Cognitive and default mode networks support developmental stability in functional connectome fingerprinting through adolescence

Maria Jalbrzikowski, Fuchen Liu, William Foran, Finnegan J. Calabro, Kathryn Roeder, Bernie Devlin, Beatriz Luna

Keywords: connectotyping, resting-state fMRI, precision medicine

Abstract

Initial studies found that individual correlational patterns from resting-state functional magnetic resonance imaging studies can accurately identify another scan from that same individual. This method is known as “connectotyping” or functional connectome “fingerprinting”. We leveraged a unique dataset of 12-30 years old (N=140) individuals who had two distinct resting state scans on the same session (Visit 1, V1), and again 12-18 months later (Visit 2, V2; henceforth 1.5 years) to assess the sensitivity and specificity of identification accuracy across different time scales (same day, 1.5 years apart) and developmental periods (youths, adults). We also used multiple statistical methods to identify the connections that enhance fingerprinting accuracy. We found that sensitivity and specificity to identify one’s own scan was high (overall average AUC: 0.94), and identifiability was significantly higher in the same session (average AUC: 0.97) in comparison to the 1.5-year comparison (average AUC:0.91). The level of fingerprinting accuracy in youths (average AUC:0.93) was not significantly different from adults (average AUC:0.96). Select connections from the Frontoparietal, Default, and Dorsal Attention networks enhanced the ability to identify an individual. Finally, we found that identification of these features generalized across datasets and use of these features improve fingerprinting accuracy in an independent longitudinal data set (N=208). These results provide a framework for understanding that features of fingerprinting accuracy are stable from adolescence through adulthood. Importantly, these features contribute to one’s uniqueness, suggesting that cognitive and default networks play a primary role in establishing one’s connectome.

Introduction

Precision medicine uses fine-grained information about individuals – their specific genetic variation, metabolic and health status, age, and so on – to identify risk level for disease, potential treatments, and probable treatment responses. This is a significant improvement from utilizing deviations from population-based average indicators of pathology, to inform treatment and expected treatment response. Precision medicine has been increasingly implemented to inform clinical practice in biomedicine, such as cancer and cardiology (Antman and Loscalzo, 2016; Ramaswami et al., 2018). Psychiatry has lagged behind these advances, in part, due to the limited availability of fine-grained information about the workings of the brain. However, recent advances in brain functional connectomics have the potential to provide the fine-grained level of individualized characterization needed for precision medicine in psychiatry. In the future, these methods have potential to inform clinical applications. A first step in this direction is to establish approaches that identify accurate, replicable fingerprinting in typical populations and importantly, through a developmental period when psychiatric illness typically emerges.

Accumulating evidence from resting-state functional magnetic resonating imaging (rsfMRI) indicates that brain network architecture is highly individualized (Gordon et al., 2017; Gratton et al., 2018; Laumann et al., 2015a, 2015b). Specifically, several studies have shown that individuals' patterns from one scan can identify another scan from that same individual at a high level of accuracy (Finn et al., 2015; Horien et al., 2019; Miranda-Dominguez et al., 2014). This method was is known as "connectotyping" (Miranda-Dominguez et al., 2014) or functional connectome "fingerprinting" (Finn et al., 2015). Using neuroimaging data for individual characterization has great challenges, however, because the inherent high level of noise in the measured signal limits within-subject reliability (Patriat et al., 2013). However, these emerging analytics abstract recurrent and unique patterns of brain functional connectivity that may allow us to use brain informatics for individualized characterization and inform the personalized medicine framework in psychiatry.

Here, we aimed to build on initial findings indicating individualized brain functional architectonics to characterize how and which brain patterns are unique to an individual and to what extent this pattern is stable or changes over time. First, we assessed the sensitivity of fingerprinting measures, i.e. the likelihood of obtaining a true positive, and the specificity of these measures, i.e. the probability of distinguishing true negatives (Florkowski, 2008; Youngstrom, 2014). We wanted to determine if fingerprinting accuracy met a sensitivity and sensitivity of ~90%, the minimal values typically required for a method to be considered clinically useful (Aamir and Hamilton, 2014; Glaros and Kline, 1988). To determine the specificity and sensitivity of fingerprinting, we applied a classification procedure. Scans from the same individual were viewed as positive pairs while all of the others were considered negative pairs. The area under curve (AUC)-receiver operating characteristics (ROC) curve were utilized as performance measurements to identify the sensitivity (true positive rate) and specificity (1-false positive rate) of fingerprinting at different thresholds and time scales.

The majority of psychiatric disorders emerge during adolescence (Paus et al., 2008), a period of remarkable neuroplasticity and change (Larsen and Luna, 2018; Luna et al., 2015; Murty et al., 2016). Thus, establishing individualized brain markers is important to determine fingerprinting in this period and contrast it to that of adults. Moreover, insofar as we are aware, there has not yet been a direct comparison of scans completed on the same day versus those completed much later (i.e., a year), and to what extent this direct comparison changes or remains the same across adolescence. It is possible that the neuroplasticity observed in resting state scans during adolescence (e.g., Murty, Calabro et. 2019, Jalbrzikowski et al., 2017) and/or motion artifact known to be more predominant in youth (Power et al., 2011; Satterwhaite et al., 2013); reduces the ability to accurately identify an individual's scan. Alternatively, "functional fingerprinting", i.e., identification accuracy of an individual's resting state scan could be robust to these changes. This is an open question, given the conflicted findings observed in the literature (Horien et al., 2019; Kaufmann et al., 2017).

We leveraged a unique, two-time point data set that had two resting state scans from the same individual conducted on the same day (Visit 1, V1), and two resting state scans from the same individual collected on the same day 12-18 months later (Visit 2, V2; henceforth 1.5 years). We used a principled classification approach to assess the level of sensitivity and specificity of fingerprinting accuracy; compared whether it is as stable for the same day as it is 1.5 years later; and determined if sensitivity and specificity at these different time scales were similar in youths and adults. We also used multiple statistical methods to determine connections that are "predictive" of individuals' scans, reflecting one's uniqueness, and we explored how these edges performed in an independent sample.

Participants

For the training sample, neuroimaging data were collected on 140 participants (12-30 years) recruited from the greater Pittsburgh metro area. Participants and their first-degree relatives did not have a psychiatric disorder as determined by phone screen and a clinical questionnaire. Any reported drug use within the last month, history of alcohol abuse, medical illness affecting the central nervous system function, IQ lower than 80, a first-degree relative with a major psychiatric disorder, or any MRI contraindications were exclusion criteria.

To test the generalizability of the predictive edges identified in our training sample, we tested the extent to which the previously identified features from each method improved fingerprinting accuracy in an independent sample (test sample) with longitudinal data (N=208, 1-3 visits). Participant and MR data acquisition information for the test sample is detailed in the Supplementary Text and Supplementary Tables S1-S2.

MR Data Acquisition: Training Sample

Data were acquired using a Siemens 3 Tesla mMR Biograph with a 12-channel head coil. Subjects' heads were immobilized using pillows placed inside the head coil, and subjects

were fitted with earbuds for auditory feedback to minimize scanner noise. For each rsfMRI run, we collected eight minutes of resting-state data, eyes open. Resting state data were collected using an echo-planar sequence sensitive to BOLD contrast (T_2*). rsfMRI parameters were Repetition Time/Echo Time=1500/30.0 ms; flip angle=50°; voxel size = 2.3×2.3×2.3 mm. Structural images were acquired using a T1 weighted magnetization-prepared rapid gradient-echo (MPRAGE) sequence (TR/TE = 2300/2.98 ms; flip angle=9°; voxel size = 1.0×1.0×1.0 mm).

Participants completed a unique two-visit scan protocol. In the first visit, individuals participated in a MRI protocol (Visit 1, V1) that included two rsfMRI runs (Pre-Task, Post-Task), with an fMRI reward learning task (~40 minutes) conducted between these two runs. Approximately 1.5 years later, the same individuals returned and completed an identical MRI protocol (Visit 2, V2), which also included two rsfMRI runs (Pre-Task, Post-Task) separated by the same fMRI task. A visual depiction of the scan protocol, along with the respective names given to each run or scan, are presented in Figure 1.

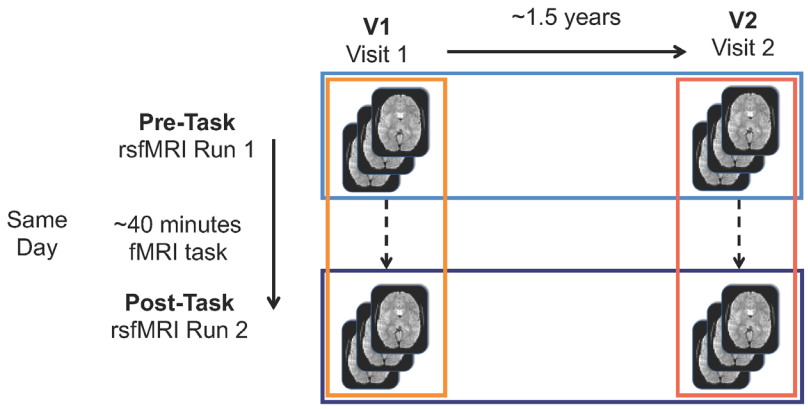

Figure 1. Visualization of protocol set-up. The colors refer to comparisons that will be made throughout the manuscript (oranges: same day comparisons, blues: 1.5 year comparisons).

MR Data Acquisition: Test Sample

Scan parameters for the test sample are detailed in Supplementary text.

rsfMRI Processing

Functional images were warped into MNI standard space using a series of affine and nonlinear transforms. Normalization based on global mode was then calculated on the functional images. Next, all functional images were spatially smoothed using a 5-mm full width at half maximum Gaussian kernel. Removal of non-stationary events in the fMRI time series was conducted using wavelet despiking (Patel and Bullmore, 2015). To control nuisance-related variability (Hallquist et al., 2013), we then conducted simultaneous multiple regression of nuisance variables and bandpass filtering at 0.009 Hz $< f <$ 0.08. Nuisance regressors included were non-brain tissue (NBT), average white matter signal, average ventricular signal, six head realignment parameters obtained by rigid body head motion correction, and the derivatives of these measures. NBT, average white matter, and average ventricular signal nuisance regressors were extracted using MNI template tissue probability masks (>95% white mater, >98% cerebral spinal fluid, (Fonov et al., 2009). ICA-Aroma was implemented to remove motion artifacts (Pruim et al., 2015a, 2015b). For all subjects, we calculated a quality control measure with respect to head motion, namely volume-to-volume frame displacement. Subjects were removed from rsfMRI analyses if the average frame displacement across the run was > 0.5mm. The resulting data showed no effects of motion by age.

Functional Network Parcellation

We applied a previously-defined, functional connectome parcellation of 333 functional regions of interest (ROIs) across cortical structures (Gordon et al., 2016) to each participant's rsfMRI data (Figure 2A). This parcellation consists of 13 reliable rsfMRI networks, many of which have been identified in other studies, including the Frontoparietal, Default, and Visual networks

(Glasser et al., 2016; Power et al., 2011; Shen et al., 2013). See Supplementary Table S3 for a list of 13 networks and details about them (for each network: number of nodes, number of within-connectivity edges, and number of between-connectivity edges).

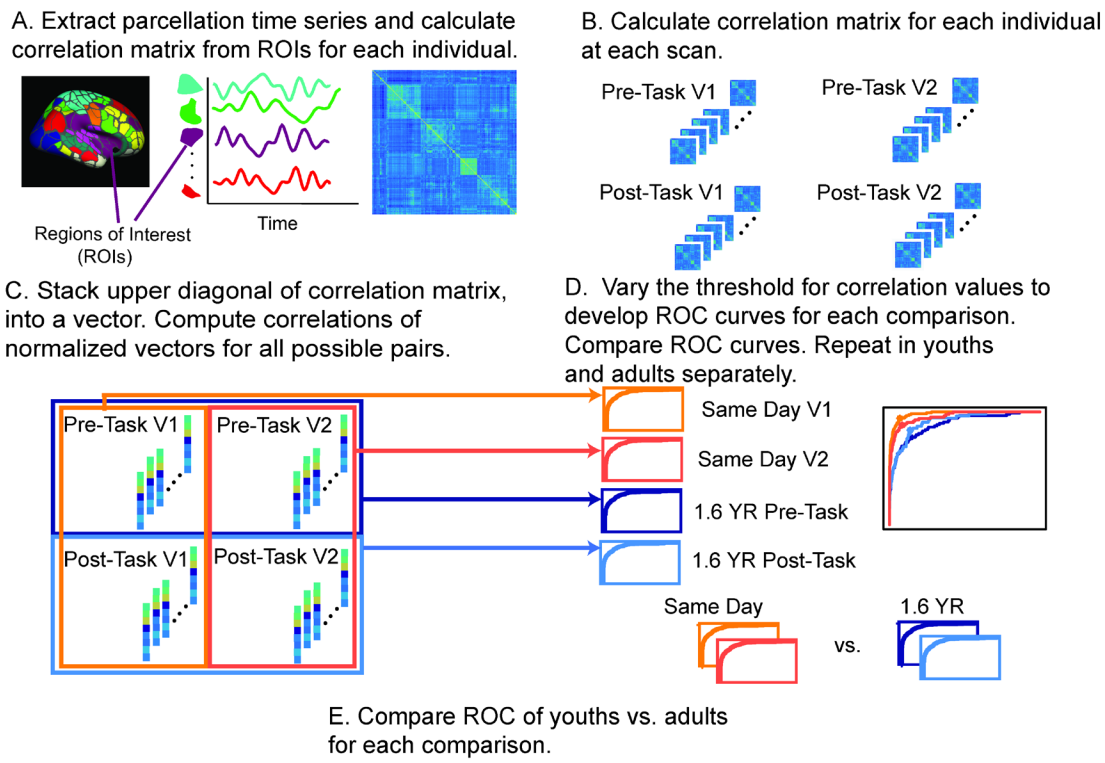

Figure 2. (A) After resting-state fMRI data were processed, we extracted out the time series from an established parcellation and (B) calculated a correlation matrix for each individual and their respective scan visit. (C) For each scan at each visit, we stacked a vector from the upper diagonal of the correlation matrix. Each stacked vector represents a scan from one person. The stacked vectors could be a separate scan from the same individual or a separate scan from a different individual. We computed correlations between each vector for all possible pairs and non-pairs. (D) By varying the threshold of the correlation values to determine what was a true-or false-positive, we developed ROC curves for each comparison and then used DeLong's method to compare the ROC curves. (E) We then compared the ROC curves of youths vs. adults for each comparison.

For each participant, we computed Pearson's correlation of each ROI's time series with that of every other ROI, producing a 333 x 333 correlation matrix (Figure 2B). The upper diagonal of the correlation matrix for each individual was stacked into a vector (55278 edges), and each vector was normalized to a mean of zero and variance of one (Figure 2C). We performed this procedure for each subject's rsfMRI run, resulting in four normalized vectors for the majority of participants. The correlation between any two of these normalized vectors, V and U, was their dot

product:

$$corr(V, U) = \sum_{e=1}^{E} V(e)U(e)\,/E,$$

where E is the total number of edges and e is an individual edge.

Identification Accuracy

Next, we sought a classifier to identify resting state fMRI-measured connectomes that "match"; ideally these would be connectomes from the same subject. To do so, we first computed the correlation of normalized vectors, V and U, for all possible pairs of subjects, as described above (Figure 2C). The classifier seeks a threshold *t* for these correlation values that yields a high rate of true positive identification, namely connectomes from the same subject, while minimizing the number of false positive identifications, connectomes from different subjects labeled as from the same subject. We varied *t* from interval of zero to one to create a receiver operating characteristic (ROC) curve. We then estimated the area under the curve to determine the accuracy of the classifier (Figure 2D). The *t* that maximized true positive rate-false positive rate (TPR-FPR) was chosen for reporting. ROC curves were generated for the entire sample for 1) same day identification accuracy (Pre-Task vs. Post-Task) and 2) identification accuracy 1.5 years apart (V1 vs. V2). To compare identification accuracy for same day vs. 1.5 years later, we compared ROC curves to each using DeLong's test for two ROC curves ((DeLong et al., 1988; Robin et al., 2011), Figure 2D).

To determine whether identification accuracy was affected by age, we split the entire sample of 12-30 year olds by the median age (20.4 years), considering the participants "youths" if they were under the median age (Visit 1 N=70) or "adults" if they were over the median age (Visit 1 N=70). We then calculated ROC curves for youths and adults separately for 1) same day fingerprinting (Pre-Task vs. Post-Task) and 2) fingerprinting 1.5 years apart (V1 vs. V2). To test for significant differences in identification between youths and adults, we used DeLong's test for two ROC curves (Figure 2E).

Identification of Predictive Edges

Next, we sought to determine which edges contribute most to identification accuracy. We took four comparisons for each subject (V1, pre versus post; V2, pre versus post; V1 pre versus V2 pre and V1 post versus V2 post) and the same comparisons among pairs (two scans from the same subject), resulting in 98790 comparisons (same subject pairs=532, non-pairs=98258). Because there was an over-representation of non-pairs (two scans, one from individual s and another from individual j) , we used synthetic minority over-sampling technique (Chawla et al., 2002) and selected 532 pairs and 1596 non-pairs. This method uses Euclidean distance to select non-pairs closest to the pairs (Chawla et al., 2002). We then split the data into training (2/3 of the data: 355 pairs, 1064 non-pairs) and test sets (1/3 of the data: 177 pairs, 532 non-pairs). For each method, described below, we used 5-fold cross validation to identify the number of edges that gives the highest AUC. The terms of the dot product (i.e., from 52,278 edges) for each comparison were the input features.

Finn Method: Finn et al. described a method to calculate the most predictive edges (Finn et al., 2015). Briefly, consider normalized vectors $V_s^{(1)}$ and $V_s^{(2)}$ from two scans from the same individual s and a normalized vector from another individual j, $V_j^{(k)}$. The product of $V_s^{(1)}(e)V_s^{(2)}(e)$ is considered the within-subject edge product. The product of $V_s^{(1)}(e)V_j^{(k)}(e)$ is the between-subject edge product. For each edge e, we count the number of times that the within-subject product is greater than every between-subject product comparison:

$$n_s(e) = \sum_{j \neq s, k \in \{1,2\}} I\left(V_s^{(1)}(e)V_s^{(2)}(e) > V_s^{(1)}(e)V_j^{(k)}(e)\right) + \sum_{j \neq s, k \in \{1,2\}} I\left(V_s^{(1)}(e)V_s^{(2)}(e) > V_j^{(k)}(e)V_s^{(2)}(e)\right)$$

where $I(\cdot)$ is an indicator function. The "edge power" for one edge, $p_s(e)$, is defined as $(n_s(e))$ divided by the total number of all possible comparisons N. Larger $p_s(e)$ means that there is higher

predictive power for edge e. We obtained the edge power for each edge e by calculating the sum of $p_s(e)$ over all subjects. We then ranked all edges by their edge power (from most predictive to least predictive), and we successively decreased the threshold of "most predictive" edges to develop the ROC curve. Through cross-validation in the training set, we determined that the optimal TPR-FPR rate was when we included the five percent "most predictive" edges. Then, in the test data set, we selected only these predictive edges and calculated a correlation between each scan and all other scans (sum of the dot product). We then used the correlation threshold to develop the ROC curve in the test data set.

Support Vector-Machine Learning and Elastic Net Regression Methods: Both of these methods are common tools for model selection. Here, the goal was to choose a set of predictive edges that differentiated same-subject pairs from other pairs. For both methods, the input information was the terms of the dot product between two scans. We developed optimal tuning parameters in the training data set and obtained weights for the selected edges. Elastic net regression was implemented using R package glmnet (Friedman et al., 2009). The support-vector machine (SVM) analysis was implemented in R package (sparseSVM, Yi and Zeng, 2018), with an elastic net penalty). We then applied the model developed in the training set to the test data set. By weighting the individual product of the selected edges, we obtained a predicted value, $\hat{Y}$, for which $Y - \hat{Y}$ ranges from -1 to 1. We developed ROC curves by changing the threshold of $Y - \hat{Y}$ to determine what constituted a "pair" in the test data.

Assessing over-representation of network connectivity in predictive edges

To identify patterns in the predictive edges for each method identified above, we conducted Chi-square tests to assess whether there was over-representation of 1) between (e.g. Frontoparietal-Default edges; the off-diagonal edges) or within-network connectivity (e.g., Frontoparietal-Frontoparietal edges; the block structure on the diagonal) and 2) specific within-network connectivity networks, and/or 3) distinct between-network connections. We used the

standardized residuals (>3.0) to determine networks that contributed to one's uniqueness.

Performance of predictive edges in test sample

To test the generalizability of the predictive edges identified in our training sample, we tested the extent to which the previously identified features from each method improved fingerprinting accuracy in an independent sample with longitudinal data (test sample).

Effects of Possible Confounds

Because "same day" scans were acquired within the same MRI session, improved same day fingerprinting accuracy could be due to better-quality registration from the same scan session. A subset of the participants (ages 18-30 years, N=76) participated in an additional scan on the same day, but in a different scan session (i.e, the participant came out of the MRI scanner and a few hours later participated in another MRI session). This subset of participants participated in a position emission tomography study that included an additional MRI scan with each visit. (See Supplementary Table S4 for details on this subset of participants.) Classification as detailed previously was performed, using the Pre- and Post-task scans at each visit, to predict identification accuracy from this additional MRI session.

Results

Participant information for both visits is presented in Table 1. Participant information for youth and adult groups are reported in Supplementary Table S5.

Table 1. Participant Information

	Pre-task					Post-task				
	N	M/F	Mean Age	Age Range	Mean FD	N	M/F	Mean Age	Age Range	Mean FD
Visit 1	140	67/73	19.7 [5.0]	12.0-31.0	0.24 [0.1]	126	57/69	20.1 [4.9]	12.0-31.0	0.25 [0.1]
Visit 2	93	48/45	22.2 [5.0]	13.5-32.6	0.22 [0.1]	86	45/41	22.3 [4.9]	13.5-32.6	0.23 [0.1]

Same Day vs. 1.5 years later

Across the entire sample, identifiability on the same day was quite high (Average AUC: 0.94). Identification accuracy between scans 1.5 years apart was also high (Average AUC: 0.91, Figure 3A). However, same day accuracy was significantly higher than identification of scans 1.5 years apart (Figure 3A, Table 2A).

Table 2. A) Across the entire sample, same-day identification accuracy was significantly higher than identification accuracy 1.5 years apart. The D-value represents the test statistic for the respective comparison. This pattern remained when youth (B) and adults (C) were assessed separately.

Group	Same Day	1.5 years apart			
		1.6 YR Pre-Task		1.6 YR Post-Task	
		D	*P*	D	*P*
A. Entire Sample	Same Day V1	4.1	2.70E-05	4.1	3.80E-05
	Same Day V2	2.7	0.02	2.9	0.003
B. Youths	Same Day V1	3.2	0.001	2	0.04
	Same Day V2	2.7	0.008	1.1	0.24
C. Adults	Same Day V1	2.8	0.008	3.2	0.001
	Same Day V2	1.8	0.07	2	0.04

Differences in fingerprinting accuracy on the same day compared to scans 1.5 years apart were observed for youth (Figure 3B, Table 2B) and adult groups (Figure 3C, Table 2C).

Fingerprinting accuracy of youth vs. adults

Table 3. A) Same-day fingerprinting accuracy is similar for youth and adults. B) Fingerprinting accuracy 1.5 years apart is similar for youth and adults in three out of four comparisons.

Comparison		Adults vs. Youth	
		D	*P*
A.	Same Day V1	-1.7	0.08
	Same Day V2	0.16	0.87
B.	1.6 YR: Pre-task V1 vs. Pre-task V2	-2.7	0.006
	1.6 YR: Pre-task V1 vs. Post Task V2	-0.42	0.67
	1.6 YR: Post-task V1 vs. Post-task V2	-0.7	0.46
	1.6 YR: Post-task V1 vs. Pre-Task V2	-1.7	0.09

Same day accuracy was not statistically different between youths and adults. The two groups exhibited similar levels of same-day identification accuracy (Table 3A). The same pattern emerged when comparing youths and adults on fingerprinting accuracy 1.5 years apart for three out of four of the comparisons (Table 3B). ROC Curves are presented in Figure 3.

All model selection methods tested improve identification accuracy in the test portion of training sample

As seen in Figure 4A, the fingerprinting accuracy significantly improved when we used predictive edges selected by the Finn method, SVM, or Elastic Net to predict identification accuracy. The three model selection techniques performed similarly to one another (Supplementary Table S6). Each method improved fingerprinting accuracy in all previous comparisons but did not change results reported above.

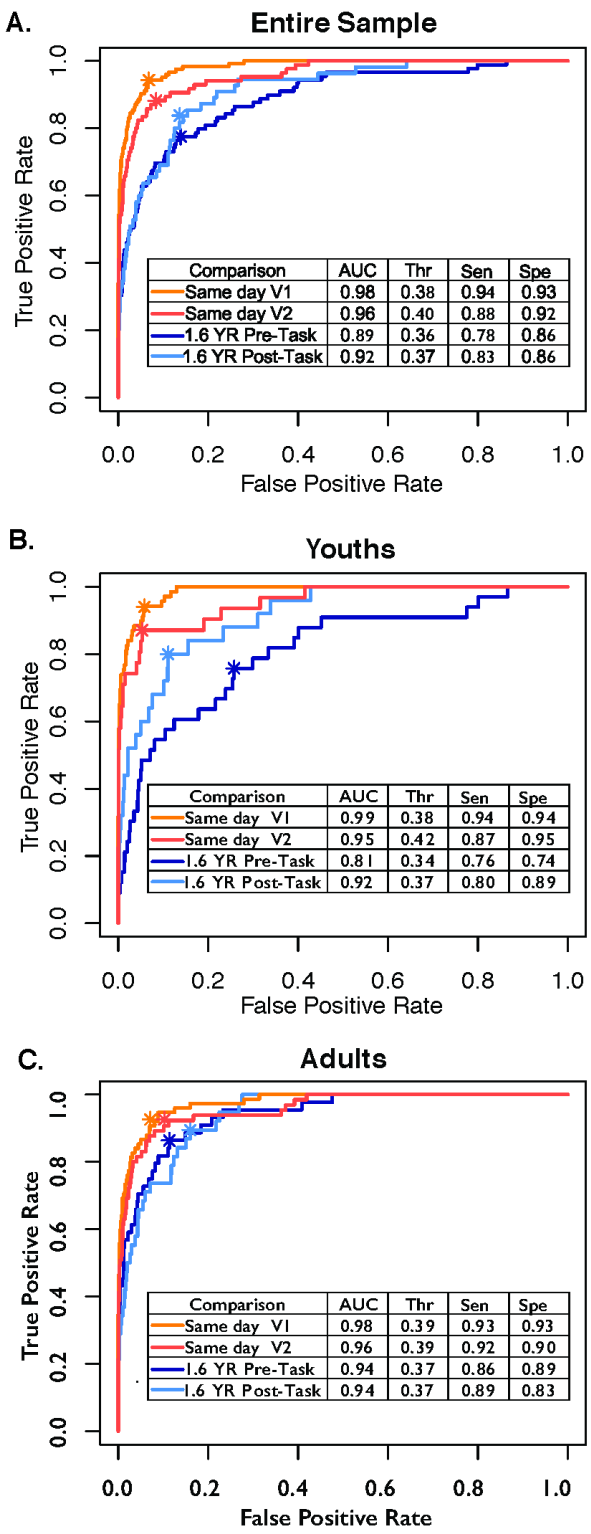

Figure 3. A) Across the entire sample, identification accuracy was higher for same day (orange) vs. 1.5 years later (blue). This pattern remained when youth (B) and adults (C) were assessed separately. AUC= Area under the curve, Thr=threshold at which optimal true positive rate was obtained. Sen=sensitivity, Spec=specificity

Predictive edges from one sample improves accuracy in independent sample

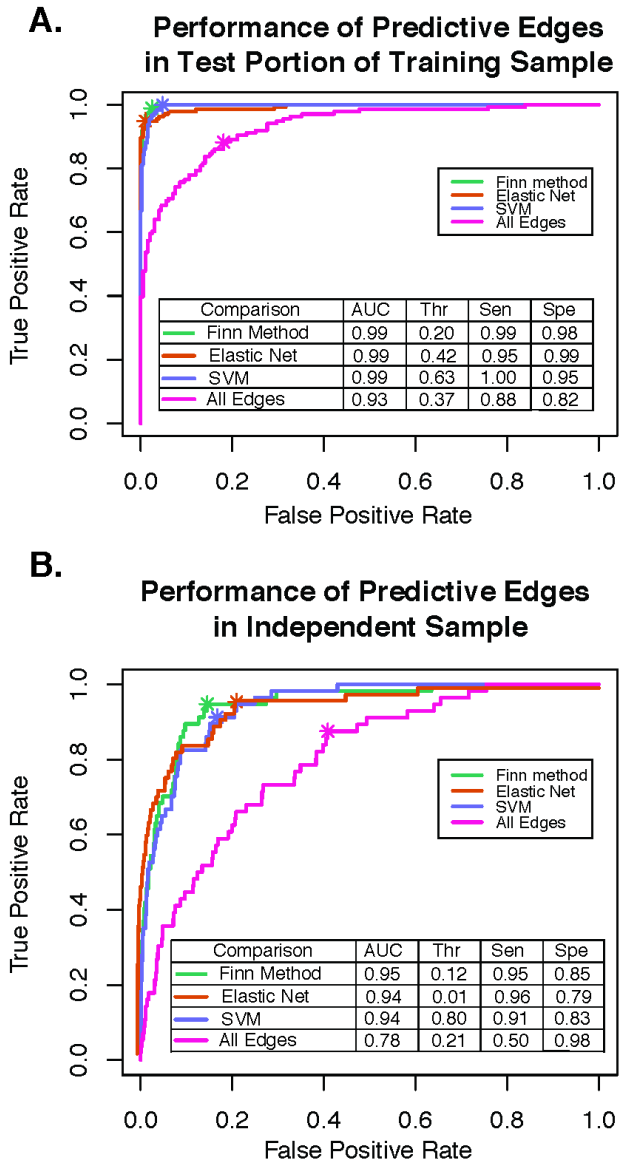

Figure 4. A). When edge selection was performed via various methods (Finn method (green), Elastic Net (orange), and SVM (purple)), identification accuracy was significantly improved in comparison to using all edges (pink) for identification accuracy. B. When we applied predictive edges previously identified in the training sample to an independent sample, all methods significantly improved identification accuracy.

Similar to previous results (Finn et al., 2015; Waller et al., 2017), we found lower identification accuracy in a longitudinal test sample with more "standard" MRI parameters (e.g., longer TR, shorter length of scan, fewer number of head coils, Figure 4B, pink ROC curve). When we applied the most predictive edges from previously applied model selection techniques, however, fingerprinting accuracy significantly improved (Figure 4B, Supplementary Table S7). When we omitted pairs from scans conducted on the same day and focused our analyses on identifying predictive edges from pairs that were 1.5 years apart, we obtained the same pattern of results (Supplementary Table S8). The weights for predictive edges from Elastic Net and SVM are provided in Supplementary Tables S9-10.

Predictive edges are over-represented in Frontoparietal, Default, and Dorsal Attention Networks

Figure 5A-C shows the relative contribution, normalized for number of edges in each network (or off/diagonal edge group) and number of edges determined to be predictive by each method. Across all methods, in comparison to between-connectivity edges (e.g., Frontoparietal-

Visual connections, Supplementary Table S11), within-connectivity edges (e.g., Frontoparietal-Frontoparietal connections) were relatively more important in predicting identification accuracy.

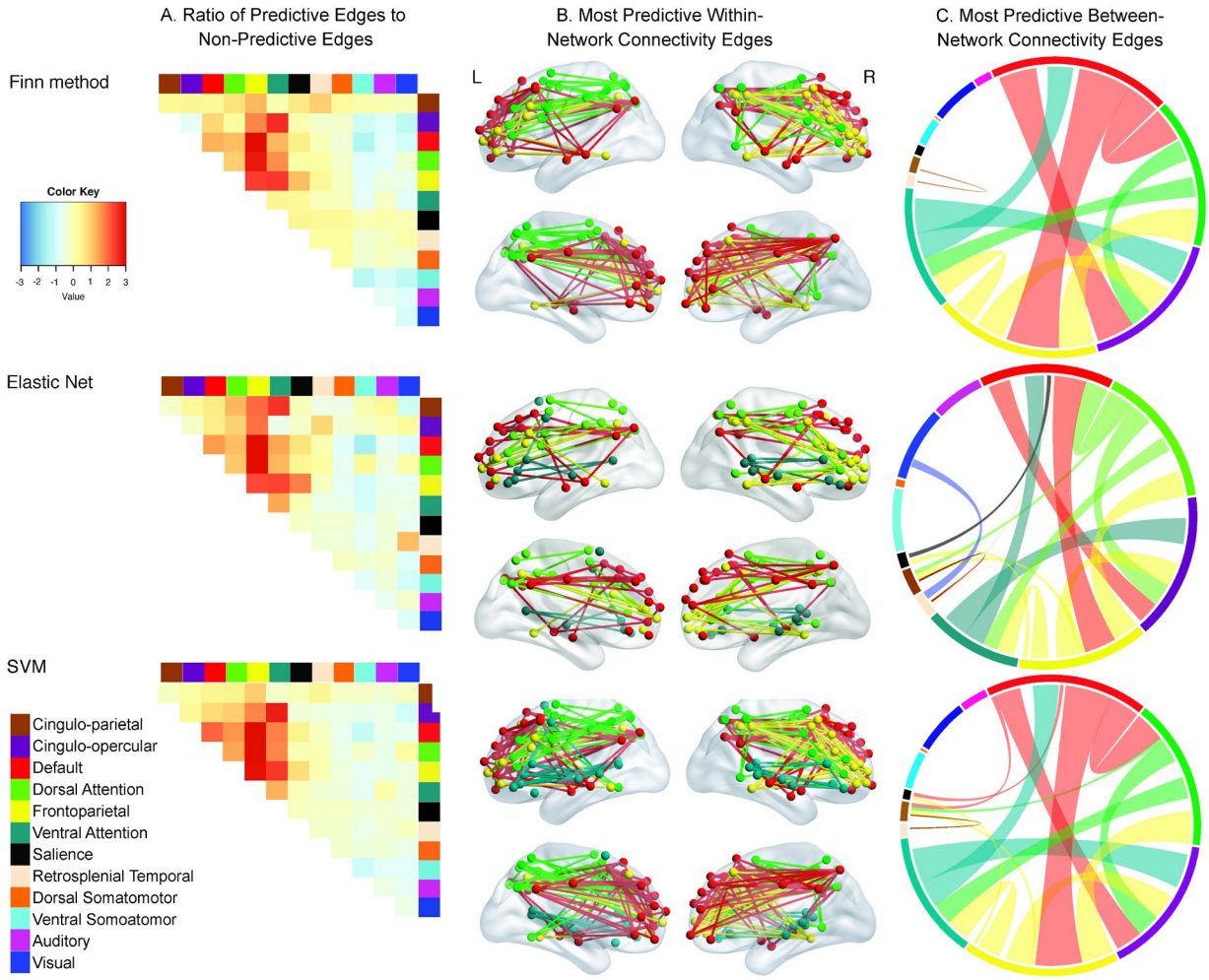

Figure 5. A) The ratio of predictive edges to non-predictive edges in each network connection, normalized for total number of edges in each connection for the three different methods. Warmer colors on the heatmap indicate that edges from a particular network are more important for identification accuracy. B) Within-network connections that are particularly important for identification accuracy in all three methods examined. In all methods, within connectivity edges in Frontoparietal (yellow), Default (red), and Dorsal Attention (bright green) networks are considered predictive. In the Finn method and SVM, within connectivity connections in the Ventral Attention network were also predictive of identification accuracy. C) Between-network connections that are particularly important for identification accuracy. The colors around the circle reflect the different networks examined. Thicker bands of color indicate a greater number of edges from that particular network were considered predictive. The between connectivity edges (lines going across the circle) were randomly chosen from one of the two connected networks (e.g., Between network connectivity between the Default and Cingulo-Opercular network is red, but between network connectivity between the Default and Ventral Attention network is green).

More specifically, within-connectivity edges from the frontoparietal, default, and dorsal attention networks drove this finding (Figure 5B). With SVM and Elastic Net, edges from the Ventral Attention network were also considered important for fingerprinting accuracy. These

networks had standardized residuals greater than 3.0 in all comparisons (Supplementary Table S12). A similar pattern emerged when we examined same-day and 1.5-year comparisons separately (Supplementary Table S13) and youth and adults separately (Supplementary Tables S14-15).

We also examined which between-network connectivity edges were relatively important for fingerprinting accuracy. Across all three methods, connections between the Frontoparietal-Default, Frontoparietal- Dorsal Attention, Ventral Attention-Cingulo-opercular, and Frontoparietal-Ventral Attention networks were over-represented in comparison to other connections (Figure 5C). See Supplementary Table S16 for standardized residuals from chi-square test of between-network connectivity and Supplementary Table S17 for standard residuals from chi-square test of over-representation of all network connections. Similar patterns emerged when we examined same-day and 1.5-year comparisons separately (Supplementary Table S18) and youth and adults separately (Supplementary Table S19).

Testing Effects of Possible Confounds.

To ensure that the improved same-day accuracy was not because individuals were in the same scan session and benefitting from improved MRI registration, we examined a subset of individuals (N=76) who completed an additional MRI session on V1 and V2 after being taken out of the scanner and repositioned. Supplementary Table S20 reports similar AUC, threshold levels, and sensitivity/specificity for identification accuracy within the same MRI session (V1 Pre-Task predicting accuracy of V1-Post-Task) and different MRI session on the same day (e.g., V1 Pre-Task predicting identification of extra session V1).

To ensure that our results were not driven by parcellation choice, we also re-ran all analyses after extracting ROIs from two separate parcellations (Power et al., 2011; Shen et al., 2013) and obtained a similar pattern of results.

Discussion

This is the first study to show that identification accuracy of one's resting state scan – how much it reflects a "functional fingerprint" – depends on the amount of time in between assessments. We provide supporting evidence that adolescents have similar levels of fingerprinting accuracy to adults (Horien et al., 2019; Miranda-Dominguez et al., 2018) and extend this literature to show that this pattern is consistent on a same-day visit and visits 1.5 years apart. Using multiple methods, we also found that a small number of edges, which are more likely to be in the Frontoparietal, Default, and Dorsal Attention networks, are consistently predictive of an individual's scan. We identified these edges in a training sample and then used these edges to improve identification accuracy in an independent sample. We propose that particular edges in the Frontoparietal, Default, and Dorsal Attention networks contribute to an individual's "uniqueness" and are stable from late childhood through adulthood. These results bring us a fuller understanding of the features of identification accuracy, both in terms of stability and relative contribution of network components.

Stability of identification accuracy across time

Our results indicate a high level of subject identification accuracy even after 18 months. although the greater time interval incurred a significant decrease in identification. This result provides compelling evidence that there are extant foundational properties to individualized resting state network organization that are persistent and specific to each individual. The significant degradation in identification after 18 months may reflect inherent changes on the level of acquisition beyond issues of registration, which we tested and found to not be a contributor. For example, low temporal signal-to-noise ratio is a feature of rsfMRI scans (Noble et al., 2017) the reduced fingerprinting accuracy may reflect greater noise between two scans as time increases. Alternatively, or in addition to, the small but significant degradation in identification may reflect inherent plasticity in this foundational network organization, particularly

with regards to higher-order processing (cognitive networks and DMN) that may be more dynamic over the long-term and present across developmental stages.

These findings have important implications to understanding stability of physiological or psychological status. In support of this notion, individuals at risk for and with psychiatric disorders have “reduced stability” or lower identification accuracy in comparison to typically developing youth and healthy controls (Kaufmann et al., 2018, 2017). There is also preliminary evidence showing that individuals with similar cognitive and behavioral profiles have more similar functional connectomes (Biazoli et al., 2017). These initial studies suggest that biologically informed methods like individual level connectome fingerprinting accuracy could be used to identify those at risk for or with psychiatric disorders. In the future, we plan to use the statistical framework laid out in this manuscript to test the specificity and sensitivity of case-control status with rsfMRI identification accuracy, identifying which edges are responsible for psychiatric liability. Furthermore, we plan to assess the extent to which the predictive edges that contribute to an individual’s uniqueness are linked to stable, trait-like features of an individual, such as personality.

Networks that Underlie Prediction

We found evidence that identification was driven by edges particularly in the Frontoparietal, Dorsal attention, and Default mode networks which was present across analytic approaches. This is a striking result that identifies networks critical for higher-order cognitive processing and endogenous self-referential processing. Thus, these results provide suggestive evidence that how we engage foundational cognitive and endogenous processes may contribute to demarcating individuality. This finding also has the potential to inform our understanding of impaired development in psychopathology that may be particularly represented in these precise networks.

Identification accuracy is similar in youths and adults

We did not observe differences in the fingerprinting accuracy between youth and adults, for both same day and 1.5-year comparisons. Indeed, others have found that identification accuracy is stable in both youth and adults (Horien et al., 2019). We extend these findings by showing accuracy is high both for same day and longer-term (1.5 years) intervals. Furthermore, because identification accuracy in both youth and adults was lower across a longer period of time and similar edges contributed to same day and 1.5-year identification accuracy, our results suggest that this reduction is not due to known developmental changes. There is significant cognitive development through adolescence (Steinberg et al., 2005; Murty, Calabro et al., 2016; Larsen & Luna, 2018), in the context of evidence for stability at the group level in network properties (Hwang et al., 2013; Jalbrzikowski et al., 2019; Marek et al., 2015). The stability in identification accuracy across development further supports that implication that network properties contain individualized foundational properties that define uniqueness.

In one case, we did find that youths had significantly worse identification accuracy in comparison to adults (Table 3B: 1.5 YR: Pre-task V1 vs. Pre-task V2) when the scans were 1.5 year apart. However, in three out of four similar relevant comparisons, we did not observe this pattern. In this particular comparison (i.e., Pre-Task V1 to Pre-Task V2), reduced accuracy in these youths was driven by the lower identification accuracy in the pre-task scans: we speculate that youths are more variable and excitable when first getting in a scanner, perhaps because they have had less experience with life events akin to an MRI scan than do adults. Furthermore, we know that identification accuracy is reduced with increased head motion (Horien et al., 2018), and youth have greater levels of head motion in comparison to adults (Satterthwaite et al., 2012). However, in our sample, we did not see a statistical difference between head motion in these two groups and our results remain stable when we use more conservative framewise displacement thresholds (<.3).

An Improved Statistical Framework for Identification Accuracy

We show that edges important for identification accuracy are similar across different methods used to identify them. Furthermore, predictive edges identified in one sample can be applied to an independent sample to improve identification accuracy. The robustness of these results demonstrates that only particular edges are important for fingerprinting accuracy. We also provide a statistical framework that can be used to assess the clinical utility of identification accuracy and to assess specific connections within and among brain networks. Together, our work shows that fingerprinting accuracy has some of the features required for use within a precision medicine framework.

Viewing identification accuracy of rsfMRI scans as a classification problem is useful to answer questions relevant for precision medicine. To improve early identification and detection of those at risk for psychiatric disorders, we need to answer questions such as, do people with similar connectivity profiles share common psychiatric features? Or, do those who go on to develop a psychiatric disorder have reduced accuracy in identifiability? These questions all fall within the realm of a classification problem, and the framework that we use in this paper can be applied to relevant data to answer these questions. Indeed, multiple research groups suggest that individual-based identification accuracy of resting state scans can be used to improve upon the current, non-biological based psychiatric diagnoses (Finn et al., 2015; Miranda Domniguez et al 2014). Importantly, the ability to identify networks that drive predictions in psychopathology can inform treatment efficacy.

We believe this statistical framework also improves upon previous methods used in this area. Previous methods, for instance, presume there is a “match” for each respective scan in the data set (i.e., each individual has at least two scans in the pool of available data) and do not consider false positives. Furthermore, given that research shows that identifiability of individuals decreases as sample size increases (Waller et al., 2017), it is important to account for sample size in the model. Additionally, while many studies show that the identification test metric for

individual identifiability is significantly greater than would be expected by chance, it is difficult to know how meaningful this metric is when identification accuracy is in the range of ~40-60% (e.g., Horien et al., 2019). Thus, in this study we viewed fingerprinting accuracy as a classification problem and showed that identification accuracy is within the ranges necessary for use in precision medicine. Finally, we assured that our statistical procedures were both replicable and generalizable to an independent data set. We 1) trained a portion of our data to identify predictive edges (i.e., 75% of training data), 2) assessed the performance of the training set in a test portion of the training data (i.e., 25% of training data), and 3) determined the generalizability of our results in a fully-independent sample (test data).

Limitations

As with any study, there are limitations present in our design. First, we split our sample by the median age and identified the two groups as “youths” and “adults” and this approach could obscure subtle differences in identification accuracy that occur across adolescent development. We chose this approach because identification accuracy increases with smaller sample sizes, and our sample size would be quite small if we split our sample into more age groups, as we have done in previous publications. Another approach could be to view age as a continuous variable; however, this imposes a strong linear assumption, when we know that development through adolescence is curvilinear in nature as stability is reached. In the future, examining identification accuracy in large samples of youth (i.e., the Adolescent Brain Cognitive Development Study, Casey et al., 2018) in comparison to large samples of adults (e.g., Human Connectome Project, Van Essen et al., 2013) may prove to be the most fruitful in terms of more fully understanding identification accuracy across development.

Conclusions and Future Directions

In this study, we showed that identification accuracy is high in both youths and adults even after extended periods of time. Importantly, our results suggest that networks properties may have an individualized foundational characterization that may be inherent to individuation

with some room for flexibility in expression. Given these implications and the sensitivity and specificity of this approach, it supports the ability to move forward into precision medicine in psychiatry. We foresee that simultaneously incorporating both group similarities and individual differences into identification accuracy will become the sine qua non for identifying those at risk for psychiatric disorders.

References

Aamir, A., Hamilton, R.G., 2014. Predictive Value Model for Laboratory Tests: Diagnostic Sensitivity, Diagnostic Specificity, Positive and Negative Predictive Value, Efficiency, Likelihood Ratio ([positive and negative]), Incidence and Prevalence. Encyclopedia of Medical Immunology. https://doi.org/10.1007/978-1-4614-9194-1_312

Antman, E.M., Loscalzo, J., 2016. Precision medicine in cardiology. Nat. Rev. Cardiol. 13, 591–602.

Biazoli, C.E., Salum, G.A., Pan, P.M., Zugman, A., Amaro, E., Rohde, L.A., Miguel, E.C., Jackowski, A.P., Bressan, R.A., Sato, J.R., 2017. Commentary: Functional connectome fingerprint: identifying individuals using patterns of brain connectivity. Front. Hum. Neurosci. 11, 47.

Casey, B.J., Cannonier, T., Conley, M.I., Cohen, A.O., Barch, D.M., Heitzeg, M.M., Soules, M.E., Teslovich, T., Dellarco, D.V., Garavan, H., Orr, C.A., Wager, T.D., Banich, M.T., Speer, N.K., Sutherland, M.T., Riedel, M.C., Dick, A.S., Bjork, J.M., Thomas, K.M., Chaarani, B., Mejia, M.H., Hagler, D.J., Jr, Daniela Cornejo, M., Sicat, C.S., Harms, M.P., Dosenbach, N.U.F., Rosenberg, M., Earl, E., Bartsch, H., Watts, R., Polimeni, J.R., Kuperman, J.M., Fair, D.A., Dale, A.M., ABCD Imaging Acquisition Workgroup, 2018. The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. Dev. Cogn. Neurosci. 32, 43–54.

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. J. Artif. Intell. Res. 16, 321–357.

DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L., 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44, 837–845.

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat. Neurosci. 18, 1664–1671.

Florkowski, C.M., 2008. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. Clin. Biochem. Rev. 29 Suppl 1, S83–7.

Fonov, V.S., Evans, A.C., McKinstry, R.C., Almli, C.R., 2009. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. NeuroImage.

Friedman, J., Hastie, T., Tibshirani, R., 2009. glmnet: Lasso and elastic-net regularized generalized linear models. R package version 1.

Glaros, A.G., Kline, R.B., 1988. Understanding the accuracy of tests with cutting scores: the sensitivity, specificity, and predictive value model. J. Clin. Psychol. 44, 1013–1023.

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

Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M., Petersen, S.E., 2016. Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. Cereb. Cortex 26, 288–303.

Gordon, E.M., Laumann, T.O., Gilmore, A.W., Newbold, D.J., Greene, D.J., Berg, J.J., Ortega, M., Hoyt-Drazen, C., Gratton, C., Sun, H., Hampton, J.M., Coalson, R.S., Nguyen, A.L., McDermott, K.B., Shimony, J.S., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., Nelson, S.M., Dosenbach, N.U.F., 2017. Precision Functional Mapping of Individual Human Brains. Neuron 95, 791–807.e7.

Gratton, C., Laumann, T.O., Nielsen, A.N., Greene, D.J., Gordon, E.M., Gilmore, A.W., Nelson, S.M., Coalson, R.S., Snyder, A.Z., Schlaggar, B.L., Dosenbach, N.U.F., Petersen, S.E., 2018. Functional Brain Networks Are Dominated by Stable Group and Individual Factors,

Not Cognitive or Daily Variation. Neuron 98, 439–452.e5.
Hallquist, M.N., Hwang, K., Luna, B., 2013. The nuisance of nuisance regression: spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. Neuroimage 82, 208–225.
Horien, C., Noble, S., Finn, E.S., Shen, X., Scheinost, D., Constable, R.T., 2018. Considering factors affecting the connectome-based identification process: Comment on Waller et al. Neuroimage 169, 172–175.
Horien, C., Shen, X., Scheinost, D., Constable, R.T., 2019. The individual functional connectome is unique and stable over months to years. Neuroimage 189, 676–687.
Hwang, K., Hallquist, M.N., Luna, B., 2013. The development of hub architecture in the human functional brain network. Cereb. Cortex 23, 2380–2393.
Jalbrzikowski, M., Liu, F., Foran, W., Roeder, K., Devlin, B., Luna, B., 2019. Resting-State Functional Network Organization Is Stable Across Adolescent Development for Typical and Psychosis Spectrum Youth. Schizophr. Bull. https://doi.org/10.1093/schbul/sbz053
Kaufmann, T., Alnæs, D., Brandt, C.L., Bettella, F., Djurovic, S., Andreassen, O.A., Westlye, L.T., 2018. Stability of the Brain Functional Connectome Fingerprint in Individuals With Schizophrenia. JAMA Psychiatry 75, 749–751.
Kaufmann, T., Alnæs, D., Doan, N.T., Brandt, C.L., Andreassen, O.A., Westlye, L.T., 2017. Delayed stabilization and individualization in connectome development are related to psychiatric disorders. Nat. Neurosci. 20, 513–515.
Larsen, B., Luna, B., 2018. Adolescence as a neurobiological critical period for the development of higher-order cognition. Neurosci. Biobehav. Rev. 94, 179–195.
Laumann, T.O., Gordon, E.M., Adeyemo, B., Snyder, A.Z., Joo, S.J., Chen, M.-Y., Gilmore, A.W., McDermott, K.B., Nelson, S.M., Dosenbach, N.U.F., Schlaggar, B.L., Mumford, J.A., Poldrack, R.A., Petersen, S.E., 2015a. Functional System and Areal Organization of a Highly Sampled Individual Human Brain. Neuron 87, 657–670.
Laumann, T.O., Koyejo, O., Gregory, B., Hover, A., Chen, M.-Y., Gorgolewski, K.J., Luci, J., Joo, S.J., Boyd, R.L., Hunicke-Smith, S., Simpson, Z.B., Caven, T., Sochat, V., Shine, J.M., Gordon, E., Snyder, A.Z., Adeyemo, B., Petersen, S.E., Glahn, D.C., McKay, D.R., Curran, J.E., Ring, H.H.H.G.O., Carless, M.A., Blangero, J., Dougherty, R., Leemans, A., Handwerker, D.A., Frick, L., Marcotte, E.M., Mumford, J.A., Poldrack, R.A., 2015b. Long-term neural and physiological phenotyping of a single human. Nat. Commun. 6, 1–15.
Luna, B., Marek, S., Larsen, B., Tervo-Clemmens, B., Chahal, R., 2015. An integrative model of the maturation of cognitive control. Annu. Rev. Neurosci. 38, 151–170.
Marek, S., Hwang, K., Foran, W., Hallquist, M.N., Luna, B., 2015. The Contribution of Network Organization and Integration to the Development of Cognitive Control. PLoS Biol. 13, e1002328.
Miranda-Dominguez, O., Feczko, E., Grayson, D.S., Walum, H., Nigg, J.T., Fair, D.A., 2018. Heritability of the human connectome: A connectotyping study. Network Neuroscience 2, 175–199.
Miranda-Dominguez, O., Mills, B.D., Carpenter, S.D., Grant, K.A., Kroenke, C.D., Nigg, J.T., Fair, D.A., 2014. Connectotyping: model-based fingerprinting of the functional connectome. PLoS One 9, e111048.
Murty, V.P., Calabro, F., Luna, B., 2016. The role of experience in adolescent cognitive development: Integration of executive, memory, and mesolimbic systems. Neurosci. Biobehav. Rev. 1–32.
Patel, A.X., Bullmore, E.T., 2015. A wavelet-based estimator of the degrees of freedom in denoised fMRI time series for probabilistic testing of functional connectivity and brain graphs. Neuroimage.
Patriat, R., Molloy, E.K., Meier, T.B., Kirk, G.R., Nair, V.A., Meyerand, M.E., Prabhakaran, V., Birn, R.M., 2013. The effect of resting condition on resting-state fMRI reliability and

consistency: A comparison between resting with eyes open, closed, and fixated. Neuroimage 78, 463–473.
Paus, T., Keshavan, M., Giedd, J.N., 2008. Why do many psychiatric disorders emerge during adolescence? Nat. Rev. Neurosci. 9, 947–957.
Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., Petersen, S.E., 2011. Functional network organization of the human brain. Neuron 72, 665–678.
Pruim, R.H.R., Mennes, M., Buitelaar, J.K., Beckmann, C.F., 2015a. Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. Neuroimage 112, 278–287.
Pruim, R.H.R., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J.K., Beckmann, C.F., 2015b. ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. Neuroimage 112, 267–277.
Ramaswami, R., Bayer, R., Galea, S., 2018. Precision Medicine from a Public Health Perspective. Annu. Rev. Public Health 39, 153–168.
Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., Müller, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 12, 77.
Satterthwaite, T.D., Wolf, D.H., Loughead, J., Ruparel, K., Elliott, M.A., Hakonarson, H., Gur, R.C., Gur, R.E., 2012. Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. Neuroimage 60, 623–632.
Shen, X., Tokoglu, F., Papademetris, X., Constable, R.T., 2013. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. Neuroimage 82, 403–415.
Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., WU-Minn HCP Consortium, 2013. The WU-Minn Human Connectome Project: an overview. Neuroimage 80, 62–79.
Waller, L., Walter, H., Kruschwitz, J.D., Reuter, L., Müller, S., Erk, S., Veer, I.M., 2017. Evaluating the replicability, specificity, and generalizability of connectome fingerprints. Neuroimage 158, 371–377.
Yi, C., Zeng, Y., 2018. sparseSVM: Solution Paths of Sparse High-Dimensional Support Vector Machine with Lasso or Elastic-Net Regularization.
Youngstrom, E.A., 2014. A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: we are ready to ROC. J. Pediatr. Psychol. 39, 204–221.