diaz et al., 2010). To date, the population-based morphological covariance network methods have been widely used as an important tool to study the human brain, including but not limited to parsing of organizational principles of healthy brains, characterization of trajectories during development and aging and identification of abnormalities in various brain diseases (see Alexander-Bloch et al., 2013a; Evans, 2013 for two excellent reviews).

However, the population-based morphological covariance networks suffer from several noticeable issues, such as neglect of interindividual variability, requirement of a large sample size and introduction of complicated models for subsequent statistical inference. All these issues limit the universal application of morphological brain networks, in particular in uncovering their neurobiological significance and clinical diagnostic and prognostic value. Recently, the advent of individual-level morphological similarity networks has overcome, to a great extent, these issues and has thus attracted considerable attention (Jiang et al., 2017; Kong et al., 2015; Li et al., 2017; Seidlitz et al., 2018; Tijms et al., 2012; Wang et al., 2016; Yu et al., 2018). With

these individual-level methods, several studies show that morphological similarity networks can capture known cortical cytoarchitecture and related gene expression (Seidlitz et al., 2018), account for interindividual differences in cognition (Li and Kong, 2017; Seidlitz et al., 2018; Tijms et al., 2014), and distinguish patients with schizophrenia from healthy controls (Zhao et al., 2020) and predict clinical progression of patients with Alzheimer's disease (Tijms et al., 2018). These findings provide strong evidence that individual-level morphological similarity networks are biologically meaningful and are of great value in helping clinical diagnosis and prognosis.

Technically, individual-level morphological similarity networks can be divided into two subcategories. The first estimates interregional morphological similarity by computing Pearson correlation coefficients in regional mean signals across different morphological indices (Li et al., 2017; Seidlitz et al., 2018). This type of method, however, not only neglects intraregional morphological distributions but also may lead to unstable similarity estimation due to a limited number of morphological indices available (< 10 in previous studies), which are treated as samples for the correlation analysis. In contrast, the second subcategory estimates interregional morphological similarity at a more refined level by taking intraregional morphological distributions into account from different perspectives (Jiang et al., 2017; Kong et al., 2015; Tijms et al., 2012; Yu et al., 2018). In this regard, we previously constructed single-subject morphological similarity networks by estimating interregional morphological similarity in the distribution of regional gray matter volume in terms of

the Kullback-Leibler divergence (*KLD*; Wang et al., 2016). We found that this method can reveal structured organization of morphological similarity networks with high test-retest (TRT) reliability (Wang et al., 2016). It should be noted that the Pearson correlation-based method and the *KLD*-based method are essentially different: the former estimates interregional morphological similarity in intraregional mean of multiple morphological indices, while the latter estimates interregional morphological similarity based on intraregional distribution of a single morphological index. The Pearson correlation method cannot be used to estimate interregional morphological similarity based on a single morphological index. This is not only because there are different numbers of vertices but also because there is no one-to-one correspondence of the vertices between two regions. Thus, these two types of methods cannot be compared directly.

In the present contribution, we extended our previous study in several aspects. First, we constructed single-subject morphological similarity networks in cerebral cortical surface rather than in volume space. The greatest benefit from this switch is that spherical registration of cortical surface meshes increases the accuracy of brain registration and thus allows more precise localization (Desai et al., 2005). In addition, inflation of cortical surface meshes promotes better visualization of buried sulci by raising them to the surface. Second, there are more forms of morphological indices available for surface mesh-based analysis, with each capturing different aspects of cerebral morphology. Thus, in contrast to a single morphological index (i.e., gray matter volume), used in our previous study, this work examined common and specific

123 degrees of organization among different types of morphological similarity networks
124 derived from different morphological indices. These analyses are of great significance
125 for comprehensive understanding of morphological similarity network architecture.
126 Third, in addition to evaluating different brain parcellation atlases as in our previous
127 study, this work examined effects of different similarity measures on morphological
128 similarity networks. Despite accumulating evidence for significant impacts of
129 different similarity measures on structural and functional brain networks (Liang et al.,
130 2012; Sarwar et al., 2019; Zalesky et al., 2010), the extent to which morphological
131 similarity networks depends on choices of similarity measure is largely unknown.
132 Finally, this work evaluated stability with respect to different sample sizes and TRT
133 reliability of morphological similarity networks.

134 We hypothesize that quantitative descriptions of surface-based single-subject
135 morphological similarity networks are dependent on choices of morphological index,
136 brain parcellation atlas and similarity measure, are robust to variation in sample size
137 and have high TRT reliability.

138

139 **Materials and Methods**

140 **General analytical pipeline**

141 In this study, using a large-scale dataset, we first derived 4 surface-based, vertexwise
142 morphological brain maps (fractal dimension, FD, gyrification index, GI, sulcal depth,
143 SD, and cortical thickness, CT) for each participant, based on which single-subject
144 morphological similarity networks were constructed using two surface atlases for

brain parcellation (a2009s atlas and a2005s atlas) and two similarity measures for interregional similarity estimation (*KLD*-based similarity, *KLDs*, and Jensen-Shannon divergence-based similarity, *JSDs*). Therefore, we obtained 16 morphological similarity networks in total for each participant (4 morphological indices × 2 parcellation atlases × 2 similarity measures). Then, we examined effects of the different analytical strategies as well as sample sizes on interregional similarity, global network organization and nodal centrality of the resultant morphological similarity networks. Finally, we utilized an independent, TRT dataset to explore reliability of surface-based single-subject morphological similarity networks under different analytical strategies, aiming to provide methodological recommendations for future studies. Figure 1 presents a flowchart of the overall pipeline of data processing.

Participants

Two publicly available datasets were used in this study: the Human Connectome Project (HCP) S900 dataset (www.humanconnectome.org) (Van Essen et al., 2013) and the Beijing Normal University (BNU) TRT dataset (http://fcon_1000.projects.nitrc.org/indi/CoRR/html/bnu_1.html). The former was used to characterize the topological organization of morphological similarity networks under different analytical strategies and effects of different sample sizes, and the latter was used to evaluate TRT reliabilities of morphological similarity networks.

HCP S900 dataset. The HCP S900 dataset includes a total of 897 healthy adult participants who completed structural MRI scans. Twenty-one participants were

167 excluded according to our quality control procedures (see below), resulting in 876
168 participants in the final analyses (male/female: 386/490; main age: 22-35 years).
169 ***BNU TRT dataset***. The BNU TRT dataset contains a total of 57 healthy young
170 participants (male/female: 30/27; age: 19-30 years) who each completed two MRI
171 scan sessions within an interval of approximately 6-weeks (40.94 ± 4.51 days). All
172 participants were right-handed and had no history of neurological or psychiatric
173 disorders.
174
175 **MRI data acquisition**
176 ***HCP S900 dataset.*** T1w images from the HCP S900 dataset were obtained using
177 a customized 3T Siemens Magnetom Connectome scanner with a 32-channel head
178 coil. Main imaging parameters were: repetition time (TR) = 2400 ms, echo time (TE)
179 = 2.14 ms, inversion time (TI) = 1000 ms, flip angle (FA) = 8°, field of view (FOV) =
180 224 × 224 mm^2, matrix = 320 × 320, thickness = 0.7 mm with no gap, and 256 sagittal
181 slices.
182 ***BNU TRT dataset***. T1w images from the BNU TRT dataset were obtained using
183 a 3T Siemens Tim Trio scanner with a 12-channel head coil. Main imaging
184 parameters were: TR = 2530 ms, TE = 3.39 ms, TI = 1100 ms, FA = 7°, FOV = 256 ×
185 256 mm^2, matrix = 256 ×192, slice thickness = 1.33 mm; interslice gap = 0.65 mm,
186 and 144 sagittal slices.
187
188 **Quality control procedures**

The HCP S900 dataset underwent multiple levels of quality control, ranging from real-time oversight during acquisition to post-acquisition manual and automated image review. Each of these procedures is specified in a formal standard operating procedure and integrated into the internal database system (Marcus et al., 2013). For the BNU TRT dataset, the quality control procedures included visual inspection of severe motion artefacts or any other apparent artefacts, followed by calculation of a series of quality evaluation metrics, such as signal-to-noise ratio, foreground to background energy ratio, ghost to signal ratio and artifact detection (Lin et al., 2015). In addition to the dataset specific quality control procedures, in this study we further visually checked the results of image segmentation via the modules “Slice Display” and “Surface Data Homogeneity” in the CAT12 toolbox. Twenty-one participants were excluded due to failed segmentation or poor image quality for the HCP S900 dataset.

MRI data preprocessing

All structural images underwent standard processes using Computation Anatomy Toolbox 12 (CAT12, version r1113, http://www.neuro.uni-jena.de/cat/), based on Statistical Parametric Mapping 12 (SPM12, version 6685, https://www.fil.ion.ucl.ac.uk/spm/software/spm12). The CAT12 offers a volume-based approach for estimating cerebral surface morphology without extensive reconstruction of cortical surface and thus is timesaving. Specifically, the CAT12 contains a processing pipeline for computing four morphological indices, including

FD, GI, SD and CT. All image preprocessing and morphological parameter computation described below were conducted in subject native space. After obtaining individual morphological maps of FD, GI, SD and CT, they were finally resampled into the common fsaverage template and smoothed using a Gaussian kernel. Specifically, individual CT maps were smoothed using a Gaussian kernel with 15-mm full width at half maximum, while individual FD, GI and SD maps were smoothed using a Gaussian kernel with 25-mm full width at half maximum. Based on the recommendations of the CAT12 manual, the usage of larger filter sizes for FD, GI and SD is due to the underlying nature of these folding measures that reflect contributions from both sulci and gyri. Therefore, the filter size should exceed the distance between a gyral crown and a sulcal fundus.

CT calculation. CT was estimated using a fast and reliable projection-based thickness (PBT) method, which requires no extensive reconstruction of the cortical surface. First, following tissue segmentation, a white matter (WM) distance map was derived by estimating the distance from the inner gray matter (GM) boundary for each GM voxel (Dahnke et al., 2013). Values at the outer GM boundary in the WM distance map (i.e., GM thickness) were then projected back to the inner GM boundary to generate a GM thickness map. Subsequently, a central surface was created at the 50% level of the percentage position between the WM distance and GM thickness maps. For the resultant central surface, a topology correction based on spherical harmonics was used to account for topological defects (Yotter et al., 2011a). Furthermore, the central surface was reparameterized into a common coordinate system via spherical

mapping (Yotter et al., 2011c), and spherical registration adopted the volume-based diffeomorphic DARTEL algorithm (Ashburner, 2007) to the surface.

FD calculation. FD reflects cortical surface folding complexity. It was estimated based on spherical harmonic reconstructions and calculated as the slope of a logarithmic plot of surface area versus the maximum l-value, where the maximum l-value is a measure of the bandwidth of frequencies used to reconstruct the surface shape (Yotter et al., 2011b).

GI calculation. GI is an indicator of cortical folding. Based on the spherical harmonic reconstructions, GI was calculated as absolute mean curvature (Luders et al., 2006). Mean curvature is an extrinsic surface measure, which provides information about the change in normal direction along the surface.

SD calculation. SD measures the depth of sulci and was calculated as the Euclidean distance between the central surface and its convex hull based on the spherical harmonic reconstructions.

Construction of individual morphological similarity networks

A network is made up of nodes and edges between the nodes. In this study, we constructed large-scale morphological similarity networks with nodes denoting brain regions and edges denoting interregional similarity in intraregional distributions of morphological indices.

Definition of network nodes. To define network nodes, we employed two widely used surface atlases, the a2009s atlas (Destrieux et al., 2010) and the DK40 atlas

(termed a2005s atlas in this study) (Desikan et al., 2006), which divided the cerebral cortex into 148 and 68 regions of interest (ROIs), respectively.

Definition of network edges. To estimate interregional morphological similarity, we utilized two measures, the *KLDs* and its variant, the *JSDs*, to estimate the similarity in intraregional distribution of morphological indices between regions. In mathematical statistics, the *KLD* is a measure of how one probability distribution is different from a second, reference probability distribution (Kullback and Leibler, 1951). This measure has been widely used in the fields of image processing and machine learning. First, we extracted values of all vertices within each ROI for each morphological index. Then, a probability density estimate was computed for each ROI and each morphological index based on a normal kernel function (MATLAB function, ksdensity). The resultant probability density functions (four per region) were further converted to probability distribution functions (PDFs). The *KLDs* between two PDFs *P* and *Q* is computed as:

$$KLDs(P,Q) = e^{-D_{KL}(P,Q)},$$

where *e* is the natural base and $D_{KL}(P,Q) = KLD(P||Q) + KLD(Q||P) = \sum_{i=1}^{n} P(i) log \frac{P(i)}{Q(i)} + \sum_{i=1}^{n} Q(i) log \frac{Q(i)}{P(i)}$, with *n* being the number of sample points (2^8 in this study) (Wang et al., 2016). For the *JSDs* between *P* and *Q*, the formula is:

$$JSDs = 1 - D_{JS} = 1 - \sqrt{JSD(P||Q)},$$

where $JSD(P||Q) = \frac{1}{2} KLD(P||\frac{1}{2}(P+Q)) + \frac{1}{2} KLD(Q||\frac{1}{2}(P+Q))$. The value range for both *KLDs* and *JSDs* is [0, 1], with 0 and 1 denoting that two PDFs are absolutely different or exactly the same, respectively.

277

278 **Network analysis**

279 Prior to graph-based topological characterization of morphological similarity
280 networks derived above, a thresholding procedure was used to convert each network
281 to a series of binary graphs. All topological analyses were performed with the
282 GRETNA toolbox (Wang et al., 2015).

283 ***Threshold selection***. Consistent with our previous study (Wang et al., 2016), we
284 employed a sparsity-based thresholding procedure, where sparsity was defined as the
285 ratio of the number of actual edges divided by the maximum possible number of
286 edges in a graph. Given the lack of a conclusive method for selecting a single sparsity,
287 a consecutive sparsity range was used in this study: [0.034 0.4] for the a2009s atlas
288 and [0.063 0.4] for the a2005s atlas (interval = 0.02). The lower limits of the sparsity
289 ranges were chosen to ensure that the resultant graphs would be estimable for the
290 small-world attributes (Watts and Strogatz, 1998), that is, the average nodal degree
291 (nodal degree is defined as the number of edges linked to a node) over all nodes of
292 each thresholded graph is larger than $2 \times \log(N)$, with N denoting the number of nodes
293 (i.e., 68 or 148 in this study). This criterion guarantees that the generated random
294 networks are connected and thus the small world parameters can be successfully
295 calculated (Watts and Strogatz, 1998). For the upper limits of the sparsity ranges, they
296 were determined to guarantee that the thresholded graphs have sparse properties (He
297 et al., 2007; Wang et al., 2009). All subsequent topological analyses were conducted
298 at each sparsity level, and thus each network measure calculated below is a function

or curve of sparsity. To provide sparsity-independent summary scalars, we computed the area under the curve (AUC; i.e., the integral over the entire sparsity range) for each network measure (Zhang et al., 2011), which were used to simplify the TRT reliability and statistical analyses.

Global and nodal network measures. For each graph, we calculated both global (clustering coefficient, C_p; characteristic path length, L_p; local efficiency, E_{loc}; global efficiency, E_{glob}; and modularity, Q) and nodal (nodal degree, k_i; nodal efficiency, e_i; and nodal betweenness, b_i) measures. Detailed formulas and interpretations of these measures can be found elsewhere (Rubinov and Sporns, 2010; Wang et al., 2011) and are summarized in Table S1. To test whether the morphological similarity networks were non-randomly organized, all global measures were further normalized by dividing them by corresponding measures averaged over 100 matched random networks. The random networks were generated using a topological rewiring method (Maslov and Sneppen, 2002), which guaranteed the same degree distributions between real and random networks. Typically, a small-world, highly efficient and modular network should fulfill the following conditions: normalized $C_p > 1$ and normalized $L_p \sim 1$, normalized $E_{loc} > 1$ and normalized $E_{glob} \sim 1$ and normalized $Q > 1$.

TRT reliability

We calculated the intraclass correlation coefficient (*ICC*) (Shrout and Fleiss, 1979) to quantify TRT reliability of morphological similarity networks. Formally, for a given

measure repeatedly observed *k* times, the *ICC* was calculated as:

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

where MS_b is the between-subject sum of squares and MS_w is the within-subject sum of squares. ICC is close to 1 for reliable measures and 0 (negative) otherwise. In accordance with our previous studies (Wang et al., 2016; Wang et al., 2011), the TRT reliability scores were categorized as poor (*ICC* < 0.25), low (0.25 < *ICC* < 0.4), fair (0.4 < *ICC* < 0.6), good (0.6 < *ICC* < 0.75) and excellent (0.75 < *ICC* < 1).

In this study, the *ICC* was calculated for each interregional morphological similarity and each network measure (global and nodal) under all possible combinations of 4 morphological indices, 2 atlases and 2 similarity measures. This resulted in a total of 16 *ICC* matrices (148 × 148 or 68 × 68) for interregional morphological similarity, 160 *ICC* values for global network measures and 48 *ICC* vectors (148 or 68) for nodal network measures.

Statistical analysis

Effects of different analytical strategies on morphological similarity networks. We explored effects of different choices of morphological index, brain parcellation atlas and similarity measure on morphological similarity networks at multiple levels. First, at the level of interregional morphological similarity, we first examined differences in the mean morphological similarity over all possible pairs of regions between the two parcellation atlases regardless of morphological indices and similarity measures (t-test). Then, for each brain parcellation atlas, we performed

two-way repeated ANOVA on the morphological similarity between any pair of regions (10,878 ANOVA under the a2009s parcellation atlas and 2,278 ANOVA under the a2005s parcellation atlas). Second, at the level of global network organization, we performed three-way repeated ANOVA on each global network measure (10 ANOVA). Finally, at the level of local network organization, we first compared differences in the mean nodal centrality across all regions for each nodal centrality measure between the two parcellation atlases regardless of morphological indices and similarity measures (2 t-tests; nodal degree was excluded from this analysis because the mean nodal degree was equal among all participants in terms of the sparsity-based thresholding procedure). Then, under each brain parcellation atlas, we performed two-way repeated ANOVA on each nodal measure of each region (444 ANOVA under the a2009s parcellation atlas and 204 ANOVA under the a2005s parcellation atlas). For significant effects, simple effects were further examined (paired t-test). All results were corrected for multiple comparisons using the false discovery rate (FDR) procedure at the level of $q < 0.05$, when applicable.

Effects of different analytical strategies on TRT reliability of morphological similarity networks. Effects of different analytical strategies on the TRT reliability of morphological similarity networks were also examined at multiple levels. First, at the level of interregional morphological similarity, we first tested differences in the mean connectional *ICC* over all possible pairs of regions between the two parcellation atlases regardless of morphological indices and similarity measures (t-test). Then, under each brain parcellation atlas, we performed two-way repeated ANOVA on all

connectional *ICC* values (2 ANOVA). Second, at the level of global network organization, we performed three-way repeated ANOVA on *ICC* values for all global network measures. Finally, at the level of local network organization, we first examined differences in the mean nodal *ICC* over all regions between the two parcellation atlases regardless of morphological indices and similarity measures (3 t-tests). Then, under each brain parcellation atlas, we performed two-way repeated ANOVA on all nodal *ICC* values for each nodal measure (6 ANOVA). For significant effects, simple effects were further examined (paired t-test). All results were corrected for multiple comparisons using the FDR procedure at the level of $q < 0.05$, when applicable.

In addition, we utilized Z-tests (McGraw and Wong, 1996) to locate edges, global measures and regions, whose *ICC* values were significantly affected by different analytical strategies. Specifically, for two given *ICC* values ICC_1 and ICC_2, we first transformed the difference between ICC_1 and ICC_2 into z:

$$z = \frac{1}{2}\ln\frac{1+(k+1)(ICC_1 - ICC_2)}{1-(ICC_1 - ICC_2)},$$

where k is the number of repeated observations (2 in this study). The z statistic has a mean of 0 with variance (McGraw and Wong, 1996):

$$\sigma^2 = \frac{k}{2(N-2)(k-1)},$$

where N is the number of participants. Then, we transformed the z into a Z score with standard normal distribution:

$$Z = \frac{z - 0}{\sqrt{\frac{k}{2(N-2)(k-1)}}} = \frac{\sqrt{N-2}}{2} \ln \frac{1 + (ICC_1 - ICC_2)}{1 - (ICC_1 - ICC_2)}$$

$$= \sqrt{N-2} \tanh^{-1}(ICC_1 - ICC_2).$$

The Z-test was performed for each interregional morphological similarity and each nodal centrality measure for each region (FD/GI/SD/CT: *KLDs* versus *JSDs*; *KLDs*/*JSDs*: FD versus GI, FD versus SD, FD versus CT, GI versus SD, GI versus CT and SD versus CT), and each global network measure (FD/GI/SD/CT - a2005s/a2009s: *KLDs* versus *JSDs*; FD/GI/SD/CT - *KLDs*/*JSDs*: a2005s versus a2009s; a2005s/a2009s - *KLDs*/*JSDs*: FD versus GI, FD versus SD, FD versus CT, GI versus SD, GI versus CT and SD versus CT). All results were corrected for multiple comparisons using the FDR procedure at the level of $q < 0.05$, when applicable.

Effects of sample size on morphological similarity networks

To test effects of different sample sizes on morphological similarity networks, we correlated the cross-subject mean nodal centrality map (concatenated across the three nodal measures) from a subgroup of participants with that from all participants. The subgroup of participants was randomly selected from all participants with sample size varying from 10 to 870 (interval = 10). This subgroup sampling procedure was repeated 1,000 times to generate 1,000 correlation coefficients at any fixed sample size, whose means and standard errors were calculated.

Validation analysis

Twin subjects. The HCP S900 dataset includes twin subjects, which may yield bias for our results. Thus, we re-analyzed the HCP S900 data to estimate the reproducibility of our results by excluding twin subjects (338).

High-resolution parcellation atlas. In this study, we utilized the a2005s atlas and a2009s atlas for network node definition because they are two of the most commonly used surface parcelllation schemes in previous brain network studies (e.g., Buchanan et al., 2020; Rodríguez-Cruces et al., 2020; Seibert et al., 2011; Zhang et al., 2019). However, these two atlases are relatively coarser (68 and 148 regions, respectively), which may be insufficient to represent local characteristics of finer-level brain regions. Thus, we further re-analyzed the data by constructing individual morphological similarity networks with a high-resolution atlas, which divides the cerebral surface into 360 regions (termed MMP atlas; Glasser et al., 2016). It should be noted that we mainly focused on a more practical question in guiding studies of morphological similarity networks: whether brain parcellation atlases with higher resolutions will give rise to higher test-retest reliabilities. We did not re-analyzed the HCP S900 dataset because of the huge amount of computation, exponentially increased computing time with the number of network nodes and a growing body of evidence for parcellation-dependent human brain networks (Arslan et al., 2018).

Smoothing size of CT maps. In this study, individual CT maps were smoothed using a Gaussian kernel with 15-mm full width at half maximum, which was smaller than those for the other three morphological indices (Gaussian kernel with 25-mm full width at half maximum). To test whether the differences in smoothing size could lead

to relatively poor performance in TRT reliability for CT-based morphological similarity networks (see Results), we smoothed individual CT maps again (BNU TRT dataset) using a Gaussian kernel with 25-mm full width at half maximum, followed by network parameter and ICC calculation.

Brain size and regional size. Brain size is an important confounding factor for analysis of brain morphology. In addition, previous functional (Wang et al., 2009) and morphological (Seidlitz et al., 2018) brain network studies showed the existence of relationships between regional size and nodal centrality. Thus, in this study we calculated cross-subject Pearson correlation between each global network measure and global morphological values, and cross-node, person-level Pearson correlation between each nodal centrality measure and regional size (defined as the number of vertex in a region) for each type of morphological similarity networks.

Results

Interregional similarities of morphological similarity networks

Similarity matrices. Unless stated otherwise, all results reported are exemplified using the a2009s atlas for brain parcellation and the *JSD*s for interregional similarity estimation.

Figure 2 shows the mean morphological similarity matrices derived from the four morphological indices of FD, GI, SD and CT. In general, the cross-subject mean interregional similarity was large with relatively small variance for each edge of each type of morphological similarity networks (FD: 0.725 ± 0.074; GI: 0.724 ± 0.069; SD:

447 0.717 ± 0.121; CT: 0.737 ± 0.072). Nonetheless, morphological index-dependent
448 similarity patterns were evident. For example, the SD-based morphological similarity
449 networks were clearly different from the others by visual inspection. This was further
450 confirmed by the relatively low Spearman's rank correlations in the mean similarity
451 matrix between any pair of morphological similarity networks ($r_{FD\text{-}GI}$ = 0.213; $r_{FD\text{-}SD}$ =
452 0.136; $r_{FD\text{-}CT}$ = 0.156; $r_{GI\text{-}SD}$ = 0.156; $r_{GI\text{-}CT}$ = 0.304; $r_{SD\text{-}CT}$ = 0.289). Interestingly, we
453 consistently found that the mean interregional similarity for edges linking
454 geometrically corresponding regions between two hemispheres (i.e., homotopic
455 connections) was significantly higher than that for edges linking nonhomotopic
456 regions (i.e., heterotopic connections) regardless of the types of morphological
457 similarity networks (t-test; FD: $P \approx 0$; GI: $P \approx 0$; SD: $P \approx 0$; CT: $P \approx 0$).

458 Both descriptive and inferential statistical results mentioned above were
459 consistently observed regardless choices of brain parcellation atlas and similarity
460 measure.

461 ***Effects of morphological index, brain parcellation and similarity measure***.
462 Significant differences were observed in the mean interregional morphological
463 similarity across all edges when using either the a2009s or the a2005s atlases for brain
464 parcellation ($P \approx 0$). Further analyses of the interregional morphological similarity of
465 each edge under both atlases consistently revealed that all edges were significantly
466 affected by choices of at least one factor of morphological index or similarity measure
467 ($P < 0.05$, FDR corrected). These findings jointly suggest widespread influence of
468 different analytical strategies on interregional similarity patterns of morphological

similarity networks. Notably, consistent with our previous study (Wang et al., 2016), no post hoc analyses were conducted since comparisons of numerical values have no utility with respect to the selection of optimal analytical strategies.

Global organization of morphological similarity networks

Small-worldness, efficiency and modularity. Each individual morphological similarity network exhibited small-world organization, high parallel efficiency and modular structure over the entire sparsity range regardless of the morphological indices. That is, compared with matched random networks, individual morphological similarity networks had higher C_p and approximately equal L_p, higher E_{loc} and approximately equal E_{glob}, and higher Q (Fig. 2). These organizational principles were consistently observed under other combinations of brain parcellation atlases and similarity measures.

Effects of morphological index, brain parcellation and similarity measure. Despite the common organizational principles, three-way repeated ANOVA revealed that quantitative values of all global network measures were significantly affected by choices of morphological index, brain parcellation atlas and/or similarity measure ($P < 0.05$, FDR corrected). Again, no post hoc analyses were further conducted, as explained above.

Local organization of morphological similarity networks

Hubs. Visual inspection indicated that the overall patterns of nodal centrality

491 measures were spatially heterogeneous and depended on different analytical strategies.
492 Thus, we first calculated Pearson correlation coefficients between each pair of nodal
493 centrality maps (4 morphological indices × 2 similarity measures × 3 nodal centrality
494 measures) averaged across all participants under each parcellation atlas. The
495 significance level for each correlation coefficient was estimated using a recently
496 proposed approach that corrects for spatial autocorrelation of brain maps (Burt et al.,
497 2020). Specifically, for each pair of nodal centrality maps used in the correlation
498 analysis, one of them was randomly selected as a target map and 10,000 surrogate
499 maps were generated that matched with the target map with respect to spatial
500 autocorrelation. Each of the surrogate maps was then used to re-compute the Pearson
501 correlation coefficient with the other map, yielding a null distribution for the expected
502 value of Pearson correlation by chance. Subsequently, a *P*-value was estimated as the
503 fraction of surrogate maps which generated a Pearson correlation equal to or greater
504 than the real Pearson correlation. Of note, variability into the estimate of the *P*-value
505 introduced by finite sampling size was accounted for by using a binomial distribution
506 to estimate the size of sampling fluctuations. As shown in Figure 3, we found that
507 under each parcellation atlas different nodal centrality measures exhibited highly
508 similar spatial patterns regardless of the choices of similarity measure (a2009s: r =
509 0.891 ± 0.087; a2005s: r = 0.714 ± 0.214), while the spatial similarities became
510 dramatically low when the correlations were calculated between different
511 morphological indices (a2009s: r = 0.306 ± 0.053; a2005s: r = 0.312 ± 0.134). These
512 findings indicate that the factor of morphological index dominates spatial distributions

of nodal centrality measures. Notably, the correlation pattern in regional centrality profiles was obviously different from the correlation pattern in regional mean morphological values (Table S2). For example, compared with the moderate negative correlations in regional mean values between GI and CT (a2009s: r = -0.476 ± 0.083; a2005s: r = -0.501 ± 0.083), low positive correlations were observed in regional centrality profiles between GI-based and CT-based morphological similarity networks (a2009s/*JSDs*: r = 0.363 ± 0.048; a2009s/*KLDs*: r = 0.366 ± 0.047; a2005s/JSDs: r = 0.276 ± 0.087; a2005s/KLDs: r = 0.278 ± 0.087). The discrepancies suggest that our findings are not affected by the existence of some spatial correlations in regional mean values between different morphological indices.

Furthermore, given central roles of highly connected regions (i.e., hubs), we identified hubs, which were defined as regions with values in the top 10% of each nodal centrality measure. We found several common features of the identified hubs, despite differential spatial distributions across different morphological indices and nodal centrality measures (Fig.S1). First, the majority of hubs was located at bilaterally homologous regions, with the ratio of such hubs to all hubs varying between 40.0% and 80.0% (mean = 62.2%) among different morphological indices and nodal centrality measures. Second, most hubs were located at brain sulci, especially at the junction between different lobes, with the ratio of such hubs to all hubs varying between 60.0% and 86.7% (mean = 75.0%) among different morphological indices and nodal measures. Finally, several regions were consistently identified hubs that were independent of choices of morphological index and nodal

centrality measure. These features also existed under other combinations of brain parcellation atlases and similarity measures (except for the second feature when regional parcellation was based on the a2005s atlas, which was composed of only gyral-based regions).

Finally, we defined a consistent hub score (*CHs*), which was calculated as the number of times a region was identified as a hub under all combinations of morphological index, similarity measure and nodal centrality measure under each parcellation atlas. Figure 4 shows regions with the top 10% highest values of *CHs*.

Effects of morphological index, brain parcellation and similarity measure. Between-atlas comparisons of the mean centrality value across all regions revealed significant differences in nodal efficiency and betweenness ($P \approx 0$). Further two-way repeated ANOVA revealed that all nodal centrality measures of all regions were significantly affected by choices of morphological index and/or similarity measure under both the a2009s and a2005s atlases ($P < 0.05$, FDR corrected).

TRT reliability of interregional similarities of morphological similarity networks

ICC of similarity matrices. Based on the BNU TRT dataset, extremely high correlations across edges were observed in the mean similarity matrices between the two sessions regardless of different morphological indices (all $r > 0.966$, Fig. S2). Further TRT reliability analysis on the interregional similarity of each edge revealed that 99.8%, 88.8%, 96.6% and 80.6% of all edges exhibited good or above TRT reliability (i.e., $ICC > 0.6$) for the FD-, GI-, SD- and CT-based morphological

similarity networks, respectively. The high TRT reliability was also observed under other combinations of brain parcellation atlases and similarity measures.

Effects of morphological index, brain parcellation and similarity measure. First, between-atlas comparison of all connectional *ICC* values (concatenated across different morphological indices and similarity measures) revealed that the TRT reliabilities were significantly higher for the a2009s than a2005s atlas (ICC_{a2009s} = 0.801 ± 0.135, ICC_{a2005s} = 0.782 ± 0.149; t_{102902} = 17.378, $P < 0.001$). Then, under each brain parcellation atlas, two-way repeated ANOVA consistently revealed that the mean *ICC* values across all edges were significantly modulated by choices of morphological index and similarity measure with a significant interaction (all $P \approx 0$). Post hoc analyses revealed that: 1) the *ICC* differed significantly among brain networks constructed using different morphological indices no matter the choices of brain parcellation atlas and similarity measure (FD > SD > GI > CT; all $P < 0.05$, FDR corrected across 24 t-tests in total); and 2) the *ICC* differed significantly among morphological similarity networks constructed using different similarity measures regardless of choices of brain parcellation atlases (*JSDs* > *KLDs* for FD-, GI- and SD-based morphological similarity networks, and *KLDs* > *JSDs* for CT-based morphological similarity networks; all $P < 0.05$, FDR corrected across 8 t-tests in total). Finally, we located edges whose TRT reliabilities were affected by choices of morphological index and similarity measure. The results showed that: 1) no edges exhibited significant differences in TRT reliabilities when morphological similarity networks were constructed using *JSDs* or *KLDs* for interregional similarity estimation

579 ($P > 0.05$, FDR corrected); and 2) the TRT reliabilities of up to 7.3% of the edges for
580 the a2009s atlas and of up to 19.2% of the edges for the a2005s atlas were
581 significantly different when morphological similarity networks were constructed
582 using different morphological indices ($P < 0.05$, FDR corrected). Interestingly, the
583 TRT reliabilities of the identified edges exhibited a common pattern of FD > SD > GI >
584 CT (Fig. S3).
585
586 **TRT reliability of global network measures of morphological similarity networks**
587 ***ICC of global network measures***. Individual morphological similarity networks
588 exhibited small-world organization, high parallel efficiency and modular structure
589 regardless of the data sessions (Fig. S4). The *ICC* analysis further revealed that most
590 global network measures exhibited fair to excellent TRT reliabilities (Fig. S5).
591 ***Effects of morphological index, brain parcellation and similarity measure***.
592 Three-way repeated ANOVA revealed that the *ICC* of global measures were
593 significantly modulated by the factors of morphological index ($F_{3,159} = 105.728$, $P \approx 0$)
594 and similarity measure ($F_{1,159} = 5.296$, $P = 0.047$), and the modulation depended on
595 choices of brain parcellation atlas ($F_{3,159} = 8.349$, $P < 0.001$ for the interaction
596 between morphological index and brain parcellation atlas, and $F_{1,159} = 15.351$, $P =$
597 0.004 for the interaction between similarity measure and brain parcellation atlas) (Fig.
598 5). Further post hoc analyses revealed that: 1) the TRT reliabilities of global network
599 measures were significantly higher for *JSDs*-based than for *KLDs*-based
600 morphological similarity networks under the a2005s ($t_{39} = 2.834$, $P = 0.007$) but not

the a2009s (t_{39} = 0.093, P = 0.926) atlas; and 2) the TRT reliabilities of global network measures differed significantly between any pair of morphological indices with a pattern of SD > FD > CT > GI under the a2005s atlas (all P < 0.05, FDR corrected), while they were significantly higher for the FD- and SD- than for the GI- and CT-based morphological similarity networks under the a2009s atlas (P < 0.05, FDR corrected). Finally, we identified specific global measures whose TRT reliabilities were affected by choices of morphological index, brain parcellation atlas and similarity measure. The results showed that: 1) no global measures exhibited significant differences in TRT reliabilities when morphological similarity networks were constructed using the a2009s or a2005s atlas for brain parcellation (P > 0.05, FDR corrected); 2) no global measures exhibited significant differences in TRT reliabilities when morphological similarity networks were constructed using *JSDs* or *KLDs* for interregional similarity estimation (P > 0.05, FDR corrected); and 3) the TRT reliabilities of many global measures were significantly affected by choices of morphological index (P > 0.05, FDR corrected) with a common pattern of FD- and SD- > GI- and CT-based morphological similarity networks (Fig. S6).

TRT reliability of nodal network measures of morphological similarity networks

ICC of nodal centralities. Similar to nodal centrality maps, the overall patterns of nodal *ICC* were spatially heterogeneous and depended on different analytical strategies. Thus, we also calculated the Pearson correlation between any pair of *ICC* maps under each atlas (4 morphological indices × 2 similarity measures × 3 nodal

623 measures). Again, the significance levels of the resultant correlation coefficients were
624 estimated using the approach from (Burt et al., 2020) to correct for spatial
625 autocorrelation of brain maps. The results showed high correlation coefficients
626 between different nodal centrality measures regardless of the choice of similarity
627 measure (a2009s: $r = 0.869 \pm 0.139$; a2005s: $r = 0.574 \pm 0.326$). However, the
628 correlation coefficients between different morphological indices were dramatically
629 low (a2009s: $r = 0274 \pm 0.145$; a2005s: $r = 0.258 \pm 0.120$) (Fig. S7).
630 We further divided all brain regions into five categories in terms of their *ICC*
631 values. As shown in Figures 6 and S8, a considerable proportion of brain regions
632 exhibited good to excellent TRT reliabilities for most morphological indices and
633 nodal centrality measures. Nevertheless, it was evident that the proportions were
634 higher for the FD- and SD- than for the GI- and CT-based morphological similarity
635 networks and were higher for nodal degree and efficiency than for nodal betweenness.
636 Finally, we identified regions under each brain parcellation atlas that consistently
637 exhibited high TRT reliabilities regardless of different analytical strategies by
638 defining a consistent excellent reliability score (*CERs*) as the number of times a
639 region exhibited excellent TRT reliability ($ICC > 0.75$) under all combinations of
640 morphological indices, similarity measures and nodal metrics. Figure 7 shows the
641 regions with the top 10% highest values of *CERs*.
642 ***Effects of morphological index, brain parcellation and similarity measure***.
643 First, comparison of all nodal *ICC* values (concatenated across different
644 morphological indices and similarity measures) between the two parcellation atlases

showed that the TRT reliabilities were significantly higher for the a2009s than for the a2005s atlas for each nodal measure (k_i: ICC_{a2009s} = 0.803 ± 0.135, ICC_{a2005s} = 0.778 ± 0.149; t_{1710} = 3.557, $P \approx 0$; e_i: ICC_{a2009s} = 0.803 ± 0.136, ICC_{a2005s} = 0.773 ± 0.154; t_{1710} = 4.081, $P \approx 0$; b_i: ICC_{a2009s} = 0.665 ± 0.189, ICC_{a2005s} = 0.624 ± 0.207; t_{1710} = 4.023, $P \approx 0$).

Then, under each brain parcellation atlas, two-way repeated ANOVA revealed that choices of morphological index and similarity measure significantly affected the mean *ICC* of nodal degree (all $P < 0.004$) and nodal efficiency (all $P < 0.006$) with nonsignificant interaction effects (all $P > 0.05$). For nodal betweenness, significant effects were only observed for the factor of morphological index under both the a2009s and a2005s atlases (both $P \approx 0$). Further post hoc analyses revealed that 1) the TRT reliabilities of nodal degree and nodal efficiency were higher for *JSDs*-based than *KLDs*-based morphological similarity networks under each brain parcellation atlas (all $P < 0.05$, FDR corrected across 4 t-tests in total); and 2) the TRT reliabilities of nodal degree and efficiency exhibited a pattern of FD- > SD- > GI- > CT-based morphological similarity networks under the a2009s atlas and a pattern of FD- and SD- > GI- > CT-based morphological similarity networks under the a2005s atlas; and nodal betweenness exhibited a pattern of SD- > FD- > GI- > CT-based morphological similarity networks for both atlases (all $P < 0.05$, FDR corrected across 36 t-tests in total).

Finally, we identified regions whose TRT reliabilities were affected by choices of morphological index and similarity measure. We found that: 1) no regions exhibited

667 significant differences in TRT reliabilities for any nodal centrality measure when
668 morphological similarity networks were constructed using *JSDs* or *KLDs* for
669 interregional similarity estimation ($P > 0.05$, FDR corrected); and 2) the TRT
670 reliabilities of up to 43.2% of the regions under the a2009s atlas and of up to 61.8% of
671 the regions under the a2005s atlas were significantly different when morphological
672 similarity networks were constructed using different morphological indices ($P < 0.05$,
673 FDR corrected). Interestingly, the TRT reliabilities of the identified regions exhibited
674 a common pattern of FD- and SD- > GI- and CT-based morphological similarity
675 networks (Figs. S9-S12).

676 ***Differences in TRT reliability between hub and non-hub regions.*** We compared
677 the TRT reliability between hub and non-hub regions under each analytical
678 combination of morphological index, brain parcellation atlas and similarity measure
679 (permutation test, 10,000 times). No significant differences were found regardless of
680 the analytical strategies ($P > 0.05$, FDR corrected).

681

682 **Effects of sample size on individual morphological similarity networks**

683 Highly positive correlations were observed between the nodal centrality profiles
684 (concatenated across nodal measures) averaged over 10 randomly selected samples
685 and the mean nodal centrality profile across all participants ($r_{FD} = 0.944 \pm 0.007$; $r_{GI} =$
686 0.936 ± 0.007; $r_{SD} = 0.970 \pm 0.005$; $r_{CT} = 0.936 \pm 0.008$). With increasing sample size,
687 the correlations continually increased and the variability naturally decreased, with a
688 tipping point observed when samples exceeded ~70 participants ($r_{FD} = 0.992 \pm 0.001$;

689 r_{GI} = 0.991 ± 0.001; r_{SD} = 0.996 ± 0.001; r_{CT} = 0.991 ± 0.001) (Fig. 8). More
690 importantly, this feature was observed under other combinations of brain parcellation
691 atlases and similarity measures.
692
693 **Reproducibility of our results**
694 ***Effects of twin subjects.*** After excluding twin subjects from the HCP S900
695 dataset, we found that all results reported remained largely unchanged, indicating
696 little effects of twin subjects on our results. The new results are briefly summarized
697 below.
698 1) *Edge-level analysis*. we found: a) relatively low Spearman's rank correlations
699 in the mean similarity matrix between any pair of morphological similarity networks
700 ($r_{FD\text{-}GI}$ = 0.240; $r_{FD\text{-}SD}$ = 0.154; $r_{FD\text{-}CT}$ = 0.170; $r_{GI\text{-}SD}$ = 0.167; $r_{GI\text{-}CT}$ = 0.304; $r_{SD\text{-}CT}$ =
701 0.304); b) higher mean interregional similarity for edges linking homotopic regions
702 than edges linking heterotopic regions regardless of the types of morphological
703 similarity networks (FD: t = 13.833, $P \approx 0$; GI: t = 12.321, $P \approx 0$; SD: t = 20.844, $P \approx$
704 0; CT: t = 15.821, $P \approx 0$); c) significant differences in the mean interregional
705 morphological similarity between the a2005s atlas and a2009s atlas (t = 91.119, $P \approx$
706 0); and d) significant effects of morphological index and/or similarity measure on the
707 interregional morphological similarity of each edge under the two parcellation atlases
708 (two-way repeated ANOVA; $P < 0.05$, FDR corrected).
709 2) *Global-level analysis*. We found that all global network measures were
710 significantly affected by choices of morphological index, brain parcellation atlas

and/or similarity measure (three-way repeated ANOVA, $P < 0.05$, FDR corrected).

3) *Nodal-level analysis*. We found: a) highly similar spatial patterns among different nodal centrality measures for the same morphological indices (a2009s: $r = 0.869 \pm 0.084$; a2005s: $r = 0.844 \pm 0.139$) but dramatically low spatial similarities between different morphological indices (a2009s: $r = 0.305 \pm 0.053$; a2005s: $r = 0.267 \pm 0.144$); b) significant differences in the mean centrality across all regions between the a2005s atlas and a2009s atlas (nodal efficiency: $t = 230.836$, $P \approx 0$; nodal betweenness: $t = 560.849$, $P \approx 0$); and c) significant effects of morphological index and/or similarity measure on each nodal centrality measure of each region under the two atlases (two-way repeated ANOVA; $P < 0.05$, FDR corrected).

4) *Analysis of sample size*. We found that the correlation of mean nodal centrality profiles between randomly selected samples and all participants continually increased with increasing sample size, with a tipping point observed when samples exceeded ~70 participants ($r_{FD} = 0.992 \pm 0.001$; $r_{GI} = 0.991 \pm 0.001$; $r_{SD} = 0.996 \pm 0.001$; $r_{CT} = 0.991 \pm 0.001$).

Effects of high-resolution parcellation atlas. Based on randomly selected 50 participants from the HCP S900 dataset, we found the existence of small-worldness, high parallel efficiency, modularity and hubs for surface-based single-subject morphological similarity networks. Moreover, these network measures derived from the MMP atlas were significantly different from those derived from the a2005s atlas and a2009s atlas, and were dependent on choices of morphological index and/or similarity measure. All these findings are as expected and in line with the main

results.

Based on the BNU TRT dataset, we found that compared with the a2005s atlas and a2009s atlas, the MMP atlas was associated with significantly higher TRT reliabilities in interregional morphological similarities (a2005s: 0.782±0.149; a2009s: 0.801±0.135; MMP: 0.817±0.125), global (a2005s: 0.644±0.208; a2009s: 0.644±0.140; MMP: 0.673±0.130) and local (a2005s: 0.725±0.186; a2009s: 0.757±0.169; MMP: 0.773±0.152) network measures. It should be noted that this general conclusion does not always hold for global network measures when considering the other two factors of morphological index and similarity measure. All TRT reliability results in this study are summarized in Table 1.

Effects of smoothing size. Based on the BNU TRT dataset, we found that the usage of a larger smoothing kernel (25-mm full width at half maximum) resulted in significantly higher TRT reliabilities for interregional morphological similarities and nodal centralities but lower TRT reliabilities for global network measures for the CT-based morphological similarity networks (paired t-test; $P < 0.05$, Bonferroni corrected; Table S3). Nonetheless, the differences did not affect our general findings that CT-based morphological similarity networks performed worse that the other three types of morphological similarity networks with respect to TRT reliability. That is, no matter which smoothing size (15-mm or 25-mm full width at half maximum) was used, significantly lower TRT reliabilities were found for the CT-based than FD-, GI- and SD-based morphological similarity networks for interregional morphological similarities, global network measures and nodal centralities (paired t-test; $P < 0.05$,

Bonferroni corrected; Table S3). These findings indicate that differences in smoothing kernel among different morphological indices contribute little to the observed differences in TRT reliabilities among different types of morphological similarity networks.

Relationships between morphological similarity networks and brain/regional size. Very low correlations were found between each global network measure and global morphological values regardless of the type of morphological similarity networks (Table S4). These findings suggest a lack of relationships between the morphological similarity networks and brain size. For nodal centrality measures, weak-moderate correlations were observed with regional size that were dependent on morphological index, brain parcellation atlas, similarity measure and nodal centrality type (Table S5). These findings indicate obvious effects of regional size on nodal centrality for the surface-based single-subject morphological similarity networks, an issue that should be taken into account in future studies. The findings of brain size-independent but regional size-dependent morphological similarity networks are in accordance to a previous study that constructed single-subject morphological similarity networks with different methods (Seidlitz et al., 2018).

Discussion

In this study, we constructed and topologically characterized surface-based single-subject morphological similarity networks, and systematically evaluated their reproducibility with respect to effects of different analytical strategies, sample

size-varying stability and test-retest reliability. We found that the morphological similarity networks exhibited nontrivial organizational principles, including small-worldness, high parallel efficiency, modularity and hubs, regardless of the analytical strategies used. Nevertheless, quantitative values of these organizational principles largely depended on the choices of morphological index, brain parcellation atlas and similarity measure. Moreover, the morphological similarity networks varied with the number of participants and approached stability until the sample size exceeded ~70. Finally, both interregional similarities and topological properties of the morphological similarity networks presented fair to good, even excellent, TRT reliabilities. Interestingly, significantly higher reliabilities were observed when interregional morphological similarity was estimated with the *JSDs* than with the *KLDs* and when the morphological similarity networks were based on the FD and SD rather than the GI and CT. Altogether, these findings suggest that surface-based single-subject morphological similarity networks provide a reliable approach for large-scale brain network studies.

Specifically organized and reliable morphological similarity networks

The human brain is powerful in both modular information processing across local regions and distributed information processing over the entire brain, allowing complicated cognition and behavior. A large number of studies have indicated that such powerful features of the human brain benefit from nontrivial wiring layouts of both functional and structural brain networks as revealed by graph-based network

799 approaches, such as small-worldness, high parallel efficiency, modularity and hubs
800 (Bullmore and Sporns, 2012; Liao et al., 2017; Sporns and Betzel, 2016; van den
801 Heuvel and Sporns, 2013). Therefore, in this study we also utilized graph-based
802 metrics to examine whether the surface-based single-subject morphological similarity
803 networks constructed with our method are governed by nontrivial organizational
804 principles rather than are wired in a random manner. We found that the surface-based
805 single-subject morphological similarity networks also exhibited the optimized
806 organization. That is, compared with matched random networks, the surface-based
807 single-subject morphological similarity networks showed higher clustering coefficient,
808 local efficiency and modularity, approximately equal characteristic path length and
809 global efficiency and the existence of hubs. Moreover, the optimized organization was
810 consistently observed regardless of different choices of morphological index, brain
811 parcellation atlas and similarity measure. The robust and inherent characteristic
812 indicates that the surface-based single-subject morphological similarity networks are
813 specifically organized rather than are wired in a random manner. It should be noted
814 that the interpretation of graph-based network findings largely depends on how
815 network nodes and edges are defined. In this study, network edges were defined as
816 divergence-based similarity in the distribution of intraregional morphological values.
817 This type of similarity is very different from routinely used functional connectivity
818 (i.e., statistical interdependence in functional signal fluctuations between regions) and
819 structural connectivity (i.e., fiber pathways between regions). Thus, our findings are
820 not directly comparable with previous functional and structural brain networks. It is

interesting to investigate the relationships between the divergence-based morphological similarity and functional/structural connectivity in the future, which can aid in understanding the extent to which different modalities of brain networks are driven by similar organizational principles.

An interesting finding in this study was that consistent hub regions were mainly in cortical sulci under the a2009s atlas. Compared with cortical gyri, cortical sulci have been found to show greater physical strain in various kinds of injuries in studies from animals, computational models and human postmortem data (Cloots et al., 2008; Ghajari et al., 2017; Goldstein et al., 2012). Accordingly, cortical sulci are frequently reported to be more vulnerable to pathological tau protein deposition and associated brain atrophy in patients with traumatic brain injury and chronic traumatic encephalopathy (Cole et al., 2018; Johnson et al., 2012; McKee et al., 2013). On the other hand, brain network hubs are also found to be susceptible to various brain disorders (Crossley et al., 2014). Thus, our results provide a possible bridge to link the two sets of previous findings, although a direct examination is needed on the relationships among brain sulci and gyri, hubs and diseases. Finally, the surface-based single-subject morphological similarity networks presented fair to good, even excellent, reliability in both interregional similarities and network measures under all analytical combinations. This is consistent with previous studies that constructed individual-level morphological similarity networks with similar methods in brain volume space (Kong et al., 2015; Wang et al., 2016) and via different methods (Jiang et al., 2017; Li et al., 2017; Yu et al., 2018).

Taken together, our findings suggest that surface-based single-subject morphological similarity networks can provide an important alternative to reliably uncover organizational principles of large-scale brain networks. Nevertheless, it should be pointed out that although optimized topological organizations and high TRT reliabilities were consistently observed for the surface-based single-subject morphological similarity networks regardless of different analytical strategies, quantitative values depended largely on how the morphological similarity networks were constructed, which is discussed below.

Effects of different morphological indices on morphological similarity networks

We found that different morphological indices resulted in significantly different values for interregional morphological similarities and network measures. This is consistent with previous studies of population-based morphological similarity networks showing morphological index specific topological organization in health and disease (Collantoni et al., 2017; Sanabria-Diaz et al., 2010). Given the highly complex, folded nature of the human cerebral cortex, these findings are not surprising because different indices characterize morphological architecture from different aspects. For example, CT reflects the size, density and arrangement of cells (neurons, neuroglia and nerve fibers), whereas GI measures the ratio between the total area of the cortex and the area that is visible in a circular region of interest. Thus, the observed differences may indicate that different morphological indices capture distinct processes of interregional interactions or different aspects of the same

interactive process (e.g., mechanical, neurochemical, and/or axonal connections). Currently, understanding mechanisms behind different morphological indices and their interrelationships is an ongoing research field. Evidence from genetic and developmental studies has shown that different morphological indices are associated with distinct genetic influences (Panizzon et al., 2009; Strike et al., 2019; Winkler et al., 2010) and exhibit differential developmental/aging trajectories (Hogstrom et al., 2013; Raznahan et al., 2011; Wierenga et al., 2014). Accordingly, it is plausible to speculate that genetic, developmental/aging, and environmental factors may partly contribute to the distinct network topology among different morphological indices. It should be noted that the explanations above are speculative and more studies are needed to provide direct empirical evidence for mechanistic understanding of network differences among different morphological indices.

Furthermore, different morphological indices were associated with significantly different values of TRT reliability of the morphological similarity networks. Specifically, TRT reliability was significantly higher for FD- and SD- than for GI- and CT-based morphological similarity networks. The discrepancy may be due to differences in computational complexity, sensitivity to noise, and/or developmental rates among the morphological indices. For example, in contrast with the intuitional CT, FD is an extremely obscure and compact measure of shape complexity, which condenses all details into a single numeric value. Such summary measures may be more resistant to noise than those that are dominated by a single aspect of brain morphology. In addition, given distinct age-related trajectories among different

887 morphological indices (Raznahan et al., 2011; Wierenga et al., 2014), those indices
888 showing faster age-related changes are typically related to lower values of TRT
889 reliability due to greater within-subject variance. Overall, our findings suggest that
890 future studies should consider utilizing different morphological indices, in particular
891 SD and FD from the perspective of TRT reliability, to provide a finer-grained
892 characterization of surface-based single-subject morphological similarity networks in
893 typical and atypical populations.
894
895 **Effects of different brain parcellation atlases on morphological similarity**
896 **networks**
897 We found that different choices of the a2009s and a2005s atlases resulted in
898 significantly different surface-based single-subject morphological similarity networks
899 in terms of the values of interregional morphological similarities and network
900 measures. This is consistent with numerous previous studies showing brain
901 parcellation-dependent functional (Ren et al., 2019; Wang et al., 2009), structural
902 (Wei et al., 2017; Zalesky et al., 2010) and morphological (Sanabria-Diaz et al., 2010;
903 Wang et al., 2016) brain networks (for recent reviews, see (Arslan et al., 2018; Qi et
904 al., 2015; Yao et al., 2015)). Accordingly, our findings together with previous studies
905 collectively suggest that the dependence on regional brain parcellation atlases is a
906 universal characteristic of large-scale brain networks, and future studies must consider
907 the influence of this factor regardless of the data modalities from which the brain
908 networks are obtained. Several potential sources may account for the observed

differences. First, the composition is different between the two atlases: the a2009s atlas is composed of sulco-gyral structures, while the a2005s atlas is made up of gyral-based neuroanatomical regions (Desikan et al., 2006; Destrieux et al., 2010). Given substantial differences between cerebral sulci and gyri in genetics, morphology, axonal pathways and function (Ge et al., 2018; Hilgetag and Barbas, 2005; Li et al., 2015; Liu et al., 2017; Zeng et al., 2015; Zhang et al., 2018), it is not surprising to observe differential network organization between the two atlases. Second, the number of regions is different between the two atlases: 148 in the a2009s atlas versus 68 in the a2005s atlas. Previous studies have shown that differences in network size (i.e., the number of nodes) significantly affect the topological organization of functional and structural brain networks (Wang et al., 2009; Zalesky et al., 2010). Finally, the two atlases differ in regional size, which is thought to be associated with nodal centralities of functional and structural brain networks (Hagmann et al., 2008; Wang et al., 2009). Together, all these factors may at least partially account for the observed brain parcellation-related differences in surface-based single-subject morphological similarity networks.

Further TRT reliability analyses revealed that the a2009s atlas was associated with higher reliability than the a2005s atlas for the surface-based single-subject morphological similarity networks. When using the high-resolution MMP atlas, even higher TRT reliability was observed. These findings indicate that brain parcellation atlases with higher resolutions may give rise to more reliable morphological similarity networks. The discrepancy may be attributable to the more sophisticated methods

used to generate higher-resolution atlases: most structures in the a2005s atlas were defined using a relatively coarse 'sulcal' approach (manual tracing from the depth of one sulcus to another, thus incorporating the gyrus within) (Desikan et al., 2006); the entire cortical surface of the a2009s atlas was classified as gyral or sulcal at a vertex level during the generation (Destrieux et al., 2010); and the MMP atlas is based on a multi-modal method that integrates multidimensional information from different imaging modalities (e.g., architectural measures derived from T1-weighted and T2-weighted structural images, cortical function measured using task functional MRI, and functional connectivity estimated from resting-state functional MRI). Currently, obtaining accurate and reliable parcellation atlases of the human brain is an important goal for brain science. In this regard, individualized, multidimensional information guided and phylogeny and ontogeny inspired parcellation methods may be promising directions in the future (Eickhoff et al., 2018). Overall, our findings suggest that relative to the a2005s atlas and a2009s atlas, the MMP atlas may be a better choice for surface-based single-subject morphological similarity network studies from the perspective of TRT reliability.

Effects of different similarity measures on morphological similarity networks

Although the morphological similarity networks constructed via *KLDs* and *JSDs* exhibited largely similar patterns, quantitative values differed significantly between the two sets of networks. This is consistent with numerous previous studies showing that functional and structural brain networks depend on the means to estimate

interregional connectivity (Liang et al., 2012; Sarwar et al., 2019; Zalesky et al., 2012). Accordingly, it seems that the dependence on interregional connectivity estimation methods in addition to regional brain parcellation atlases is another universal characteristic of large-scale brain networks, and thus future studies should carefully choose suitable measures and methods for estimating interregional connectivity in terms of their research themes, purposes and contents. The observed differences may be attributable to different mathematical properties between the *KLD* and *JSD*, although the latter is based on the former. First, the *KLD* and *JSD* have different value ranges: the *KLD* is nonnegative without an upper limit, while the *JSD* is bounded by 0 and 1. Second, the *KLD* is not symmetric, while the *JSD* is symmetric. Thus, compared with the *JSD*, the *KLD* undergoes additional processing steps (e.g., exponential transform) to generate symmetric and bounded surface-based single-subject morphological similarity networks. This may introduce extra noise or unknown disturbance in estimating interregional morphological similarities. This might also be the reason why the *JSDs*-based morphological similarity networks showed higher reliability for interregional morphological similarities and network measures. Overall, given the mathematical advantages of the *JSD* relative to the *KLD* and higher TRT reliability of the *JSDs*-based than *KLDs*-based morphological similarity networks, we recommend using the *JSDs* as a measure for estimating interregional morphological similarity in future studies. In the future, it will be necessary to systematically compare the *JSDs* with other similarity measures and methods for a possible better choice, such as multivariate Euclidean distance (Yu et

975 al., 2018) and Fréchet distance.

976

977 **Effects of different sample sizes on morphological similarity networks**

978 We found that surface-based single-subject morphological similarity networks varied
979 along with the number of participants and approached stability until the sample size
980 exceeded ~70. Moreover, the critical value was largely independent of different
981 analytical strategies. Thus, we recommend at least 70 participants for future studies of
982 surface-based single-subject morphological similarity networks. However, such a
983 sample size may be challenging to attain for innovative clinical and translational
984 research due to cost and feasibility concerns. Although small sample size undermines
985 the reliability (Button et al., 2013), it can produce more projected scientific value per
986 dollar spent than larger sample size for studies of new ideas (Bacchetti et al., 2011).
987 Interestingly, we found that even when the sample size was only 10, the mean nodal
988 centrality profile of the surface-based single-subject morphological similarity
989 networks was largely similar to that from all 876 participants. This suggests limited
990 effects of sample size on surface-based single-subject morphological similarity
991 networks. This feature makes the approach proposed here a potential tradeoff between
992 sample size and research cost and feasibility for future brain network studies.
993 Presumably, the weak influence of sample size may reflect small interindividual
994 differences in surface-based single-subject morphological similarity networks. An
995 interesting topic for future exploration is the association of interindividual differences
996 in surface-based single-subject morphological similarity networks with interindividual

997 differences in cognition and behavior.

998

999 **Limitations and Future Directions**

1000 First, consistent with numerous functional and structural brain network studies, the
1001 surface-based single-subject morphological similarity networks were largely
1002 dependent on choices of regional brain parcellation atlases and similarity estimation
1003 methods. Thus, systematic research on different processing pipelines is warranted in
1004 the future for establishing a "better" methodological framework in the construction of
1005 surface-based single-subject morphological similarity networks. Second, although the
1006 surface-based single-subject morphological similarity networks exhibited high TRT
1007 reliability under different analytical strategies, their repeatability across multiple sites,
1008 scanners, scanning parameters and magnetic fields should be further examined. Third,
1009 similar to functional and structural brain networks, surface-based single-subject
1010 morphological similarity networks constructed here also exhibited optimized
1011 topological organizations (e.g., small-worldness and hubs). It will be interesting in the
1012 future to quantify the similarities and differences in organizational principles between
1013 morphological and functional/structural brain networks. Fourth, this study extended
1014 our previous work of single-subject morphological similarity networks from volume
1015 space to cerebral cortical surface. A previous functional MRI study showed that
1016 surface-based computation can increase TRT reliability of local short-range functional
1017 connectivity (Zuo et al., 2013). Whether surface-based morphological similarity
1018 networks outperform volume-based morphological similarity networks in TRT

reliability is thus an interesting topic in the future. Fifth, only four surface-based morphological indices that were computationally available for the CAT12 toolbox were used to construct morphological similarity networks in this study. Future studies can examine the feasibility of our method on the basis of other morphological indices, such as surface area and surface normal. Sixth, in this study, we found that different morphological indices were associated with distinct interregional similarity patterns and topological organization of surface-based single-subject morphological similarity networks. It is important to develop network models to integrate the complementary information for a holistic view of morphological similarity networks. At the current stage, multilayer network methods may be a good solution for such a requirement (De Domenico, 2017; Vaiana and Muldoon, 2018). Seventh, we utilized binary rather than weighted network model to characterize morphological similarity networks in this study because weighted networks are computationally expensive. This is particularly important for this study because of the huge amount of computation. It's essential for future studies to employ weighted network model, which can provide more information on the topological organization of morphological similarity networks, and may be more sensitive to capture morphological network alterations under conditions where interregional morphological similarities alter profoundly, such as diseases, development and aging. Finally, after demonstrating the reliable, nonrandom organization of the divergence-based single-subject morphological similarity networks, the next important thing is to uncover biological meaning of the networks. For example, to what extent are the morphological similarity networks under genetic

control and to what extent do the morphological similarity networks determine individual behavioral and cognitive performance? Such studies are crucial to speed up future application of the divergence-based single-subject morphological similarity networks in health and disease.

Conclusion

In conclusion, this study constructed and evaluated surface-based single-subject morphological similarity networks and demonstrated that the morphological similarity networks possessed nontrivial topological organization, were affected by different analytical strategies but largely independent of sample size, and exhibited high TRT reliability. Based on these findings, we conclude that the surface-based single-subject morphological similarity networks can serve as a reliable way to characterize large-scale brain networks in future studies.

Conflict of Interests

The authors declare no competing interests.

Data Availability Statement

All data that support the findings of this study are from publicly available datasets.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos.

1063 81922036 and 81671764), Key Realm R&D Program of Guangdong Province (No.
1064 2019B030335001) and Key Realm R&D Program of Guangzhou (No.
1065 202007030005). Data were provided [in part] by the Human Connectome Project,
1066 WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil;
1067 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH
1068 Blueprint for Neuroscience Research; and by the McDonnell Center for Systems
1069 Neuroscience at Washington University.

References

Alexander-Bloch, A., Giedd, J.N., Bullmore, E., 2013a. Imaging structural co-variance between human brain regions. Nature Reviews Neuroscience 14, 322.

Alexander-Bloch, A., Raznahan, A., Bullmore, E., Giedd, J., 2013b. The convergence of maturational change and structural covariance in human cortical networks. Journal of Neuroscience 33, 2889-2899.

Arslan, S., Ktena, S., Makropoulos, A., Robinson, E., Rueckert, D., Parisot, S., 2018. Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex. Neuroimage 170, 5-30.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. Neuroimage 38, 95-113.

Bacchetti, P., Deeks, S.G., McCune, J.M., 2011. Breaking free of sample size dogma to perform innovative translational research. Science translational medicine 3, 87ps24-87ps24.

Bassett, D.S., Bullmore, E., Verchinski, B.A., Mattay, V.S., Weinberger, D.R., Meyer-Lindenberg, A., 2008. Hierarchical organization of human cortical networks in health and schizophrenia. Journal of Neuroscience 28, 9239-9248.

Buchanan, C.R., Bastin, M.E., Ritchie, S.J., Liewald, D.C., Madole, J.W., Tucker-Drob, E.M., Deary, I.J. and Cox, S.R., 2020. The effect of network thresholding and weighting on structural brain networks in the UK Biobank. Neuroimage 211, 116443.

Bullmore, E., Sporns, O., 2012. The economy of brain network organization. Nature Reviews Neuroscience 13, 336-349.

Burt, J.B., Helmer, M., Shinn, M., Anticevic, A. and Murray, J.D., 2020. Generative modeling of brain maps with spatial autocorrelation. NeuroImage 220, 117038.

Button, K.S., Ioannidis, J.P., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S., Munafò, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nature Reviews Neuroscience 14, 365-376.

Cloots, R., Gervaise, H., Van Dommelen, J., Geers, M., 2008. Biomechanics of traumatic brain injury: influences of the morphologic heterogeneities of the cerebral cortex. Annals of biomedical engineering 36, 1203.

Cole, J.H., Jolly, A., de Simoni, S., Bourke, N., Patel, M.C., Scott, G., Sharp, D.J., 2018. Spatial patterns of progressive brain volume loss after moderate-severe

traumatic brain injury. Brain 141, 822-836.

Collantoni, E., Meneguzzo, P., Tenconi, E., Manara, R., Santonastaso, P., Favaro, A,. 2017. Structural Covariance Networks in Anorexia Nervosa (AN): A Multimodal Graph Theoretical Analysis. European Psychiatry 41, S282-S282.

Crossley, N.A., Mechelli, A., Scott, J., Carletti, F., Fox, P.T., McGuire, P., Bullmore, E.T., 2014. The hubs of the human connectome are generally implicated in the anatomy of brain disorders. Brain 137, 2382-2395.

Dahnke, R., Yotter, R.A., Gaser, C., 2013. Cortical thickness and central surface estimation. Neuroimage 65, 336-348.

De Domenico, M., 2017. Multilayer modeling and analysis of human brain networks. Giga Science 6, gix004.

Desai, R., Liebenthal, E., Possing, E.T., Waldron, E., Binder, J.R., 2005. Volumetric vs. surface-based alignment for localization of auditory cortex activation. Neuroimage 26, 1019-1029.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31, 968-980.

Destrieux, C., Fischl, B., Dale, A., Halgren, E., 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 53, 1-15.

Eickhoff, S.B., Yeo, B.T. and Genon, S., 2018. Imaging-based parcellations of the human brain. Nature Reviews Neuroscience 19, 672-686.

Evans, A.C., 2013. Networks of anatomical covariance. Neuroimage 80, 489-504.

Ge, F., Li, X., Razavi, M.J., Chen, H., Zhang, T., Zhang, S., Guo, L., Hu, X., Wang, X., Liu, T., 2018. Denser growing fiber connections induce 3-hinge gyral folding. Cerebral cortex 28, 1064-1075.

Ghajari, M., Hellyer, P.J., Sharp, D.J., 2017. Computational modelling of traumatic brain injury predicts the location of chronic traumatic encephalopathy pathology. Brain 140, 333-343.

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M. and Smith, S.M., 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171-178.

1136 Goldstein, L.E., Fisher, A.M., Tagge, C.A., Zhang, X.-L., Velisek, L., Sullivan, J.A.,
1137 Upreti, C., Kracht, J.M., Ericsson, M., Wojnarowicz, M.W., 2012. Chronic traumatic
1138 encephalopathy in blast-exposed military veterans and a blast neurotrauma mouse
1139 model. Science translational medicine 4, 134ra160-134ra160.

1140 Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J.,
1141 Sporns, O., 2008. Mapping the structural core of human cerebral cortex. PLoS
1142 biology 6.

1143 He, Y., Chen, Z.J., Evans, A.C., 2007. Small-world anatomical networks in the human
1144 brain revealed by cortical thickness from MRI. Cerebral cortex 17, 2407-2419.

1145 Hilgetag, C.C., Barbas, H., 2005. Developmental mechanics of the primate cerebral
1146 cortex. Anatomy and embryology 210, 411.

1147 Hogstrom, L.J., Westlye, L.T., Walhovd, K.B., Fjell, A.M., 2013. The structure of the
1148 cerebral cortex across adult life: age-related patterns of surface area, thickness, and
1149 gyrification. Cerebral cortex 23, 2521-2530.

1150 Jiang, J., Zhou, H., Duan, H., Liu, X., Zuo, C., Huang, Z., Yu, Z., Yan, Z., Initiative,
1151 A.s.D.N., 2017. A novel individual-level morphological brain networks constructing
1152 method and its evaluation in PET and MR images. Heliyon 3, e00475.

1153 Johnson, V.E., Stewart, W., Smith, D.H., 2012. Widespread tau and amyloid□beta
1154 pathology many years after a single traumatic brain injury in humans. Brain pathology
1155 22, 142-149.

1156 Kong, X., Liu, Z., Huang, L., Wang, X., Yang, Z., Zhou, G., Zhen, Z., Liu, J., 2015.
1157 Mapping individual brain networks using statistical similarity in regional morphology
1158 from MRI. PloS one 10, e0141840.

1159 Kullback, S., Leibler, R.A., 1951. On information and sufficiency. The annals of
1160 mathematical statistics 22(1), 79-86.

1161 Li, G., Liu, T., Ni, D., Lin, W., Gilmore, J.H., Shen, D., 2015. Spatiotemporal patterns
1162 of cortical fiber density in developing infants, and their relationship with cortical
1163 thickness. Human brain mapping 36, 5183-5195.

1164 Li, W., Yang, C., Shi, F., Wu, S., Wang, Q., Nie, Y., Zhang, X., 2017. Construction of
1165 individual morphological brain networks with multiple morphometric features.
1166 Frontiers in Neuroanatomy 11, 34.

1167 Liang, X., Wang, J., Yan, C., Shu, N., Xu, K., Gong, G., He, Y., 2012. Effects of
1168 different correlation metrics and preprocessing factors on small-world brain
1169 functional networks: a resting-state functional MRI study. PloS one 7.

1170 Liao, X., Vasilakos, A.V., He, Y., 2017. Small-world human brain networks:
1171 perspectives and challenges. Neuroscience & Biobehavioral Reviews 77, 286-300.

1172 Lin, Q., Dai, Z., Xia, M., Han, Z., Huang, R., Gong, G., Liu, C., Bi, Y., He, Y., 2015.
1173 A connectivity-based test-retest dataset of multi-modal magnetic resonance imaging
1174 in young healthy adults. Scientific data 2, 1-10.

1175 Liu, H., Jiang, X., Zhang, T., Ren, Y., Hu, X., Guo, L., Han, J., Liu, T., 2017.
1176 Elucidating functional differences between cortical gyri and sulci via sparse
1177 representation HCP grayordinate fMRI data. Brain research 1672, 81-90.

1178 Luders, E., Thompson, P.M., Narr, K., Toga, A.W., Jancke, L., Gaser, C., 2006. A
1179 curvature-based approach to estimate local gyrification on the cortical surface.
1180 Neuroimage 29, 1224-1230.

1181 Marcus, D.S., Harms, M.P., Snyder, A.Z., Jenkinson, M., Wilson, J.A., Glasser, M.F.,
1182 Barch, D.M., Archie, K.A., Burgess, G.C., Ramaratnam, M., Hodge, M., Horton,
1183 W., Herrick, R., Olsen, T., McKay, M., House, M., Hileman, M., Reid, E., Harwell,
1184 J., Coalson, T., Schindler, J., Elam, J.S., Curtiss, S.W., Van Essen, D.C., WU-Minn
1185 HCP Consortium, 2013. Human Connectome Project informatics: quality control,
1186 database services, and data visualization. Neuroimage 80, 202-219.

1187 Maslov, S., Sneppen, K., 2002. Specificity and stability in topology of protein
1188 networks. Science 296, 910-913.

1189 McGraw, K.O., Wong, S.P., 1996. Forming inferences about some intraclass
1190 correlation coefficients. Psychological methods 1, 30.

1191 McKee, A.C., Stein, T.D., Nowinski, C.J., Stern, R.A., Daneshvar, D.H., Alvarez, V.E.,
1192 Lee, H.-S., Hall, G., Wojtowicz, S.M., Baugh, C.M., 2013. The spectrum of disease in
1193 chronic traumatic encephalopathy. Brain 136, 43-64.

1194 Panizzon, M.S., Fennema-Notestine, C., Eyler, L.T., Jernigan, T.L., Prom-Wormley,
1195 E., Neale, M., Jacobson, K., Lyons, M.J., Grant, M.D., Franz, C.E., 2009. Distinct
1196 genetic influences on cortical surface area and cortical thickness. Cerebral cortex 19,
1197 2728-2735.

1198 Qi, S., Meesters, S., Nicolay, K., ter Haar Romeny, B.M., Ossenblok, P., 2015. The
1199 influence of construction methodology on structural brain network measures: A
1200 review. Journal of neuroscience methods 253, 170-182.

1201 Raznahan, A., Lerch, J.P., Lee, N., Greenstein, D., Wallace, G.L., Stockman, M.,
1202 Clasen, L., Shaw, P.W., Giedd, J.N., 2011. Patterns of coordinated anatomical change
1203 in human cortical development: a longitudinal neuroimaging study of maturational
1204 coupling. Neuron 72, 873-884.

1205 Ren, Y., Guo, L., Guo, C.C., 2019. A connectivity-based parcellation improved
1206 functional representation of the human cerebellum. Scientific reports 9, 1-12.

1207 Rodríguez-Cruces, R., Bernhardt, B.C., Concha, L., 2020. Multidimensional
1208 associations between cognition and connectome organization in temporal lobe
1209 epilepsy. Neuroimage 213, 116706.

1210 Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity:
1211 uses and interpretations. Neuroimage 52, 1059-1069.

1212 Sanabria-Diaz, G., Melie-García, L., Iturria-Medina, Y., Alemán-Gómez, Y.,
1213 Hernández-González, G., Valdés-Urrutia, L., Galán, L., Valdés-Sosa, P., 2010. Surface
1214 area and cortical thickness descriptors reveal different attributes of the structural
1215 human brain networks. Neuroimage 50, 1497-1510.

1216 Sarwar, T., Ramamohanarao, K., Zalesky, A., 2019. Mapping connectomes with
1217 diffusion MRI: deterministic or probabilistic tractography? Magnetic resonance in
1218 medicine 81, 1368-1384.

1219 Seibert, T.M., Brewer, J.B., 2011. Default network correlations analyzed on native
1220 surfaces. Journal of neuroscience methods 198, 301-311.

1221 Seidlitz, J., Váša, F., Shinn, M., Romero-Garcia, R., Whitaker, K.J., Vértes, P.E.,
1222 Wagstyl, K., Reardon, P.K., Clasen, L., Liu, S., 2018. Morphometric similarity
1223 networks detect microscale cortical organization and predict inter-individual cognitive
1224 variation. Neuron 97, 231-247.

1225 Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater
1226 reliability. Psychological bulletin 86, 420.

1227 Sporns, O., Betzel, R.F., 2016. Modular brain networks. Annual review of psychology
1228 67, 613-640.

1229 Strike, L.T., Hansell, N.K., Couvy-Duchesne, B., Thompson, P.M., de Zubicaray, G.I.,
1230 McMahon, K.L., Wright, M.J., 2019. Genetic complexity of cortical structure:
1231 differences in genetic and environmental factors influencing cortical surface area and
1232 thickness. Cerebral cortex 29, 952-962.

1233 Tijms, B.M., Seriès, P., Willshaw, D.J., Lawrie, S.M., 2012. Similarity-based
1234 extraction of individual networks from gray matter MRI scans. Cerebral cortex 22,
1235 1530-1541.

1236 Tijms, B.M., ten Kate, M., Gouw, A.A., Borta, A., Verfaillie, S., Teunissen, C.E.,
1237 Scheltens, P., Barkhof, F. and van der Flier, W.M., 2018. Gray matter networks and
1238 clinical progression in subjects with predementia Alzheimer's disease. Neurobiology

of Aging 61, 75-81

Tijms, B.M., Yeung, H.M., Sikkes, S.A., Möller, C., Smits, L.L., Stam, C.J., Scheltens, P., van der Flier, W.M. and Barkhof, F., 2014. Single-Subject Gray Matter Graph Properties and Their Relationship with Cognitive Impairment in Early- and Late-Onset Alzheimer's Disease. Brain Connectivity 4, 337-346.

Vaiana, M., Muldoon, S.F., 2018. Multilayer brain networks. Journal of Nonlinear Science, 1-23.

van den Heuvel, M.P., Sporns, O., 2013. Network hubs in the human brain. Trends in cognitive sciences 17, 683-696.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.-M.H., 2013. The WU-Minn human connectome project: an overview. Neuroimage 80, 62-79.

Wang, H., Jin, X., Zhang, Y., Wang, J., 2016. Single□subject morphological brain networks: connectivity mapping, topological characterization and test–retest reliability. Brain and behavior 6, e00448.

Wang, J., Wang, L., Zang, Y., Yang, H., Tang, H., Gong, Q., Chen, Z., Zhu, C., He, Y., 2009. Parcellation□dependent small□world brain functional networks: A resting□state fMRI study. Human brain mapping 30, 1511-1523.

Wang, J., Wang, X., Xia, M., Liao, X., Evans, A., He, Y., 2015. GRETNA: a graph theoretical network analysis toolbox for imaging connectomics. Frontiers in human neuroscience 9, 386.

Wang, J., Zuo, X., Gohel, S., Milham, M.P., Biswal, B.B., He, Y., 2011. Graph theoretical analysis of functional brain networks: test-retest evaluation on short-and long-term resting-state functional MRI data. PloS one 6.

Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of ‘small-world’networks. nature 393, 440.

Wei, K., Cieslak, M., Greene, C., Grafton, S.T., Carlson, J.M., 2017. Sensitivity analysis of human brain structural network construction. Network Neuroscience 1, 446-467.

Wierenga, L.M., Langen, M., Oranje, B., Durston, S., 2014. Unique developmental trajectories of cortical thickness and surface area. Neuroimage 87, 120-126.

Winkler, A.M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P.T., Duggirala, R., Glahn, D.C., 2010. Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. Neuroimage 53, 1135-1146.

1273 Yao, Z., Hu, B., Xie, Y., Moore, P., Zheng, J., 2015. A review of structural and
1274 functional brain networks: small world and atlas. Brain informatics 2, 45-52.

1275 Yotter, R.A., Dahnke, R., Thompson, P.M., Gaser, C., 2011a. Topological correction
1276 of brain surface meshes using spherical harmonics. Human brain mapping 32,
1277 1109-1124.

1278 Yotter, R.A., Nenadic, I., Ziegler, G., Thompson, P.M., Gaser, C., 2011b. Local
1279 cortical surface complexity maps from spherical harmonic reconstructions.
1280 Neuroimage 56, 961-973.

1281 Yotter, R.A., Thompson, P.M., Gaser, C., 2011c. Algorithms to improve the
1282 reparameterization of spherical mappings of brain surface meshes. Journal of
1283 Neuroimaging 21, e134-e147.

1284 Yu, K., Wang, X., Li, Q., Zhang, X., Li, X., Li, S., 2018. Individual morphological
1285 brain network construction based on multivariate euclidean distances between brain
1286 regions. Frontiers in human neuroscience 12.

1287 Zalesky, A., Cocchi, L., Fornito, A., Murray, M.M., Bullmore, E., 2012. Connectivity
1288 differences in brain networks. Neuroimage 60, 1055-1062.

1289 Zalesky, A., Fornito, A., Harding, I.H., Cocchi, L., Yücel, M., Pantelis, C., Bullmore,
1290 E.T., 2010. Whole-brain anatomical networks: does the choice of nodes matter?
1291 Neuroimage 50, 970-983.

1292 Zeng, T., Chen, H., Fakhry, A., Hu, X., Liu, T., Ji, S., 2015. Allen mouse brain atlases
1293 reveal different neural connection and gene expression patterns in cerebellum gyri and
1294 sulci. Brain Structure and Function 220, 2691-2703.

1295 Zhao, W., Guo, S., Linli, Z., Yang, A., Lin, C., Tsai, S., 2020. Functional, Anatomical,
1296 and Morphological Networks Highlight the Role of Basal Ganglia-Thalamus-Cortex
1297 Circuits in Schizophrenia. Schizophrenia Bulletin 46, 422-443.

1298 Zhang, J., Wang, J., Wu, Q., Kuang, W., Huang, X., He, Y., Gong, Q., 2011. Disrupted
1299 brain connectivity networks in drug-naive, first-episode major depressive disorder.
1300 Biological psychiatry 70, 334-342.

1301 Zhang, T., Chen, H., Razavi, M.J., Li, Y., Ge, F., Guo, L., Wang, X., Liu, T., 2018.
1302 Exploring 3□hinge gyral folding patterns among HCP Q3 868 human subjects.
1303 Human brain mapping 39, 4134-4149.

1304 Zhang, Z., Allen, G.I., Zhu, H., Dunson, D., 2019. Tensor network factorizations:
1305 Relationships between brain structural connectomes and traits. Neuroimage 197,
1306 330-343.

1307 Zuo, X., Xu, T., Jiang, L., Yang, Z., Cao, X., He, Y., Zang, Y., Castellanos, F.X.,
1308 Milham M.P., 2013. Toward reliable characterization of functional homogeneity in the
1309 human brain: preprocessing, scan duration, imaging resolution and computational
1310 space. Neuroimage 65, 374-386.

Figure Legends

Figure 1. Flow chart of imaging processing, network construction, topological characterization and statistical analysis in this study. ① For each structural image from the HCP and BNU datasets, four vertexwise morphological maps (FD, GI, SD and CT) were first extracted. ② Each morphological map was then divided into different numbers of regions according to two brain atlases (a2009s and a2005s). ③ Subsequently, the probability distribution function was estimated for each region in terms of signal distribution of each morphological index and was used to estimate interregional similarity with *JSDs* and *KLDs*, respectively. ④ This formed a total of 16 similarity matrices for each image (4 morphological indices × 2 brain parcellation atlases× 2 similarity measures). ⑤ Before topological characterization of the resultant similarity matrices, a sparsity-based procedure was further used to threshold each of them into a series of binary networks. ⑥ Finally, 10 global and 3 nodal graph theory-based network metrics were calculated to characterize topological organization of each binary network. FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT, cortical thickness; *JSDs*, Jensen-Shannon divergence-based similarity; *KLDs*, Kullback-Leibler divergence-based similarity; TRT, test-retest.

Figure 2. Mean similarity matrices and global network measures of morphological similarity networks. Upper panel: mean similarity matrices of morphological similarity networks (brain parcellation: a2009s; similarity estimation: *JSDs*). Although the similarity strength was large for each connection regardless of the

choice of morphological index, differential similarity patterns were evident. Lower panel: global organization of morphological similarity networks (brain parcellation: a2009s; similarity estimation: *JSDs*). Compared with matched random networks, individual morphological similarity networks exhibited higher clustering coefficient (i.e., normalized $C_p > 1$) and approximately equal characteristic path length (i.e., normalized $L_p \sim 1$), higher local efficiency (i.e., normalized $E_{loc} > 1$) and approximately equal global efficiency (i.e., normalized $E_{glob} \sim 1$), and higher modularity (i.e., normalized $Q > 1$), indicating small-world, highly efficient and modular organization. FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT, cortical thickness.

Figure 3. Spatial correlations of mean nodal centrality maps under different analytical strategies. Under each parcellation atlas, high correlations were observed between different nodal metrics and between different similarity measures; however, dramatically low correlations were observed between different morphological indices. *JSDs*, Jensen-Shannon divergence-based similarity; *KLDs*, Kullback-Leibler divergence-based similarity; FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT, cortical thickness; k_i, nodal degree; e_i, nodal efficiency; b_i, nodal betweenness; *, $P < 0.05$, Bonferroni corrected.

Figure 4. Consistent hubs under different analytical strategies. A specific set of regions were consistently identified as hubs, among which regions with the top 10%

of *CHs* were listed under each parcellation atlas. *CHs*, consistent hub scores.

Figure 5. TRT reliability values for global metrics under different analytical strategies. TRT reliability of global network metrics was significantly affected by choices of morphological index, brain parcellation atlas and similarity measure, with a general pattern of FD and SD > GI and CT and *JSDs* > *KLDs*. *ICC*, intraclass correlation; *JSDs*, Jensen-Shannon divergence-based similarity; *KLDs*, Kullback-Leibler divergence-based similarity; FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT, cortical thickness.

Figure 6. Proportions of regions at different levels of nodal TRT reliability (brain parcellation: a2009s; similarity estimation: *JSDs*). The proportions of regions showing excellent TRT reliability were significantly higher for nodal degree and efficiency than for nodal betweenness, and for FD- and SD-based than for GI- and CT-based morphological similarity networks. FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT, cortical thickness; k_i, nodal degree; e_i, nodal efficiency; b_i, nodal betweenness.

Figure 7. Regions consistently showing excellent TRT reliability under different analytical strategies. A specific set of regions were consistently identified to show excellent TRT reliability, among which regions with the top 10% of *CERs* are listed under each parcellation atlas. *CERs*, consistent excellent reliability scores.

Figure 8. Effect of different sample sizes on morphological similarity networks (brain parcellation: a2009s; similarity estimation: *JSD*s). With increasing sample size, a continuous increase was observed in the Pearson correlation coefficients between nodal centrality profiles that were separately averaged for 1000 subgroups of participants and nodal centrality profiles that were averaged over all participants, while the variability naturally decreased, with a tipping point observed when samples exceeded ~70 participants. FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT, cortical thickness.

Figure S1. Hubs with the top 10% of nodal centrality of morphological similarity networks (brain parcellation: a2009s; similarity estimation: *JSDs*). FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT, cortical thickness; k_i, nodal degree; e_i, nodal efficiency; b_i, nodal betweenness.

Figure S2. Mean similarity matrices and their intersession correlations and TRT reliability values for morphological similarity networks (brain parcellation: a2009s; similarity estimation: *JSDs*). No matter the choices of morphological index, the mean similarity matrices derived from session 1 (first row) and session 2 (second row) were highly correlated with each other (third row), with most connections showing good to excellent TRT reliability (i.e., *ICC* > 0.6, last row). *JSDs*, Jensen-Shannon divergence-based similarity; S1, session 1; S2, session 2; *ICC*, intraclass correlation;

1399 FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT, cortical thickness.

1400

1401 **Figure S3.** Connections showing significantly different values of TRT reliability
1402 between different morphological indices. Varied proportions of edges (numbers in
1403 white) showed morphological index-dependent TRT reliability, with a general pattern
1404 of FD > SD > GI > CT. *JSDs*, Jensen-Shannon divergence-based similarity; *KLDs*,
1405 Kullback-Leibler divergence-based similarity; FD, fractal dimension; GI, gyrification
1406 index; SD, sulcal depth; CT, cortical thickness.

1407

1408 **Figure S4**. Global organization of morphological similarity networks derived from
1409 the BNU TRT dataset (brain parcellation: a2009s; similarity estimation: *JSDs*). For
1410 both sessions, individual morphological similarity networks exhibited small-world,
1411 highly efficient and modular organization. FD, fractal dimension; GI, gyrification
1412 index; SD, sulcal depth; CT, cortical thickness.

1413

1414 **Figure S5**. TRT reliability values for global network metrics of morphological
1415 similarity networks (brain parcellation: a2009s; similarity estimation: *JSDs*). Most
1416 global metrics exhibited fair to excellent TRT reliability (i.e., *ICC* > 0.6) regardless of
1417 choices of morphological index. FD, fractal dimension; GI, gyrification index; SD,
1418 sulcal depth; CT, cortical thickness.

1419

1420 **Figure S6**. Global network measures showing significantly different TRT reliability

values between different morphological indices. The TRT reliability of many global network measures depended on the choices of morphological index, with a general pattern of FD and SD > GI and CT. Squares/circles indicate significant/non-significant differences, and color bars indicate t-statistics. *JSDs*, Jensen-Shannon divergence-based similarity; *KLDs*, Kullback-Leibler divergence-based similarity; FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT, cortical thickness.

Figure S7. Spatial correlations of nodal TRT reliability maps under different analytical strategies. Under each parcellation atlas, high correlations were observed between nodal degree and efficiency and between different similarity measures; however, dramatically low correlations were observed between nodal betweenness and the other two nodal metrics and between different morphological indices. *ICC*, intraclass correlation; *JSDs*, Jensen-Shannon divergence-based similarity; *KLDs*, Kullback-Leibler divergence-based similarity; FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT, cortical thickness; k_i, nodal degree; e_i, nodal efficiency; b_i, nodal betweenness; *, $P < 0.05$, Bonferroni corrected.

Figure S8. Regional maps at different levels of nodal TRT reliability (brain parcellation: a2009s; similarity estimation: *JSDs*). Most regions exhibited excellent TRT reliability for FD- and SD-based morphological similarity networks regardless of nodal measures. FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT,

1443 cortical thickness; k_i, nodal degree; e_i, nodal efficiency; b_i, nodal betweenness.

1444

1445 **Figure S9**. Regions showing significantly different TRT reliability values between
1446 different morphological indices (brain parcellation: a2009s; similarity estimation:
1447 *JSDs*). Varied proportions of regions (numbers in white) were identified to show
1448 morphological index-dependent TRT reliability, with a general pattern of FD > SD >
1449 GI > CT. Similar patterns were also found under other combinations of brain
1450 parcellation atlas and similarity measure (Figure 18, 19 and 20). *JSDs*,
1451 Jensen-Shannon divergence-based similarity; FD, fractal dimension; GI, gyrification
1452 index; SD, sulcal depth; CT, cortical thickness; k_i, nodal degree; e_i, nodal efficiency;
1453 b_i, nodal betweenness.

1454

1455 **Figure S10**. Regions showing significantly different TRT reliability values between
1456 different morphological indices (brain parcellation: a2009s; similarity estimation:
1457 *KLDs*). *KLDs*, Kullback-Leibler divergence-based similarity; FD, fractal dimension;
1458 GI, gyrification index; SD, sulcal depth; CT, cortical thickness; k_i, nodal degree; e_i,
1459 nodal efficiency; b_i, nodal betweenness.

1460

1461 **Figure S11**. Regions showing significantly different TRT reliability values between
1462 different morphological indices (brain parcellation: a2005s; similarity estimation:
1463 *JSDs*). *JSDs*, Jensen-Shannon divergence-based similarity; FD, fractal dimension; GI,

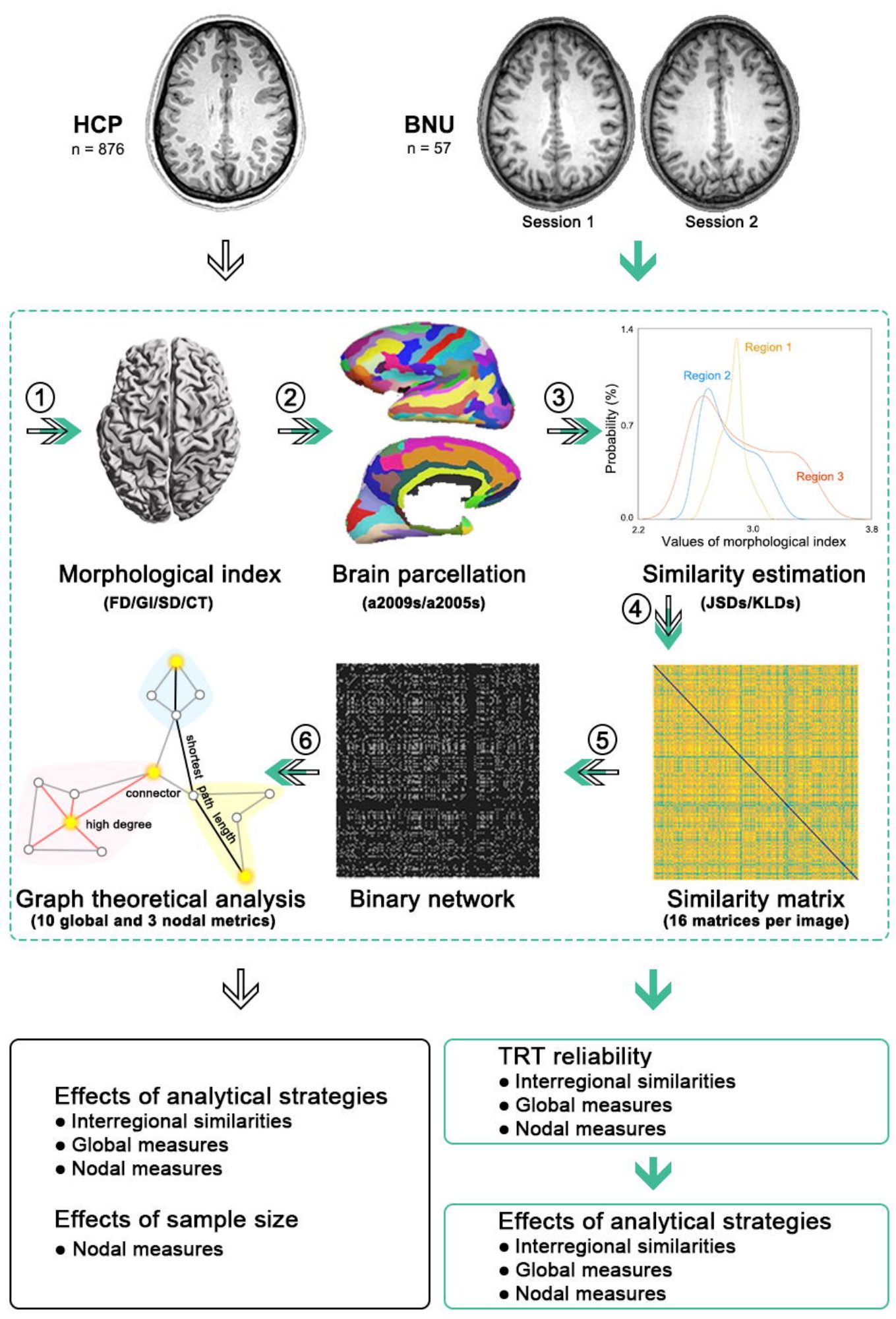
gyrification index; SD, sulcal depth; CT, cortical thickness; k_i, nodal degree; e_i, nodal efficiency; b_i, nodal betweenness.

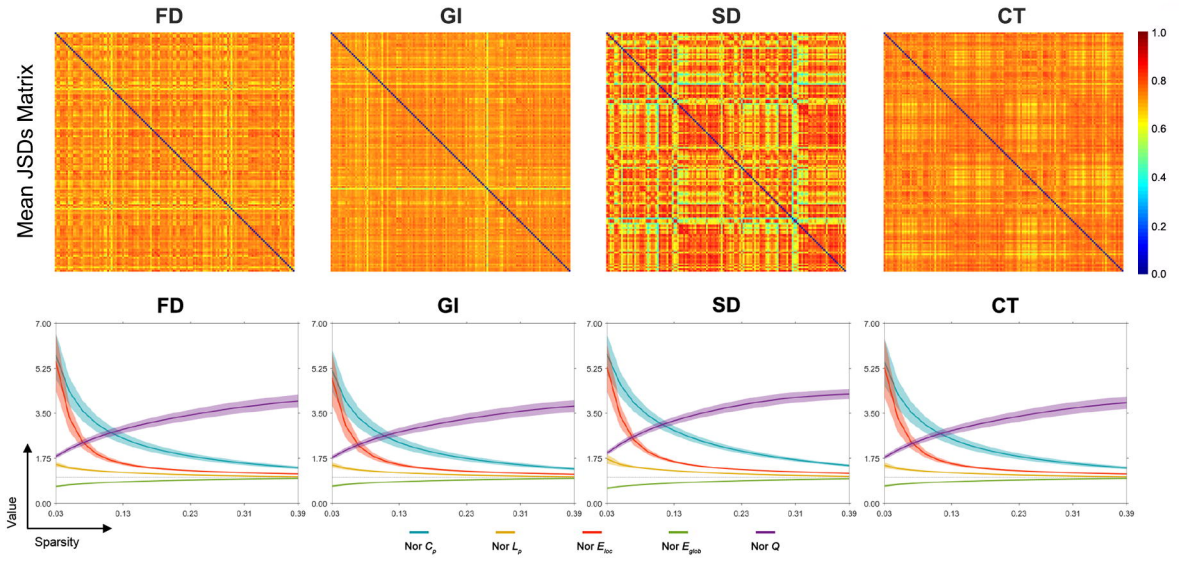
Figure S12. Regions showing significantly different TRT reliability values between different morphological indices (brain parcellation: a2005s; similarity estimation: *KLDs*). *KLDs*, Kullback-Leibler divergence-based similarity; FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT, cortical thickness; k_i, nodal degree; e_i, nodal efficiency; b_i, nodal betweenness.

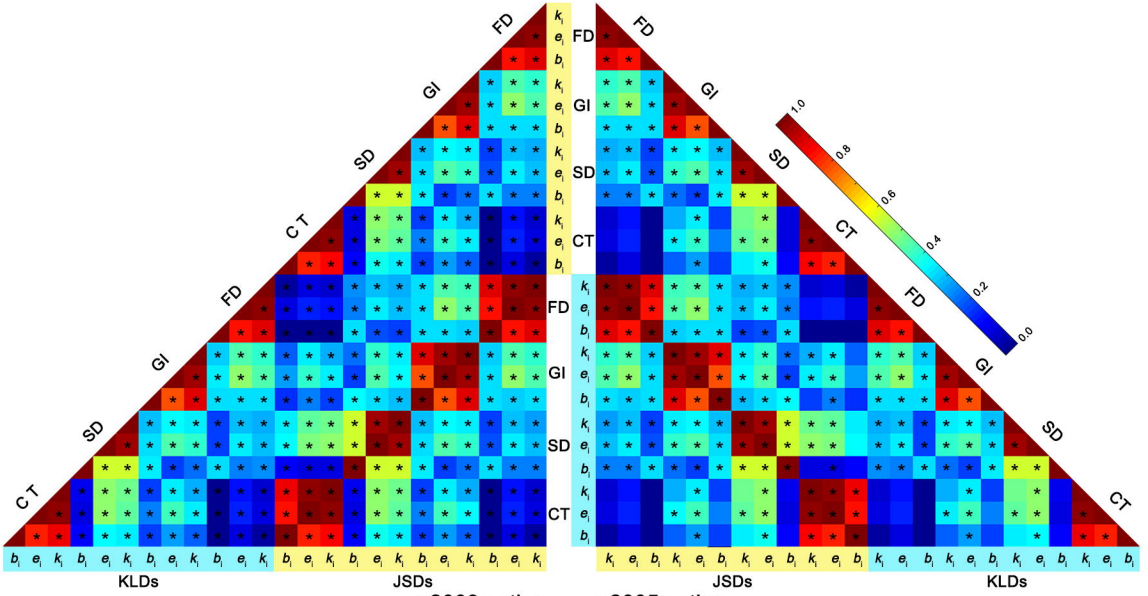
Table 1. Summary of TRT reliability of morphological brain networks under different analytical strategies

Parcellation atlas	Similarity measure	Morphological index	Interregional similarity	Global organization	Local organization
a2005s	*KLDs*	FD	0.876±0.128	0.768±0.064	0.830±0.111
		GI	0.734±0.141	0.413±0.093	0.663±0.166
		SD	0.830±0.167	0.873±0.050	0.834±0.135
		CT	0.636±0.161	0.514±0.152	0.572±0.174
	JSDs	FD	**0.877±0.127**	0.772±0.060	0.829±0.112
		GI	0.735±0.141	0.418±0.087	0.664±0.166
		SD	0.842±0.153	**0.877±0.045**	**0.835±0.135**
		CT	0.634±0.161	0.519±0.146	0.572±0.174
a2009s	*KLDs*	FD	0.883±0.098	0.766±0.086	0.842±0.115
		GI	0.749±0.132	0.520±0.087	0.704±0.159
		SD	0.853±0.150	0.766±0.037	0.840±0.145
		CT	0.695±0.134	0.525±0.064	0.641±0.154
	JSDs	FD	**0.883±0.098**	**0.770±0.084**	**0.842±0.114**
		GI	0.749±0.133	0.516±0.087	0.705±0.158
		SD	0.863±0.132	0.761±0.037	0.841±0.144
		CT	0.693±0.136	0.531±0.067	0.641±0.155
MMP	*KLDs*	FD	0.892±0.077	0.743±0.073	0.852±0.101
		GI	0.767±0.111	0.606±0.073	0.728±0.140
		SD	0.886±0.121	0.807±0.074	0.844±0.130
		CT	0.711±0.121	0.536±0.079	0.668±0.148
	JSDs	FD	0.891±0.076	0.740±0.076	**0.852±0.101**
		GI	0.766±0.109	0.605±0.075	0.728±0.140
		SD	**0.892±0.106**	**0.813±0.070**	0.844±0.130
		CT	0.709±0.121	0.537±0.078	0.668±0.148

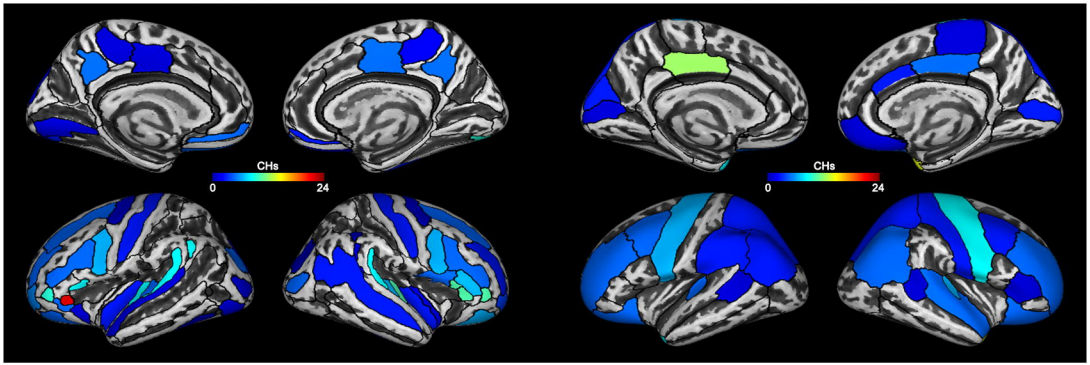
The highest TRT reliability is highlighted in bold under each brain parcellation atlas and each analytical level. *KLDs*, Kullback-Leibler divergence-based similarity; *JSDs*, Jensen-Shannon divergence-based similarity; FD, fractal dimension; GI, gyrification index; SD, sulcal depth; CT, cortical thickness.


HCP
n = 876
BNU
n = 57
Session 1
Session 2
1
2
3
Region 1
Region 2
Region 3
Probability (%)
1.4
0.7
0.0
2.2
3.0
3.8
Values of morphological index
Morphological index
(FD/GI/SD/CT)
Brain parcellation
(a2009s/a2005s)
Similarity estimation
(JSDs/KLDs)
4
5
6
shortest path length
connector
high degree
Graph theoretical analysis
(10 global and 3 nodal metrics)
Binary network
Similarity matrix
(16 matrices per image)
Effects of analytical strategies
• Interregional similarities
• Global measures
• Nodal measures
Effects of sample size
• Nodal measures
TRT reliability
• Interregional similarities
• Global measures
• Nodal measures
Effects of analytical strategies
• Interregional similarities
• Global measures
• Nodal measures


FD
GI
SD
CT
Mean JSDs Matrix
1.0
0.8
0.6
0.4
0.2
0.0
FD
GI
SD
CT
7.00
5.25
3.50
1.75
0.00
0.03
0.13
0.23
0.31
0.39
Value
Sparsity
Nor C_p
Nor L_p
Nor E_{loc}
Nor E_{glob}
Nor Q

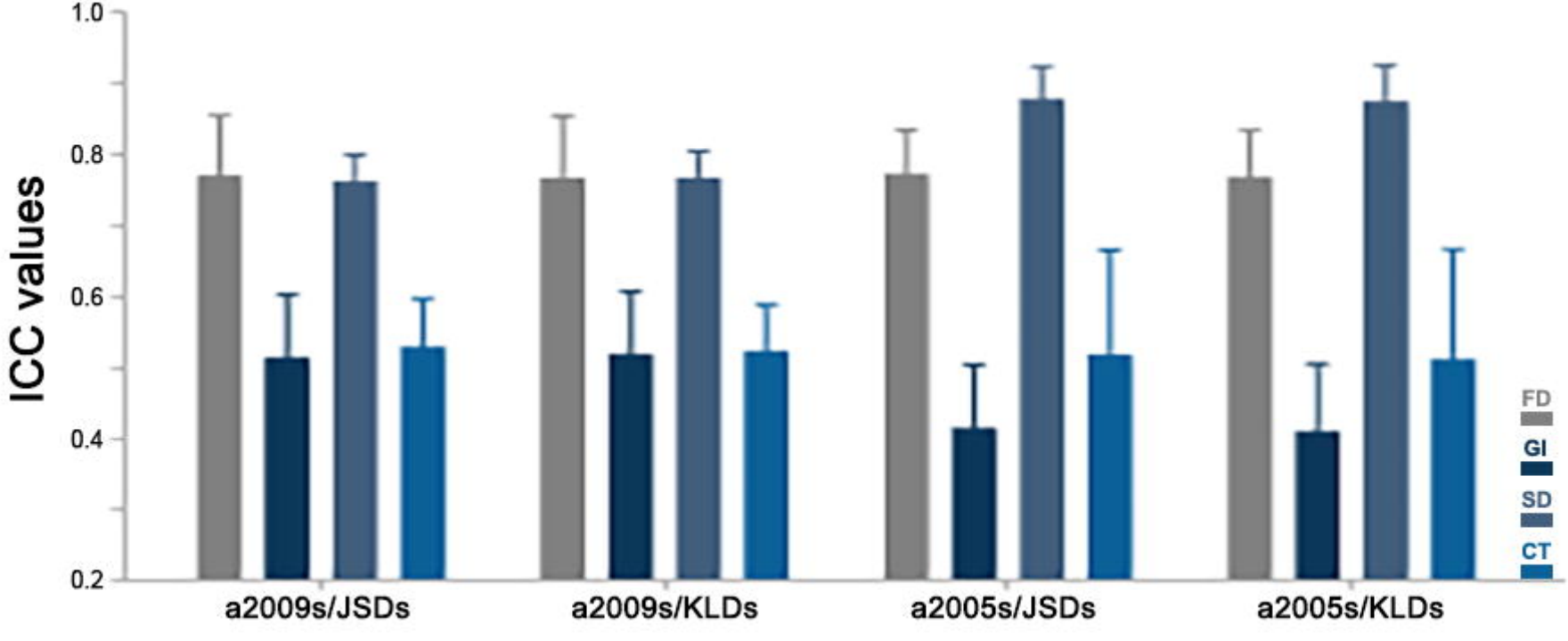

FD
GI
SD
CT
KLDs
JSDs
a2009s atlas
a2005s atlas
1.0
0.8
0.6
0.4
0.2
0.0

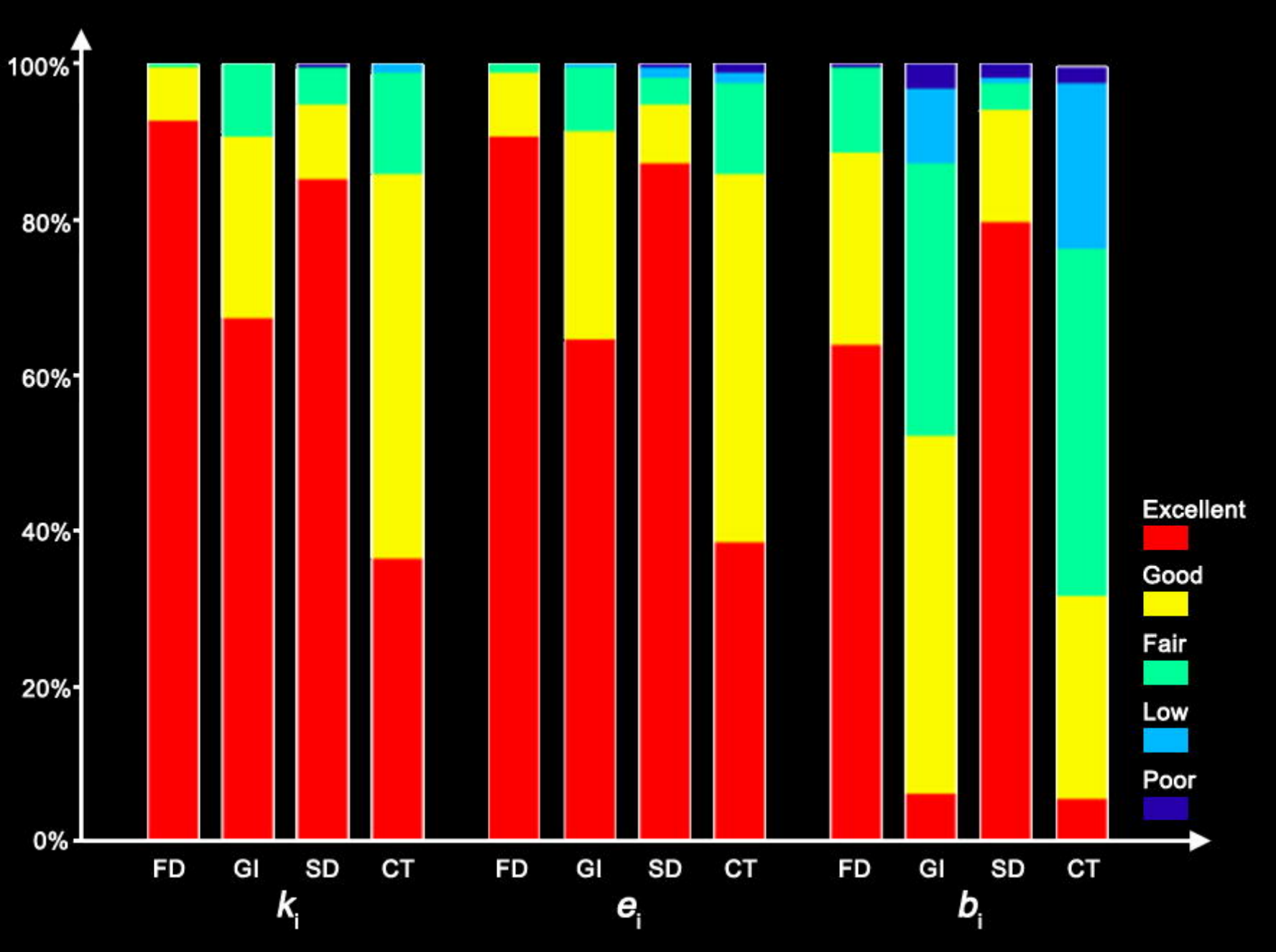
A



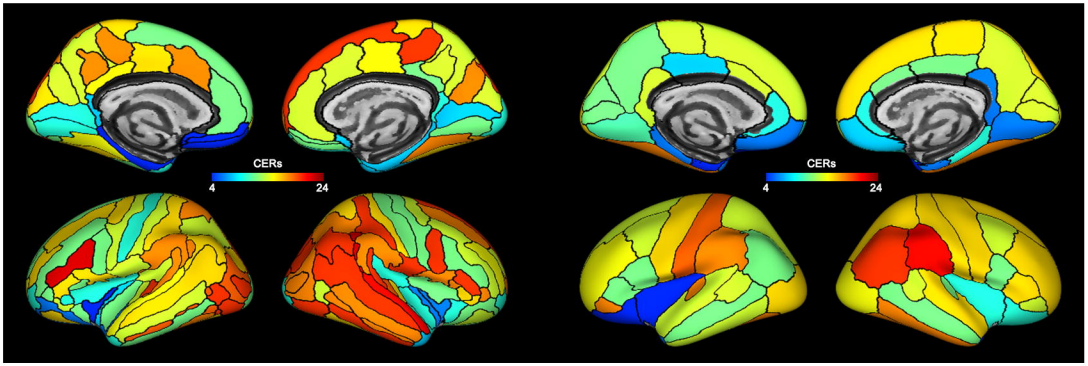
B

	Left hemisphere		Right hemisphere	
	CHs	Name	CHs	Name
a2005s atlas	14	*Middle temporal gyrus*	16	*Paracentral lobule*
	10	*Lateral occipital cortex*	10	*Lateral occipital cortex*
	8	*Inferior temporal gyrus*	8	*Inferior temporal gyrus*
a2009s atlas	24	*Horizontal ramus of the anterior segment of the lateral sulcus (or fissure)*	12	*Horizontal ramus of the anterior segment of the lateral sulcus (or fissure)*
	11	*Lateral orbital sulcus*	12	*Anterior transverse temporal gyrus (of Heschl)*
	10	*Anterior transverse temporal gyrus (of Heschl)*	12	*Vertical ramus of the anterior segment of the lateral sulcus (or fissure)*
	10	*Planum temporale or temporal plane of the superior temporal gyrus*	12	*Lateral orbital sulcus*
	10	*Vertical ramus of the anterior segment of the lateral sulcus (or fissure)*	11	*Posterior transverse collateral sulcus*
	10	*Sulcus intermedius primus (of Jensen)*	10	*Planum temporale or temporal plane of the superior temporal gyrus*
	8	*Inferior part of the precentral sulcus*	8	*Triangular part of the inferior frontal gyrus*
			8	*Orbital sulci (H-shaped sulci)*
			8	*Inferior part of the precentral sulcus*
			8	*Transverse temporal sulcus*


ICC values
1.0
0.8
0.6
0.4
0.2
a2009s/JSDs
a2009s/KLDs
a2005s/JSDs
a2005s/KLDs
FD
GI
SD
CT

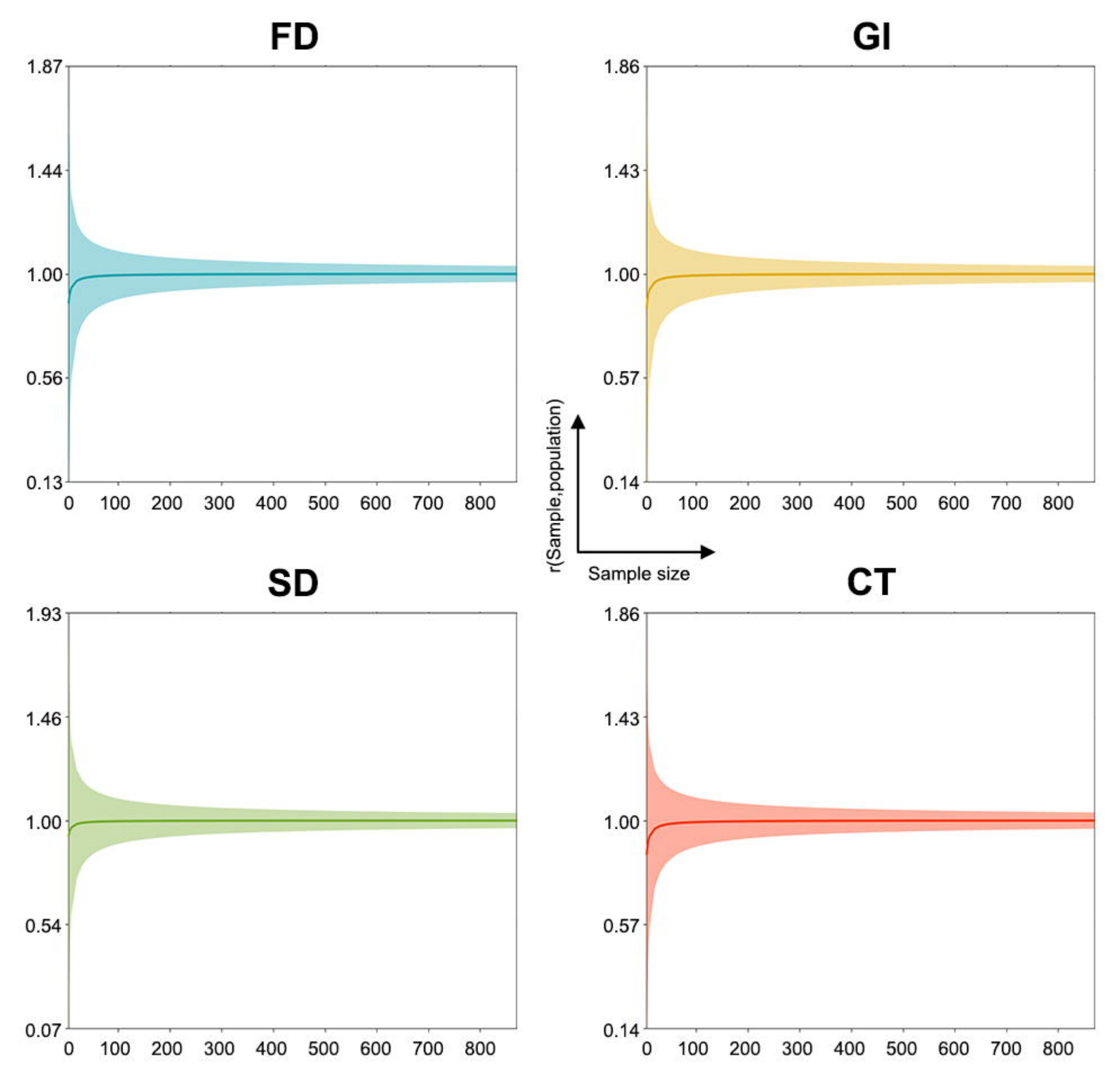

100%
80%
60%
40%
20%
0%
FD
GI
SD
CT
FD
GI
SD
CT
FD
GI
SD
CT
k_i
e_i
b_i
Excellent
Good
Fair
Low
Poor

A



B

	Left hemisphere		Right hemisphere	
	CERs	Name	CERs	Name
a2005s atlas	19	*Postcentral gyrus*	21	*Supramarginal gyrus*
	18	*Fusiform gyrus*	20	*Inferior parietal cortex*
	18	*Pars orbitalis*	18	*Fusiform gyrus*
	18	*Supramarginal gyrus*	18	*Inferior temporal gyrus*
	18	*Transverse temporal cortex*		
a2009s atlas	22	*Inferior frontal sulcus*	22	*Sulcus intermedius primus (of Jensen)*
	20	*Inferior occipital gyrus (O3) and sulcus*	21	*Lateral aspect of the superior temporal gyrus*
	20	*Superior occipital gyrus (O1)*	20	*Inferior occipital gyrus (O3) and sulcus*
	20	*Middle occipital sulcus and lunatus sulcus*	20	*Subcentral gyrus (central operculum) and sulci*
	20	*Anterior occipital sulcus and preoccipital notch (temporo-occipital incisure)*	20	*Triangular part of the inferior frontal gyrus*
	20	*Transverse temporal sulcus*	20	*Superior frontal gyrus (F1)*
			20	*Superior occipital gyrus (O1)*
			20	*Angular gyrus*
			20	*Planum temporale or temporal plane of the superior temporal gyrus*
			20	*Inferior temporal gyrus (T3)*
			20	*Marginal branch (or part) of the cingulate sulcus*
			20	*Superior occipital sulcus and transverse occipital sulcus*
			20	*Postcentral sulcus*
			20	*Inferior part of the precentral sulcus*
			20	*Superior temporal sulcus (parallel sulcus)*


FD
1.87
1.44
1.00
0.56
0.13
0 100 200 300 400 500 600 700 800
GI
1.86
1.43
1.00
0.57
0.14
0 100 200 300 400 500 600 700 800
r(Sample,population)
Sample size
SD
1.93
1.46
1.00
0.54
0.07
0 100 200 300 400 500 600 700 800
CT
1.86
1.43
1.00
0.57
0.14
0 100 200 300 400 500 600 700 800