bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The Impact of Graph Construction Scheme and Community Detection Algorithm on the Repeatability of Community and Hub Identification in Structural Brain Networks

Stavros I. Dimitriadis[1-6,8*], Eirini Messaritaki[1,2,3,7], Derek K. Jones[1,3]

1. *Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, Maindy Road, Cardiff, CF24 4HQ, UK*
2. *Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, CF24 4HQ, UK*
3. *School of Psychology, Cardiff University, 70 Park Place, Cardiff, CF10 3AT, UK*
4. *Neuroinformatics Group, Cardiff University Brain Research Imaging Centre, School of Psychology, Cardiff University, Cardiff, United Kingdom*
5. *Neuroscience and Mental Health Research Institute, Cardiff University, Cardiff, United Kingdom*
6. *MRC Centre for Neuropsychiatric Genetics and Genomics, School of Medicine, Cardiff University, Cardiff, United Kingdom*
7. *BRAIN Biomedical Research Unit, Cardiff University, Maindy Road, Cardiff, CF24 4HQ, UK*

***Corresponding author:**
Stavros I. Dimitriadis
Contact:DimitriadisS@cardiff.ac.uk / stidimitriadis@gmail.com
Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, Maindy Road, Cardiff, CF24 4HQ, UK

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract:

A critical question in network neuroscience is how nodes cluster together to form communities, to form the mesoscale organization of the brain. Various algorithms have been proposed for identifying such communities, each identifying different communities within the same network. Here, (using test-retest data from the Human Connectome Project), the repeatability of 33 community detection algorithms, each paired with 7 different graph construction schemes was assessed.

Repeatability of community partition depended heavily on both the community detection algorithm and graph construction scheme. Hard community detection algorithms (in which each node is assigned to only one community) outperformed soft ones (in which each node can be belong to more than one community). The highest repeatability was observed for the fast multi-scale community detection algorithm paired with a graph construction scheme that combines 9 white matter metrics. This pair also gave the highest similarity between representative group community affiliation and individual community affiliation. Connector hubs had higher repeatability than provincial hubs. Our results provide a workflow for repeatable identification of structural brain networks communities, based on optimal pairing of community detection algorithm and graph construction scheme.

Keywords: Structural Brain network, Community detection, Permutation test, Normalized Mutual Information (NMI), Diffusion Magnetic Resonance Imaging (dMRI), hard community detection, soft community detection, overlapping communities.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1. Introduction

The human brain can be modelled as a network (Bassett and Sporns, 2017) and summarised as a graph. In structural networks, the nodes of the graph are small volumes of tissue which are interconnected via white matter tracts (edges). Graph theory can provide novel insights into healthy human brain function (Bassett et al., 2011; Braun et al., 2015) and its alteration in various diseases (Braun et al., 2016, Baker et al. (2015); Collin et al. (2016); Drakesmith et al. (2015); Aerts et al. (2016); Nelson et al. (2017); Vidaurre et al. (2018); Imms et al. (2019)).

An open question in network neuroscience is how neural units cluster together to form inter-connected groups and provide the coordinated brain activity that gives rise to action, perception and behaviour (Bassett and Mattar, 2017; Pessoa et al., 2018). Modularity is a quintessential concept in network neuroscience, wherein neural units are densely connected to one another, forming clusters or *modules* (Meunier et al., 2010). This is an efficient architecture allowing a complex network to integrate information locally, while maintaining its adaptability to any external stimulus. Networks in nature often show hierarchical, modular organization (Blondel et al., 2008; Fortunato 2010; Fortunato and Castellano 2012; A. Lancichinetti and Fortunato 2009a,b; Newman,2004 2012; Meunier et al., 2009). In the brain, such hierarchical modularity could support segregated neuronal interactions and their integration at the global level. Networks with such structure (Fortunato, 2010) are more complex than those with random structure (Sporns et al., 2000), and have been well demonstrated in functional brain networks (Sporns and Betzel, 2016).

Commonly studied graph theoretical metrics, such as global and local efficiency, clustering coefficient, shortest path length, and small-worldness (Rubinov and Sporns 2010), provide useful information related to the global and local properties of the network, but to investigate *mesoscale* network organization community (or modularity) detection techniques are more appropriate (Meunier et al. 2010; Giusti et al., 2016). Following community partition, a frequently used methodology to identify the community structure uses two modular network metrics (Guimera et al., 2005; van den Heuvel and Sporns, 2013). The participation coefficient, P_i, of a node quantifies the distribution of its links among the modules of the network, while the within-module z-score, $z_{i,}$ measures the connectedness of a node to other nodes in the module. These metrics in turn can be used to classify hubs as either provincial or connector hubs (Guimera and Amaral, 2005). Please refer to Appendix 1 for more detail on the computation of P_i and z_i and how this classification is made.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Complex network analysis of multimodal structural and functional brain connectivity has identified a subset of brain areas that play a key role for an efficient neural signalling and communication (van den Heuvel and Sporns, 2013). These brain areas, called hubs, support dynamic functional coupling within and between functional subnetworks. In empirical structural brain networks, the term 'rich-club' characterises brain areas/nodes with high degree that are more densely interconnected between each other compared to the rest of the network.

Various methodologies for structural network generation have been proposed, involving tractography with different algorithms and assigning edge weights using different diffusion MRI-based metrics. The resulting graphs are quite different from each other and have different levels of robustness and repeatability (Smith et al., 2015; Owen et al., 2013, 2013; Zhong et al., 2015; Dimitriadis et al., 2017b; Yuan et al., 2019). We recently explored the repeatability of structural brain graphs, their edge weights and graph-theoretical metrics, for twenty-one different edge-weighting schemes (Messaritaki et al., 2019a). We demonstrated that integrating several metrics as edge weights is very good at capturing differences between populations, and is interesting from the perspective of developing biomarkers (Dimitriadis et al 2017, Clarke et al 2020).

We constructed structural brain networks from a set of test-retest diffusion MRI scan data from the Human Connectome project using the b=2000 s/mm^2 data and the seven most reproducible graph-construction schemes as derived from our previous study on the same data (Messaritaki et al., 2019b). We then applied thirty-three community detection algorithms. The 'hard' algorithms assign every node to only one community, while the 'soft' algorithms can assign a node to multiple communities. For every pair of community detection algorithm and graph construction scheme, we estimated the reproducibility of nodal P_i and z_i and of provincial and connector hubs, based on both modular network metrics. Our aim was to identify the combination of graph construction scheme and community detection algorithm with the highest agreement of individual communities between the two repeat scan sessions.

The quality criterion for the estimated community partitions was also important in our study. To this end, we compared the quality index of the community partitions estimated over the original graphs with quality indexes of the community partitions computed over surrogate null versions of the original graph (Guimera et al., 2004). We previously reported a statistical procedure for performing condition and group comparisons in terms of brain communities (Dimitriadis et al., 2012). Here, we applied a similar approach to assess between-scan pairwise community similarity for every pair of graph construction scheme and community

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

detection algorithm. We adopted a proper community partition distance metric, the Normalized Mutual Information (NMI) (Lancichinetti and Fortunato,2009a ; Alexander-Bloch et al. 2012). Finally, we derived a consensus cluster across participants and repeat scans (Dong et al., 2014; Ozdemir et al., 2015). The agreement of consensus cluster with individual communities adopting NMI was also used as an objective criterion of the optimal combination of graph construction scheme and community detection algorithm.

We note that the analysis presented here does not aim to assess how well these structural networks represent the organisation of the human brain. The accuracy of these networks and of the metrics used as edge-weights in representing the functional organisation of the brain has been validated in recent work by Messaritaki et al. (2020). Additionally, the metrics used as edge-weights are routinely used in network analyses in the literature (for example Taylor et al., 2015; Nigro et al., 2016; Caeyenberghs et al., 2016). Our analysis does, however, address one aspect of the accuracy of the partition of the structural connectome. If a partition of the structural connectome is not repeatable in the absence of changes resulting from maturation or intervention, then that partition is not an accurate representation of the modular organisation of the structural connectome. Only partitions that are repeatable can convey reliable information about the structural organisation of the human brain. In other words, even though the repeatability of a partition is not a sufficient condition for it to be representative of the brain's structural organisation, it is a necessary one.

The rest of this manuscript is organised as follows: Section 2 describes the graph-construction schemes, community detection algorithms, community partition similarity, the methodology for detecting connector and provincial hubs and their repeatability. Section 3 reports our results in terms of repeatable community partitions across the 2D space of graph-construction schemes / community detection algorithms, the repeatability of nodal P_i/ z_i and the detection of connector/provincial hubs. The Discussion summarises the main outcome of our study explaining its advantages, limitations, and suggestions for future directions.

2. Methods

All analyses were performed using MATLAB (2019a; The Mathworks, Inc, Massachussets, United States).

2.1. Data

We analysed the test-retest MRI and diffusion-MRI dataset from the multimodal neuroimaging database of the Human Connectome Project (HCP) (Sotiropoulos et al.,

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2013b; Glasser et al., 2013 ; van Essen et al., 2013). We used the data from the 37 participants for whom there were 90 gradient directions for each *b*-value. The participants on this test-retest dataset were scanned twice with the between-scan time interval ranging between 1.5 and 11 months. The age-range of the participants was 22–41 years. The test-retest time interval is shorter than the expected time over which maturation-induced structural changes can be measured with diffusion MRI (dMRI).

The diffusion-weighted images (DWIs) had a resolution of (1.25×1.25×1.25) mm^3 and were acquired for three different diffusion weightings (*b*-values: 1000□s/mm^2, 2000□s/mm^2 and 3000□s/mm^2) across 90 gradient directions for each *b*-value. The HCP acquisition details and pre-processing are described in Sotiropoulos et al. (2013a,b), Feinberg et al. (2010), Moeller et al. (2010), Setsompop et al. (2012), Xu et al. (2012), Glasser et al. (2013). Specifically, the diffusion images were corrected for EPI distortions, eddy-current distortions, participant movement and gradient nonlinearities. The diffusion data were also registered to the structural data. We performed the following analyses using the b=2000□s/mm^2 data.

2.2. Tractography

We performed whole-brain tractography using ExploreDTI-4.8.6 (Leemans et al., 2009), estimating the fiber orientation distribution function (fODF) using constrained spherical deconvolution (CSD) (Tournier et al., 2004). Tracking was initiated on a 2×2×2 mm grid, with a 1mm step size, angular threshold of 30° and fiber length range of 50−500 mm.

2.3. Graph generation

2.3.1. Node definition

The Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002) was used to define 90 cortical and subcortical areas (45 areas per hemisphere) as nodes of the structural brain graphs. Structural brain networks (SBN) were generated for each participant using ExploreDTI-4.8.6 (Leemans et al., 2009).

2.3.2. Edge weights

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Edges were weighted using the seven most reproducible graph-construction schemes identified previously with the same dataset (Messaritaki et al., 2019b), and which were based on different combinations of the nine metrics listed in Table 1 (see Section 2.3.4). Each graph was normalised to have a maximum edge weight of 1, while the elements in the main diagonal were set to zero (see Figure 1).

Table 1. Metrics used in connectivity matrices.

Metric	Abbreviation
Fractional anisotropy	FA
Mean diffusivity	MD
Radial diffusivity	RD
Number of streamlines	NS
Percentage of streamlines	PS
Streamline density	SLD
Tract volume	TV
Tract length	TL
Euclidean distance between nodes	ED

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

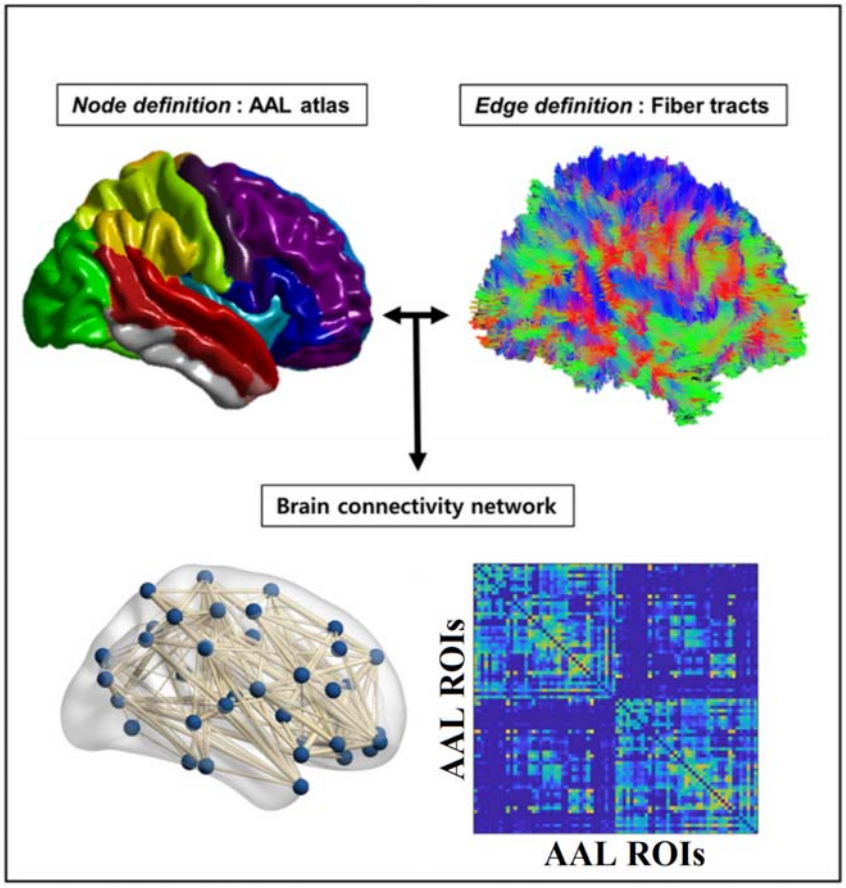


Figure 1. Flowchart of the construction of a structural brain network based on tractography and diffusion metrics (see Table 1).

2.3.3 Integrated Edge-Weights

Combining multiple metrics into an *integrated* edge weight is supported by the fact that each metric conveys information about different tissue properties, while at the same time topological properties of SBNs are affected by more than one metric. Here, using the data-driven algorithm described in our previous work (Dimitriadis et al., 2017b,c) the nine metrics in Table 1 were used to form an integrated SBN for each participant and scan session.

The algorithm down-weights the more similar metrics and up-weights the most dissimilar metrics to enhance the integration of complementary topological information across the nine metrics. An orthogonal-minimal-spanning-tree (OMST) algorithm was then applied to the resulting networks, selecting the edges that preserve connectivity between nodes, while guaranteeing that the overall network efficiency is maximised. More details on the OMST algorithm and its implementation can be found in our previous work (Dimitriadis et al., 2017a,b,c; Messaritaki et al., 2019) and the related code is freely available at https://github.com/stdimitr/multi-group-analysis-OMST-GDD.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2.3.4 Graph Construction Schemes

Seven graph construction schemes were used in this study, summarised in Table 2 and falling broadly into two categories. We briefly explain their construction methodologies here.

The first category includes graphs constructed via the data-driven algorithm (Dimitriadis et al., 2017b, c). A) NS-OMST: apply the OMST filtering algorithm (Dimitriadis et al., 2017b,c) to the NS-weighted matrix. B) NS+FA - OMST: Integrate the NS-weighted and FA-weighted matrices with the data-driven algorithm. C) 9m-OMST: Integrate all nine diffusion metrics (as originally reported in Dimitriadis et al. 2017b, see Table 2).

The second category includes SBNs with edges weighted by the NS or the FA and applying a threshold to remove edges with the lowest weights. The threshold was determined by imposing the constraint that the graphs exhibit the same sparsity as the OMST graphs that exhibited the highest reproducibility (Messaritaki et al., 2019b). Once the topology of each of those graphs was specified, the weights of the edges were either kept as they were or re-weighted with one of the remaining two metrics. These graphs are as follows (see Table 2). D) NS-thr: Keep the highest-NS edges. E) NS-t/FA-w: Threshold to keep the highest-NS edges, then reweight those edges with their FA. F) NS-t/MD-w: Keep the highest-NS edges, then reweight those edges with their MD. G) FA-t/NS-w: Keep the highest-FA edges, then reweight those edges with their NS.

As we have shown previously, these seven schemes exhibit different values of similarity between them, from 0.99 to 0.42 (Messaritaki et al. (2019b), Table 3), motivating their inclusion in a study on the repeatability of community detection.

Table 2. Summary of the graph-construction schemes

Abbreviation	Initial Edge Weights	Topology	Final Edge Weights	Symbol
NS – OMST	NS	OMST	Unchanged	A
NS + FA OMST	lin. comb. of NS and FA	OMST	Unchanged	B
9-m OMST	lin. comb. of all 9 metrics in Table 1	OMST	Unchanged	C
NS-thr	NS	keep highest-NS edges	Unchanged	D

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abbreviation	Initial Edge Weights	Topology	Final Edge Weights	Symbol
NS-t/FA-w	NS	keep highest-NS edges	re-weight with FA	E
NS-t/MD-w	NS	keep highest-NS edges	re-weight with MD	F
FA-t/NS-w	FA	keep highest-FA edges	re-weight with NS	G

2.3.5 Community detection algorithms

Communities or modules are defined as subgroups of nodes that are more interconnected with each other compared to the rest of the network (Newman and Girvan 2004; Radicchi et al. 2004). In the present study, we compared thirty-three different community detection algorithms, comprising twenty-six with hard clustering and seven with soft clustering. (see Figure 2). In hard clustering, community membership can be represented as a vector that encapsulates the assignments of every brain area to every detected graph cluster (community). In our case, clustering has a dimension of 1 x 90, equalling the number of brain regions in the AAL parcellation. In soft clustering, the outcome is a matrix that encapsulates how many soft clusters a given node (brain area) belongs to. A more detailed description of the adopted community detection algorithms is provided in Appendix 2 & Supplementary Material.

In the present study, we considered for the very first time in brain network neuroimaging a large number of community detection algorithms. We adopted thirty-three graph partition algorithms further divided into twenty-six hard clustering algorithms and seven soft clustering algorithms.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

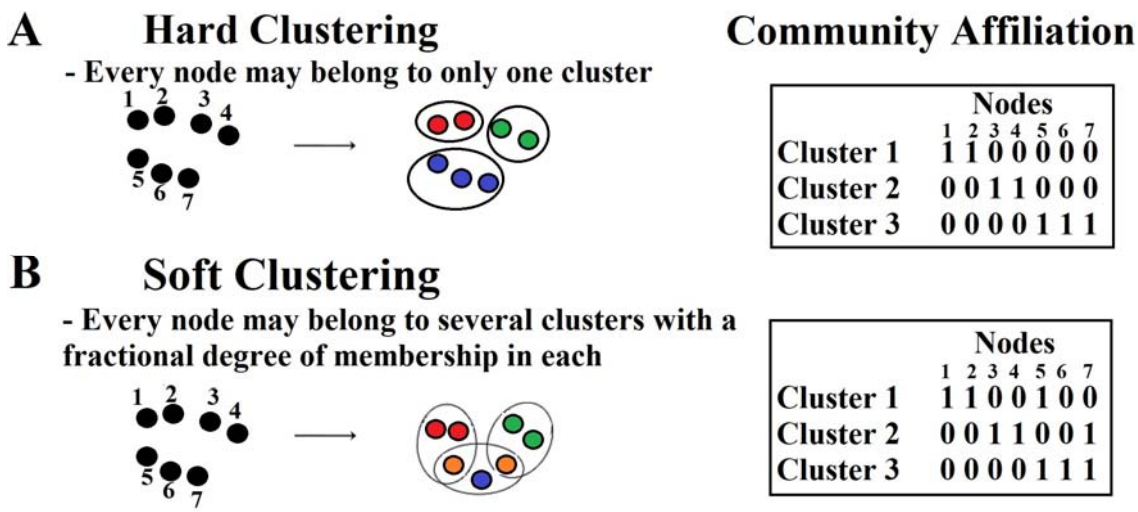


Figure 2. **An example of hard and soft clustering in a toy example containing 7 nodes.**

A. Hard clustering: A node can only belong to one cluster. The table shows the community assignment to every node.

B. Soft clustering: Five out of seven nodes are clustered in a single cluster/community {nodes 1,2,3,4,6} while nodes {5 and 7} belong to two communities: node 5 belongs to communities 1 and 3 while node 7 belongs to communities 2 and 3. The table shows the community assignment to every node.

2.3.6 Permutation Test on Quality Modular Indexes

For every participant, scan and graph construction method, we produced 1000 surrogate null graph models by randomizing the weighted connections while preserving both the degree and strength of every node and the overall connectedness of the network (Rubinov and Sporns,2010).

All of the hard clustering algorithms (no.s 1-26) involved a Q quality index for the communities detected. For further details of Q quality indexes see Le Martelot and Hankin, (2011,2012 a,b).

For the soft clustering algorithms (no. 27-33), we estimated the Normalized Mutual Information (NMI; see Appendix 3) between the original community affiliation and the surrogate null communities produced via the application of every algorithm to the surrogate graph model.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2.3.7 Between-Scan Community Detection Agreement

We quantified the graph-partition distance with the normalized mutual information (NMI; see Appendix 3).

2.3.8 Consensus Clustering

A consensus matrix was constructed for every pair of {graph construction scheme – community detection algorithm} that showed high test-retest reproducibility across the cohort (group-averaged NMI > 0.9). We quantified how many times two nodes across the 74 SBNs (37 participants x 2 scans) were classified in the same community and this entry, $t_{i,j}$, was assigned to the relevant pair of nodes. The consensus matrix has the same dimensions as the original SBN (90 x 90 in our case) with entries assuming integer values between 0 and 74 {37 participants x 2 scans}, which were then transformed to denote the probability of occurrence of a pair of nodes (brain areas) being classified as belonging to the same community across the cohort and scan sessions. We converted the consensus matrix into a probability one by dividing each entry by 74.

In order to get a consensus or group representative community per graph-construction scheme and community detection algorithm, consensus matrices should be iteratively thresholded and clustered with a community detection algorithm (Lancichinetti and Fortunato, 2012). This algorithm uses an absolute arbitrary threshold to eliminate weak connections and iteratively apply a graph partition technique. Instead of an arbitrary filtering scheme, we adopted our OMST algorithm (Dimitriadis et al., 2017a,b,c) to topologically-filter the consensus matrix in a data-driven way. We then extracted the consensus – group representative community by applying the community detection algorithms across the graph construction schemes (Newman,2006). See Figure 3.E for an example of a consensus matrix.

2.3.9 Agreement of consensus representative community with individual community structures

An important criterion of our analysis is the high similarity between the consensus clustering and individual clustering for every graph construction scheme that showed high

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

group-averaged community similarity (NMI > 0.9). To this end, we estimated this community similarity for every case.

Figure 3 illustrates the various steps of the analysis.

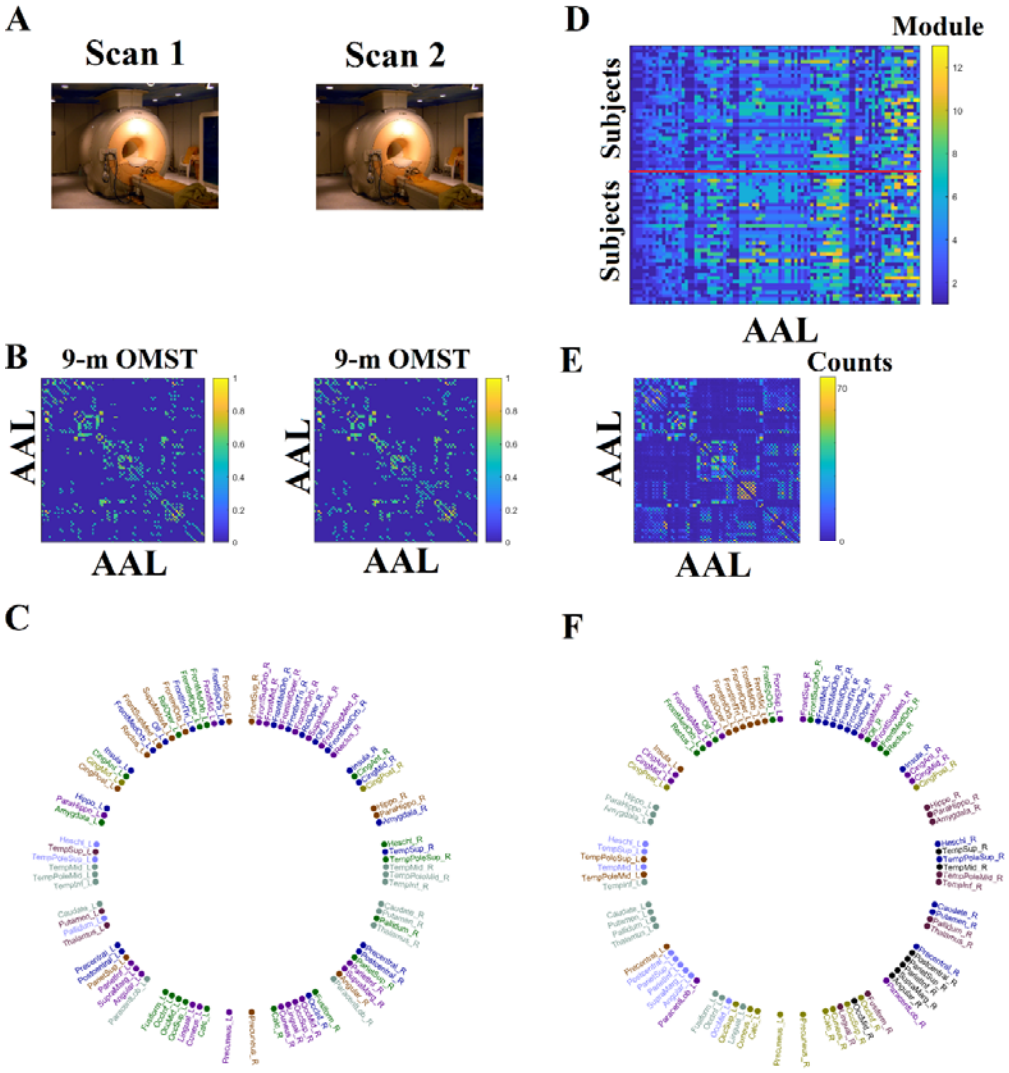


Figure 3. Outline of the presented methodology.

The demonstration based on 9-m OMST graph-construction scheme and gso-discrete mode community detection algorithm.

A. Repeat – Scan Sessions
B. Structural brain networks from participant 1 from both sessions using 9-m OMST graph-construction scheme
C. Individual community affiliation of participant 1 from scan session 1. Each colour represents one community.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

D. Vectorised community affiliations of the whole cohort from scan sessions 1 and 2 separated with a red line. Every module is coded with a different colour.
E. Consensus matrix is built over group community affiliations across both sessions as presented in D. Weights in the consensus matrix refer to the total number of times two brain areas are grouped together across the cohort and scan sessions with the maximum value being (number of participants) x (scan sessions) = 74.
F. Representative community affiliation after graph partitioning the consensus matrix presented in E. Each community is encoded to a different colour. Similarity NMI distance has been estimated between representative community affiliation presented in F and individual community affiliations presented in C.

2.3.10 Evaluating the Combined Graph Construction Schemes - Community Detection Algorithms

As mentioned previously, we first identified the combinations of graph construction schemes and community detection algorithms with higher group-averaged between-scan community affiliation agreement (NMI > 0.9) with a $p < 0.05$ based on the bootstrapping procedure. We then adopted a criterion of highest community similarity between the consensus clustering with individual community affiliation (clustering). It is important that consensus clustering expresses the inter-subject variability and acts as a vector median for the whole group (Dimitriadis et al., 2012). The final ranking of pairs of graph construction schemes and community detection algorithms was based on: 1) high between-scan group-averaged community similarity with a $p < 0.05$ (bootstrapping), 2) Q quality index with $p < 0.05$ based on surrogate null brain models and 3) high community similarity between consensus clustering and individual community affiliations (clusterings) assessed via a two-way ANOVA (see section 2.7).

2.4 Modular Driven Structural Brain Hub Detection

Here, we applied the aforementioned methodology solely on the graph construction schemes and community detection algorithm that fulfil the evaluation criteria of section 2.3.10.

2.5 Reliability of Nodal Participation Coefficient P_i and Within-Module Z-Score z_i,

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We also explored the intra-class correlation coefficient (ICC) of nodal participation coefficient P_i and within-module z-score z_i. As a main outcome of this hub detection approach, we quantified the consistency of connector/provincial hub detection first within participant between scans, and secondly across the cohort.

2.6 Assessing a Reproducible Structural Core of the Human Brain

We detected structural hubs for every participant, scan session and graph construction scheme by applying an absolute threshold to the participation coefficient and within-module z-score (Guimera and Amaral,2005; Hagmann et al., 2007,2008); van den Heuvel and Sporns,2013). We estimated an agreement index that quantifies the percentage of connector/provincial hubs that were detected in both scans. This agreement index is defined as:

$$Agreement = \sum_{k=1}^{no\ of\ subjects} \frac{CH_1 \cap CH_2}{no\ of\ subjects} \quad (1)$$

where $CH_{1,2}$ are two vectors of size 1x90 with ones in positions where a brain area is detected as connector or provincial hub. This agreement index is normalized by the total number of participants and takes the absolute value of 1 when a node/ROI is detected as either connector or provincial hub across all participants and in both scans.

2.7 Statistical Analysis

Firstly, we determined pairs of graph construction schemes and community detection algorithms that fulfilled the evaluation criteria presented in 2.3.10. The first two criteria were evaluated via bootstrapping and surrogate null models, while for the third, based on the group-averaged similarity of individual community partitions with consensus community partition, a two-way ANOVA was adopted ($p < 0.05$).

The detection of reproducible brain structural hubs employing community-based hub detection network metrics require reproducible communities. For that reason, we followed hub detection analysis over the best pairs of graph construction schemes and community detection algorithms. Then, we adopted a two-way ANOVA ($p < 0.05$) to detect the best pair of graph construction scheme and community detection algorithms over hub detection analysis, using as input the ICC for the participation index and the within-module z-score across nodes and the agreement-index.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

3. Results

3.1 Quality of the Detected Communities

Q original values were transformed into p-values by comparing them with the 1000 surrogate (permuted) Q values (see sup.material). P-values ranged between 0.013 and 0.021 across participants, scans and graph construction schemes. For the soft clustering algorithms (no. 27-33) that do not include a Q quality index, we measured the NMI between the original community affiliation and the 1000 surrogate-based communities. The group-averaged NMI ranged between 0.08 and 0.12 which supported the quality of the derived soft communities with the whole set of soft clustering algorithms. These findings support the quality of the extracted graph partitions and allow us to include all the participants, scans, graph construction schemes and community detection algorithms in our analysis.

3.2 Group-Averaged Between-Scan Agreement of Communities Affiliations

Figure 4 shows group-averaged between-scan agreement of community affiliations across graph constructions schemes and community detection schemes. Based on the highest group-averaged NMI values and the detected p-values (p = 0.0001) derived from the permutation test, we detected the following four community detection algorithms across the seven graph construction schemes as having the highest between-scan agreement:

1. (mscd_afg): Fast multi-scale community detection algorithm using the criterion from Arenas et al. (2008)
2. (mscd_rb): Fast multi-scale community detection algorithm using the criterion from Reichardt & Bornholdt, (2006)
3. (mscd_rn): Fast multi-scale community detection algorithm using the criterion from Ronhovde & Nussinov (2009)
4. (mscd_so): Fast multi-scale community detection algorithm using stability optimisation (Le Martelot and Hankin, 2012).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

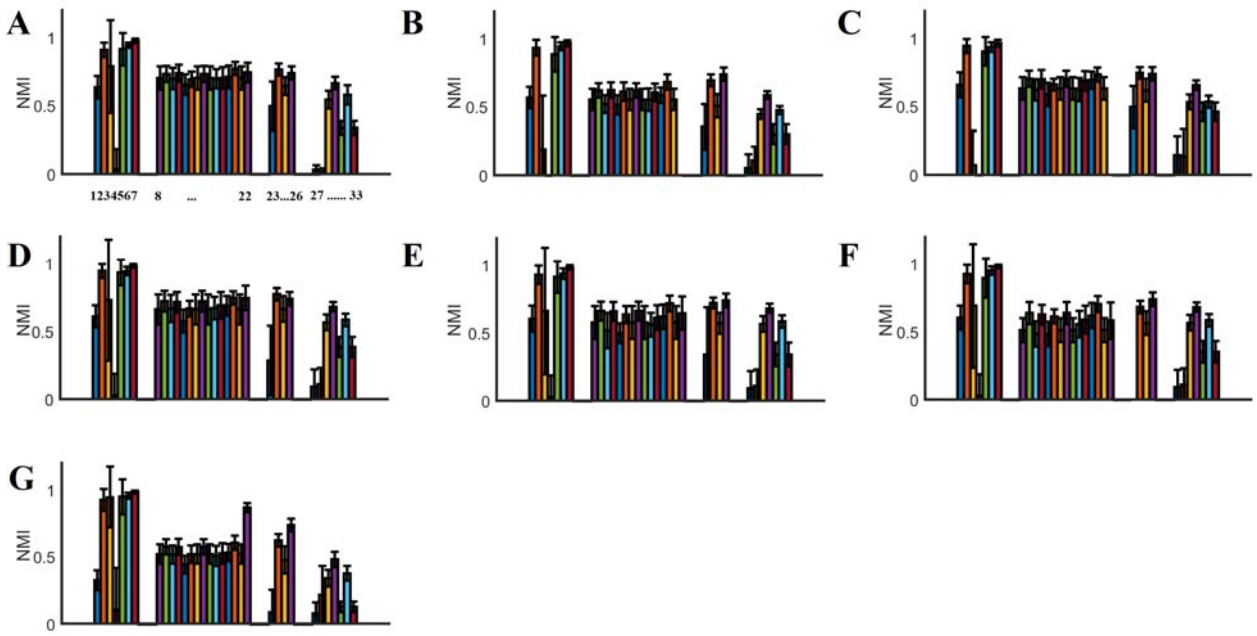


Figure 4. Between-Scan Agreement of Communities Affiliations across graph construction schemes and community detection algorithms.

Every subplot refers to one of the 7 graph-construction schemes. The bars define the group-averaged between-scan agreement of community affiliations. Numbers below the plot in A refer to the number list of community detection algorithms represented in section 2.3.3. Community detection algorithms with the highest agreement between the two scans (NM1 > 0.9) were: mscd_afg,mscd_rb, mscd_rn and mscd_so. For the abbreviations of the graph-construction schemes please see Table 2.

3.3 Similarity of Individual Community Partitions with Consensus Community Partition

The highest similarity between individual community partitions and consensus community partition was detected for the combination of 9-m-OMST graph construction scheme and mscd_so community detection algorithm. The second highest similarity was detected for mscd_afg and 9-m OMST. MSCD with *rb* and *rn* criterions failed to produce an acceptable community similarity between a group representative community estimated via consensus clustering and the individual community affiliations (see Table 3). Two-way ANOVA reported an effect of group-averaged similarity of individual community partitions with consensus community partition across the four community detection algorithms (df =3, F = 54.14, p = 0.783 x 10^{-9}, corrected for multiple comparison) but there was no effect of the

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

graph construction or their interaction. Based on the mean group-averaged similarity across the seven graph construction schemes, we ranked the four community-detection algorithms. The Mscd_so community detection algorithm produced the highest group-averaged similarity of individual community partitions with consensus community partition compared to the four community detection algorithms.

Table 3. Group-averaged similarity of individual community partitions with consensus community partition. Similarities are expressed in NMI scale. We assigned with bold the top ranked values.

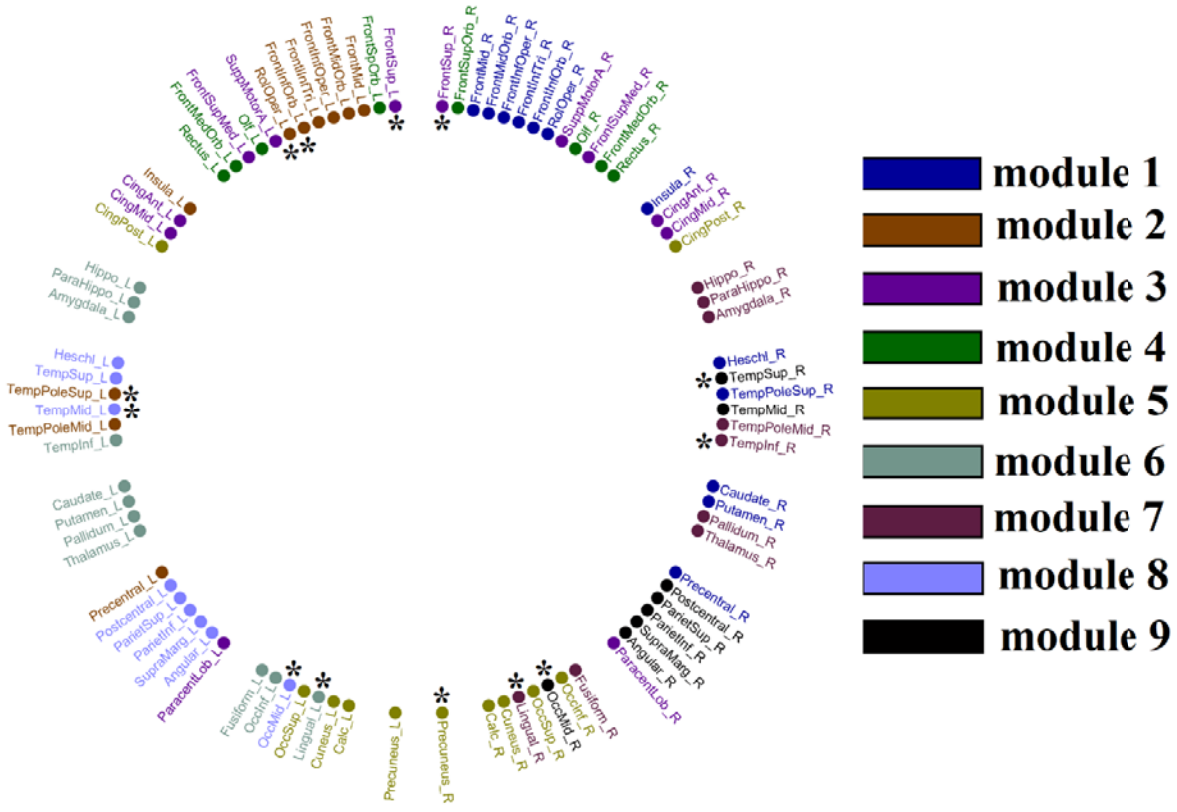
	NS-OMST	NS+FA OMST	9-m OMST	NS-thr	NS-t/FA-w	NS-t/MD-w	FA-t/NS-w
mscd_afg	0.64	0.57	**0.70**	0.61	0.61	0.61	0.39
mscd_rb	0.53	0.39	0.39	0.32	0.31	0.30	0.18
mscd_rn	0	0	0.36	0	0	0	0
mscd_so	**0.70**	0.65	**0.75**	**0.68**	**0.68**	**0.68**	0.64

3.4 Evaluation of the Best Combination of Graph Construction Scheme and Community Detection Algorithm

Based on the evaluation criteria presented in section 2.4 for the final ranking of the pairs of graph construction schemes and community detection algorithms, and taking into account the individual scores, we ranked as the best pair the combination of mscd_so with 9-m OMST. This pair showed a high between-scan group-averaged community similarity estimated with NMI and evaluated via bootstrapping ($p = 0.0001$). The Q quality index has a value of 0.67 with $p = 0.001$. The Mscd_so community detection algorithm also produced the highest group-averaged similarity of individual community partitions with consensus community partition compared to the four community detection algorithms supported by two-way ANOVA.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 5 illustrates the topology of the nine communities across the 90 AAL brain areas based on the combination of mscd_so with 9-m OMST. Importantly, modules number 1, 7 and 9 group together brain areas located within the left hemisphere, modules number 2, 6 and 8 group together brain areas located within the right hemisphere, while modules 3, 4 and 5 involve areas from both hemispheres. Specifically, they involve bilateral ROIs from the fronto-parietal, cingulo-opercular and default mode networks like rectus, anterior and middle gyrus, frontal superior gyrus, frontal superior medial gyrus, supplementary motor area, precuneus, cuneus, calcarine and occipital superior gyrus. The bilateral superior temporal gyrus, superior temporal pole and middle temporal pole play an inter-hemispheric modular connector role (see * in Figure 5). Five out of thirteen consistent connector hubs are located within inter-hemispheric modules. Interestingly, eight homologous brain areas are grouped together in either left (module 8) or right hemisphere (module 9). These areas are: hippocampus, parahippocampal gyrus, amygdala, inferior temporal gyrus, thalamus, pallidum, fusiform and lingual gyrus.



bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 5. Topological Layout of Modular Assignment into the 90 AAL brain areas based on the community affiliation extracting from the consensus matrix related to 9-m OMST graph construction scheme and mscd_so community detection algorithm.
With '*', we denoted the connector hubs detected consistently across participants and repeat scans from the same combination of {mscd-so, 9-m OMST} (see section 3.5).
This circular plot illustrates the 90 AAL brain areas into 45 of the left hemisphere on the left semi-circle and 45 of the right hemisphere on the right semi-circle. Our analysis gave nine communities/modules where each one is encoded with a different colour.

3.5 ICC of Nodal Participation Coefficient Index and Within-Module Z-score

Table 4 shows the group-averaged ICC of nodal Participation Coefficient P_i. ICC values were first estimated per node and were then averaged across the 90 nodes for every pair of graph construction scheme and community detection algorithms. The highest ICC score was detected for the mscd_so algorithm combined with both the 9-m OMST and NS-thr graph construction schemes. On average across the seven graph construction schemes, the mscd_so algorithm also demonstrated the highest average ICC score. However, these findings did not reach statistical significance. Two-way ANOVA revealed no main effect of network-averaged of Participation Coefficient index across the four community detection algorithms (df =3, F = 2.86, p = 0.05) no main effect of the graph construction and no significant interaction term.

Figure 6 shows the nodal ICC for the {mscd_so,9-m OMST} and {mscd_afg, NS-thr} pairs. Applying a Wilcoxon Rank-Sum-test between the two sets of 90 ICCs, we detected a significantly higher ICC for the mscd_so – 9-m OMST pair (p-value = 0.041).

Table 4. ICC of Nodal Participation Coefficient index across every combination of graph construction scheme with community detection algorithms. We denote the top ranked values in bold letters.

	NS-OMST	NS+FA OMST	9-m OMST	NS-thr	NS-t/FA-w	NS-t/MD-w	FA-t/NS-w
mscd_afg	0.54	0.50	0.62	**0.67**	0.58	0.60	0.38
mscd_rb	0.61	0.55	0.68	0.73	0.67	0.65	0.46
mscd_rn	0.58	0.55	0.68	0.67	0.61	0.61	0.43
mscd_so	0.69	0.70	**0.80**	**0.80**	**0.77**	**0.76**	0.43

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

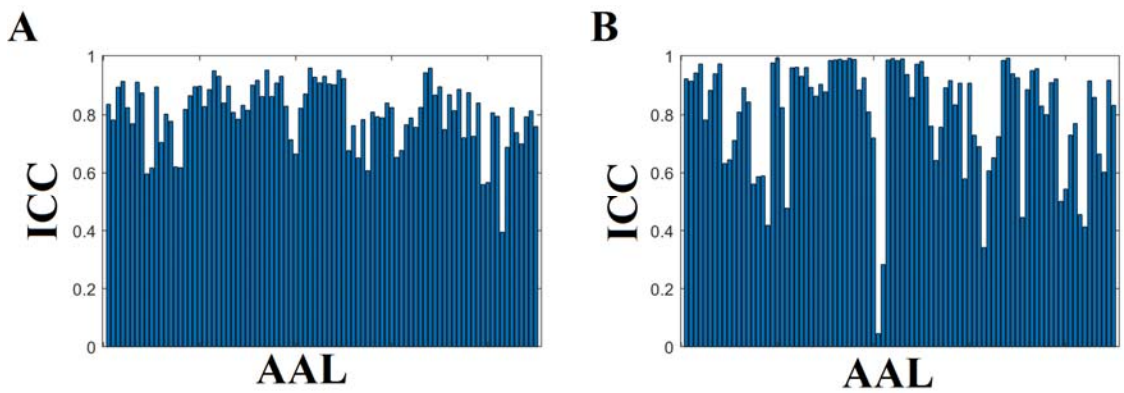


Figure 6. ICC of Nodal Participation Coefficient index for the best combinations of graph construction scheme and community detection algorithm

A. mscd_so – 9-m OMST
B. mscd_afg – NS-thr

Table 5 shows the group-averaged ICC of nodal within-module Z-scores, z_i. ICC values were first estimated per node and were then averaged across the 90 brain areas for every pair of graph construction scheme and community detection algorithms. The highest ICC scores were found for the mscd_rb and mscd_so mode algorithms combined with the NS-thr graph construction scheme. On average across the seven graph construction schemes, the mscd_so mode algorithm demonstrated the highest average ICC score. These findings reached statistical significance. Two-way ANOVA revealed a main effect of network-averaged z_i across the four community detection algorithms (df =3, F = 5.89, p = 0.0037, corrected for multiple comparisons) but no main effect of the graph construction scheme, nor an interaction term. The mscd_rb community detection algorithm produced the highest ICC.

Figure 7 shows the nodal ICC for the {mscd_rb, NS-thr} and {mscd_so, NS-thr}. Applying a Wilcoxon Rank-Sum-test between the two sets of 90 ICCs, we found the mscd_rb algorithm had significantly higher ICC (p-value = 0.0335×10^{-9}).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

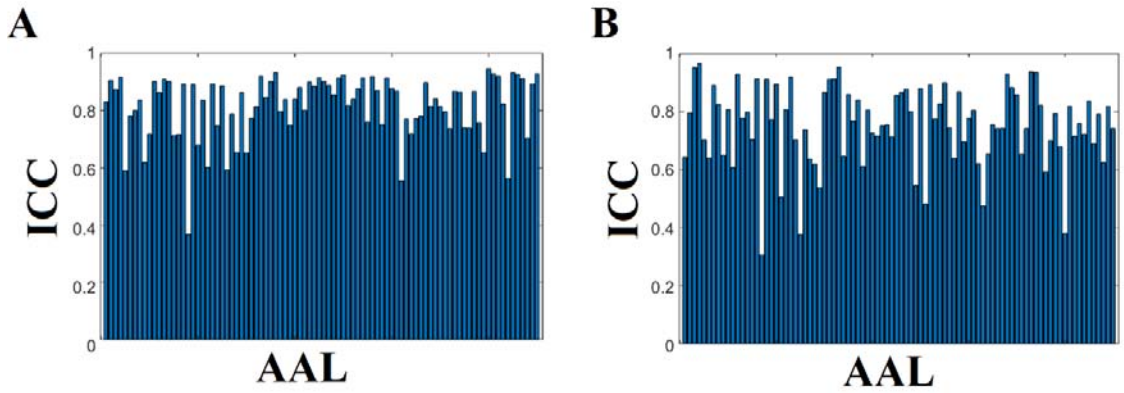


Figure 7. ICC of Nodal Within-Module Z-score for the best combinations of graph construction scheme and community detection algorithm

A. mscd_rb – NS-thr

B. mscd_so – NS-thr

Table 5. ICC of Nodal Within-Module Z-score across every combination of graph construction scheme with community detection algorithms. We denote the top ranked values in bold letters.

	NS-OMST	NS+FA OMST	9-m OMST	NS-thr	NS-t/FA-w	NS-t/MD-w	FA-t/NS-w
mscd_afg	0.71	0.66	0.65	0.79	0.69	0.69	0.69
mscd_rb	**0.74**	0.66	0.64	**0.81**	0.72	0.72	0.73
mscd_rn	0.56	0.57	0.60	0.68	0.55	0.57	0.61
mscd_so	0.73	0.59	0.73	**0.75**	0.60	0.62	0.52

3.6 Reproducibility of Structural Hubs Detection based on Participation Coefficient Index and Within-Module Z-score

We estimated the Agreement Score of both connector and provincial hub detection across the cohort. The Agreement Score of connector hub detection across the cohort was higher than the Agreement Score for provincial hubs (Table 6 versus Table 7).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The highest Agreement Score for provincial hub detection was found for {mscd_so, 9-m OMST} (Table 6). On average across the seven graph construction schemes, the mscd_so algorithm demonstrated the highest average Agreement Score for provincial hub detection. Two-way ANOVA revealed a main effect of network-averaged Within-Module Z-score across the four community detection algorithms (df =3 ,F = 35.55, p = 0.542 x 10^{-6}, corrected for multiple comparison) but no main effect of the graph construction schemes nor an interaction. The mscd_so community detection algorithm produced the highest Agreement Score for provincial hub detection.

The highest Agreement Score for connector hubs detection was found for {mscd_rb, FA-t/NS-w} (Table 7). On average across the seven graph construction schemes, the mscd_rb algorithm demonstrated the highest average Agreement Score of connector hub detection. These findings are supported statistically. Two-way ANOVA revealed a main effect of network-averaged Within-Module Z-score across the four community detection algorithms (df =3, F = 24.3, p = 0.188 x 10^{-4}) but no effect of the graph construction schemes nor their interaction. The mscd_rb community detection algorithm produced the highest agreement score for connector hub detection.

Figure 8 shows the Agreement Score for connector hubs for the pair {mscd_rb, FA-t/NS-w}. The group of connector hubs is indicated alongside modular representation of consensus modules illustrated in Figure 5 and also in Table 8. Interestingly, five out of thirteen consistent connector hubs are located within the inter-hemispheric modules (see Figure 5). Our conclusion is that the combination of modular network metrics P_i and z_i succeeded in uncovering a consistent core of connector hubs but failed to detect provincial hubs consistently.

Table 6. Agreement Scores of Provincial Hubs across every combination of graph construction scheme and community detection algorithm. We assigned with bold the top ranked values.

	NS-OMST	NS+FA OMST	9-m OMST	NS-thr	NS-t/FA-w	NS-t/MD-w	FA-t/NS-w
mscd_afg	0.69	**0.72**	0.75	0.60	0.58	0.51	0.52
mscd_rb	0.46	0.23	0.28	0.30	0.31	0.30	0.21
mscd_rn	0.00	0.00	0.27	0.00	0.00	0.00	0.00
mscd_so	0.69	0.73	**0.83**	**0.80**	0.59	0.61	0.24

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

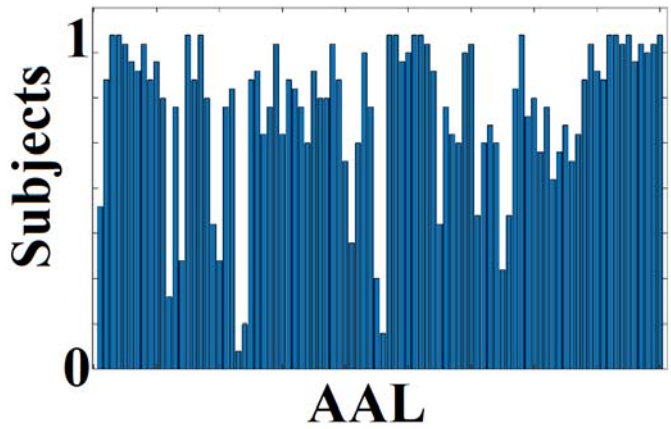


Figure 8. Agreement Score of Connector Hubs for the best pair of {mscd_rb,FA-t/NS-w}.

Table 7. Agreement Scores of Connector Hubs across every combination of graph construction scheme and community detection algorithm. We assigned with bold the top ranked values.

	NS-OMST	NS+FA OMST	9-m OMST	NS-thr	NS-t/FA-w	NS-t/MD-w	FA-t/NS-w
mscd_afg	0.40	0.13	0.23	0.38	0.32	0.31	0.43
mscd_rb	**0.65**	0.42	0.45	**0.67**	0.57	0.58	**0.80**
mscd_rn	0.13	0.05	0.09	0.19	0.15	0.14	0.49
mscd_so	0.17	0.03	0.05	0.12	0.12	0.12	0.17

Table 8. Consistent connector hubs aligned with the detected module number illustrated in Figure 5.

Connector Hubs	Module Number
FrontSup_L	3
FrontSup_R	3
FrontInfOrb_L	2
RolOper_L	2
Lingual_L	6
Lingual_R	7
OccMid_L	8
OccMid_R	9
Precuneus_R	5
TempSup_R	9

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

TempPoleSup_L	2
TempMid_L	8
TempInf_R	7

4. Discussion

We have presented the first extensive study in the literature on the robustness of community detection in structural brain networks by exploring different graph-construction schemes (previously shown to exhibit high repeatability themselves) and various community detection algorithms. Our main findings have direct implications for longitudinal studies and studies comparing healthy controls versus diseased populations.

The key findings of our analysis can be summarized as follows:

1. The repeatability of community affiliations depends heavily on the combination of graph-construction scheme and community detection algorithm. All previously reported studies of network communities adopted a specific pair of graph-construction scheme and community detection algorithm, with the majority of them focused on Newman's modularity objective criterion (Newman, 2006; Sporns and Betzel, 2016; Betzel et al., 2017). Based on our first criterion of high repeatability of community affiliation between the two scans and across the cohort (NMI > 0.9) supported statistically via bootstrapping ($p = 0.0001$), we identified four community detection algorithms as the best choices:

A. mscd_agb

B. mscd_rb

C. mscd_rn

D. mscd_so

These four algorithms gave excellent repeatability across the entire set of graph-construction schemes (see Table 2 and Figure 4).

2. Complementary to the repeatability of the community partitions, we assessed their quality (second criterion). The majority of the graph partition algorithms used here (28 out of 33) estimated a quality index in combination with the community partition. By comparing the quality index of the original community partition to that derived from 1000 surrogate null graph versions, we computed their significance values to be in the range $0.013 \leq P < 0.021$. For the soft clustering algorithms, where this quality

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

index was absent, we estimated the mean clustering distance (mean NMI) between the original partition and partitions derived from 1000 surrogate null graph versions of the original graph. The mean NMI was below 0.1. Our findings indicated community detection at the mesoscale is repeatable and, as the Q quality for the four selected algorithms reached statistically significant levels, is of high quality.

3. Our third criterion was the high similarity of consensus community affiliation with individual community affiliations. The majority of the neuroimaging studies that employed a single graph-construction scheme and one community detection algorithm did not evaluate their findings under this framework (Ryali et al., 2015; Rosero et al., 2017; Akiki and Abdallah, 2019). The consensus matrix integrates community affiliations across the entire cohort counting the number of times two nodes are assigned to the same community. To reveal a representative community affiliation, we have to apply a community detection algorithm. However, almost all previous studies report this representative community affiliation was determined without estimating its similarity with *individual* community affiliations. Following a two-way ANOVA, we observed an effect of group-averaged similarity of individual community partitions with consensus community partition across the four community detection algorithms. Of the four community-detection algorithms, the mscd_so community detection algorithm produced the highest group-averaged similarity of individual community partitions with consensus community partition. We found the highest similarity between individual community partitions and representative community partition for the {mscd_so, 9-m OMST} pair. The second highest similarity was detected for {mscd_afg, 9-m OMST} pair.
4. Two-way ANOVA reported an effect of group-averaged similarity of individual community partitions with consensus community partition across the four community detection algorithms (df =3, F = 54.14, p = 0.783 x 10^{-9}, corrected for multiple comparisons) but there was no effect of the graph construction scheme or their interaction. The mean group-averaged similarity across the seven graph construction schemes allowed us to rank the four graph construction schemes.
5. An important result of our analysis is that soft clustering community detection algorithms gave the least repeatable results. Therefore, we recommend the use of hard-clustering algorithms for the detection of brain communities, at least when using the AAL template.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

6. The best combination of graph-construction scheme (9-m OMST) and community detection algorithm (mscd_so) revealed nine distinct modules (as illustrated topologically in Figure 5) with interesting findings:
   A. Modules 1, 7 and 9 group together brain areas located exclusively within the left hemisphere while modules 2, 6 and 8 group together brain areas located exclusively within the right hemisphere. Modules 3, 4 and 5 group brain areas from both hemispheres together.
   B. Modules 3 – 5 that integrate brain areas from both hemispheres involve bilateral ROIs from fronto-parietal, cingulo-opercular and default mode network like rectus, anterior and middle gyrus, frontal superior gyrus, frontal superior medial gyrus, supplementary motor area, precuneus, cuneus, calcarine and occipital superior gyrus. A multi-tasking fMRI study has previously suggested the flexible role of the fronto-parietal network in cognitive control and adaptive demands of cognitive tasks (Cole et al., 2013).
   C. Five out of thirteen consistent connector hubs are located within inter-hemispheric modules 3 – 5 supporting their inter-connecting role (Figure 5).
   D. Interestingly, eight homologous brain areas were grouped together in either left (module 8) or right hemisphere (module 9). These areas are: hippocampus, parahippocampal gyrus, amygdala, inferior temporal gyrus, thalamus, pallidum, fusiform and lingual. Lesions of hippocampus, parahippocampal gyrus, amygdala and fusiform gyrus in participants with temporal lobe epilepsy caused an impaired associative memory in learning tasks that require learning and recall of objects and faces (Weniger et al., 2004). These four brain areas including the thalamus are those most consistently implicated in neurodegenerative dementias, especially in Alzheimer's Disease, even at an early stage (Manuello et al., 2018).
   E. The bilateral superior temporal gyrus, superior temporal pole and middle temporal pole play an inter-hemispheric integration role. Inter-hemispheric functional connections between temporal lobes predict language impairment in adolescents born preterm (Northam et al., 2012). Phonological awareness, a key factor in reading acquisition was positively correlated with radial diffusivity of the interhemispheric pathways connecting temporal lobes (Dougherty et al. 2007). This bilateral temporal module could play a key role in many functions and dysfunctions.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

7. The core of our study was an extensive analysis to identify the optimal pair of graph-construction scheme and community detection algorithm. The choice of this pairing will also affect repeatability of connector and provincial hub detection based on the participation coefficient score P_i and the within-module z-score z_i. Our results revealed a high repeatability of nodal P_i with the mscd_so algorithm across the seven graph construction schemes. The highest ICC score was reached for the {mscd_so, 9-m OMST} pair. A significantly higher repeatability of nodal z_i was found for mscd_so algorithm compared to the rest of the community detection algorithms. The highest ICC was achieved for the pairs {mscd_rb, NS-thr} and {mscd_so, NS-thr} .
8. The Agreement Score was high only for connector and not for provincial hubs using both modular network metrics. The Agreement Score was statistically higher for the {mscd_so, 9-m OMST} pair than for the rest of the community detection algorithms. We detected a group of thirteen repeatable connector hubs across the cohort, but no provincial hubs. Based on our results, we therefore recommend to not use these modular network metrics for the detection of provincial hubs, at least when using the AAL atlas. The designation of a brain node as a hub depends also on the scale at which brain networks are constructed. Many brain areas in a basic atlas template group together functionally heterogenous subareas and it is possible that a finer-grained parcellation may affect the nodes' classification as a hub or not. For example, the thalamus, despite comprising 50 – 60 specialized sub-nuclei (Herrero et al., 2002) is in many studies, including ours, treated as a single node.

In our previous study on the same cohort, we focused on the repeatability of network topologies focusing on edge weights and graph theoretical metrics. We demonstrated that network topology and edge weights are repeatable, but the repeatability depends on the graph-construction scheme (Messaritaki et al., 2019b). The important finding in this work is that the repeatability of network topologies and edge weights does not guarantee the repeatability of community detection at the mesoscale. In the present study, we focused on this important tool for mesoscale network topological investigations, and the detection of robust communities in structural brain networks over the same participants. To the best of our knowledge, this is the first study in the literature that explores the robustness of community detection over a large set of graph-construction schemes (seven) and community detection algorithms (thirty-three). Our analysis detected an optimal pair of {mscd_so, 9-m OMST} that fulfills the three basic criteria: high repeatability of community affiliations between the

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

two scan-sessions, quality over surrogate null graph partitions and high similarity of group community affiliations with the individual community affiliations. To the best of our knowledge, this is the first time that the second and third criterion were used for the validation of representative consensus community affiliation (this includes studies using a single graph-construction scheme and community detection algorithm).

Running the comparison study for the whole set of thirty-three graph partition algorithms (including graph partition of the original graph and 1000 surrogate null models) takes a few hours on a personal computer. We suggest to the neuroscience community to always run such an analysis over an in-house test-retest dataset acquired with the same settings as in the targeted dataset. Optimizing the set of algorithms over the test-retest study will increase the chance of repeatability of findings over the single-scan dataset. This process will increase the reproducibility of research findings, especially important for cross-sectional studies.

Our study has a few limitations. This dataset involves a specific data acquisition protocol and a specific tractography algorithm. We recommend following our analysis for every study because such an investigation could improve the repeatability and reproducibility of the findings at the mesoscale while also increasing the power of the study at the nodal and network level (Messaritaki et al., 2019b). Additionally, in our study we used only one of the three available b-values to perform the tractography. This was mainly done in order to supplement our previous work (Messaritaki et al., 2019), and we chose the b-value of 2000 s/mm^2, because this value provides a balance between the b-value being sufficiently high to resolve crossing fibers with CSD, while at the same time being ensuring sufficient SNR in the signal for robust measurements, and that higher-order effects of the diffusion do not need to be taken into account when calculating the diffusion metrics. Using one b-value also reflects acquisition protocols routinely used in other studies, and therefore makes our work more applicable to the general literature. At the same time, tractography results could be improved by combining data from all avaialble b-values, and the implications of using different community detection algorithms in those cases should be explored as well. Moreover, three or more scan sessions would be also desirable to get a more robust assessment of repeatability. Scanning the same participant on different scanners and /or with different protocols would also allow assessment of reproducibility as well as repeatability. Lastly, the reproducibility of estimates of structural brain networks is affected by the resolution of the MR data (Vaessen et al., 2010), the parcellation scheme used (Bassett et al., 2011), the interval time between the scan sessions and others.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

5. Conclusions

In this study, we compared several graph-construction schemes and community detection algorithms for the detection of reproducible communities in structural brain networks. Our extensive analysis showed that every choice in both groups of algorithms exhibits different reproducibility in community detection algorithms, as well as in connector/provincial hubs detection. Our analysis indicates that our analytic pathway should be adopted and performed in every study in order to extract reliable results at the mesoscale of structural brain networks.

Declarations of interest

None.

Author Contributions

Conceptualization: SID

Methodology:SID

Software:SID

Validation:SID

Formal analysis:SID, EM

Formal Analysis: SID

Investigation:SID

Resources: The original diffusion MRI are free available from the Human Connectome project. The analysis of tractography and the construction of structural brain networks has been realized by EM.

Data curation:EM,SID

Roles/Writing - original draft: SID,

Writing - review & editing: EM,DJ

Funding acquisition:EM,SID

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Data and code availability

The HCP test-retest data is freely available as listed above.

The code used to generate the graphs for the structural brain networks with the OMST schemes is available at: h*ttps://github.com/stdimitr/multi-group-analysis-OMST-GDD*. The structural brain networks and the code used to perform the reproducibility analysis will be released as soon as the paper will be accepted from author's github website https://github.com/stdimitr.

Source code of community detection algorithms are provided on the Dr.Le Martelot's personal homepage, author's homepage and also were implemented by our team. The collection of the whole set of the algorithms will be reported in our github homepage.

Acknowledgements

We are grateful to the Human Connectome Project for making the test-retest data freely available.

"Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University."

The work was partly funded under the BRAIN Biomedical Research Unit, which is funded by the Welsh Government through Health and Care Research Wales. DKJ is supported by a Wellcome Trust Investigator Award (096646/Z/11/Z) and a Wellcome Trust Strategic Award (104943/Z/14/Z). SID was supported by MRC grant MR/K004360/1 and a Marie Sklodowska-Curie COFUND EU-UK Research Fellowship. EM is supported by a Wellcome Trust ISSF Research Fellowship at Cardiff University (204824/Z/16/Z).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

Aerts H, W. Fias, K. Caeyenberghs, D.,2016. MarinazzoBrain networks under attack: robustness properties and the impact of lesions Brain, 139, 3063-3083

Ahn YY , Bagrow JP , Lehmann S.,2010. Link communities reveal multiscale complexity in networks. Nature 466, 761–764

Akiki, TJ, Abdallah, CG.,2019. Determining the hierarchical architecture of the human brain using subject-level clustering of functional networks. Scientific Reports volume 9, Article number: 19290

Arenas, A., Fernandez, A., and Gomez, S., 2008. Analysis of the structure of complex networks at different resolution levels. New Journal of Physics, 10, 053039.

Avena-Koenigsberger, A., Misic, B. & Sporns, O.,2017. Communication dynamics in complex brain networks. Nat. Rev. Neurosci. 19, 17–33

Baker,S.T.E., Lubman, D.I., Yucel,M., Allen, N.B. , Whittle, S., Fulcher, B.D., Zalesky, A., Fornito, A.,2015. Developmental changes in brain network hub connectivity in late adolescence.J. Neurosci., 35 (24), 9078-9087

Bassett, D.S., Wymbs, N.F., Porter, M.A., Mucha, P.J., Carlson, J.M., Grafton, S.T., 2011. Dy1297 namic reconfiguration of human brain networks during learning. Proc. Natl. Acad. 1298 Sci. U. S. A. 108, 7641–7646.

Bassett, D.S., Brown, J.A., Deshpande, V., Carlson, J.M., Grafton, S.T., 2011. Conserved and variable architecture of human white matter connectivity. Neuroimage 54:1262–1279

Bassett, D. S. & Sporns, O.,2017. Network neuroscience. Nat. Neurosci. 20, 353–364

Bassett, D.S. & Mattar M.G.,2017.A Network Neuroscience of Human Learning: Potential To Inform Quantitative Theories of Brain and Behavior.Trends Cogn Sci. 21(4), 250–264.

Betzel, R.F., Medaglia, J.D., Papadopoulos, L., Baum, G., Gur, R., Gur, R., Roalf, D., Satterthwaite, T.D., Bassett, D.S., 2017. The modular organization of human anatomical brain networks: Accounting for the cost of wiring. Network Neuroscience, 1(1), 42–68

Blondel, Vincent D; Guillaume, Jean-Loup; Lambiotte, Renaud; Lefebvre, Etienne (9 October 2008). "Fast unfolding of communities in large networks". Journal of Statistical Mechanics: Theory and Experiment. 2008 (10): P10008

Borgatti, S.P., Everett, M.G.,2000.Models of core/periphery structures.Social. Netw., 21 (4),375-395

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Braun, U., Schäfer, A., Walter, H., Erk, S., Romanczuk-Seiferth, N., Haddad, L., Schweiger, J.I., 610 Grimm, O., Heinz, A., Tost, H., Meyer-Lindenberg, A., Bassett, D.S., 2015. Dynamic 611 reconfiguration of frontal brain networks during executive cognition in humans. Proc. Natl. 612 Acad. Sci. U.S.A. 112, 11678–83.

Braun,U., Schäfer,A., Bassett,D.S., Rausch,F., Schweiger, J.I., Bilek,E., Erk,S., RomanczukSeiferth,N., Grimm, O., Geiger,L.S., Haddad,L., Otto,K., Mohnke, S., Heinz, A., Zink, M., Walter, H., Schwarz, E. , Meyer-Lindenberg, A., Tost, H. , 2016.Dynamic brain network reconfiguration as a potential schizophrenia genetic risk mechanism modulated by NMDA receptor function . Proc. Natl. Acad. Sci. USA, 113 (44), 12568-12573

Caeyenberghs K, Metzler-Baddeley C, Foley S, Jones DK. Dynamics of the Human Structural Connectome Underlying Working Memory Training. J Neurosci. 2016 Apr 6;36(14):4056-66.

Clarker H, Messaritaki E, Dimitriadis SI, Metzler-Baddeley C (2021).Dementia risk factors modify hubs but leave other connectivity measures unchanged in asymptomatic individuals: a graph theoretical analysis. doi: https://doi.org/10.1101/2020.10.08.331025

Cole M.W., Reynolds J.R., Power J.D., Repovs G., Anticevic A., Braver T.S.Multi-task connectivity reveals flexible hubs for adaptive task control.Nat. Neurosci., 16 (9) (2013), pp. 1348-1355,

Colizza V, A. Flammini, M.A. Serrano, A.,2006. Vespignani. Detecting rich-club ordering in complex networks.Nat. Phys., 2 (2), 110-115

Collin G, M.P. van den Heuvel, L. Abramovic, A. Vreeker, M.A. de Reus, N.E. M van Haren, M.P.M. Boks, R.A. Ophoff, R.S.,2016. KahnBrain network analysis reveals affected connectome structure in bipolar disorder Hum. Brain Mapp., 27,122-134

Danon, L., D´ıaz-Guilera, A., Arenas.A.,2006. The effect of size heterogeneity on community identification in complex networks. Journal of Statistical Mechanics: Theory and Experiment, (11),P11010

Dimitriadis, S. I., Laskaris, N. A., Del Rio-Portilla, Y., and Koudounis, G. C. (2009). Characterizing dynamic functional connectivity across sleep stages from EEG. Brain Topogr. 22, 119–133.

Dimitriadis, S. I., Kanatsouli, K., Laskaris, N. A., Tsirka, V., Vourkas, M., and Micheloyannis, S. (2012). Surface EEG shows that Functional Segregation via phase coupling contributes to the neural substrate of mental calculations. Brain Cogn. 80, 45–52.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Dimitriadis,S.I., Antonakakis,M., Simos,P., Fletcher,J.M., A.C. Papanicolaou,A.C.,2017a.Data-driven Topological Filtering Based on Orthogonal Minimal Spanning Trees: Application to Multi-Group MEG Resting-State Connectivity.Brain Connectivity 7(10),661-670.

Dimitriadis,S.I., Drakesmith,M., Bells,S., Parker,G.D., Linden,D.E., D.K. Jones,D.K.,2017b. Improving the reliability of network metrics in structural brain networks by integrating different network weighting strategies into a single graph.Front. Neurosci., 11, p. 694

Dimitriadis, S.I., Salis,C., Tarnanas,I., Linden,D.E.,2017c.Topological filtering of dynamic functional brain networks unfolds informative chronnectomics: a novel data-driven thresholding scheme based on orthogonal minimal spanning trees (OMSTs).Front. Neuroinf., 28,p.11

Ding C, Li T, Peng W, and Park H. Orthogonal nonnegative matrix tri-factorizations for clustering. In Proc. of the 12th ACM SIGKDD Int'l Conf., pages 126–135, 2006.

Dong, X., Frossard, P., Vandergheynst, P., Nefedov, N., 2014. Clustering on multi-layer graphs via subspace analysis on grassmann manifolds. Signal Processing, IEEE Transactions on 62(4),905–918

Dougherty, R. F., Ben-Shachar, M., Deutsch, D. K., Hernandez, A., Fox, G. R., and Wandell, B. A., 2007. Temporal-callosal pathway diffusivity predicts phonological skills in children. Proc. Natl. Acad. Sci. U.S.A. 104, 8556–8561

Drakesmith,M, K. Caeyenberghs, A. Dutt, S. Zammit, C.J. Evans, A. Reichenberg, G. Lewis, A.S. Dvid, D.K. Jones,2015.Schizophrenia-like topological changes in the structural connectome of individuals with subclinial psychotic experiences.Hum. Brain Mapp., 36, 2629-2643

Feinberg DA et al., Multiplexed echo planar imaging for sub-second whole brain fMRI and fast diffusion imaging.PLoS One, 5 (12) (2010), e15710

Fortunato, S., 2010. Community detection in graphs. Phys Rep 486,(3-5),75–174

Fortunato, S., Castellano, C., 2012. Community structure in graphs Computational Complexity. Springer, 490–512.

Friston, K.J., 2011. Functional and effective connectivity: a review. Brain Connect 1(1),13–36.

Giusti, G., Ghrist,R., Bassett,D.S.,2016.Two's company, three (or more) is a simplex: algebraic-topological tools for understanding higher-order structure in neural data.J. Comput. Neurosci., 41(1),1-14

Glasser et al, The minimal preprocessing pipelines for the Human Connectome Project, *NeuroImage* 80:105-124.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Guimerà R, Sales-Pardo M, Amaral LA. Modularity from fluctuations in random graphs and complex networks. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2004;70(2 Pt 2):025101. doi:10.1103/PhysRevE.70.025101

Guimera, R. and Amaral, L.A.N., 2005. Functional cartography of complex metabolic networks. Nature 433, 895–900

Hagmann, P., Kurant, M., Gigandet, X., Thiran, P., Wedeen, V.J., et al., 2007. Mapping human whole-brain structural networks with diffusion MRI. PLoS One 2, e597

Herrero, M.T., Barcia,C., Navarro.,J.M.,2002.Functional anatomy of thalamus and basal ganglia.Child's Nerv. Syst., 18,386-404

Huang, J., Sun, H., Liu, Y., Song, Q., and Weninger, T., 2011. Towards online multiresolution community detection in large-scale networks. PLoS ONE, 6, e23829.

Kernighan, B. W.; Lin, Shen,1970. An efficient heuristic procedure for partitioning graphs". Bell System Technical Journal. 49: 291–307

Imms,P., Clemente,A., Cook,M., D'Souza,W., Wilson,P.H., Jones,D.K., Caeyenberghs,K.,2019.The structural connectome in traumatic brain injury: a meta-analysis of graph metrics.Neuroscience and Behavioural Reviews, 99, 128-137

Lancichinetti A, Fortunato S (2009a) Community detection algorithms: a comparative analysis. Physical Review E 80(5 Pt 2):056117.

Lancichinetti, A., Fortunato, S. Kertesz, J.,2009b. Detecting the overlapping and hierarchical community structure in complex networks. N. J. Phys. 11, 033015

Lancichinetti,A., & Fortunato,S., 2012. Consensus clustering in complex networks. Scientific Reports 2, Article number: 336.

Leemans, A., Jeurissen, B., Sijbers, J., Jones, D., 2009. ExploreDTI: a graphical toolbox for processing, analyzing, and visualizing diffusion MR data. Proceedings of the International Society for Magnetic Resonance in Medicine, Honolulu, Hawaii.

Le Martelot, E., Hankin, C.,2011. Multi-scale community detection using stability as optimisation criterion in a greedy algorithm. KDIR, 216–225.

Le Martelot E, Hankin C,2012a. Multi-scale community detection using stability optimisation. Int J Web Based Communities, 9,323–348.

Le Martelot E, Hankin C,2012b. Fast multi-scale detection of relevant communities in large-scale networks. Comput J , 56, 1136–1150.

Manuello, J., Nani, A., Premi, E., Borroni, B., Costa, T., Tatu, K.,et al.,2018. The pathoconnectivity profile of alzheimer's disease: a morphometric coalteration network analysis. Front Neurol 8,739

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Messaritaki, E., Dimitriadis,S.I., Jones,D.K.,2019a.Assessment of the reproducibility of structural brain networks derived using different edge-weighting strategies Proceedings of the 27th Annual Meeting of the ISMRM, Montreal, Canada

Messaritaki, E.,Dimitriadis,S.I., D.K. Jones.,D.K.,2019b. Optimization of graph construction can significantly increase the power of structural brain network studies. NeuroImage.199(1),495-511

Messaritaki E, Foley S, Schiavi S, Magazzini L ,Bethany Routley B, Jones DK, Singh KD (2021).Predicting MEG resting-state functional connectivity from microstructural information. *Network Neuroscience* 1–42.

Meunier, D., Lambiotte, R., Fornito, A., Karen, D.E., Bullmore, E.T.,2009. Hierarchical modularity in human brain functional networks. Frontiers in neuroinformatics,3,37

Meunier, D., Lambiotte, R., Bullmore, E.T.,2010. Modular and hierarchically modular organization of brain networks. Front Neurosci 8;4:200.

Moeller et al, Multiband multiscale GE-EPI at 7 Tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI, Magn. Reson. Med., 63 (5) (2010), pp. 1144-1153

Nelson, B.G., Bassett,D.S., Camchong, J., Bullmore,E.T., Lim,K.O.,2017.Comparison of large-scale human brain function and anatomical networks in schizophrenia.Neuroimage: Clinic, 15,439-448

Newman, M.E., Girvan, M.,2004. Finding and evaluating community structure in networks. Physical Review E 69(2 Pt 2): 026113

Newman, M.E.J., 2006. Modularity and community structure in networks. Proc Natl Acad Sci U S A 103,8577–8582.

Newman, M.E.J., 2012. Communities, modules and large-scale structure in networks. Nat Phys 8(1),25–31

Nigro S, Riccelli R, Passamonti L, Arabia G, Morelli M, Nisticò R, Novellino F, Salsone M, Barbagallo G, Quattrone A. Characterizing structural neural networks in de novo Parkinson disease patients using diffusion tensor imaging. Hum Brain Mapp. 2016 Dec;37(12):4500-4510.

Northam, G.B., Liégeois, F., Tournier, J.D., et al.,2012. Interhemispheric temporal lobe connectivity predicts language impairment in adolescents born preterm. Brain,135(Pt 12),3781–3798

Pavan,M., Pellilo.,M.,2017.Dominant sets and pairwise clustering.IEEE Trans on PAMI, 29,167-17

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Owen, J.P., Ziv,E., Bukshpun,P., Pojman,N., Wakahiro,M., Berman,J.I., Roberts, T.P.,Friedman,E.J., Sherr, E.H., Mukherjee,P.,2013.Test-retest reliability of computational network measurements derived from the structural connectome of the human brain.Brain Connect., 3, 160-176

Ozdemir, A., Bolanos, M., Bernat, E., Aviyente, S., 2015. Hierarchical Spectral Consensus Clustering for Group Analysis ofFunctional Brain Networks. IEEE Trans Biomed Eng 62(9),2158–2169

Palla, G., Derényi, I., Farkas, I., Vicsek,T.,2005. Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society. Nature 435, 814–818

Park, H.J., Friston, K.J.,2013. Structural and Functional Brain Networks: From Connections to Cognition. Science 342(6158),1238411

Pessoa, L.,2018.Understanding emotion with brain networks. Curr Opin Behav Sci.,19,19–25.

Psorakis, I., Roberts,S., Ebden,M., Sheldon. B.,2011.Overlapping Community Detection using Nonnegative Matrix Factorization. Physical Review E 83, 066114

Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., Parisi, D., 2004. Defining and identifying communities in networks. ProcNatl Acad Sci U S A 101(9),2658–2663.

Reichardt, J.,Bornholdt, S., 2006. Statistical mechanics of community detection. Physical Review E, 74, 016110.

Ronhovde, P., Nussinov, Z., 2009. Multiresolution community detection for megascale networks by information-based replica correlations. Physical Review E, 80, 016109.

Rasero, J., Pellicoro, M., Angelini, L., Cortes, J.M., Marinazzo, D., Stramaglia, S.,2017. Consensus clustering approach to group brain connectivity matrices. Netw Neurosci., 1(3),242–253

Reichardt, J., Bornholdt, S., 2006. Statistical mechanics of community detection. Phys. Rev. E 74(1):016110

Rosvall, M., Bergstrom, C.T., 2008. Maps of random walks on complex networks reveal community structure. Proc Natl Acad Sci USA 105,1118–1123.

Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. NeuroImage 52(3),1059–1069.

Ryali, S., Chen, T., Padmanabhan, A., Cai, W., Menon, V.,2015. Development and validation of consensus clustering-based framework for brain segmentation using resting fMRI. J Neurosci Methods. 240,128–140.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Setsompop et al, Blipped-controlled aliasing in parallel imaging for simultaneous multiscale echo planar imaging with reduced g-factor penalty Magn. Reson. Med., 67 (5) (2012), pp. 1210-1224

Shi, J., Malik, J.,2000. Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 22(8), 888–905

Smith, R.E., Tournier, J.D., Calamante, F., Connelly,A.,2015.The effects of SIFT on the reproducibility and biological accuracy of the structural connectome. Neuroimage, 104, 253-265

Sotiropoulos,S.N., Jbabdi,S., Xu,J., Andersson,J.L., Moeller,S., Auerbach,E.J., Glasser, M.F., Hernandez, M., Sapiro,G., Jenkinson,M., D.A. Feinberg,D.A., Yacoub,E., Lenglet,C ., Essen,D.C.V., Ugurbil,K., Behrens,T.E.,2013a.Advances in diffusion MRI acquisition and processing in the human connectome project.Neuroimage, 80,125-143

SotiropoulosSN, S. Moeller, S. Jbabdi, J. Xu, J.L. Anderson, E.J. Auerbach, E. Yacoub, D. Fe inberg, K. Setsompop, L.L. Wald, T.E.J. Behrend, K. Ugurbil, C. Lenglet,2013b. Effects of image reconstruction on fibre orientation mapping from multichannel diffusion MRI: reducing the noise floor using SENSE. Magn. Reson. Med., 70 (6) (2013b), pp. 1682-1689

Sporns, O, Tononi, G, Edelman, G M. Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. Cerebral cortex. 10(2):127–141.2000

Sporns, O., 2013. Network attributes for segregation and integration in the human brain. Curr Opin Neurobiol 23(2),162–171.

Sporns, O. & Betzel, R.F.,2016. Modular brain networks. Annu. Rev. Psychol. 67, 613–640.

Taylor PN, Cheol E. Han, Jan-Christoph Schoene-Bake, Bernd Weber, Marcus Kaiser.Structural connectivity changes in temporal lobe epilepsy: Spatial features contribute more than topological measures.NeuroImage: Clinical,8,2015,pp322-328.

TzourioMazoyer,N., Landeau,B., Papathanassiou,D., Crivello,F., Etard,O., Delcroix,N., Maz oyer,B., Joliot,M.,2002.Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcelation of the MNI MRI single-subject brain. Neuroimage, 15,273-289

Yuan,J.P., HenjeBlom,E., Flynn,T., Chen,Y., Ho,T.T., Connolly,C.G., DumontWalter,R.A., Yang,T.T., Xu,D.,Tymofiyeva,O.,2019.Test-retest Reliability of Graph Theoretical Metrics in Adolescent Brains. Brain Connect. 9(2),144-154

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Vaessen, M.J., Hofman, P.A.M., Tijssen, H.N., Aldenkamp, A.P., Jansen, J.F.A., Backes W.H.,2010. The effect and reproducibility of different clinical DTI gradient sets on small world brain connectivity measures. NeuroImage 51,1106–1116

van den Heuvel, M.P. and Sporns, O., 2011. Rich-club organization of the human connectome. J. Neurosci. 31, 15775–15786

van den Heuvel, M. P. & Sporns, O.,2013. Network hubs in the human brain. Trends Cogn. Sci. 17, 683–696

Van Essen,D.C., Smith,S.M., Barch,D.M., Timothy, E.J., Behrens, Yacoub, E., Ugurbil,K., for the WU-Minn HCP Consortium.,2013. The WU-Minn Human Connectome Project: An overview. NeuroImage 80, 62-79.

Vidaurre, D.R., Becker,A.R., Quinn, A.J., Alfaro-Almagro, F., Smith, S.M., Woolrich, M.W., 2018.Discovering dynamic brain networks from big data in rest and task.Neuroimage, 180 (B), 646-656

Welton, T., Kent, D.A. , Auer, D.P., Dineen,R.A.,2015. Reproducibility of graph-theoretic brain network metrics: a systematic review. Brain Connect., 5 (4),193-202

Weniger, G., Boucsein, K., Irle, E.,2004.Impaired associative memory in temporal lobe epilepsy subjects after lesions of hippocampus, parahippocampal gyrus, and amygdala. Hippocampus. 14(6),785-96.

Xu et al, Highly accelerated whole brain imaging using aligned-blipped-controlled-aliasing multiband EPI.Proceedings of the 20th Annual Meeting of ISMRM (2012), p. 2306

Zhong, S., He,Y., Gong.G,2015.Convergence and divergence across construction methods or human brain matter networks: an assessment based on individual differences.Hum. Brain Mapp., 36,1995-2013

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Appendix 1. Detecting Structural Hubs based on Nodal Participation Coefficient P_i and Within-Module Z-Score z_i

By recovering the community partition and estimating the participation coefficient, we can classify brain hubs into provincial and connector hubs (Guimera and Amaral,2005). 'Provincial hubs' are high-degree nodes that primarily connect to nodes in the same module. 'Connector hubs' are high-degree nodes that show a diverse connectivity profile by connecting to several different modules within the network. Brain hubs are important brain areas that are vulnerable and susceptible to disconnection and dysfunction in brain disorders. Rich club organization of structural hubs supports the robustness of inter-hub connections and promotes the efficient information exchange between brain areas and its integration across the brain (van den Heuvel and Sporns,2011).

The distinction of nodes into hubs and non-hubs by a combination of network topology and community affiliation is supported by a pair of network metrics called: participation coefficient P_i and within-module z-score, z_i. This definition has been first reported by Guimera and Amaral (2005). Here, we first reported the reliability of these nodal metrics in structural brain networks.

The degree of a node i is defined as , where A_{ij} is the adjacency matrix of the graph. Within-module z-score for node i is defined as

$$z_i = \frac{\kappa_i - \bar{\kappa}_{s_i}}{\sigma_{s_i}}, \quad (1)$$

where κ_i is the number of edges of node i to other nodes in its module s_i s_i is the average of κ over all the nodes in s_i , and σ_{si} is the standard deviation of κ in s_i .

The participation coefficient P_i, for node i is defined as

$$P_i = 1 - \sum_{s=1}^{N_M} \left(\frac{\kappa_{is}}{\kappa_i} \right)^2 \quad (2)$$

where k_{is} is the number of links of node i to nodes in module s, and k_i is the total degree of node i. The participation coefficient of a node is therefore close to 1 if its links are uniformly distributed among all the modules and zero if all its links are within its own module.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Provincial Hubs: If a node with a large degree, $k \gg 1$, has at least 5/6 of its links within the module, then it follows that $P = 1 - (5/6)^2 - (k/6)(1/k^2) = 0.31 - 1/(6k) \approx 0.30$. i.e. , $P \leq 0.3$

Connector Hubs: If a node with a large degree has at least half of its links within the module, then it follows that $P = 1 - 1/4 - (k/2)(1/k^2) = 0.75 - 1/(2k)$. Since $k \gg 1$, $P < 0.75$ for such nodes.

Both provincial and connector hubs demonstrate a high within-module z-score which means that they have many within-module edges. In this work, we used the threshold originally proposed by Guimera and Amaral, (2005) for the z_i dimension as $z_i > 2.5$ for both types of studied hubs (see Figure 5 in Guimera and Amaral, (2005)). In the P_i dimension, we defined a node as provincial hub if $P_i <= 0.3$ and as connector hub if $0.3 < P_i < 0.75$.

The intra-class correlation coefficient (ICC) was estimated for every nodal participation coefficient index, P_i, and within-module z-score z_i across the cohort and for every selected pair of graph construction scheme and community detection algorithm that showed higher group-averaged community similarity (NMI > 0.9).

Appendix 2. Graph partition Algorithms

We described briefly the thirty-three graph-partition algorithms used in the present study.

Hard clustering algorithms are divided into three groups:

A. Fast multi-scale community detection algorithms:

1. (fast_mo): Fast greedy modularity optimisation algorithm based on the multi-scale algorithm but optimised for modularity (mono-scale) (Le Martelot and Hankin, 2012a,b).
2. (mscd_afg): Fast multi-scale community detection algorithm using the criterion from Arenas et al. (2008)
3. (mscd_hslsw): Fast multi-scale community detection algorithm using the criterion from Huang et al. (2011)
4. (mscd_lfk): Fast multi-scale community detection algorithm using the criterion from Lancichinetti et al. (2009b)
5. (mscd_rb): Fast multi-scale community detection algorithm using the criterion from Reichardt & Bornholdt, (2006)

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

6. (mscd_rn): Fast multi-scale community detection algorithm using the criterion from Ronhovde & Nussinov (2009)
7. (mscd_so): Fast multi-scale community detection algorithm using stability optimisation (Le Martelot and Hankin, 2012).

B. **Multi-Scale Community Detection algorithms using Stability as Optimisation Criterion in a Greedy Algorithm and Multi-Scale Community Detection using Stability Optimisation as described in (Le Martelot and Hankin,2011):**

8. (gso_discrete): Greedy stability optimisation using the Markov chain model.
9. (gso_continuous): Greedy stability optimisation using a time-continuous Markov process.
10. (msgso_discrete): Multi-step greedy stability optimisation using the Markov chain model.
11. (msgso_continuous): Multi-step greedy stability optimisation using a time-continuous Markov process.
12. (rgso_discrete): Randomised greedy stability optimisation using the Markov chain model.
13. (rgso_continuous): Randomised greedy stability optimisation using a time-continuous Markov process.
14. (gso_discrete_markovian): Front-end function calling any of the aforementioned algorithms based on discrete Markovian model. We specified a pre-processing phase removing nodes with a single neighbour and a post-processing phased based on the Kernighan-Lin algorithm adapted for stability (Kernighan and Lin,1970).
15. (gso_discrete_continuous): Same as above but based on continuous Markovian model.
16. (Newman): Greedy modularity optimisation with Newman's fast algorithm.
17. (Danon): Greedy optimisation of Danon et al.'s criterion (Danon et al., 2006).
18. (Louvain): Uses the Louvain method (Blondel et al.,2008)
19. (Louvain_modularity): Uses the Louvain method to optimise modularity criterion with gamma=0.5.
20. (Louvain_so): Uses the Louvain method to optimise ‘stability’ criterion with Markovian time

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

21. (reichardt): Greedy optimisation of Reichardt & Bornholdt's criterion (Reichardt and BornHoldt, 2006).
22. (ronhovde): Greedy optimisation of Ronhovde & Nussinov's criterion (Reichardt and BornHoldt, 2006)

C. Hard community detection algorithms involving state-of-the-art graph partition algorithms

23. (shi_malik): From tens of available spectral clustering algorithms, we adopted the algorithm from Shi and Malik, (2000)
24. (dominant_sets): Dominant sets (Pavan and Pelillo, 2007). We have adopted this algorithm in our previous studies (Dimitriadis et al., 2009,2012)
25. (modularity): The algorithm by Newman, 2006
26. (infomap): An important algorithm for identifying communities in large graph-oriented problems and synthetic graphs (Rosvall M, Bergstrom CT, 2008))

The soft clustering algorithms we considered are:

27. (ogso_discrete): Greedy stabilisation optimisation with overlapping communities for discrete Markov Model. The network is first converted to an edge-graph and then community detection is performed on the edge-graph
28. (ogso_continuous): Greedy stabilisation optimisation with overlapping communities for continuous Markov Model. The network is first converted to an edge-graph and then community detection is performed on the edge-graph
29. (link): Link Communities: The network is first converted to an edge-graph and then community detection is performed on the edge-graph. In an edge-graph, nodes refer to edges of the original SBN while edges denote edges between pair of connected edges of the original SBN. This algorithm uncovers overlapping community structure via hierarchical clustering of network links (Ahn et al., 2010). We used single-linkage criterion.
30. (link_complete): Link Communities: Same as above but we used complete-linkage criterion.
31. (nnmf): Non-negative Matrix Factorization (NNMF): (Psorakis et al., 2011)

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

32. (k-cliques): This algorithm partitions a graph into soft communities using the simplest subgraph forms, the k-cliques. Here, we used 3-cliques (Palla et al., 2005)

33. (orthogonal nonnegative matrix t-factorizations): Orthogonal nonnegative matrix t-factorizations (Ding et al., 2006)

For the first set of multiscale community detection algorithms, we considered the following set of resolution parameter ps = [0.1 0.15, …, 0.55, 0.60]. For the discrete and continuous Markovian time series, we searched over ts = [0.1, 0.02, …, 0.9,1.0]. Finally, for Louvain's modularity method and Reichardt's method, we considered γ = [0.05, 0.1, 0.15, …,0.95,1.5]. The aforementioned parameters were optimized over the highest group-averaged community affiliation similarity between the two scans estimated with NMI. For further details, see the following section.

For the Louvain's methods, we ran the algorithms 1000 times and we followed the construction of consensus matrix approach. We described this approach in section 2.3.8 for the construction of consensus matrix across participants and scans.

Appendix 3. Normalized Mutual Information: Graph – partition similarity

To assess the reproducibility of the thirty-three community detection techniques across the seven graph-construction schemes and repeat-scan sessions, we first quantified the similarity between the community partitions from the two scan sessions separately for every participant using the Normalized Mutual Information (NMI) (Alexander-Bloch et al. 2012), defined as follows (Lancichinetti and Fortunato,2009a):

$$NMI(A,B) = \frac{-2\sum_{i=1}^{C_A}\sum_{j=1}^{C_B} N_{ij}\log\left(\frac{N_{ij}N}{N_{i.}N_{.j}}\right)}{\sum_{i=1}^{C_A} N_{i.}\log\left(\frac{N_{i.}}{N}\right) + \sum_{j=1}^{C_B} N_{.j}\log\left(\frac{N_{.j}}{N}\right)} \quad (1)$$

where A and B are the community partitions of two SBNs from the two scan sessions while C_A, C_B are the number of communities in partition A and B, correspondingly. N denotes the number of nodes (here 90), while N_{ij} is the overlap between A's and B's communities i and j which practically means the number of common nodes between the two partitions. N_i and N_j are the total number of nodes in A's and B's communities i and j respectively. The NMI ranges from 0 to 1 where 0 corresponds to two independent partitions and 1 to identical

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.07.082271; this version posted April 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

partitions. This definition was used for hard community partition comparisons while for soft community partition, we adopted the homologue definition of NMI tailored to soft graph clustering (Lancichinetti et al., 2009a).

We calculated NMI values between every possible pair of scans and for each of the seven graph-construction schemes and the thirty-three community detection algorithms giving an space of {7 graph-construction schemes x 33 community detection algorithms}. The NMI was then averaged across the 37 participants, to create a group-averaged NMI, and ranked the community detection algorithms with high test-retest reproducibility (group-averaged values NMI > 0.9) in at least one of the seven graph construction schemes (see Figure 4).

To quantify the statistical significance of group-mean community similarity quantified with NMI, we adopted a nonparametric test via bootstrapping procedure (Dimitriadis et al., 2012). Practically, the outcome of every community detection algorithm in our cohort is a matrix (dimensions {(no of participants) x (no of nodes)}) with elements containing integer numbers assigned to every detected community. Analysing the scan-rescan data gives two such matrices, one per scan session. The bootstrapping procedure comprises:

a) The two matrices (dimensions {(no of participants) x (no of nodes)}) tabulating the community affiliations across the cohort are integrated into a single matrix (dimensions
{2 x no of participants x no of nodes})
b) The rows of this integrated matrix are shuffled, mixing the community affiliations of participants across scan sessions
c) The integrated matrix is divided into two matrices of equal size {no of participants x no of nodes} and the group mean NMI between those two shuffled matrices is estimated.
d) Steps (b) and (c) are repeated 10,000 times
e) A P-value is assigned to every graph construction scheme - community detection algorithm pair by counting the number of times the permuted between-scan group mean similarity exceeds the actual between-scan group mean similarity, divided by the number of permutations (here 10,000)