# **Research Paper:** Prediction of Brain Connectivity Map in Resting-State fMRI Data Using Shrinkage Estimator



Atiye Nazari' 💿, Hamid Alavimajd'\* 💿, Nezhat Shakeri' 💿, Mohsen Bakhshandeh' 💿, Elham Faghihzadeh' 💿, Hengameh Marzbani<sup>3, 4</sup> 💿

- 1. Department of Biostatistics, Faculty of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
- 2. Department of Radiology Technology, Faculty of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
- 3. Department of Biomedical Engineering and Medical Physics, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran.
- 4. Neural Engineering Research Center, Noorafshar Hospital, Tehran, Iran.



**Citation** Nazari, A., Alavimajd, H., Shakeri, N., Bakhshandeh, M., Faghihzadeh, E., & Marzbani, H. (2019). Prediction of Brain Connectivity Map in Resting-State fMRI Data Using Shrinkage Estimator. *Basic and Clinical Neuroscience*, 10(2), 147-156. http:// dx.doi.org/10.32598/bcn.9.10.140

doi http://dx.doi.org/10.32598/bcn.9.10.140

# $\odot$ $\odot$

# Article info:

Received: 07 Oct 2017 First Revision: 10 Nov 2017 Accepted: 27 Feb 2018 Available Online: 01 Mar 2019

#### **Keywords:**

Resting-State fMRI, Functional connectivity, Shrinkage estimator, Mean Squared Error, Seed-based correlation analysis

# ABSTRACT

**Introduction:** In recent years, brain functional connectivity studies are extended using the advanced statistical methods. Functional connectivity is identified by synchronous activation in a spatially distinct region of the brain in resting-state functional Magnetic Resonance Imaging (MRI) data. For this purpose there are several methods such as seed-based correlation analysis based on temporal correlation between different Regions of Interests (ROIs) or between brain's voxels of prior seed.

**Methods:** In the current study, test-retest Resting State functional MRI (rs-fMRI) data of 21 healthy subjects were analyzed to predict second replication connectivity map using first replication data. A potential estimator is "raw estimator" that uses the first replication data from each subject to predict the second replication connectivity map of the same subject. The second estimator, "mean estimator" uses the average of all sample subjects' connectivity to estimate the correlation map. Shrinkage estimator is made by shrinking raw estimator towards the average connectivity map of all subjects' first replicate. Prediction performance of the second replication correlation map is evaluated by Mean Squared Error (MSE) criteria.

**Results:** By the employment of seed-based correlation analysis and choosing precentral gyrus as the ROI over 21 subjects in the study, on average MSE for raw, mean and shrinkage estimator were 0.2169, 0.1118, and 0.1103, respectively. Also, percent reduction of MSE for shrinkage and mean estimator in comparison with raw estimator is 49.14 and 48.45, respectively.

**Conclusion:** Shrinkage approach has the positive effect on the prediction of functional connectivity. When data has a large between session variability, prediction of connectivity map can be improved by shrinking towards population mean.

\* Corresponding Author:

Hamid Alavimajd, PhD.

*Address:* Department of Biostatistics, Faculty of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. *Tel:* +98 (21) 22707347 *E-mail:* alavimajd@sbmu.ac.ir

.....

# Highlights

- Resting-state functional MRI scans with large session-to-session variability.
- Shrinkage estimators improve the prediction of each subject-specific functional connectivity maps.
- Using the shrinkage method has the advantage of analyzing big data with numerous variables and low observations.

Shrinkage estimator can make it possible to cluster the brain more reliably because it provides a more reliable estimation of the functional connectivity map.

## Plain Language Summary

High variability of the acquired data and short-term scans in Functional magnetic resonance imaging studies are two significant challenges of data analysis, causing unreliable results. Shrinkage approach, in functional connectivity studies, presents reliable estimators considering the mentioned challenges. Comparing the results of classic estimators to the result of shrinkage estimators, it is found that shrinkage estimators provide more reliable and precise results. Moreover, because of creating a reliable functional connectivity map, any further analysis would be better in terms of reliability.

# 1. Introduction

unctional Magnetic Resonance Imaging (fMRI) is one of the valuable instruments to discover the function of the human brain. The fMRI is a non-invasive method that uses Blood Oxygen Level Dependent (BOLD) contrast mechanism (Daliri & Behroozi, 2012). In recent years, lots of researchers studied the patterns of brain functional connectivity (Behroozi, Daliri, & Boyaci, 2011; Ghaderi et al., 2017; Sadeghi et al., 2017). Functional connectivity focuses on how brain voxels and regions interact and function with each other.Some studiesare conducted on Resting-State (rs)-fMRI data to understand patterns of brainconnectivity and their role in brain diseases and disturbances (Zhang, Guindani, & Vannucci, 2015). In the resting-state imaging, without the stimulus, the subject is requested to lie down in the scanner device and not to move until the end of the imaging time (Daliri & Behroozi, 2013). Low-frequency fluctuations (<0.1 Hz) are observed in resting-state networks (Biswal et al., 1995).

Functional connectivity implies temporal correlation of the BOLD signals between voxels or Regions of Interest (ROIs) (Zhang et al., 2015). There are various analytical methods such as clustering (Cordes et al., 2002), partial correlation (Cribben et al., 2012; Varoquaux et al., 2010), independent component analysis and principal component analysis (Andersen, Gash, & Avison, 1999; Calhoun et al., 2001; McKeown et al., 1998). Some methods explore the dynamic functional connectivity, since healthy brain function may show rich dynamics over the course of time (Borumandnia et al., 2017).

Seed-based correlation analysis is one of the most widely used methods in functional connectivity analysis that is according to the temporal correlation between ROIs or between voxels in an ROI (Biswal et al., 1995; Fox et al., 2005). In new studies, the estimation of restingstate functional connectivity in aseed-based correlation analysis is improved by shrinkage approach (Shou et al., 2014). This estimator is used in scan-rescan rs-fMRI data to predict the functional connectivity by shrinking the subject-specific estimator towards the average connectivity maps of all subjects (Shou et al., 2014).

Due to large variability in subject-specific data, the results of short-term data usually tend to be highly unstable. One way to achieve a reliable estimate for each subject is to increase the time of brain imaging from a standard time 5-10 minutes to 30-60 minutes (Cohen et al., 2008). This gives the analyst more information, but this approach has limitations. A large number of rs-fMRI scans with a 5-10 minute imaging time were collected in the past and contained valuable information that is thus excluded. The impossibility of long-term scans of children, elderly, and sick people, and high cost of the brain imaging are leading obstacles (Mejia et al., 2015). Mejia et al. (2015). evaluated the effect of shrinkage estimators in simulated data with different lengths of time. In terms of reliability, shrinkage estimation of short-term data with 200 time and subject-specific estimation of a longterm data with 1000 time points have the same results

In the current study, shrinkage approach was applied to 21 healthy subjects (7-12 years old; 11 females and 10 males) with two scanning sessions to estimate correlation maps of the second replication using the first replication correlation coefficient. Each scanning session had 74-time points and the current study aimed at investigating the advantage of shrinkage approach to improve the prediction of functional connectivity in very short rs-fMRI scans.

# 2. Methods

Let  $Y_{ij}(v, t)$  denote fMRI time series for each voxel of ROI at time t=1, ..., T, for subject i=1, ..., I, scanning session j=1, ..., J. In the current study it was as follows: I=21, J=2, T=74. Seed time course is defined as:

(1)  
$$Y_{ij}^{s}(t) = \frac{1}{|S|} \sum_{v \in S} Y_{ij}(v, t)$$

, where S and |S| are the collection of voxels and number of voxels in ROI, respectively. Seed-based correlation map is defined as the correlation between  $Y_{ij}(v, t)$  and  $Y^{s}_{ij}(t)$ :

(2)

$$W_{ij}(\mathbf{v}) = \frac{\sum_{t=1}^{T} \{Y_{ij}(\mathbf{v}, t) - \overline{Y}_{ij}(\mathbf{v}, .)\} \{Y_{ij}^{s}(t) - \overline{Y}_{ij}^{s}(.)\}}{\sqrt{\sum_{t=1}^{T} \{Y_{ij}(\mathbf{v}, t) - \overline{Y}_{ij}(\mathbf{v}, .)\}^{2} \sum_{t=1}^{T} \{Y_{ij}^{s}(t) - \overline{Y}_{ij}^{s}(.)\}^{2}}}$$

, where  $\overline{Y}_{ij}$  (v,.) is the average of  $Y_{ij}$  (v, t), and  $Y_{ij}^{s}$  (.) is the average of  $Y_{ij}^{s}$  (t) over time (Shou et al., 2014). Equation (2) shows that the connectivity map is not dependent on time and can be calculated for each subject, each replication, and all voxels in ROI.

The goal was to predict second replication connectivity map,  $W_{i2}$  (v), of each subject using the first scanning session information. The connectivity map from the first replication of each subject can be considered as the estimation of second replication for the same subject,  $\hat{W}_{i2}^{R}(v)=W_{i1}(v)$ , named "raw estimator". The second estimator is "mean estimator" that uses the mean of the first replication connectivity map of all subjects in the study: (3)

$$\hat{W}_{i2}^{M}(\mathbf{v}) = \frac{1}{I} \sum_{i=1}^{I} W_{i1}(\mathbf{v}), i = 1, ..., 21.$$

Therefore, the results are the same for all subjects. The third estimator is "shrinkage estimator" that shrinks raw estimator towards the average of the first replication connectivity map of all subjects (Shou et al., 2014).

For shrinking, Fisher Z transformation was employed as Equation (4) to normalize the correlation values,  $W_{ij}$  (V) (Shou et al., 2014).

(4)  $\hat{V}_{ij} = \frac{1}{2} \log \frac{1 + W_{ij}(v)}{1 - W_{ij}(v)}$ 

Then shrinkage estimator was performed according to Equation (5) on transformed coefficient (Shou et al., 2014).

(5)  $\hat{V}^{SH}(v)$ :

$$\hat{V}_{i2}^{SH}(v) = \lambda(v) * \hat{V}_{2}^{M}(v) + \{1 - \lambda(v)\} * \hat{V}_{i2}^{R}(v)$$

, where:

(6)

$$\hat{V}_{i2}^{R}(\mathbf{v}) = V_{i1}(\mathbf{v})$$
$$\hat{V}_{2}^{M}(\mathbf{v}) = \frac{1}{I} \sum_{i=1}^{I} V_{i1}(\mathbf{v})$$

and  $\lambda$  (v) are shrinkage parameters that can take a value in [0, 1] (James & Stein, 1961; Shou et al., 2014).

Finally, to achieve the original scale of correlation, inverse Fisher Z transformation was applied as Equation (7) to the shrinkage estimator (Shou et al., 2014).

(7)  
$$\hat{W}_{ij}^{SH}(v) = \tanh(\hat{V}_{ij}^{SH}(v))$$

If  $\lambda$  (v)=1, then raw estimator  $\hat{V}_{i2}^{R}$  (V) is completely unreliable and shrinkage estimator is reduced to the mean estimator. If  $\lambda$  (v)=0, the raw estimator is completely reliable and shrinkage estimator is equal to raw estimator, and no shrinkage occurs towards the average correlation map. Data are not usually reliable: hence, shrinking is a good option to improve prediction. Shou et al. (2013) reported  $\lambda$  (v)= $\lambda$ =0.1 based on their data for shrinkage parameters of all voxels since this value is close to the average reliability of voxels. The optimized value of  $\lambda$ 

(v) can be calculated based on the replication data. The  $\lambda$  (v) is considered as reliability or Intra-class Correlation Coefficient (ICC) for connectivity map of each voxel (Shou et al., 2014).

To estimate  $\lambda$  (v), classical measurement error model for replication study is defined as:

(8)

$$V_{ij}(v) = X_i(v) + U_{ij}(v)$$

, where  $X_i(v)$  is true unobserved correlation coefficient for subject i, and  $U_{ij}(v)$  is the measurement error for subject i, replication j (Carroll et al., 2006).

It was assumed  $X_i$  (.) and  $U_{ij}$  (.) were uncorrelated with E ( $X_i$  (v))= $\mu_x$  (v) and E ( $U_{ij}$  (v))=0, therefore, the Best Linear Unbiased Estimator (BLUE) of  $X_i$  (v) was defined as Equation (9) (Shou et al., 2014):

(9)  

$$\hat{\mathbf{V}}_{i2}^{SH}(\mathbf{v}) = \hat{X}_{i}(\mathbf{v}) = \lambda(\mathbf{v}) * \hat{V}_{2}^{M}(\mathbf{v}) + \{1 - \lambda(\mathbf{v})\} * \hat{V}_{i2}^{R}(\mathbf{v})$$

, where:

(10)

$$\lambda(v) = \frac{Var\{X_i(v)\}}{Var\{X_i(v)\} + Var\{U_{i1}(v)\}}$$

Variance of  $X_i(v)$  and  $U_{ij}(v)$  can be estimated based on the data as:

(11)

$$\begin{split} D_{i}(v) &= V_{i2}(v) - V_{i1}(v) \\ Var\{U_{ij}(v)\} &= \frac{1}{2} Var_{i}\{D_{i}(v)\} = \frac{1}{2(I-1)} \sum_{i=1}^{I} \{D_{i}(v) - \overline{D}(v)\}^{2} \\ Var\{V_{ij}(v)\} &= \frac{1}{J(J-1)} \sum_{j=1}^{J} \sum_{i=1}^{I} \{V_{ij}(v) - V_{j}(v)\} \\ Var\{X_{i}(v)\} &= Var\{V_{ij}(v)\} - Var\{U_{ij}(v)\} \end{split}$$

, where  $\overline{D}(v)$  and  $\overline{V}_i(v)$  are the average of  $D_i(v)$ and  $V_{ij}(v)$ , respectively. Estimating  $\lambda(v)$  and putting in Equation (4), shrinkage estimator is resulted (Mejia et al., 2015).

The performance of the three estimators is evaluated using Mean Square Error (MSE) criterion (Lehmann & Casella, 2006). The small value of MSE shows that the estimated value is close to true value. Therefore, the estimator with the lowest MSE among the others is the best one.

12)  
$$MSE = E \left( \hat{W}_{i,2}(\mathbf{v}) - \hat{W}_{i,2}(\mathbf{v}) \right)^2$$

(

 $\hat{W}_{i2}(v)$  is estimated as the connectivity map by each of the three methods, and  $W_{i2}(v)$  is the true value of obtained correlation in the second replication (Haman & Valenta, 2013).

The brain images used in the current study were related to a global competition called "ADHD-200 Preprocessed"<sup>1</sup> in 2011, where its pre-processed source is free available. A variety of ways, including Athena pipeline<sup>2</sup> are used to process these images. This method is the combination of processes performed by imaging software of FSL and AFNI (Bellec et al., 2017). Among this dataset, 21 healthy subjects between 7 to 12 years old and with 2 scanning sessions were randomly selected.

For pre-processing, structural images of skull bone tissue and background are removed from images, and White Matter (WM), Gray Matter (GM), and Cerebrospinal Fluid (CSF) regions are segmented. The binary mask is made up of white and gray matter regions. Skull-off images are mapped using a linear transformation to the Montreal Neurological Institute space (MNI). The resolution is changed to  $1 \times 1 \times 1$  mm, and the made masks are applied and the brightness of the images is normalized. Gaussian filter with 6 mm FWHM was used to smooth the images (Bellec et al., 2017).

Functional preprocessing steps were performed (e.g. removing the first four volumes, slice timing correction, motion correction, removing nuisance variance), data were also transformed to MNI space at 4×4×4 mm<sup>3</sup> resolution, temporally filtered using 0.009-0.08 pass-band filter, and spatially smoothed (6-mm kernel) (Bellec et al., 2017).

The current study employed seed-based correlation analysis method to evaluate the estimation of functional connectivity of the second replication connectivity map. In this method, the precentral gyrus was selected as ROI. For the first time, Biswal et al. (1995) examined the analysis of brain function in this ROI. The mask of the precentral regions were prepared using the WFU PickAtlas toolbox in MATLAB R2014a software (Maldjian, Lau-

<sup>1.</sup> http://www.nitrc.org/frs/?group\_id=383

<sup>2.</sup> http://www.nitrc.org/plugins/mwiki/index.php/neurobureau: AthenaPipeline

rienti, & Burdette, 2004; Maldjian et al., 2003). All brain images were mapped to type 2 Eve Atlas in the SPM12 (Oishi et al., 2009). Other calculations were performed in MATLAB R2014a software.

# 3. Results

Seed-based correlation map of the second replication was estimated by raw, mean, and shrinkage methods for each voxel of the precentral region and for the 21 subjects. Fig**ure 1** shows the distribution of correlation coefficients for each subject. Red points are the outliers. Mean estimation of connectivity map was the same for all subjects; hence, the corresponding boxplot for all subjects had the same pattern (Figure 1 b). As an example, the results of the subject 18 were considered. For this subject, the raw estimator predicted that the correlation coefficients were more scattered (Figure 1 b) than the values obtained from the second replicate of this subject (Figure 1 d).



#### **NEUR**SCIENCE

Figure 1. Distribution of correlation coefficients for each subject

Boxplot of raw correlation estimates (a); mean correlation estimates (b); shrinkage correlation estimates (c); and the correlation coefficient of second replication data (d); Boxplots show the distribution of correlation coefficients of voxels of the precentral region for each subject. Red points are the outlier.



#### NEURSSCIENCE

Figure 2. Connectivity maps of the specific subject with similar connectivity map to the average of all subjects using different methods

First row: Raw estimator, Second row: Mean estimator, Third row: Shrinkage estimator, Fourth row: True correlation coefficient of the second replication, Fifth row: The mask of the precentral gyrus; The white and yellow colors indicate the positive correlation and the orange and red colors represent a negative correlation. Each column shows the results in a different axial slice.



#### NEURSSCIENCE

Figure 3. Connectivity maps in sagittal, coronal, and axial views of the specific subject with similar connectivity map to the average of all subjects using different methods

(a) Raw estimator; (b) Mean estimator; (c) Shrinkage estimator; (d) Correlation coefficient of second replication; and (e) The mask of the precentral gyrus; The white and yellow colors indicate the positive correlation and the orange and red colors represent a negative correlation.



NEURSSCIENCE

**Figure 4.** The boxplots of MSE for 3 different prediction methods; raw, mean, and shrinkage correlation estimates

The shrinkage estimator showed a similar pattern along the box and the range of correlation coefficients (Figure 1 c). All three methods can predict negative skew in the distribution of truecorrelation coefficients. Figure 2 displays the correlation map in eight different axial slices for a particular subject. Connectivity map of this subject was almost similar to the average of the connectivity map of all subjects. The white and yellow colors indicate the positive correlation and the orange and red colors represent a negative correlation. The last row is the mask of the precentral region and the voxels of this region are displayed in blue (Figure 2 e). The colors showed that the shrinkage estimator (Figure 2 c) had closer correlation patterns to the map of the connection derived from the second repetition (Figure 2 d).

Figure 3 also shows the results of the specific subject in certain spatial coordinates in the three views. In other views, the shrinkage estimator (Figure 3 c) shows the nearest correlation map to the true values (Figure 3 d). Using the Mean Squares Error (MSE) criterion, the performance of the three estimators to predict the second replication connectivity map was evaluated. Table 1 reports MSE from these three methods to each sub-

**Table 1.** Average of MSEs over all voxels for raw, mean, and shrinkage estimators for each subject and the reduction in MSE for mean and shrinkage estimators in comparison with the raw estimator

MSEMSERed.%MSERed.%10.21830.134238.500.133638.3620.19290.098149.150.097549.4330.13130.085434.950.085135.1340.25200.112755.250.111455.7850.23130.113750.800.112551.0360.23170.163929.230.162929.6670.14400.098533.450.095533.6880.20520.108247.240.108147.3090.21490.124142.220.123742.43100.17250.103340.060.103040.26110.41630.157262.220.157062.28120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	Subject	Raw	Mean		Shrinkage	
10.21830.134238.500.133638.3620.19290.098149.150.097549.4330.13130.085434.950.085135.1340.25200.112755.250.111455.7850.23130.113750.800.112551.0360.23170.163929.230.162929.6670.14400.098533.450.095533.6880.20520.108247.240.108147.3090.21490.124142.220.123742.43100.17250.103340.060.103040.26110.41630.157262.220.157062.28120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99		MSE	MSE	Red. %	MSE	Red. %
20.19290.098149.150.097549.4330.13130.085434.950.085135.1340.25200.112755.250.111455.7850.23130.113750.800.112551.0360.23170.163929.230.162929.6670.14400.098533.450.095533.6880.20520.108247.240.108147.3090.21490.124142.220.123742.43100.17250.103340.060.103040.26110.41630.157262.220.157062.28120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	1	0.2183	0.1342	38.50	0.1336	38.36
30.13130.085434.950.085135.1340.25200.112755.250.111455.7850.23130.113750.800.112551.0360.23170.163929.230.162929.6670.14400.098533.450.095533.6880.20520.108247.240.108147.3090.21490.124142.220.123742.43100.17250.103340.060.103040.26110.41630.157262.220.157062.28120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	2	0.1929	0.0981	49.15	0.0975	49.43
40.25200.112755.250.111455.7850.23130.113750.800.112551.0360.23170.163929.230.162929.6670.14400.098533.450.095533.6880.20520.108247.240.108147.3090.21490.124142.220.123742.43100.17250.103340.060.103040.26110.41630.157262.220.157062.28120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	3	0.1313	0.0854	34.95	0.0851	35.13
50.23130.113750.800.112551.0360.23170.163929.230.162929.6670.14400.098533.450.095533.6880.20520.108247.240.108147.3090.21490.124142.220.123742.43100.17250.103340.060.103040.26110.41630.157262.220.157062.28120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	4	0.2520	0.1127	55.25	0.1114	55.78
60.23170.163929.230.162929.6670.14400.098533.450.095533.6880.20520.108247.240.108147.3090.21490.124142.220.123742.43100.17250.103340.060.103040.26110.41630.157262.220.157062.28120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	5	0.2313	0.1137	50.80	0.1125	51.03
70.14400.098533.450.095533.6880.20520.108247.240.108147.3090.21490.124142.220.123742.43100.17250.103340.060.103040.26110.41630.157262.220.157062.28120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	6	0.2317	0.1639	29.23	0.1629	29.66
80.20520.108247.240.108147.3090.21490.124142.220.123742.43100.17250.103340.060.103040.26110.41630.157262.220.157062.28120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	7	0.1440	0.0985	33.45	0.0955	33.68
90.21490.124142.220.123742.43100.17250.103340.060.103040.26110.41630.157262.220.157062.28120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	8	0.2052	0.1082	47.24	0.1081	47.30
100.17250.103340.060.103040.26110.41630.157262.220.157062.28120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	9	0.2149	0.1241	42.22	0.1237	42.43
110.41630.157262.220.157062.28120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	10	0.1725	0.1033	40.06	0.1030	40.26
120.19610.087655.340.087355.49130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	11	0.4163	0.1572	62.22	0.1570	62.28
130.21990.141135.830.140636.03140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	12	0.1961	0.0876	55.34	0.0873	55.49
140.26250.156638.020.155338.52150.10350.064138.070.064038.11160.28580.063077.930.062977.99	13	0.2199	0.1411	35.83	0.1406	36.03
15     0.1035     0.0641     38.07     0.0640     38.11       16     0.2858     0.0630     77.93     0.0629     77.99	14	0.2625	0.1566	38.02	0.1553	38.52
16 0.2858 0.0630 77.93 0.0629 77.99	15	0.1035	0.0641	38.07	0.0640	38.11
	16	0.2858	0.0630	77.93	0.0629	77.99
170.13170.115712.110.115112.60	17	0.1317	0.1157	12.11	0.1151	12.60
18         0.3384         0.1032         69.48         0.1022         69.79	18	0.3384	0.1032	69.48	0.1022	69.79
19         0.1861         0.0753         59.49         0.0751         59.62	19	0.1861	0.0753	59.49	0.0751	59.62
20 0.2052 0.1372 33.13 0.1366 33.42	20	0.2052	0.1372	33.13	0.1366	33.42
21 0.2025 0.1074 46.93 0.1070 47.16	21	0.2025	0.1074	46.93	0.1070	47.16
Average         0.2169         0.1118         48.45         0.1103         49.14	Average	0.2169	0.1118	48.45	0.1103	49.14

NEURSSCIENCE

ject. For all subjects, the MSE of mean estimator was lower than the raw estimator and the MSE of shrinkage estimator was the lowest. The minimum and maximum MSE of raw estimator were 0.1035 and 0.4163, respectively. By the shrinkage method, MSE decreased to a minimum of 0.0629 and a maximum of 0.1629. The average MSE of raw, mean, and shrinkage estimators was 0.2169, 0.1118, and 0.1103, respectively. Also, the decrease of MSE for mean and shrinkage estimator in comparison with MSE of the raw estimator was positive for all people indicating a better prediction of the true correlation coefficients of the second repetition using the mean and shrinkage estimators. On average, the reduction in MSE was 48.45% for the mean estimator and 49.14% for shrinkage estimator.

Figure 4 shows the MSE boxplot of all three methods. The raw estimator had larger MSE values and, with regard to box length, a more dispersed distribution. The MSE values for the mean and shrinkage estimators were almost proportional to the amount and dispersion and less than those of the raw estimator. The median MSE of raw, mean, and shrinkage estimators was 0.2053, 0.1083, and 0.1082, respectively. Due to the fact that the median line was not in the middle of the box, the distribution of MSE values for the three estimators was skew to the right. There were two outliers of 0.4163 and 0.3384 in the distribution of MSE of the raw estimator, which were respectively related to the eleventh and eighteenth subjects.

## 4. Discussion

The current study employed a shrinkage approach to improve the estimation of functional connectivity in the seed-based correlation analysis. The test-retest rs-fMRI data with 74 time points in 21 healthy subjects were applied. The precentral gyrus was selected as an ROI. On Fisher Z transformation correlation coefficient, the average MSE for raw, mean, and shrinkage estimators was 0.2169, 0.1118, and 0.1103, respectively. The prediction performance improved 49.14% and 48.45% by shrinkage and mean estimators in comparison with that of the raw estimator. The prediction of the functional connectivity of all subjects improved.

Shou et al. (2014) also performed this procedure in the 20-subject rs-fMRI test-retest data with 210 time points. Applying shrinkage and mean estimators on fisher transformed coefficients reduced MSE by 30% and 25%, respectively. The shrinkage estimator improved the prediction of 18 out of 20 subjects.

Mejia et al. (2015) used the shrinkage approach to estimate the similarity matrix. The implementation of the shrinkage estimator on the Fisher-transformed and original correlation coefficients improved the reliability of the results 29% and 26%, respectively.

The number of people in the study does not change the results of the raw estimator, since this estimation does not depend on the information of other sample subjects, but the performance of shrinkage estimator is influenced by the number of people in the study. By simulating the different sample sizes, the shrinkage estimator provides the best result with 20 subjects and no significant improvement in prediction is achieved with increasing the subjects (Mejia et al., 2015).

A shrinkage approach is used to improve the results of many classical estimators. In the current study investigated the advantage of analyzing big data with many variables and low observations. When rs-fMRI scans are very short with large session-to-session variability, the estimation of each subject-specific connectivity maps can be improved by shrinking towards the average of other subjects. Also, the correlation matrix of voxels used as a similarity matrix of clustering methods can be estimated by shrinkage method. Due to the fact that the shrinkage estimator provides reliable estimates of the correlation between the time series of the voxels, its application makes it possible to cluster the brain more reliably.

# **Ethical Considerations**

# Compliance with ethical guidelines

There is no ethical guidelines; Because the data has been downloaded.

## Funding

The Department of Biostatistics of Shahid Beheshti University of Medical Sciences financially supported the study (Grant No.: 12413).

## **Authors contributions**

Conceptualization: Atiye Nazari, Hamid Alavimajd, Nezhat Shakeri, Mohsen Bakhshandeh; Methodology: Atiye Nazari, Hamid Alavimajd, Nezhat Shakeri, Mohsen Bakhshandeh, Elham Faghihzadeh; Investigation: Atiye Nazari, Hamid Alavimajd, Nezhat Shakeri, Mohsen Bakhshandeh; Writing-original draft: Atiye Nazari, Hamid Alavimajd; Writing-review & editing: Atiye Nazari, Hamid Alavimajd, Nezhat Shakeri; Funding acquisition: Hamid Alavimajd; Supervision: Hamid Alavimajd; Software: Atiye Nazari, Hengameh Marzbani; and Visualization: Atiye Nazari, Hengameh Marzbani.

#### Conflict of interest

The authors declared no conflict of interest.

# References

- Andersen, A. H., Gash, D. M., & Avison, M. J. (1999). Principal component analysis of the dynamic response measured by fMRI: A generalized linear systems framework. *Magnetic Resonance Imaging*, 17(6), 795-815. [DOI:10.1016/S0730-725X (99)00028-4]
- Behroozi, M., Daliri, M. R., & Boyaci, H. (2011). Statistical analysis methods for the fMRI data. *Basic and Clinical Neuroscience*, 2(4), 67-74. [PMID] [PMCID]
- Bellec, P., Chu, C., Chouinard-Decorte, F., Benhajali, Y., Margulies, D. S., & Craddock, R. C. (2017). The neuro bureau ADHD-200 preprocessed repository. *NeuroImage*, 144(Part B), 275-86. [DOI:10.1016/j.neuroimage.2016.06.034]
- Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine*, 34(4), 537-41. [DOI:10.1002/mrm.1910340409]
- Borumandnia, N., Majd, H. A., Zayeri, F., Baghestani, A. R., Tabatabaee, M., & Faeghi, F. (2017). Human brain functional connectivity in Resting-State fMRI data across the range of weeks. *Middle East Journal of Family Medicine*, 7(10), 148. [DOI:10.5742/ MEWFM.2017.93068]
- Calhoun, V. D., Adali, T., Pearlson, G. D., & Pekar, J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping*, 14(3), 140-51. [DOI:10.1002/hbm.1048]
- Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. (2006). Measurement error in nonlinear models: A modern Perspective. Boca Raton, Florida: CRC Press. [DOI:10.1201/9781420010138]
- Cohen, A. L., Fair, D. A., Dosenbach, N. U., Miezin, F. M., Dierker, D., Van Essen, D. C., et al. (2008). Defining functional areas in individual human brains using resting functional connectivity MRI. *NeuroImage*, 41(1), 45-57. [DOI:10.1016/j.neuroimage.2008.01.066] [PMCID]
- Cordes, D., Haughton, V., Carew, J. D., Arfanakis, K., & Maravilla, K. (2002). Hierarchical clustering to measure connectivity in fMRI resting-state data. *Magnetic Resonance Imaging*, 20(4), 305-17. [DOI:10.1016/S0730-725X (02)00503-9]
- Cribben, I., Haraldsdottir, R., Atlas, L. Y., Wager, T. D., & Lindquist, M. A. (2012). Dynamic connectivity regression: determining state-related changes in brain connectivity. *NeuroImage*, 61(4), 907-20. [DOI:10.1016/j.neuroimage.2012.03.070] [PMCID]
- Daliri, M., & Behroozi, M. (2012). fMRI: Clinical and research applications. OMICS Journal of Radiology, 1, e112. [DOI:10.4172/2167-7964.1000e112]

- Daliri, M., & Behroozi, M. (2013). Advantages and disadvantages of Resting State functional connectivity Magnetic Resonance Imaging for clinical applications. OMICS Journal of Radiology, 3, e123. [DOI:10.4172/2167-7964.1000e123]
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27), 9673-8. [DOI:10.1073/pnas.0504136102] [PM-CID]
- Ghaderi, A. H., Nazari, M. A., Shahrokhi, H., & Darooneh, A. H. (2017). Functional brain connectivity differences between different ADHD presentations: Impaired functional segregation in ADHD-combined presentation but not in ADHD-inattentive presentation. *Basic and Clinical Neuroscience*, 8(4), 267-78. [DOI:10.18869/nirp.bcn.8.4.267] [PMCID]
- Haman, J., & Valenta, Z. (2013). Shrinkage approach for gene expression data analysis. *European Journal of Biomedical Informatics*, 9(3), 2-8.
- James, W., & Stein, C. (1961). Estimation with quadratic loss. Paper presented at the 4<sup>th</sup> Berkeley symposium on mathematical statistics and probability. Berkeley, California, United States, June 20-30 July 1960.
- Lehmann, E. L., & Casella, G. (2006). Theory of point estimation. Berlin: Springer Science & Business Media.
- Maldjian, J. A., Laurienti, P. J., & Burdette, J. H. (2004). Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *NeuroImage*, 21(1), 450-5. [DOI:10.1016/j.neuroimage.2003.09.032]
- Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19(3), 1233-9. [DOI:10.1016/S1053-8119 (03)00169-1]
- McKeown, M. J., Makeig, S., Brown, G. G., Jung, T. P., Kindermann, S. S., Bell, A. J., et al. (1998). Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*, 6(3), 160-88. [DOI:10.1002/ (sici)1097-0193 (1998)6:3<160::aid-hbm5>3.3.co;2-r]
- Mejia, A. F., Nebel, M. B., Shou, H., Crainiceanu, C. M., Pekar, J. J., Mostofsky, S., et al. (2015). Improving reliability of subjectlevel Resting-State fMRI parcellation with shrinkage estimators. *NeuroImage*, 112, 14-29. [DOI:10.1016/j.neuroimage.2015.02.042] [PMCID]
- Oishi, K., Faria, A., Jiang, H., Li, X., Akhter, K., Zhang, J., et al. (2009). Atlas-based whole brain white matter analysis using large deformation diffeomorphic metric mapping: Application to normal elderly and Alzheimer's disease participants. *NeuroImage*, 46(2), 486-99. [DOI:10.1016/j.neuroimage.2009.01.002]
- Sadeghi, A. Z., Jafari, A. H., Oghabian, M. A., Salighehrad, H. R., Batouli, S. A. H., Raminfard, S., etal. (2017). Changes in effective connectivity network patterns in drug abusers, treated with different methods. *Basic and Clinical Neuroscience*, 8(4), 285-98. [DOI:10.18869/nirp.bcn.8.4.285]
- Shou, H., Eloyan, A., Lee, S., Zipunnikov, V., Crainiceanu, A., Nebel, M., et al. (2013). Quantifying the reliability of image replication studies: The image intra-class correlation coefficient (I2C2). Cognitive, Affective & Behavioral Neuroscience, 13(4), 714-24. [DOI:10.3758/s13415-013-0196-0]

- Shou, H., Eloyan, A., Nebel, M. B., Mejia, A., Pekar, J. J., Mostofsky, S., ET AL. (2014). Shrinkage prediction of seed-voxel brain connectivity using resting state fMRI. *NeuroImage*, 102(2), 938-44. [DOI:10.1016/j.neuroimage.2014.05.043]
- Varoquaux, G., Gramfort, A., Poline, J.-B., & Thirion, B. (2010). Brain covariance selection: better individual functional connectivity models using population prior. Paper presented at the Advances in Neural Information Processing Systems, Vancouver, Canada, 6 December 2010.
- Zhang, L., Guindani, M., & Vannucci, M. (2015). Bayesian models for fMRI data analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 7(1), 21-41. [DOI:10.1002/wics.1339]