Data in Brief 43 (2022) 108454

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Dataset of brain functional connectome and its maturation in adolescents

Zack Y. Shan[a,*], Abdalla Z. Mohamed[a], Paul Schwenn[a], Larisa T. McLoughlin[a], Amanda Boyes[a], Dashiell D. Sacks[a], Christina Driver[a], Vince D. Calhoun[b], Jim Lagopoulos[a], Daniel F. Hermens[a]

[a] *Thompson Institute, University of the Sunshine Coast, Birtinya, QLD, Australia*
[b] *Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Georgia State University, Georgia Institute of Technology, Emory University, Atlanta, GA, USA*

ARTICLE INFO

Article history:
Received 16 June 2022
Revised 30 June 2022
Accepted 4 July 2022
Available online 8 July 2022

Dataset link: Dataset of brain functional connectome and its maturation in adolescents (Reference data)

Keywords:
fMRI
Functional connectivity
Adolescent
Brain developmental changes
Longitudinal study

ABSTRACT

We provided the dataset of brain connectome matrices, their similarities measures to self and others longitudinally, and Kessler's psychological distress scales (K10) including the response to each question. The dataset can be used to replicate the results of the manuscript titled "A longitudinal study of functional connectome uniqueness and its association with psychological distress in adolescence". The functional connectome (whole-brain and 13 networks) matrices were calculated from the resting-state functional MRIs (rs-fMRIs). We collected rs-fMRI and Kessler's psychological distress scale (K10) in 77 adolescents longitudinally up to 9 times from 12 years of age every four months. After removal of data with excessive motion, 262 functional connectome matrices were provided with this paper. The 300 regions of interest (ROIs) were defined using the Greene lab brain atlas. The functional connectome matrices were calculated as correlations between time series from any pair of ROIs extracted

DOI of original article: 10.1016/j.neuroimage.2022.119358
* Corresponding author.
E-mail address: zshan@usc.edu.au (Z.Y. Shan).

https://doi.org/10.1016/j.dib.2022.108454

from pre-processed fMRIs. This dataset could be potentially used to

1. Understand developmental changes in the functional brain connectivity,
2. As a normal control database of functional connectome matrices,
3. Develop and validate connectome and network-related analysing methods.

Specifications Table

Subject	Neuroscience: Developmental
Specific subject area	Developmental changes in functional connectome in resting state functional MRI of adolescents.
Type of data	Table Figure Functional connectome matrices Connectome similarity and K10 measures
How the data were acquired	Resting-state functional MRIs (rsfMRIs) were collected using a 3 Tesla Siemens Skyra (Erlangen, Germany) scanner with a 64-channel head and neck coil. Kessler's psychological distress scale was collected using the K10 questionnaire.
Data format	Analyzed Filtered
Description of data collection	The rsfMRI data were acquired using echo-planar imaging multiband sequence: TR/TE = 1600ms/30ms, spatial resolution = 3.027 × 3.027 × 3.0 mm, acquisition matrix = 74 × 74, flip angle = 65°, multiband factor = 4. A total of 300 fMRI volumes were collected with the eye closed in 8 minutes.
Data source location	Institution: Thompson Institute, University of the Sunshine Coast City/Town/Region: Birtinya/QLD Country: Australia Primary data sources: The primary data of fMRI and K10 scores from 77 patients at up to 9 time points were acquired at the same location.
Data accessibility	**The data are hosted on Mendeley Data.** **Repository name: Mendeley Data** **Data identification number:** doi:10.17632/y7pcrp63yp.3 **Direct URL to data:** https://data.mendeley.com/datasets/y7pcrp63yp/3
Related research article	Z.Y. Shan, A.Z. Mohamed, P. Schwenn, L.T. McLoughlin, A. Boyes, D.D. Sacks, C. Driver, V.D. Calhoun, J. Lagopoulos, D.F. Hermens, A longitudinal study of functional connectome uniqueness and its association with psychological distress in adolescence, Neuroimage, 2022; 258:119358. doi:10.1016/j.neuroimage.2022.119358. Note: The figures (DIB Figs. 4–7) cited by the above paper were moved to Mendeley Data for the relevance of this Data in Brief article.

Value of the Data

- These data can be used to replicate the results of the co-submitted manuscript and understand developmental changes in the functional brain connectivity.
- Researchers in developmental neuroscience and youth mental health can benefit from these data by investigating brain functional connectome maturation.
- These data can be used as a normal control database of functional connectome matrices.

1. Data Description

This manuscript provides data of functional connectome matrices (connectome_mat.zip) and their similarities (Similarities_mat.zip) in MATLAB (MATLAB R2019a, The MathWorks Inc., Natick, MA) mat files. Connectome matrices were calculated based on 262 resting-state fMRI data collected from 32 females and 45 males at up to 9 time points (TPs, aged from 12–15.5 years old). The number of scans per participant and age at each scan is summarised in Fig. 1.

The functional connectivity for each participant at a TP was provided as a 300*300 matrix which contains correlations between time series of any pair of 300 regions of interests (ROIs) defined according to Greene lab brain atlas [1]. We also provide similarity measures of a connectome of a given participant at a time point (T01, T02, …, T08) to those from all participants in the next time point. Connectomes from the same participant at two consecutive TPs were required to calculate self-similarity, thus 158 similarity matrices from 63 participants were provided. Each mat file contains a 29*n matrix in which each row represents a network (Table 1), and the first column is the self-similarity and the 2-n columns are similarities-to-others in the next time point (n-1 others).

The functional connectome uniqueness, self-similarity, similarities-to-other, age, sex, and K10 measures are provided as an excel spreadsheet. Three uniqueness indices were calculated: uniqueness_max, uniqueness_avg, and uniqueness_2nd_max is defined as the ratio of the self-similarity to the maximal, averaged, and the second maximal similarities-to-others, respectively. The similarities-to-others are averaged as similarities to others having the same sex and having different sex.

Finally, provided in this manuscript are several tables and figures of the secondary analyses of this dataset, including:

Table 2 Coefficient estimation of the frequency serving as maximal similarity-to-others pairing

Fig. 2. Scatter plot of averaged K10 scores and the number of uniqueness measurements.

Fig. 3. Node and network definitions.

Note: The figures (DIB Fig 4–7) cited by the above paper were moved to Mendeley Data for the relevance of this Data in Brief article.

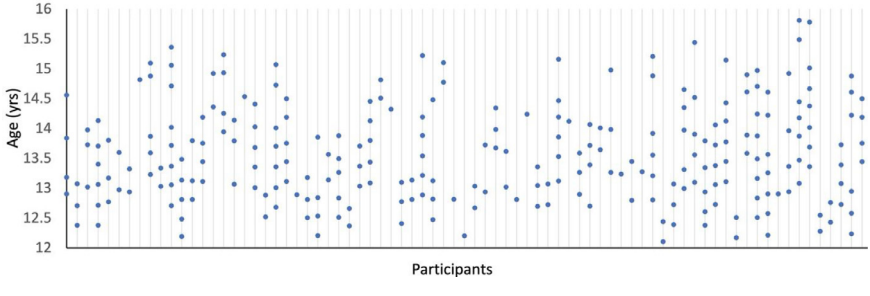

Fig. 1. The line plot of the number of scans per participant. Each line represents a participant (77 participants in total), and a dot represents an MRI scan. The y-axis indicates the age when a scan was performs.

Table 1
Networks in the similarity matrix[§]

Row number	Networks
1	whole_brain
2	Auditory_inter
3	Auditory_intra
4	Cingulo-Opercular_inter
5	Cingulo-Opercular_intra
6	Default Mode_inter
7	Default Mode_intra
8	Dorsal Attention_inter
9	Dorsal Attention_intra
10	Fronto-Parietal_inter
11	Fronto-Parietal_intra
12	Medial TemporalLobe_inter
13	Medial TemporalLobe_intra
14	Parieto-Medial_inter
15	Parieto-Medial_intra
16	Reward_inter
17	Reward_intra
18	Salience_inter
19	Salience_intra
20	Somatomotor Dorsal_inter
21	Somatomotor Dorsal_intra
22	Somatomotor Lateral_inter
23	Somatomotor Lateral_intra
24	Ventral Attention_inter
25	Ventral Attention_intra
26	Visual_inter
27	Visual_intra
28	unassigned_inter
29	unassigned_intra

[§] : The *_inter are connectivities between a network node to a node that is not in the same network. The *_intra are connectivities between any two nodes from the same network.

Table 2
Coefficient estimation of the frequency serving as maximal similarity-to-others pairing[§]

				95.0% Confidence Interval for B	
Coefficients	Standardized Beta	t	Sig.	LB	UB
(Constant)		.386	.702	-.171	.252
Age	-.046	-.301	.765	-.019	.014
Sex	.108	.731	.469	-.011	.024
AVG_mc_X	-.271	-.831	.411	-.614	.256
AVG_mc_Y	.176	.334	.740	-.195	.272
AVG_mc_Z	.182	.308	.759	-.089	.121
STD_mc_X	.551	1.528	.134	-.128	.924
STD_mc_Y	-.636	-.812	.422	-.452	.193
STD_mc_Z	.152	.208	.836	-.093	.115
V_Excluded	.118	.333	.741	-.006	.008

[§] : LB = lower bound, UB = upper bound, AVG_mc_X, AVG_mc_Y, AVG_mc_Z, STD_mc_X, STD_MC_Y, and STD_mc_Z are averaged motion and standard deviation in X, Y, Z of fMRI data. V_Excluded is the number of volumes excluded for motion correction during fMRI preprocessing.

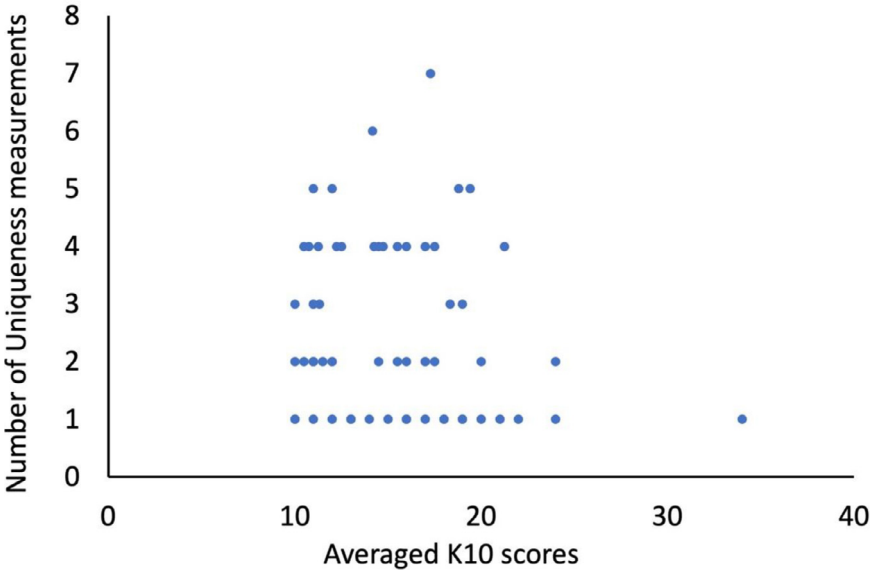

Fig. 2. Scatter plot of averaged K10 scores and the number of uniqueness measurements. The x-axis represents the K10 score averaged over multiple visits of each participant. The y-axis is the number of uniqueness measurements of each participant. No significant correlation was observed between averaged K10 scores and the number of measurements.

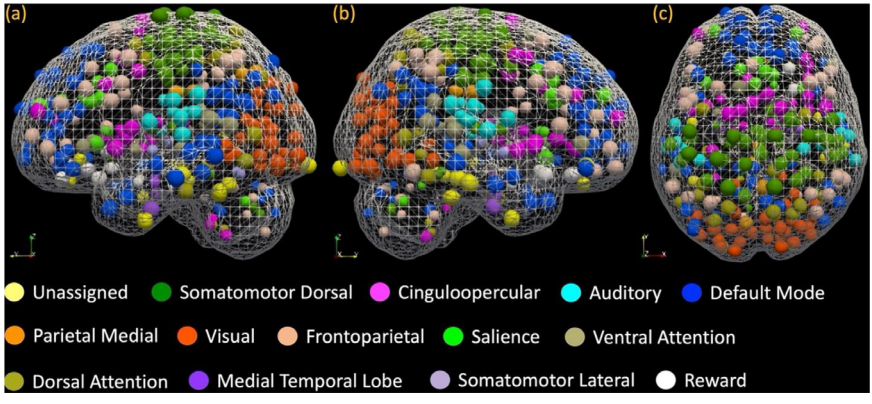

Fig. 3. Node and network definitions. The 300 regions of interest (ROIs) and an averaged brain template were surface rendered with coloured surfaces and triangle meshes respectively. ROIs were spheres with radii of 4-5mm according to the Greene lab brain atlas [1] and coloured by assignment of networks.

1.1. Secondary Analysis of Relationship Between the Number of Uniqueness Measurements and K10 Scores

A bivariate correlation analysis was performed to detect if the number of uniqueness measurements is related to the K10 scores in this sample. No significant correlation was found ($P = 0.25$, Fig. 2).

1.2. Regions of Interest (ROIs) and Network Definition

We used a 300 ROI set from Greene lab [1] for node and network definition in this study. Seitzman et al. used a semi-automated, multistep process for determination of consensus functional network communities. The unassigned network is the unlabelled regions that cannot be assigned to a network, similar to the unknown network identified in Power et al. [2].

1.3. Secondary Analysis of Multivariate Regression Analysis of the Frequency of Being Maximal Similarity-to-Others

To explore if the known factors affect a connectome serving as maximal similarity-to-other in the above uniqueness definition, we performed a multivariate regression analysis: the frequency of being maximal similarity-to-others pairing partner as the dependent variable; age, sex, averaged motion and standard deviation in X, Y, Z direction, and the number of volumes excluded during fMRI preprocessing as predictors. Note that only the TP2 similarity-to-others was used here to avoid the potential bias that some participants may appear at more TPs than others.

2. Experimental Design, Materials and Methods

2.1. Participants

The Longitudinal Adolescent Brain Study (LABS) is an ongoing project at the USC Thompson Institute, USC. In addition to MRI data, electroencephalography (EEG) and a battery of behavioural data (self-report questionnaires, cognitive assessment, and neuropsychiatric interviews) were collected as part of the LABS project. Participants were recruited from the general population in the Sunshine Coast region and enrolled at 12 years old. Participants were excluded from the study if they have/are a major neurological disorder, intellectual disability, major medical illness, not proficient in spoken and written English, or sustained a head injury that involved loss of consciousness for greater than 30 min.

The current study utilised the structural and resting-state functional MRI data and the associated K10 scores. A total of 262 datasets from 32 females and 45 males collected at 9 time points (TPs, 12–15.5 years old) were presented with this manuscript after 6 data sets were excluded during pre-processing. The uniqueness analysis was performed using MRI data collected from the same subject at two consecutive TPs, with 158 uniqueness indices estimated from 63 participants (37 males and 26 females) at 8 TPs.

2.2. Psychological Distress Assessment

The psychological distress was assessed using the K10 questionnaire [3]. The K10 questionnaire is a 10-item self-report with a five-point Likert scale asking about feelings over the past 30 days to measure depression and anxiety levels. The total score is the sum of individual item scores, with the overall score ranging from 10 (indicate no distress) to 50 (severe distress).

2.3. MRI Acquisition

A 3 Tesla Siemens Skyra (Erlangen, Germany) scanner at the Thompson Institute, USC, collected the MRI data with a 64-channel head and neck coil. The T1-weighted structural MRI data were collected using the magnetization prepared rapid acquisition gradient-echo sequence. The

specific parameters include the time of repetition of 2200ms, echo time of 1.77ms, time of inversion of 850ms, flip angle of 7 degrees, the field of view of 230 mm, and 0.9 mm3 isotropic voxels. The resting-state fMRI data were acquired using an echo-planar imaging multiband sequence. The specific parameters include time of repetition of 1600ms, echo time of 30ms, acquisition matrix of 74 by 74, 3.027 × 3.027 × 3.0 mm voxels with no gap, flip angle of 65 degrees, and multiband factor of 4. Three hundred volumes were collected with closed eyes in 8 min.

2.4. Image Pre-Processing

The MRI data were pre-processed using FMRIB's Software Library (FSL 6.0.4), SPM12 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging, London, United Kingdom), and CONN functional connectivity toolbox (CONN, version 17). The fMRI pre-processing includes (1) slice timing and motion correction using the FSL, (2) normalisation to the Montreal Neurological Institute (MNI) space using SPM12, and (3) noise removing using CONN.

The fMRI data were firstly slice-time corrected (FSL-slice timer), motion-corrected (FSL-mcflirt), skull-stripped (FSL-bet). Any fMRI datasets with head displacement $>$ 3 mm were excluded [4], and fMRI scans with frame-to-frame displacements $>$ 0.40mm were censored [5] with any dataset with $<$ 85% of volumes remaining were excluded [4]. A total of six datasets were excluded due to motion artifacts.

Following motion correction, The T1- weighted data were co-registered to their corresponding fMRI scans, corrected for field bias (N4BiasField Correction), and segmented into grey and white matter and cerebrospinal fluid (SPM-unified segmentation). The SPM-DARTEL pipeline was used to generate a study-specific template. It generates the study-specific template and individual warp files that can be used to normalise the structural and fMRI data into a common space. The generated template was registered to the MNI space, and then all the individual fMRI data were spatially normalised to the MNI space (2 mm^3 voxels). Following normalization, the fMRI data were smoothed using a 6mm Gaussian kernel of full width at half maximum.

The normalised fMRI data were further cleaned using the CONN functional connectivity toolbox denoising pipeline [6]. The ART-based motion scrubbing and outlier volume removal was performed, and the motion ART artifacts were regressed out. In addition, the physiological fluctuations were removed using the CompCor algorithm [7] to regress out white matter and cerebrospinal fluid signals. This is achieved by using principal component analysis (PCA) of the multivariate BOLD signal within the white matter and cerebrospinal fluid masks to estimate the first five orthogonal time series to regress out the effect of the physiological fluctuations. However, the global signal was not regressed out due to its controversy as a pre-processing step [8]. Finally, the data were temporally filtered with band-pass ($f = 0.007$–0.1 Hz).

2.5. Functional connectivity and connectome similarities and uniqueness analysis

All functional connectivity, connectome similarities and uniqueness analyses were performed using in-house MATLAB (MATLAB R2019a, The MathWorks Inc., Natick, MA) scripts and using the 300 regions of interest (ROI) from Greene lab atlas [1] representing 14 different functional networks. The Greene lab atlas consists of 264 cortical ROIs [2] and 36 ROIs located at subcortical and cerebellar structures. These ROIs were spheric in shape (radius of 4-5 mm) defined using a winner-take-all partitioning method to rsfMRI data considering anatomy [1]. Each ROI was assigned to 14 networks using semi-automated consensus network procedures. An information-theoretic community detection algorithm generated the InfoMap [9], then weighted using the highest weighted community, examined by experts, and cross-validated using an independent dataset.

The BOLD time course was extracted from each ROI to estimate the functional connectome. The connectivity analysis was performed by estimating the Pearson correlation coefficients be-

tween each pair of ROIs' time courses and normalised to Z scores using the Fisher transformation representing the connectome of each fMRI dataset. The network-based connectome was estimated as the connectivity matrices of any pair of ROIs within the network, while the whole-brain connectome represents the connectivity matrix of all 300 ROIs.

The Pearson correlation between edge values from two connectomes was used to estimate two connectome similarity indices. These indices include the self-similarity index representing the similarity between connectomes of the same participant at different time points and the similarity-to-others index representing the similarity between a given individual's connectome to the connectome of any other individuals at a different time point.

The connectome uniqueness is defined as the ratio of (i) self-similarity to the maximal similarities-to-others (uniqueness_max), (ii) self-similarity to the averaged similarities-to-others (uniqueness_avg), and (iii) self-similarity to the second maximal similarities-to-others. The uniqueness_avg values lie beyond 3 times of the standard deviation from the mean values were considered as outliers and removed from further analysis.

Ethics Statements

The current study was approved by the Human Research Ethics Committee at The University of the Sunshine Coast (USC) (IRB: A181064), and written informed consents were obtained from all participants and their guardian/s. The research has been carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Dataset of brain functional connectome and its maturation in adolescents (Reference data) (Mendeley Data).

CRediT Author Statement

Zack Y. Shan: Conceptualization, Methodology, Formal analysis, Writing – original draft, Visualization; **Abdalla Z. Mohamed:** Methodology, Data curation, Writing – review & editing; **Paul Schwenn:** Validation, Resources, Writing – review & editing, Visualization; **Larisa T. McLoughlin:** Data curation, Writing – review & editing; **Amanda Boyes:** Data curation, Writing – review & editing; **Dashiell D. Sacks:** Data curation, Writing – review & editing; **Christina Driver:** Data curation, Writing – review & editing; **Vince D. Calhoun:** Methodology, Writing – review & editing; **Jim Lagopoulos:** Resources, Writing – review & editing, Supervision, Funding acquisition; **Daniel F. Hermens:** Investigation, Resources, Writing – review & editing, Project administration, Funding acquisition.

Acknowledgments

We thank the Longitudinal Adolescent Brain Study (LABS) participants and their caregivers. This work was supported by a grant from the Prioritising Mental Health Initiative, Australian Commonwealth Government.

References

[1] B.A. Seitzman, C. Gratton, S. Marek, R.V. Raut, N.U.F. Dosenbach, B.L. Schlaggar, S.E. Petersen, D.J. Greene, A set of functionally-defined brain regions with improved representation of the subcortex and cerebellum, Neuroimage 206 (2020) 116290, doi:10.1016/j.neuroimage.2019.116290.

[2] J.D. Power, A.L. Cohen, S.M. Nelson, G.S. Wig, K.A. Barnes, J.A. Church, A.C. Vogel, T.O. Laumann, F.M. Miezin, B.L. Schlaggar, S.E. Petersen, Functional network organization of the human brain, Neuron 72 (2011) 665–678, doi:10.1016/j.neuron.2011.09.006.

[3] R.C. Kessler, G. Andrews, L.J. Colpe, E. Hiripi, D.K. Mroczek, S.L. Normand, E.E. Walters, A.M. Zaslavsky, Short screening scales to monitor population prevalences and trends in non-specific psychological distress, Psychol. Med. 32 (2002) 959–976, doi:10.1017/s0033291702006074.

[4] L.J. Hearne, L. Cocchi, A. Zalesky, J.B. Mattingley, Reconfiguration of brain network architectures between resting-state and complexity-dependent cognitive reasoning, J. Neurosci. 37 (2017) 8399–8411, doi:10.1523/jneurosci.0485-17.2017.

[5] J.D. Power, A. Mitra, T.O. Laumann, A.Z. Snyder, B.L. Schlaggar, S.E. Petersen, Methods to detect, characterize, and remove motion artifact in resting state fMRI, Neuroimage 84 (2014) 320–341, doi:10.1016/j.neuroimage.2013.08.048.

[6] S. Whitfield-Gabrieli, A. Nieto-Castanon, Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks, Brain Connect. 2 (2012) 125–141, doi:10.1089/brain.2012.0073.

[7] Y. Behzadi, K. Restom, J. Liau, T.T. Liu, A component based noise correction method (CompCor) for BOLD and perfusion based fMRI, Neuroimage 37 (2007) 90–101, doi:10.1016/j.neuroimage.2007.04.042.

[8] K. Murphy, M.D. Fox, Towards a consensus regarding global signal regression for resting state functional connectivity MRI, Neuroimage 154 (2017) 169–173.

[9] M. Rosvall, C.T. Bergstrom, Maps of random walks on complex networks reveal community structure, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 1118–1123.