2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Skylar J Brooks[1†], Sean M Parks[1†] and Catherine Stamoulis[1,2*], *Member, IEEE*

Big Data-Driven Brain Parcellation from fMRI: Impact of Cohort Heterogeneity on Functional Connectivity Maps

Abstract— Ongoing large-scale human brain studies are generating complex neuroimaging data from thousands of individuals that can be leveraged to derive data-driven, anatomically accurate brain parcellations. However, despite their promise and many strengths, these data are highly heterogeneous, a characteristic that may affect the anatomical accuracy and generalization of the template but has received relatively little attention. Using multiple similarity measures and thresholding approaches, this study investigated the topological intra- and inter-individual variability of resting-state (rs) functional edge maps (often used for brain parcellation), estimated from rs-fMRI connectivity in n = 5878 children from the Adolescent Brain Cognitive Development (ABCD) study. Findings from this initial investigation indicate that choosing a subject- vs cohort-based threshold for estimating edge maps from connectivity matrices does not significantly impact the map topology. In contrast, the choice of similarity measure and non-linear relationship between similarity and edge map sparsity may have a significant impact on map classification and the generation of parcellation atlases. Multi-level classification revealed multiple clusters with a potentially complex mapping onto biological variables beyond simple demographics.

***Clinical Relevance*— Case-control neuroimaging studies should use domain-specific (e.g., demographics-specific) atlases for parcellating the brain, to improve accuracy and rigor of cohort comparisons. To be generalizable, such atlases need to be derived from large datasets, which are inherently heterogeneous. In a cohort of 5878 children (age ~9-10 years), this study systematically assessed the impact of heterogeneity and similarity of edge maps, which are derived from rs-fMRI connectivity and typically used to generate parcellation atlases.**

I. Introduction

Network Neuroscience is a rapidly growing area of research that focuses on studying the brain's structural and functional circuitry as networks with specific topological properties. These properties directly reflect the brain's ability to process information from the outside world, respond to cognitive demands, and change as a result of normal (e.g., aging or development) or pathological processes [1-3]. Depending on the resolution of the modality used to measure brain activity (e.g., EEG or fMRI), a number of assumptions are made in order to represent the brain as a network of discrete regions. In modalities with low spatial resolution, such as scalp EEG, the number of nodes is limited by the spatial sampling (typically >1 cm even in high-density EEG). In this case, the number of nodes is usually assumed to be equal to the number of electrodes. However, in fMRI, with a 1-2 mm resolution, images need to be downsampled by parcellating the brain into regions, which become the network nodes.

Significant prior work has focused on the problem of neuroimage parcellation. Cohort-independent and cohort-specific, as well as atlas-based vs data-derived methods have been developed[4-9]. Each makes various assumptions and has advantages and shortcomings. Atlas-based parcellation uses a common, cohort-independent template (derived from anatomical knowledge and/or an independent sample) to divide the brain into distinct regions. The choice of atlas can significantly impact the parcellation accuracy and generalization. For example, within the adult lifespan, there are often differences in the boundaries between parcels as a result of anatomical and cortical thickness differences between individuals and/or aging. There is growing evidence that age-specific parcellations are necessary in order to better characterize variability in functional brain activity [10]. Recent efforts have also focused on *data-derived* approaches based on morphological or connectivity similarities between individuals within a cohort, which can drive groupings of anatomical or functional images for estimating parcels [5-9].

Deriving a single representative template from a heterogeneous cohort of individuals with potentially different anatomical or functional characteristics remains challenging. This is a significant hurdle in neuroimaging, since use of inaccurate and/or non-representative atlases may lead to variable, atlas-dependent and poorly reproducible results [11]. To facilitate this process, the characteristics of the data heterogeneity need to be better understood, particularly in large datasets such as those generated by the Human Connectome Project (HCP) and the Adolescent Brain Cognitive Development (ABCD) study [12]. Both are generating data from thousands of individuals. It is therefore critical to systematically investigate the structure and inherent (dis)similarity of anatomical and functional images collected as part of these studies. This may lead to improved data-driven parcellations that are biologically meaningful and representative of the data and their natural clusters.

A number of data-derived parcellation atlases have been derived based on resting-state (RS) fMRI (rs-fMRI), and it has been shown that it is possible to delineate functional brain areas based on RS connectivity patterns [4]. Many of these atlases have been generated using approaches that are agnostic to the inherent variability of the rs-fMRI dataset, which can be significant, particularly in developing brains. In 5878 children from the ABCD study [12], this study systematically investigated within- and across-brain heterogeneity and similarity of RS connectomes and their topology. Hierarchical classification of topological similarity between estimated edge maps was used to elucidate the structure of the dataset which could lead to the estimation of multiple (instead of a single) parcellation templates.

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

II. MATERIALS AND METHODS

A. Neuroimaging Data

Minimally preprocessed rs-fMRI and structural MRI (T1) data from 5878 children, age 9-10 years [2885 males (49.1%), 2993 females (50.9%)], from the ABCD study - the largest ever longitudinal study on adolescent brain development, were analyzed. Each participant had a maximum of four 5-min runs of rs-fMRI. These data are publicly available [13].

B. Preprocessing

Data were processed using the software package SPM12 and a recently developed pipeline for fMRI analysis [14] as follows: 1) A small number (5-16) of initial frames were first removed, depending on the scanner (in this study GE or Siemens) and software version, to allow for equilibration of the T1w fMRI signal; 2) Images were then co-registered to the structural MRI, slice-time corrected according to scanner type, and normalized to MNI-152 space. Head motion-related artifacts were suppressed from each voxel's time series using a General Linear Model (GLM) with filtered rigid body motion parameters as model variables; 3) Residual time series were filtered (in both directions) with a third order elliptical filter in the frequency range 0.01 - 0.25 Hz, typically assumed to be physiologically meaningful for BOLD signals [15]; 4) Frame displacement (FD) was calculated from filtered motion estimates.

Frames with FD > 0.3 mm were censored [16, 17]. Only runs with ≤10% censored frames for motion and brains with at least one such run were further analyzed. This relatively conservative threshold was necessary given spurious correlations in fMRI due to excessive motion [16]. To reduce the very high dimension (>900,000 voxels) of data normalized to MNI space, in this initial study a relatively high-resolution discretization was used, which was based on the Schaefer-1000 atlas, the Melbourne parcellation (for subcortical regions), and the probabilistic Cerebellar Atlas [9,18-19]. This reduced the spatial dimension to 1088 regions (1000 cortical, 54 subcortical and 34 cerebellar). Finally, another level of (data-driven) denoising was performed on the 1088 time series using a modified Ensemble Empirical Mode Decomposition (EEMD), in order to remove residual artifact contributions [20]. Each time series was decomposed into a set of ~4-5 narrowband components (modes). Based on their intrinsic frequencies and amplitudes some modes were eliminated and new time series were synthesized as linear superpositions of the remaining modes.

C. Estimation of topological similarity

In this initial investigation, RS connectivity was estimated as the peak cross-correlation of each pair of time series, resulting in a 1088 X 1088 symmetric matrix. Two approaches were used to assess the similarity between connectivity matrices: one based directly on the topological distance between matrix pairs and the other based on comparisons of edge maps derived from these matrices. The first approach only used the run with the lowest number of censored frames from each brain and did not require thresholding, whereas the second approach was applied for within and across-brains and used multiple thresholds.

Distance measure between connectivity matrices: The distance between these matrices was estimated as the sum of the squared difference between individual elements in each pair of matrices, resulting in a 5878 X 5878 cohort matrix, which was then used in hierarchical clustering to identify clusters of brains with similar RS connectivity topology.

C1. Edge Map Generation: The Canny edge detection algorithm [21] was used to estimate binary edge maps from connectivity matrices. The algorithm identifies local gradient peak changes in images, using strong and weak edge thresholds. These were estimated using run-specific and cohort-wide thresholding. Resulting binary maps had ones for edges at or above the strong threshold and zeros for all edges below the weak threshold. Weak edges with values between the two thresholds were kept only if they were connected to strong edges. The goal of the analysis using two sets of thresholds was to assess their respective impact on measures of topological similarity and potential advantages of global (cohort-level) thresholds.

C2. Thresholding:

a) Run-specific thresholds: For each brain and run, the connectivity matrix was z-transformed and its derivative was computed. The median and 25th percentile of pairwise differences in each column were estimated as well as their confidence intervals (CI). The upper CI of the median difference was used as the strong threshold in all cases. In separate analyses, the lower CI of the median or upper CI of the 25th percentile were used as weak thresholds.

b) Cohort-based thresholds: A common set of strong and weak edge thresholds was estimated from all brains and runs. For each run and derivative of the z-transformed connectivity matrix, the median, 25th and 75th percentile were estimated, as well as each statistic's median over the entire set of brains and corresponding CI. Separate analyses were conducted using different CI combinations as the thresholds.

C3. Similarity Measures: In addition to distance between connectivity matrices, 4 similarity coefficients (Jaccard, Sørensen-Dice, Baroni-Urbani and Simple Matching [22-24]) were also estimated from pairs of edge maps (M1, M2), based on the contingency table for edges with: a (1,1) – edge present in both maps (positive match); b (0,1) – absent in M1, present in M2; c (1,0) – present in M1, absent in M2; d (0,0) – edge absent in both maps (negative match).

The most conservative measure is the *Jaccard coefficient (J)*, estimated as the intersection of edges in two maps as:

$$J = \frac{a}{a+b+c}$$

Given that appropriately thresholded RS connectomes are sparse, this intersection is expected to be relatively low. The *Sørensen-Dice coefficient*

$$S = \frac{2a}{2a+b+c}$$

is similar, but

positive matches are more heavily weighted. The *Baroni-Urbani coefficient* $B=\frac{a+\sqrt{ad}}{a+b+c+\sqrt{ad}}$ considers both positive and negative matches but weighs them unequally. Finally, the *Simple Matching coefficient*, the least conservative measure $SM=\frac{a+d}{a+b+c+d}$, weighs positive and negative matches equally. Quantifying edge map similarity using different measures is important for assessing the impact of each measure on how connectomes are grouped for estimating group-specific parcellation templates.

Within-brain similarity: For each brain with multiple good RS runs, pairs of edge maps were compared. Brains with only one run were excluded from within-brain analyses. *Between-brain similarity*: Each brain's edge maps were also compared with every other brain's maps. Median similarity between each brain and all others as well as overall similarity within the cohort were then estimated.

D. Hierarchical clustering

Connectivity distance measures and edge map similarity coefficients were classified using hierarchical clustering [25]. This approach was selected assuming a potential underlying hierarchical organization of the cohort's RS connectivity (dis)similarities, possibly as a result of age and/or sex-related topological differences. Although in this exploratory analysis an optimum number of clusters was not estimated from the data, separate classifications using a variable number of clusters were performed.

III. Results

A. Edge map-based topological similarity

Table 1 summarizes within- and between-brain edge map similarity measures, for run-specific and cohort-wide weak and strong thresholds based on the median.

TABLE I. EDGE MAP SIMILARITY

Threshold	***Within Brain Similarity: Median (IQR)***			
	Jaccard	***Sorensen-Dice***	***Baroni-Urbani***	***Simple Matching***
Run-Specific	0.25 (0.05)	0.40 (0.07)	0.80 (0.04)	0.98 (0.01)
Common	0.22 (0.10)	0.36 (0.13)	0.75 (0.14)	0.97 (0.03)
	Across Brains Similarity: Median (IQR)			
Run-Specific	0.16 (0.03)	0.27 (0.4)	0.73 (0.04)	0.97 (0.01)
Common	0.12 (0.06)	0.22 (0.09)	0.66 (0.15)	0.96 (0.03

Table 2: Within and across-brain similarity measures based on run-specific and cohort-based edge map thresholding.

As expected, edge map similarity strongly depended on the chosen measure, with low similarity based on the Jaccard coefficient and very high similarity based on Simple Matching. Differences between these estimates as a function of run-specific vs common (cohort-wide) thresholds were modest and non-significant for both within and across brain similarity measures. This indicates that the choice of cohort-level thresholds does not negatively impact the edge map estimation at the individual run and brain, presumably due to the large sample size used to estimate these statistical thresholds and their CIs. Overall, cohort-wide thresholds resulted in slightly less sparse (higher density maps (see Figure 1). Finally, as expected, topological similarities between individual brains were substantially lower than those within brains across measures. The very low Jaccard and very high Simple Matching coefficient may, however, be misleading in deciding whether such a large cohort can be used to estimate a common parcellation based on RS connectivity. Instead, a moderate measure, such as the Baroni-Urbani coefficient, may be more appropriate.

The relationship between edge map similarity and map sparsity (based on density, reflecting the number of connections above the selected thresholds), was also assessed. Density was estimated as the ratio of non-zero edges in a map and the maximum possible number of edges. Figure 1 shows the relationship between map density and similarity (for each of the 4 coefficients) for run-specific vs cohort-wide thresholds used to estimate the edge maps. There was no clear relationship between map sparsity and the Jaccard or Sorensen-Dice coefficients when a run-specific threshold was used, but an overall trend of decreasing similarity with increasing density for the other two coefficients. There was a non-linear relationship between Jaccard, Sorensen-Dice, and Baroni-Urbani measures and density when a cohort-level threshold was used. However, following an initial decrease in similarity with increasing map density, all 3 coefficients increased, at least within the estimated range of densities, which in these maps is overall low due to the sparsity of RS connectivity. This indicates that sparsity may also impact the topological similarity between RS connectomes positively or negatively.

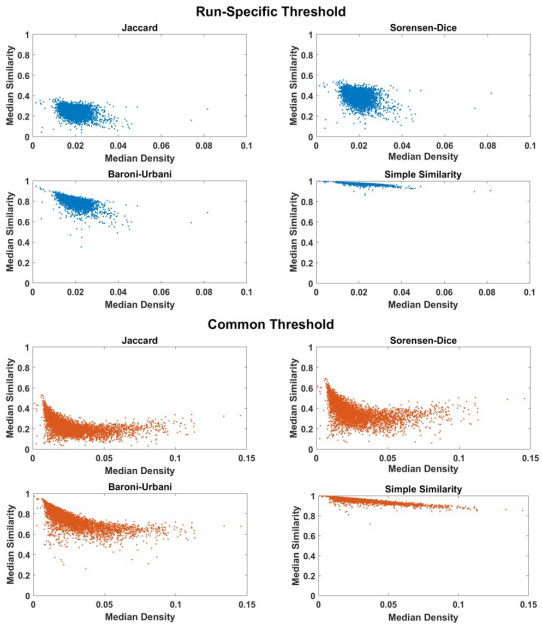

Figure 1: Map similarity measures as a function of density for run-specific (top) and cohort-wide (bottom) thresholds.

B. *Classification of connectome similarity*

Distance measures between connectivity matrices were first classified using hierarchical clustering, assuming 3-6 clusters. Given potentially significant low-level variability (noise) between raw connectivity matrices, this measure did not reveal any meaningful (inherently separable) clusters of similar topology. This was also the case when Jaccard, Sorenson-Dice, and Simple-Matching coefficients were clustered. In contrast, when using the Baroni-Urbani coefficient, clearly separable clusters were identified ($p<0.001$ for differences between clusters). Figure 2 shows the distribution of Baroni-Urbani coefficient values in each of 3 (left panel), 4 (middle panel) and 6 (right panel) clusters.

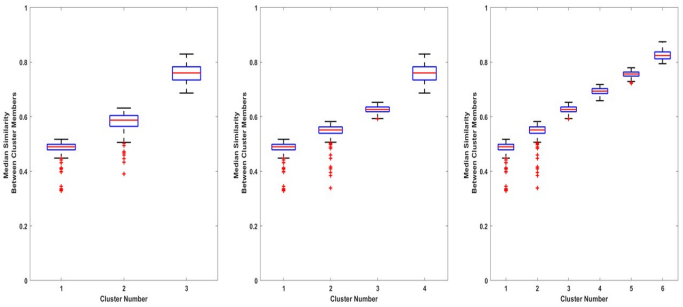

Figure 2: Distribution of Baroni-Urbani coefficients in 3, 4 or 6 clusters following hierarchical classification.

The relationship between clustering and biological variables of interest, particularly age and sex, was then assessed. There was no clear mapping between cluster membership and these variables (for any number of clusters), suggesting that the inherent topological similarity of these connectomes may be the result of multiple potentially non- separable and potentially unmeasured (latent) factors.

IV. CONCLUSION

This study investigated the impact of resting connectome heterogeneity on the utility of connectivity-based and thus data-driven brain parcellations in almost 6000 developing children from the ABCD study. A custom fMRI processing pipeline was used in the analyses, which provided increased flexibility for image processing compared to available tools (e.g., CONN-SPM). It was shown that the choice of topological (edge map) similarity measure may have a significant impact both on the interpretation of similarity within and between brains and their classification. Edge maps may appear dissimilar when using a very conservative similarity measure such as the Jaccard coefficient. Thus, deriving a single parcellation template based on it may not be appropriate. The same maps may appear highly similar when using a non-conservative measure such as the Simple Matching coefficient, leading to templates estimated from highly variable connectivity matrices. Indeed, classification of edge map similarity based on these extreme measures led to poorly separable clusters. Instead, a moderate measure such as the Baroni-Urbani coefficient may be more appropriate in this process, and indeed led to multiple well-separable edge map clusters. This suggests that more than one parcellation may be required to best represent the natural topological similarity between brains in a large cohort. Finally, clusters could not be directly mapped onto connectome differences in age or sex, suggesting that connectome similarity in large datasets may be associated with multiple factors beyond biological variables. Future analyses are planned to investigate these factors and estimate multiple sub-cohort-based parcellations.

REFERENCES

[1] Bullmore, E, Sporns, O (2009), Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci., 10: 186-198.

[2] Power, JD, Cohen, AL, Nelson, SM et al (2011), Functional Network Organization of the Human Brain, Neuron, 72 (4): 665-678.

[3] Bassett, DS, Sporns, O (2017). Network Neuroscience. Nat. Neurosci, 20: 353-364.

[4] Cohen, A et al (2008), Defining functional areas in individual human brains using resting functional connectivity MRI, NeuroImage, 41 (1): 45-57.

[5] Gordon, EM, Laumann, TO, Adeyemo, B, et al (2014), Generation and evaluation of cortical area parcellation from resting-state correlations, Cereb Cortex, 26 (1): 288-303.

[6] Wig, GS et al (2014), Parcellating an individual subject's cortical and subcortical brain structures using snowball sampling of resting-state correlations. Cereb Cortex, 24 (8): 2036-2054.

[7] Glasser, MF et al (2016). A multi-modal parcellation of human cerebral cortex. Nature, 536 (7615): 171-178.

[8] Eickhoff, S, Yeo, B, Genon, S (2018), Imaging-based parcellations of the human brain. Nat Rev Neurosci, 19 (11): 672-686.

[9] Schaefer, A et al (2018), Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. Cereb Cortex, 28 (9): 3095-3114.

[10] Han, L et al (2018), Functional Parcellation of the Cerebral Cortex Across the Human Adult Lifespan, Cereb Cortex, 28 (12): 4403–4423.

[11] Lawrence, RM et al (2020). Standardizing Human Brain Parcellations. bioRxiv, 1-16.

[12] Casey, BJ et al (2018). The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. Developmental Cognitive Neuroscience, 32: 43-54.

[13] NIMH Data Archive: https://nda.nih.gov

[14] https://github.com/cstamoulis1/Next-Generation-Neural-Data-Analysis-NGNDA-

[15] Yuen, NH, et al (2019). Intrinsic frequencies of the resting-state fMRI signal: The frequency dependence of functional Connectivity and the effect of mode mixing. Frontiers in Neuroscience, 13: 900.

[16] Power, JD, et al (2012), Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion, NeuroImage, 59 (3): 2142-2154.

[17] Siegel, JS, Power, JD, Dubis, JW, et al (2014), Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. Human Brain Mapping, 35 (5): 1981-1996.

[18] Tian, Y, Margulies, DS, Breakspear, M, et al (2020), Topographic organization of the human subcortex unveiled with functional connectivity gradients. Nature Neuroscience, 23 (11): 1421-1432.

[19] Diedrichsen, J et al (2009). A probabilistic MR atlas of the human cerebellum, NeuroImage, 46 (1): 39-46.

[20] Torres, ME, et al (2011). A Complete Ensemble Empirical Mode Decomposition with Adaptive Noise. IEEE Int. Conf. On Acoust., Speech and Signal Proc., ICASSP-11: 4144-4147.

[21] Canny, J (1986). A Computational Approach To Edge Detection. IEEE Trans Pattern Analysis and Machine Intelligence, 8(6): 679-698.

[22] Dice, LR (1945), Measures of the Amount of Ecologic Association Between Species. Ecology, 26 (3): 297-302.

[23] Baroni-Urbani, C, Buser, MW (1976), Similarity of Binary Data. Systematic Zoology, 25 (3): 251-259.

[24] Sokal, RR, Michener, CD (1958), A statistical method for evaluating systematic relationships. University of Kansas Science Bulletin, 38, 1409-1438.

[25] Kaufman, L, Roussew, P (1990), Finding Groups in Data, An Introduction to Cluster Analysis, John Wiley.