Test-Retest Reliability of Dynamic Functional Connectivity Parameters for a Two-State Model

Xiaojing Fang1,2, Michael Marxen1

1Department of Psychiatry, Technische Universität Dresden, Dresden, Germany
2Institute and Polyclinic for Diagnostic and Interventional Neuroradiology, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

*Correspondence: Michael Marxen, Section of Systems Neuroscience, Technische Universität Dresden, Würzburger Straße 35, 01187 Dresden, Germany, Email: michael.marxen@tu-dresden.de

Running title: Reliability of Dynamic Functional Connectivity Parameters

Keywords: neuroimaging, fMRI, connectomics, brain states, reproducibility, replicability, sliding-window analysis, correlation

Figure count: 3 Table count: 4

Data availability statement: Will follow with submission for peer-review.

Funding statement and acknowledgements: This research was supported by the Deutsche Forschungsgemeinschaft (DFG) grants 178833530 (SFB 940) and 402170461 (TRR 265).

Conflict of interest disclosure: All authors declare no conflict of interest.

Ethics approval and patient consent statement: The study was approved by the Ethics Review Board of the Technische Universität Dresden (EK45022016). All subjects signed informed consent forms after receiving a detailed description of the experiment.
Abstract

Dynamic functional connectivity (DFC) metrics derived from resting-state functional magnetic resonance imaging (rs-fMRI) that measure dynamic features of brain communications that change over time are gaining more and more attention. A key concern is the reliability of DFC parameters, especially given the many methodological options. The DFC parameters mean dwell time (MDT), prevalence (Prev.), inter-transition interval (ITI) and state variability (Var.) were computed in 23 participants using a two-state model, sliding window analysis, and two imaging sessions. Reliability of the connectivity states, test-retest reliability of the parameters as well as correlations between DFC parameters were investigated for different scan lengths (i.e., 16 vs 8 min.), atlases (i.e., 116 vs. 442 regions of interest) and with/without within-subject centering of DFC matrices. The results showed an integrated (I) and a segregated (S) brain state with high intra-class correlation (ICC) values of the states between sessions (0.67 ≥ ICC ≥ 0.99). Reliability of the DFC parameters is, at best, intermediate with an ICC = 0.51 for prevalence for not-centered, 16 min. data and the coarse atlas. Moreover, MDTs of both states were positively correlated for centered data. Prevalence of one state was positively correlated with MDT of the same state and frequently negatively correlated with MDT of the other state. Var was positively correlated with MDT and Prev for the same state and negatively correlated for the other state for not-centered data. This study suggests that prevalence is the most reliable DFC parameter with lower reliability for shorter scans, more atlas regions and centering. Thus, it is advantageous to formulate hypotheses related to brain dynamics in terms of prevalence, especially in small-scale studies.

Key words: dynamic functional connectivity; reliability; resting-state fMRI; sliding window analysis
1. Introduction

The analysis of functional connectivity (FC) between brain regions, a measurement of inter-regional synchrony (Karl J Friston, 1994), has become one of the primary research methods in the field of neuroscience (K. J. Friston, 2011). While many studies implicitly assumed a stationary nature of FC, the dynamic nature of brain activity (Chang & Glover, 2010) is gaining more and more attention in recent years. Dynamic functional connectivity (DFC) describes dynamic features of brain communications that change over time (Hutchison et al., 2013). By now, a large number of works have investigated the dynamic fluctuations of the connections, graphic properties of the DFC matrices and all kinds of characteristics of DFC patterns at rest, illustrating temporal changes in regional interactions, signal transfer between networks and the functional architecture of whole-brain states (Hutchison et al., 2013; Preti, Bolton, & Van De Ville, 2017; Zhang, Baum, Adduru, Biswal, & Michael, 2018). Additionally, a lot of findings revealed not only neural correlates of resting-state DFC with electroencephalography powers (Chang, Liu, Chen, Liu, & Duyn, 2013; Tagliazucchi, von Wegner, Morzelewski, Brodbeck, & Laufs, 2012; Thompson et al., 2013), but also associations of the connectivity with demographics, consciousness, behaviour, cognition (Hutchison et al., 2013; Preti et al., 2017; Shine et al., 2016) and various neurological and psychiatric disorders, such as Parkinson’s disease (Zoller et al., 2019) and schizophrenia (Calhoun, Miller, Pearlson, & Adali, 2014; Hutchison et al., 2013). All the evidence supports that DFC, a reflection of changing patterns of neural activity, is a useful approach to further our understanding of the fundamental characteristics of the brain at rest.

So far, researchers have developed many methods to investigate temporal variations of connectivity, such as time-frequency analysis (Chang & Glover, 2010), time-series modelling (Eavani, Satterthwaite, Gur, Gur, & Davatzikos, 2013; Lindquist, Xu, Nebel, & Caffo, 2014), change-point detecting (Cribben, Wager, & Lindquist, 2013) and co-activation pattern analysis (Liu & Duyn, 2013). Possibly the most popular approach to DFC is sliding window analysis (SWA), which divides a time course into small windows and computes functional connectivity measures for each window (Chang & Glover, 2010). Due to its simplicity in algorithm and application, SWA has been widely used for diverse research interests (Allen et al., 2014; Choe et al., 2017; Hutchison et al., 2013; Preti et al., 2017). However, an inevitable problem of SWA-based DFC measures compared to static FC (SFC) is that reliability is bound to decrease with smaller window size and shorter scan time due to the low signal-to-noise ratio (SNR) of resting-state functional magnetic resonance imaging (rs-fMRI) signals, artefact signals from non-neural factors (e.g., respiration, cardiac pulsation and noise) (Hutchison et al., 2013) and fluctuations of the measures on the time scale of hours.
and days (e.g. circadian effects) (Orban, Kong, Li, Chee, & Yeo, 2021). As a consequence, sensitivity and specificity of DFC approaches have to be evaluated carefully (Calhoun et al., 2014; Handwerker, Roopchansingh, Gonzalez-Castillo, & Bandettini, 2012; Hlinka & Hadrava, 2015). Hence, an immediate concern is to quantify and maximize the test-retest reliability, measured frequently through intra-class correlation (ICC), of utilized DFC parameters, which will aid to optimize processing pipelines, to register concrete hypotheses, enhance the interpretability of findings and, generally speaking, move this field forward.

Surprisingly, only a few reports of test-retest reliability for brain states were published. By hierarchically clustering the DFC between four subdivisions of the posteromedial cortex and 156 target regions of interest (ROIs), Yang and colleagues observed five clusters in one rs-fMRI session and six in the other for the same 22 participants, revealing partial reproducibility for the states and derived parameters (i.e., duration and transitions) (Z. Yang, Craddock, Margulies, Yan, & Milham, 2014). Another study with 100 subjects from the Human Connectome Project Data Release (HCP) dataset (Van Essen et al., 2013) revealed four similar k-means co-activation clusters across sessions with intermediate reliabilities: ICC ≤ 0.54 for temporal faction (prevalence) or mean dwell time (Smith, Zhao, Keilholz, & Schumacher, 2018). Choe and colleagues (Choe et al., 2017) reported similar values for two k-means connectivity clusters (e.g. ICC = 0.56 for dwell time using SWA). Although there is currently no a priori “correct” number of brain states, when maximizing reliability is the prerogative, choosing k = 2 clusters is advantageous for three reasons: 1) reliability of the state centroid connectivity matrices will be better than for larger k because the states will be most distinct (i.e., furthest away from each other in the clustering space); 2) for parameters such as mean dwell time, noise reduction through averaging should increase with the total time spent in each state (i.e., [scan time]/k, which is largest for k = 2); 3) two states are in line with the the widely accepted view that there are two fundamental principles of brain organization (K. J. Friston, 2009; Shine et al., 2016), namely integration and segregation. This allows a straightforward functional interpretation of the two states.

Naturally, other factors will affect the reliability of DFC parameters such as scan length (Hindriks et al., 2016) and other details of the processing pipelines. Consequently, after validating the interpretation of the two states for our data based on the global graph metrics modularity and global efficiency, we addressed the following questions using a two-state model and a straightforward SWA approach: 1) How reliable are the two states between scanning sessions? 2) How reliable are particular parameters that quantify the dynamics (see section 2.6) for this model? 3) What is the effect of scan length on the reliability? 4) What is the effect of choosing an atlas with a larger number of regions? and 5) What is the effect of within-subject
centering of the connectivity matrices before clustering, which will remove differences between subjects in static connectivity?

2. Methods

These methods have been described in a similar form for different participants in the preregistration https://osf.io/286fb “Linking Need for Cognition to a Two-State Model of Dynamic Brain Connectivity during Rest”.

2.1. Participants

Twenty-five healthy young adults were recruited from a volunteer database. Inclusion criteria were: no magnetic-resonance contraindications, normal or correctable-to-normal vision and an age between 20-35 years old. For the analysis, one participant was excluded because of incomplete MRI data; one participant was excluded because of severe, prefrontal signal drop out despite field map correction. Therefore, we analyzed the data of 23 subjects [mean ± standard deviation (SD) of age: 24.0 ± 3.8 years, 16 females]. Participants received a financial compensation of €40 for full participation in both sessions. The study was approved by the Ethics Committee of the Technische Universität Dresden (EK 4012016) and all participants signed informed consent forms after receiving a detailed description of the experiment.

2.2. Study Design

To evaluate reliability, a repeated-measure design was used. Study participants were scanned twice with the same scanning protocol with approximately one week between scanning sessions. To eliminate potential session order effects, we generated two sets of data (i.e., session A and B) by randomly assigning the first acquisition session from 12 participants and the second acquisition session from the remaining 11 participants to session A, while the respectively paired scans were assigned to session B.

2.3. MRI data acquisition

Imaging data were collected on a Siemens 3 T Magnetom Trio Tim scanner (Siemens, Erlangen, Germany) equipped with a 32-channel head coil. In each session, structural T1-weighted data were acquired after brief localizer scans (~1 min), using a 3D magnetization-prepared rapid gradient echo (MPRAGE) sequence with the parameters of repetition time (TR) 2400 ms, echo time (TE) 2.19 ms, voxel size 0.85 mm × 0.85 mm × 0.85 mm, inversion time 1000 ms, field of view (FOV) 272 mm × 272 mm, flip angle 8°, matrix 320 × 320, slice thickness 1.0 mm, band width (BW) 210 Hz/Px and 240 slices. T2-weighted data were acquired in each session with the parameters of TR 3200 ms, TE 565 ms, voxel size 0.85 mm × 0.85 mm × 0.85 mm, FOV 272 mm × 272 mm, flip angle 120°, matrix 320 × 320, thickness 0.85 mm, BW 744 Hz/Px and 240
slices. We acquired a B0 field map with parameters of TR 832 ms, TE-1 5.19 ms, TE-2 7.65 ms, voxel size 2.0 mm × 2.0 mm × 2.0 mm, gap 0 mm, FOV 208 mm × 208 mm, flip angle 58°, matrix 104 × 104, BW 278 Hz/Px, and 78 slices. Then, we collected rs-fMRI data based on a multi-band axial (T > C ~ -17°) 2D EPI sequence (Moeller et al., 2010) and with the parameters of TR 987 ms, TE 32.6 ms, voxel size 2.0 mm × 2.0 mm × 2.0 mm, slice gap 0 mm, FOV 192 mm × 192 mm, flip angle 62°, matrix 96 × 96, BW 1860 Hz/Px, 72 interleaved slices, and 1000 volumes. The two rs-fMRI scans lasted approximately 16:37 min. each (mean ± SD of the interval between the two scans: 8.87 ± 4.84 days). Additionally, we acquired diffusion-weighted images not relevant to this study. All participants received foam padding for head-movement reduction and earplugs for hearing protection, and were instructed to close their eyes and to try not to fall asleep.

2.4. Preprocessing

The following preprocessing pipeline using fMRIPrep 1.2.5 (zenodo.org/record/4252786#.X7TzMGhKhPZ) based on Nipype 1.1.6 (zenodo.org/record/4035081#.X7Ty32hKhPY) (Esteban et al., 2019) was employed: skull stripping, field map correction, co-registration of the structural and functional images, head-motion estimation and correction, spatial normalization to MNI152NLin2009cAsym space. All subjects met our inclusion criteria of less than 7.5% of frames with framewise displacement (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012) ≥ 0.5 mm. Then, the nuisance covariates (i.e., six head motion parameters, signals of cerebral spinal fluid and white matter) were removed from the preprocessed data. The effects of low-frequency drift and high-frequency physiological noise were reduced by using band-pass filtering (1/(w*TR) ~ 0.1 Hz) in time (Leonardi & Van De Ville, 2015). Here, we specified w (i.e., the window width in TR for the sliding-window analysis) as 40 (Zhang et al., 2018), which resulted in a lower frequency boundary of 0.025 Hz.

2.5. Extraction of functional brain states

2.5.1. Sliding-window analysis

In the current study, an instance of a DFC brain state is defined as a connectivity matrix of Fisher-z values based on Pearson’s correlation that were computed for the time window of 40*TR = 39.48s based on a previous study (Zhang et al., 2018). The sliding-window analysis was computed in DynamicBC toolbox (Liao et al., 2014) (github.com/guorongwu/DynamicBC) with a step size of 1 TR resulting in 961 brain-state instances (i.e., DFC matrices) per scanning run. Time series were extracted by spatially averaging all voxel-wise signals within each ROI.
2.5.2. Clustering of state instances and back projection

The two-state (i.e., state I and S) classification of the DFC matrices was performed in Matlab 2018b (Mathworks Inc., Natick, MA, USA) by assigning each matrix to one of \( k = 2 \) clusters in k-means clustering (Lloyd, 1982) employing cosine distance \((CD)\) to measure the similarity between each state instance and the centroids of a cluster (i.e., \( CD = 1 - \cos(\alpha) \), where \( \alpha \) is the angle between two vectorized DFC matrices. Thus, \( CD \in [0, 2] \), with 0 being the smallest distance, i.e., the highest similarity.) (Karahanoglu & Van De Ville, 2015). Moreover, \( k = 2 \) was found to be optimal according to silhouette statistics (Rousseeuw, 1987) computed for cluster numbers \( k = 2 \ldots 20 \) (See Supplemental Fig. S1). For session B, DFC matrices were assigned to state I or S by back-projecting each matrix onto the cluster centroids from session A, i.e. brain state instances from session B were assigned to the state/cluster with the smaller \( CD \) to the corresponding cluster centroid of session A. Cluster centroids were defined here as the group mean of the normalized, i.e., divided by the root of the sum of squares of all matrix elements, DFC matrices for the respective state in each session. Distances between all state instances and the two cluster centroids were plotted to validate the clustering procedure.

2.5.3. Processing pipeline variations

We investigated \( 2 \times 2 \times 2 = 8 \) different processing pipelines or strategies. Three factors were varied:

1) The length of the rs-fMRI time series: 1000 frames (16:27 min.) versus 500 frames (8:14 min.). For the analysis with 500 frames, we only considered the first 500 frames of a run.

2) The brain atlas used to define the ROIs (see below): 116 ROIs versus 442 ROIs.

3) Within-subject centering of DFC matrices versus ‘not-centered’. For the centered data, the within-subject-and-session mean connectivity matrix was removed from the DFC matrices before the clustering, which eliminates differences in SFC between subjects. As the resulting cluster centroids no longer reflect functional connectivity strength in this case, we added session-level mean connectivity matrices back to the centroid when appropriate.

2.5.4. Employed brain atlases

The Automated Anatomical Labeling atlas (AAL) (Tzourio-Mazoyer et al., 2002) and a joint atlas (Schaefer et al., 2018; Shine et al., 2016) were applied in this study to explore the reliability of dynamic properties of brain states on different spatial scales. The former consists of 116 ROIs and was grouped into nine networks / regions based on Yeo’s seven functional networks of the cerebral cortex (Yeo et al., 2011) and anatomical parcellations of the subcortical structures and the cerebellum: the visual network, the sensory-motor
network, the dorsal attention network, the ventral attention network, the limbic network, the fronto-parietal network, the default mode network, the basal ganglia network, and the cerebellar network. The joint atlas consisted of 442 ROIs: 400 cortical parcels of the Schaefer atlas (Schaefer et al., 2018), 14 subcortical regions from the Harvard-Oxford subcortical atlas, i.e. the bilateral thalamus, caudate, putamen, ventral striatum, globus pallidus, amygdala, and hippocampus, and 28 cerebellar regions from the SUIT atlas (Diedrichsen, Balsters, Flavell, Cussans, & Ramnani, 2009). This joint atlas was grouped into the same nine networks used for the AAL atlas (see Supplementary Tables S1-S3).

2.6. DFC parameters

We computed seven different dynamical parameters of the brain based on the state sequence of each subject in each session (i.e., all the dynamic parameters were specific to each of the two states and were indexed by I or S, except ITI):

a) Mean dwell time (MDT): the mean time one participant maintained state I or S continuously during a run. Note that we excluded all windows within the very first and last states from the computation of MDT since the beginning/end of the state is unknown. Consequently, only individual DFC series with at least two state transitions were used to calculate MDT for at least one of the states.

b) Prevalence (Prev.): the fraction of windows spent in state I or S concerning with respect to the total number of windows in the DFC series. Note: \( \text{Prev}_I = 1 - \text{Prev}_S \).

c) Inter-transition interval (ITI): the average time length remaining in any state before transitioning to the other state during scans, i.e., \( \text{ITI} = (\text{MDT}_I + \text{MDT}_S)/2 \) for an even number of states in a series. Note that only DFC series including at least three state transitions were considered for the computation.

d) State variability (Var): the mean Euclidean distance of each state instance of one subject to the run-average of all state instances within the same state:

\[
\text{Var}_{I,S} = \frac{1}{T_{I,S}} \sum_{t=1}^{T_{I,S}} D_{\text{Euclid}} \left( C_t, \frac{1}{T_{I,S}} \sum_{t'=1}^{T_{I,S}} C_{t'} \right),
\]

where \( C_t \) is the DFC matrix at time window \( t \) and \( D_{\text{Euclid}}(C_a, C_b) \) the Euclidean distance between the vectorized DFC matrices \( C_a \) and \( C_b \). Note that this definition does not use cosine distance or the above definition of the cluster centroid because such metrics are not invariant to translations in the clustering space.

Following the above definitions, the seven parameters are not completely independent. Nevertheless, each of them represents a different aspect of brain state dynamics and may maximize associations with other parameters. The number of subjects used to compute each parameter are given in supplementary table S5.
2.7 Data analyses

2.7.1 Reliability of cluster centroids

To evaluate the reliability of cluster centroids between sessions, we computed the intra-class correlation coefficient (ICC) (Bartko, 1966), a popular measurement that quantifies test-retest reliability and was previously used in other rs-fMRI studies (Choe et al., 2017; Zhang et al., 2018; Zuo & Xing, 2014):

\[ ICC = \frac{\sigma^2_s}{\sigma^2_s + \sigma^2_e} = \frac{\sigma^2_{total} - \sigma^2_e}{\sigma^2_{total}} = 1 - \frac{\sigma^2_e}{\sigma^2_{total}}, \]

where, \(\sigma^2_s\) indicates the between-subject variance, \(\sigma^2_e\) the within-subject variance and \(\sigma^2_{total} = \sigma^2_s + \sigma^2_e\) the total variance observed in the data.

The ICC used in this study was based on a one-way ANOVA intra-class correlation framework (Bartko, 1966). Note that ICC values may become less than zero in this framework if variability within the group is greater than across groups (i.e., “divergence” within a group). In addition, we calculated Spearman’s correlation, which is robust to outliers and non-normal distributions (de Winter, Gosling, & Potter, 2016) and would quantify reliability of within-session ranks.

2.7.2 Graph-theoretical analyses of cluster centroids

To describe the functional organization characteristics of the two brain states (i.e. cluster centroids), we computed the two global graph parameters modularity and global efficiency (Cohen & D’Esposito, 2016; Rubinov & Sporns, 2010) for the 16-min data in brain connectivity toolbox (BCT, Version 2019-03-03, nitrc.org/frs/?group_id=241):

a) Modularity \(Q = \sum_{i=1}^{m} (e_{ii} - (\sum_{j=1}^{m} e_{ij})^2)\), where \(e_{ij}\) means the proportion of edges that link a node in module \(i\) to any node in module \(j\), and \(m\) the number of all non-overlapping modules generated by Newman’s modularity algorithm (Newman & Girvan, 2004).

b) Global efficiency \(E_{global} = \frac{1}{n(n-1)} \sum_{i \neq j} d_{ij}^{-1}\), where \(n\) is the number of nodes, and \(d_{ij}\) the minimum distance between node \(i\) and \(j\). In this study, we only computed the distance based on connectivity between the nodes with positive strength (Barch et al., 2013).

We evaluated the variance within these graph measures by generating 1000 bootstrapping samples of the DFC matrices for each state/cluster of session A.
2.7.3 Reliability of dynamic parameters

We computed values of ICC and Spearman’s correlation to quantify the reliability of the dynamic parameters between the two sessions. Confidence intervals of the reliability were estimated by bootstrapping with 1000 samples.

2.7.4 Correlations between DFC parameters

To quantify the interdependence between the dynamic parameters, we calculated Spearman’s correlations between the averaged parameter values across the two sessions.
3. Results

3.1. Reliability of cluster centroids

Results of the clustering procedure (session A) and back-projection (session B) for all processing strategies are shown in Fig. 1 for the coarse atlas and Fig. S2 for the fine atlas, consistently demonstrating that the DFC matrices cluster around an overall strongly connected (i.e., integrated) state I and a weakly connected (i.e., segregated) state S (see also Section 3.2). For the centered data (Fig. 1), the centroid of state I shows correlations above the run-mean while state S shows correlations below the run-mean (see also Fig. S5). Quantitative reliability between the clusters from session A and B in terms of ICC and Spearman’s correlation are given in Table 1 and Table S6, respectively. Note that, for the not-centered data, reliabilities were also computed after removing the group-mean matrix (Table 1c and Table S6c) to eliminate the bias that not-centered centroids generally exhibit high positive ICC values even between different states (Table 1a and Table S6a). Overall, the reliabilities between corresponding states were generally high (ICC $\geq 0.67$), especially for longer scan time (ICC $\geq 0.86$), while not-corresponding states showed negative ICC values in the centered data. Cosine distances between corresponding states were generally very low (CD $\leq 0.08$) and very high for not-corresponding states (CD $\geq 1.96$) in the centered data (see Table S6). State reliability generally decreased for shorter scan lengths, more ROIs and centering.
Fig. 1. Cluster centroids (k = 2) of the DFC matrices based on the AAL atlas for different time-length and not-centered/centered data. For session A, k-means clustering was used for the state assignment, while back projection results are shown for session B. Displayed are the cluster centroids (see Methods; N1: visual network; N2: sensory-motor network; N3: dorsal attention network; N4: ventral attention network; N5: limbic network; N6: fronto-parietal network; N7: default mode network; N8: basal ganglia network; N9: cerebellum network).
a) 16 min data (116 ROIs)

1) Not-centered

Clustering results based on session A

State S

State I

Back-projection results in session B

State S

State I

2) Centered

Clustering results based on session A

State S

State I

Back-projection results in session B

State S

State I

Rescaled z values

Rescaled differential z values
b) 8 min data (116 ROIs)

1) Not-centered

Clustering results based on session A
State S

State I

Back-projection results in session B
State S

State I

2) Centered

Clustering results based on session A
State S

State I

Back-projection results in session B
State S

State I

Rescaled z values

Rescaled differential z values
Table 1. Reliability (measured by ICC) between the two clusters for the two sessions

|                | 116-ROI atlas | 442-ROI atlas |
|----------------|--------------|--------------|
| **a) Clusters based on not-centered data** |             |              |
| 16-min data    | 8-min data   | 16-min data  | 8-min data   |
| Session B      | Session B    | Session B    | Session B    |
| State S        | State I      | State S      | State I      |
| State S        | ICC = 0.97   | ICC = 0.96   | ICC = 0.96   |
| State I        | ICC = 0.67   | ICC = 0.66   | ICC = 0.55   |
| **b) Clusters based on centered data** |             |              |
| 16-min data    | 8-min data   | 16-min data  | 8-min data   |
| Session B      | Session B    | Session B    | Session B    |
| State S        | State I      | State S      | State I      |
| State S        | ICC = 0.93   | ICC = 0.86   | ICC = 0.86   |
| State I        | ICC = -0.13  | ICC = -0.12  | ICC = -0.06  |
| **c) Clusters based on not-centered data & removing session mean matrices** |             |              |
| 16 min data    | 8 min data   | 16 min data  | 8 min data   |
| Session B      | Session B    | Session B    | Session B    |
| State S        | State I      | State S      | State I      |
| State S        | ICC = 0.93   | ICC = 0.86   | ICC = 0.89   |
| State I        | ICC = -0.47  | ICC = -0.32  | ICC = -0.22  |
3.2. *Graph-theoretical Analyses*

The two global graph parameters modularity and efficiency are shown in Fig. 5 for the 16-min data including results of the bootstrapping procedure. The results demonstrate that the graphs of the two states were distinct with respect to these two parameters: the state with low connectivity, i.e., the “segregated” state (S), had low global efficiency and high modularity, while the state with strong connectivity, i.e., the “integrated” (I) state, had high global efficiency and low modularity. Bootstrapping for session A showed that these clusters did not overlap within this two-parameter space. Additionally, the modularity and global efficiency values of the cluster centroids in session B were in the range of the bootstrapping results for session A (Fig. 5; Table S4). Note that individual centering reduced modularity and increased global efficiency for both states (Table S4). Additionally, the two clouds moved closer to each other for the centered data (Fig. 5).
Fig. 2. Global graph-theoretical measures of cluster centroids for 16-min data. Point clouds were generated by bootstrapping DFC matrices from session A and recalculation of the cluster centroid (Btsp. - bootstrapping; Ses. – session). For centered data, the session mean was added back. Note that session B centroids were computed as mean values of the respective state after back projection (see Methods).
3.3. **Reliability of dynamic parameters**

ICCs, $p$-values and confidence intervals based on bootstrapping for all parameters and pipelines are given in Table 2. Spearman’s correlation are given in Table S7. The highest ICC value of 0.55 was obtained for $Var_I$ for the 8-min data, however, it appears to be driven by outliers as the Spearman correlation is far from significant and the confidence interval is very wide. The next highest ICC value of 0.51 was obtained for prevalence for the 16 min. data with the coarse atlas and without centering (Fig. 3). It is the only parameter that shows consistent values between pipelines at an intermediate level ($0.22 \leq ICC \leq 0.51$) with a number of significant values (uncorrected).
Fig. 3. Scatter plots of prevalence between sessions for different pipelines to illustrate reliability. Solid line: $p$ value < 0.05; dotted line: $p$ value > 0.05.
Table 2. ICC values of the dynamic parameters for all pipelines. CI: confidence interval based on bootstrapping (1000 times). Bold: significant reliability (uncorrected).

| Mean dwell time | 16-min data | 8-min data | 16-min data | 442 ROIs | 8-min data |
|-----------------|-------------|------------|-------------|----------|------------|
| State I         | Not-centered | ICC = 0.05 | ICC = -0.08 | ICC = 0.17 | ICC = -0.34 | ICC = 0.31 |
|                 | Centered     | ICC = 0.23 | ICC = -0.17 | ICC = -0.31 | ICC = 0.21 | ICC = 0.31 |
|                 | p = 0.4042   | p = 0.1506 | p = 0.2187  | p = 0.9192 | p = 2.148   | p = 0.0802 |
| CI: [-0.27, 0.13] | CI: [-0.36, 0.58] | CI: [-0.11, 0.41] | CI: [-0.59, -0.06] | CI: [-0.54, 0.11] | CI: [-0.45, 0.35] | CI: [-0.09, 0.56] |
| State S         | ICC = 0.09   | ICC = 0.09 | ICC = 0.08  | ICC = 0.13 | ICC = -0.02 | ICC = 0.44 |
| Prevalece       | CI: [-0.34, 0.59] | CI: [-0.40, 0.47] | CI: [-0.30, 0.53] | CI: [-0.43, 0.16] | CI: [-0.08, 0.38] | CI: [-0.24, 0.33] |
|                 | ICC = 0.51   | ICC = 0.35 | ICC = 0.33  | ICC = 0.42 | ICC = 0.30 | ICC = 0.27 |
|                 | p = 0.0036   | p = 0.0382 | p = 0.0149  | p = 0.0696 | p = 0.0767 | p = 0.4901 |
| CI: [0.18, 0.76] | CI: [-0.30, 0.59] | CI: [-0.01, 0.62] | CI: [-0.22, 0.52] | CI: [0.04, 0.72] | CI: [-0.07, 0.49] | CI: [-0.10, 0.60] |
| Int. transition interval | ICC = -0.04 | ICC = 0.25 | ICC = 0.27  | ICC = -0.05 | ICC = -0.25 | ICC = 0.20 |
| State S         | p = 0.5645   | p = 0.1156 | p = 0.1058  | p = 0.5827 | p = 0.8646 | p = 2.189 |
| CI: [-0.41, 0.17] | CI: [-0.20, 0.61] | CI: [-0.17, 0.56] | CI: [-0.55, 0.03] | CI: [-0.43, 0.07] | CI: [-0.18, 0.43] | CI: [-0.19, 0.62] |
| State I         | ICC = 0.20   | ICC = 0.21 | ICC = 0.00  | ICC = 0.26 | ICC = 0.11 | ICC = 0.17 |
| State variability | p = 0.1762   | p = 0.1606 | p = 0.5030  | p = 0.4948 | p = 0.1109 | p = 0.3369 |
| State S         | CI: [-0.04, 0.60] | CI: [-0.21, 0.58] | CI: [-0.11, 0.08] | CI: [-0.06, 0.49] | CI: [-0.11, 0.49] | CI: [-0.09, 0.42] |
| CI: [-0.23, 0.45] | ICC = 0.35 | ICC = -0.19 | ICC = 0.29  | ICC = 0.22 | ICC = 0.55 | ICC = 0.19 |
| CI: [0.00, 0.52] | p = 0.0540   | p = 0.1306 | p = 0.0451  | p = 0.0879 | p = 0.1543 | p = 0.0034 |
| State I         | CI: [0.12, 0.51] | CI: [-0.09, 0.44] | CI: [-0.08, 0.57] | CI: [-0.60, 0.23] | CI: [0.00, 0.52] | CI: [-0.14, 0.44] |
| CI: [0.05, 0.84] | p = 0.1581   | p = 0.1581 |

CI: confidence interval based on bootstrapping (1000 times). Bold: significant reliability (uncorrected).
3.4. Correlations between DFC parameters

Spearman’s correlations between different dynamic parameters are given in Tables 3 and 4 for the coarse atlas and tables S8 and S9 for the fine atlas. With some consistency over all pipelines (but not always significant), MDTs of both states were positively correlated for centered data. Prevalence of one state was positively correlated with MDT of the same state and frequently negatively correlated with MDT of the other state. Var was positively correlated with MDT and Prev of the same state and negatively correlated with MDT and Prev of the other state for not-centered data. As to be expected, ITI was positively associated with the MDTs, especially with MDTs and in the 16-min data.
**Table 3.** Spearman’s correlations between DFC parameters (mean of both sessions) for not-centered data and 116 ROIs (AAL atlas). Bold: significant correlation (uncorrected).

| 16-min data | Mean dwell time | Prevalence S | Inter-transition interval | State variability |
|-------------|----------------|--------------|---------------------------|-------------------|
|              | State S | State I | State S | State S | State S | State I |
| Mean dwell time  | State S | State I | State S | State S | State S | State I |
| State I | -       | -       | -       | -       | -       | -       |
| State S | $\rho = -0.25$ | -       | -       | -       | -       | -       |
| $p = 0.3059$ | -       | -       | -       | -       | -       | -       |
| Prevalence S | State S | State I | State S | State S | State S | State I |
| $\rho = 0.71$ | $\rho = -0.74$ | -       | -       | -       | -       | -       |
| $p = 0.0099$ | $p = 0.0005$ | -       | -       | -       | -       | -       |
| Inter-transition interval | State S | State I | State S | State S | State S | State I |
| $\rho = 0.63$ | $\rho = 0.47$ | $\rho = 0.09$ | -       | -       | -       | -       |
| $p = 0.0044$ | $p = 0.0448$ | $p = 0.7209$ | -       | -       | -       | -       |
| State variability | State S | State I | State S | State S | State S | State I |
| $\rho = -0.55$ | $\rho = 0.58$ | $\rho = -0.74$ | $\rho = -0.11$ | $\rho = -0.12$ | -       | -       |
| $p = 0.0156$ | $p = 0.0110$ | $p = 0.0005$ | $p = 0.6517$ | $p = 0.6361$ | -       | -       |

| 8-min data | Mean dwell time | Prevalence S | Inter-transition interval | State variability |
|-------------|----------------|--------------|---------------------------|-------------------|
|              | State S | State I | State S | State S | State S | State I |
| Mean dwell time  | State S | State I | State S | State S | State S | State I |
| State I | -       | -       | -       | -       | -       | -       |
| State S | $\rho = -0.3$ | -       | -       | -       | -       | -       |
| $p = 0.235$ | -       | -       | -       | -       | -       | -       |
| Prevalence S | State S | State I | State S | State S | State S | State I |
| $\rho = 0.44$ | $\rho = -0.69$ | -       | -       | -       | -       | -       |
| $p = 0.0762$ | $p = 0.003$ | -       | -       | -       | -       | -       |
| Inter-transition interval | State S | State I | State S | State S | State S | State I |
| $\rho = 0.78$ | $\rho = 0.25$ | $\rho = 0.14$ | -       | -       | -       | -       |
| $p = 0.0004$ | $p = 0.3367$ | $p = 0.5854$ | -       | -       | -       | -       |
| State variability | State S | State I | State S | State S | State S | State I |
| $\rho = 0.15$ | $\rho = -0.36$ | $\rho = 0.72$ | $\rho = -0.45$ | $\rho = -0.52$ | -       | -       |
| $p = 0.5659$ | $p = 0.1588$ | $p = 0.0017$ | $p = 0.7297$ | -       | -       | -       |
| State I | $\rho = -0.62$ | $\rho = 0.51$ | $\rho = -0.72$ | $\rho = -0.52$ | -       | -       |
| $p = 0.0094$ | $p = 0.0366$ | $p = 0.0016$ | $p = 0.0346$ | -       | -       | -       |
Table 4. Spearman’s correlations between DFC parameters (mean of both sessions) for centered data and 116 ROIs (AAL atlas). Bold: significant correlation (uncorrected).

|                      | Mean dwell time | Prevalence | Inter-transition interval | State variability |
|----------------------|----------------|------------|----------------------------|-------------------|
| 16-min data          |                |            |                            |                   |
| State S              |                |            |                            |                   |
| State I              |                |            |                            |                   |
| rho = 0.30           |                |            |                            |                   |
| p = 0.1649           |                |            |                            |                   |
| rho = 0.56           |                |            |                            |                   |
| p = 0.0063           |                |            |                            |                   |
| rho = 0.80           |                |            | rho = 0.37                 |                   |
| p = 0.0000           |                |            | p = 0.0001                 |                   |
| rho = -0.06          |                |            | rho = 0.21                 | rho = 0.22        |
| p = 0.7881           |                |            | p = 0.3452                 | p = 0.3064        |
| rho = 0.00           |                |            | rho = 0.20                 | rho = 0.10        |
| p = 0.9874           |                |            | p = 0.3571                 | p = 0.6365        |
|                      |                |            |                            | p = 0.0025        |
| 8-min data           |                |            |                            |                   |
| State S              |                |            |                            |                   |
| State I              |                |            |                            |                   |
| rho = 0.58           |                |            |                            |                   |
| p = 0.0040           |                |            |                            |                   |
| rho = 0.67           |                |            | rho = 0.20                 |                   |
| p = 0.0005           |                |            | p = 0.3561                 |                   |
| rho = 0.94           | rho = 0.80     | rho = 0.59 |                            |                   |
| p = 0.0000           | p = 0.0000     | p = 0.0029 |                            |                   |
| rho = 0.09           | rho = 0.14     | rho = 0.15 | rho = 0.11                 |                   |
| p = 0.6767           | p = 0.5350     | p = 0.4823 | p = 0.6057                 |                   |
| rho = -0.01          | rho = -0.22    | rho = 0.17 | rho = -0.09                | rho = 0.29        |
| p = 0.9714           | p = 0.3035     | p = 0.4312 | p = 0.6965                 | p = 0.1756        |
4. **Discussion**

To evaluate the potential utility of DFC parameters as personality traits or biomarkers for further correlation analyses, knowing and maximizing their reliability can be crucial. In the introduction, we made an argument for a two-state model as a good starting point. Our graph-based analysis confirmed our expectation that clustering results in an integrated brain state with low modularity and high global efficiency and a segregated brain state with opposite properties. Subsequently, we investigated the reliability of the resulting state centroids and seven parameters using eight different pipelines. We used rs-fMRI data of 23 subjects from two sessions using an in-house fMRI sequence protocol with multi-band factor 6, an isotropic resolution of 2 mm and a TR = 0.987s. This sequence was also used in other studies of our center, which is one reason for us to evaluate reliability of parameters derived from this particular sequence. Because of the relatively small sample size, some of our observations need to be regarded with care, as the differences between pipelines, for example, are not statistically significant.

**Reliability of cluster centroids**

Consistent with Choe and colleagues’ study (Choe et al., 2017), we found high reliability of the paired clusters between sessions for both states based on all the strategies. However, shorter scan times and a finer atlas both reduced the ICC reliability from 0.93 to 0.67 (Table 1) for the centered data. This effect can also be observed when looking at the similarity of the clusters in terms of cosine distance (Table S6). For the not-centered data after removing the session mean, the effect was not as severe dropping ICC from 0.93 to 0.79. These effect needs confirmation in a larger study. As to be expected, removing the session mean from the not-centered data increased cosine distances of the unpaired clusters, because the vectors are moved towards the center of the clustering space.

Furthermore, the structure of the two states based on the not-centered data is visually consistent with dynamic two-state patterns in the whole brain at rest for not only healthy people (Choe et al., 2017) with eyes opened/closed (Weng et al., 2020) and different ages (young as 20-35 years old and older as 55-80 years old)(X. Yang et al., 2022), but also clinic patients of Parkinson’s disease (Fiorenzato et al., 2019; Kim et al., 2017), major depressive disorder (Zheng et al., 2022) and dementia with Lewy bodies (Ma et al., 2019). This consistency in state structure between studies is encouraging because, if the structure had varied, statements about brain organization would have
lacked a common ground for comparison. However, this comparability may be severely reduced for short scan lengths and fine atlases. Also, note that, for the not-centered data without removing the session mean, ICC values overestimate reliability in the sense that ICC values are even high between the two different states (ICC ≤ 0.72), which indicates that the two states differ primarily in strength, not the direction of the correlations. Figures S3 and S6 show the effect of centering on the distances to each centroid. Centering increases the dissimilarity between the unpaired states. It also shifts the prevalence of either state towards 0.5 (see Fig. 3) and, thus, reduces the number of subjects that remain in the same state for the whole scanning run, i.e., never change states.

**Reliability of dynamic parameters**

*Prev* showed the highest reliability across analyses (i.e., 4/8 pipelines), which is consistent with existing findings (Abrol et al., 2017; Choe et al., 2017; Smith et al., 2018) and our theoretical consideration that it is computed based on the complete scan time, while, for example, *MDT* is derived in average from only half the scan time. Thus, the *Prev* parameter is recommended for the formulation of hypotheses related to correlations between dynamic measures of functional brain integration and other subject characteristics.

Reliability of the DFC parameters is, at best, intermediate with an ICC = 0.51 for prevalence for not-centered, 16 min. data and the coarse atlas. Shorter scans, more atlas regions, and centering reduced this value. The reliability of the *MDTs* was less than of prevalence with a maximum of ICC = 0.44 for the centered, 8 min. data with the fine atlas, which was expected for the above reason. The finding that centering increased the reliability of *MDT* for all pipelines indicates that between-session variations in SFC have a negative impact on the reliability of *MDTs*, which seems to be reversed for prevalence where between-subject variations in static connectivity (Zurich, 2017) appear to help with re-identification of a subject in the second session.

Several reports suggested a higher reproducibility (Abrol et al., 2017) and higher reliabilities (ICC = 0.56 (Choe et al., 2017) and ICC ≤ 0.47 (Smith et al., 2018)) of the *MDTs* than observed by us. This inconsistency suggests a high sensitivity of *MDT* to differences in methodological (e.g., direct clustering versus two-level clustering and ROIs selection based on independent component analysis versus parcellation atlas) (Abrol et al., 2017; Choe et al., 2017; Smith et al., 2018). In addition to the MRI sequence parameters, both the number of regions as well as the ICA-based noise filtering may have an impact here. The reliabilities of the *ITI* and *Var* parameters were low
and not significant for many pipelines and not systematically related to the pipelines. More power would be needed to make further claims about these parameters.

Several considerations are relevant with respect to centering. To obtain the highest reliability for prevalence, centering is not recommended. It will eliminate a large portion of the between-subject variance, which is a disadvantage for cross-sectional analyses. It will also affect the interpretation of the brain state and may make the labels “integrated” versus “segregated” less meaningful, i.e., a subject that displays more integration (high connectivity) over all frames may lose this label when such between-subject differences in overall (static) connectivity are removed. Not to center, however, also has considerable disadvantages: 1) as seen, it reduces reliability of the MDTs; 2) without centering, static and dynamic effects are mixed, i.e. between-subject differences may be driven by differences in static connectivity rather than by the dynamics; and 3) centering could reduce noise if differences in static connectivity are fluctuating on the time scale of hours to days and do not effect the dynamics. Thus centering may be especially recommended for longitudinal studies over longer time periods.

Correlations between DFC parameters

MDTs of both states were positively correlated for centered data between the two states and with ITI, which indicates a tendency for subjects towards less state switching with longer MDTs rather than a compensatory or anti-correlated pattern that a longer duration of MDT would be associated with shorter MDT. The result revealed that prevalence of one state was positively correlated with MDT of the same state and frequently negatively correlated with MDT of the other state. This indicates that statements relating to brain integration may be either put in terms of prevalence or MDT, i.e. a more integrated brain would have both a higher prevalence and longer MDT. Var was positively correlated with MDT and Prev for the same state and negatively correlated for the other state for non-centered data. However, we need to carefully interpret these results given the low reliability of the Var parameter.

Limitations and future works

There are some limitations of the current work that should be addressed in the future: 1) as mentioned above, the primary limitation of the current study is the small sample size, which does not allow us to observe significant differences between pipelines and parameters. Thus, we are preparing to repeat the analysis in a larger cohort of Human Connectome Project
(www.humanconnectome.org) subjects. 2) We employed back-projection in the second session instead of re-clustering to avoid mixing effects of clustering due to noise on the group level and differences in connectivity. It is unclear whether re-clustering would introduce noise or remove it. To answer this question, one would need to repeat the analysis with re-clustering. 3) The two-state model may not be optimal to maximize reliability of a specific state or identify group differences in such specific states. Thus, models with more states and other dynamic methods, e.g. co-activation analyses (Karahanoglu & Van De Ville, 2015), could show superiority in detecting specific differences in different networks. 4) Our clustering approach does not allow for mixed states. Several studies suggested that a subject may simultaneously be in multiple overlapping states (Leonardi, Shirer, Greicius, & Van De Ville, 2014; Miller et al., 2016). 5) Many more processing pipelines are imaginable that may improve reliability, e.g. more atlases with different granularities, different noise filters, different MR-sequences, co-activation versus temporal correlation.

Conclusions

In summary, we found high reliability of clustering results with a two-state model but only medium to low reliability of the dynamic parameters. We recommend state prevalence as the most reliable parameter to formulate hypotheses. Shorter scan lengths and a finer atlas decrease reliability while centering reduces reliability of prevalence but improves reliability of MDTs. Larger samples are needed to substantiate our findings.
References

Abrol, A., Damaraju, E., Miller, R. L., Stephen, J. M., Claus, E. D., Mayer, A. R., & Calhoun, V. D. (2017). Replicability of time-varying connectivity patterns in large resting state fMRI samples. *Neuroimage, 163*, 160-176. doi:10.1016/j.neuroimage.2017.09.020

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb Cortex, 24*(3), 663-676. doi:10.1093/cercor/bhs352

Barch, D. M., Burgess, G. C., Harms, M. P., Petersen, S. E., Schlaggar, B. L., Corbetta, M., . . . Consortium, W. U.-M. H. (2013). Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage, 80*, 169-189. doi:10.1016/j.neuroimage.2013.05.033

Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological Reports.*

Calhoun, V. D., Miller, R., Pearlson, G., & Adali, T. (2014). The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron, 84*(2), 262-274. doi:10.1016/j.neuron.2014.10.015

Chang, C., & Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *Neuroimage, 50*(1), 81-98. doi:10.1016/j.neuroimage.2009.12.011

Chang, C., Liu, Z., Chen, M. C., Liu, X., & Duyn, J. H. (2013). EEG correlates of time-varying BOLD functional connectivity. *Neuroimage, 72*, 227-236. doi:10.1016/j.neuroimage.2013.01.049

Choe, A. S., Nebel, M. B., Barber, A. D., Cohen, J. R., Xu, Y., Pekar, J. J., . . . Lindquist, M. A. (2017). Comparing test-retest reliability of dynamic functional connectivity methods. *Neuroimage, 158*, 155-175. doi:10.1016/j.neuroimage.2017.07.005

Cohen, J. R., & D’Esposito, M. (2016). The Segregation and Integration of Distinct Brain Networks and Their Relationship to Cognition. *J Neurosci, 36*(48), 12083-12094. doi:10.1523/JNEUROSCI.2965-15.2016

Cribben, I., Wager, T. D., & Lindquist, M. A. (2013). Detecting functional connectivity change points for single-subject fMRI data. *Front Comput Neurosci, 7*, 143. doi:10.3389/fncom.2013.00143

de Winter, J. C., Gosling, S. D., & Potter, J. (2016). Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological methods, 21*(3), 273.

Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E., & Ramnani, N. (2009). A probabilistic MR atlas of the human cerebellum. *Neuroimage, 46*(1), 39-46. doi:10.1016/j.neuroimage.2009.01.045

Eavani, H., Satterthwaite, T. D., Gur, R. E., Gur, R. C., & Davatzikos, C. (2013). Unsupervised learning of functional network dynamics in resting state fMRI. *Inf Process Med Imaging, 23*, 426-437. doi:10.1007/978-3-642-38868-2_36

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., . . . Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat Methods, 16*(1), 111-116. doi:10.1038/s41592-018-0235-4

Fiorenzato, E., Strafella, A. P., Kim, J., Schifano, R., Weis, L., Antonini, A., & Biundo, R. (2019). Dynamic functional connectivity changes associated with dementia in Parkinson's disease. *Brain, 142*(9), 2860-2872. doi:10.1093/brain/awz192

Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: a synthesis. *Human brain mapping, 2*(1 - 2), 56-78.

Friston, K. J. (2009). Modalities, modes, and models in functional neuroimaging. *Science, 326*(5951), 399-403. doi:10.1126/science.1174521

Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain Connect, 1*(1), 13-36. doi:10.1089/brain.2011.0008
Handwerker, D. A., Roopchansingh, V., Gonzalez-Castillo, J., & Bandettini, P. A. (2012). Periodic changes in fMRI connectivity. *Neuroimage, 63*(3), 1712-1719. doi:10.1016/j.neuroimage.2012.06.078

Hindriks, R., Adhikari, M. H., Murayama, Y., Ganzetti, M., Mantini, D., Logothetis, N. K., & Deco, G. (2016). Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI? *Neuroimage, 127*, 242-256. doi:10.1016/j.neuroimage.2015.11.055

Hlinka, J., & Hadrava, M. (2015). On the danger of detecting network states in white noise. *Front Comput Neurosci, 9*, 11. doi:10.3389/fncom.2015.00011

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., . . . Chang, C. (2013). Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage, 80*, 360-378. doi:10.1016/j.neuroimage.2013.05.079

Karahanoglu, F. I., & Van De Ville, D. (2015). Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks. *Nat Commun, 6*, 7751. doi:10.1038/ncomms8751

Kim, J., Criaud, M., Cho, S. S., Diez-Cirarda, M., Mihaescu, A., Coakeley, S., . . . Strafella, A. P. (2017). Abnormal intrinsic brain functional network dynamics in Parkinson's disease. *Brain, 140*(11), 2955-2967. doi:10.1093/brain/awx233

Leonardi, N., Shirer, W. R., Greicius, M. D., & Van De Ville, D. (2014). Disentangling dynamic networks: Separated and joint expressions of functional connectivity patterns in time. *Hum Brain Mapp, 35*(12), 5984-5995. doi:10.1002/hbm.22599

Liao, W., Wu, G. R., Xu, Q., Ji, G. J., Zhang, Z., Zang, Y. F., & Lu, G. (2014). DynamicBC: a MATLAB toolbox for dynamic brain connectome analysis. *Brain Connect, 4*(10), 780-790. doi:10.1089/brain.2014.0253

Lindquist, M. A., Xu, Y., Nebel, M. B., & Caffo, B. S. (2014). Evaluating dynamic bivariate correlations in resting-state fMRI: a comparison study and a new approach. *Neuroimage, 101*, 531-546. doi:10.1016/j.neuroimage.2014.06.052

Liu, X., & Duyn, J. H. (2013). Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proc Natl Acad Sci U S A, 110*(11), 4392-4397. doi:10.1073/pnas.1216856110

Lloyd, S. P. (1982). Least-Squares Quantization in Pcm. *Ieee Transactions on Information Theory, 28*(2), 129-137. doi:10.1109/Tit.1982.1056489

Ma, W. Y., Yao, Q., Hu, G. J., Xiao, C. Y., Shi, J. P., & Chen, J. (2019). Dysfunctional Dynamics of Intra- and Inter-network Connectivity in Dementia With Lewy Bodies. *Front Neurol, 10*, 1265. doi:10.3389/fneur.2019.01265

Miller, R. L., Yaesoubi, M., Turner, J. A., Mathalon, D., Preda, A., Pearlson, G., . . . Calhoun, V. D. (2016). Higher Dimensional Meta-State Analysis Reveals Reduced Resting fMRI Connectivity Dynamism in Schizophrenia Patients. *PLoS One, 11*(3), e0149849. doi:10.1371/journal.pone.0149849

Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., & Ugurbil, K. (2010). Multiband multislice 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), 1144-1153. doi:10.1002/mrm.22361

Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Phys Rev E Stat Nonlin Soft Matter Phys, 69*(2 Pt 2), 026113. doi:10.1103/PhysRevE.69.026113

Orban, C., Kong, R., Li, J., Chee, M. W. L., & Yeo, B. T. T. (2021). Correction: Time of day is associated with paradoxical reductions in global signal fluctuation and functional connectivity. *PLoS Biol, 19*(5), e3001258. doi:10.1371/journal.pbio.3001258

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage,
Preti, M. G., Bolton, T. A., & Van De Ville, D. (2017). The dynamic functional connectome: State-of-the-art and perspectives. *Neuroimage, 160*, 41-54. doi:10.1016/j.neuroimage.2016.12.061

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics, 20*, 53-65.

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage, 52*(3), 1059-1069. doi:10.1016/j.neuroimage.2009.10.003

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X. N., Holmes, A. J., . . . Yeo, B. T. T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex, 28*(9), 3095-3114. doi:10.1093/cercor/bhx179

Shine, J. M., Bissett, P. G., Bell, P. T., Koyejo, O., Balsters, J. H., Gorgolewski, K. J., . . . Poldrack, R. A. (2016). The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance. *Neuron, 92*(2), 544-554. doi:10.1016/j.neuron.2016.09.018

Smith, D. M., Zhao, Y., Keilholz, S. D., & Schumacher, E. H. (2018). Investigating the Intersession Reliability of Dynamic Brain-State Properties. *Brain Connect, 8*(5), 255-267. doi:10.1089/brain.2017.0571

Tagliazucchi, E., von Wegner, F., Morzelewski, A., Brodbeck, V., & Laufs, H. (2012). Dynamic BOLD functional connectivity in humans and its electrophysiological correlates. *Front Hum Neurosci, 6*, 339. doi:10.3389/fnhum.2012.00339

Thompson, G. J., Merritt, M. D., Pan, W. J., Magnuson, M. E., Grooms, J. K., Jaeger, D., & Keilholz, S. D. (2013). Neural correlates of time-varying functional connectivity in the rat. *Neuroimage, 83*, 826-836. doi:10.1016/j.neuroimage.2013.07.036

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., . . . Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15*(1), 273-289. doi:10.1006/nimg.2001.0978

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., & Consortium, W. U.-M. H. (2013). The WU-Minn Human Connectome Project: an overview. *Neuroimage, 80*, 62-79. doi:10.1016/j.neuroimage.2013.05.041

Weng, Y., Liu, X., Hu, H., Huang, H., Zheng, S., Chen, Q., . . . Huang, R. (2020). Open eyes and closed eyes elicit different temporal properties of brain functional networks. *Neuroimage, 222*, 117230. doi:10.1016/j.neuroimage.2020.117230

Yang, Z., Craddock, R. C., Margulies, D. S., Yan, C. G., & Milham, M. P. (2014). Common intrinsic connectivity states among posteroomedial cortex subdivisions: Insights from analysis of temporal dynamics. *Neuroimage, 93 Pt 1*, 124-137. doi:10.1016/j.neuroimage.2014.02.014

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., . . . Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol, 106*(3), 1125-1165. doi:10.1152/jn.00338.2011

Zheng, R., Chen, Y., Jiang, Y., Zhou, B., Li, S., Wei, Y., . . . Cheng, J. (2022). Abnormal dynamic functional connectivity in first-episode, drug-naive adolescents with major depressive disorder. *J Neurosci Res, 100*(7), 1463-1475. doi:10.1002/jnr.25047

Zoller, D., Sandini, C., Karahanoglu, F. I., Padula, M. C., Schaer, M., Eliez, S., & Van De Ville, D. (2019). Large-Scale Brain Network Dynamics Provide a Measure of Psychosis and Anxiety in 22q11.2 Deletion Syndrome. *Biol Psychiatry Cogn Neurosci Neuroimaging, 4*(10), 881-892. doi:10.1016/j.bpsc.2019.04.004
Zuo, X. N., & Xing, X. X. (2014). Test-retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: a systems neuroscience perspective. *Neurosci Biobehav Rev, 45*, 100-118. doi:10.1016/j.neubiorev.2014.05.009

Zurich, B. T. (2017). Dynamic Functional Connectivity tutorial: solutions.