Original Article:

Reliability of functional connectivity in resting-state functional MRI

Atiye Nazari [1], Hamid Alavimajd [2,]*, Nezhat Shakeri [1], Mohsen Bakhshandeh [2]

[1]Department of Biostatistics, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[2]Department of Radiology Technology, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

*Corresponding author: email address: alavimajd@sbmu.ac.ir (H. Alavimajd)

ABSTRACT

Functional MRI is a noninvasive method in brain imaging. Localization, classification, prediction and connectivity are the most common issues. Functional connectivity is a branch of fMRI that focuses on connectivity between voxels and ROIs. There are several methods for investigating functional connectivity such as correlation analysis. In any field, it is very important that results of any research have reliability according to the experiment. Any methods and measurement instruments need to be reliable. Without reliability, results are meaningless and our research is not trustworthy. Brain imaging can be used as a valuable tool for pre-surgical planning, so the results should be highly reproducible. Test-retest reliability can be explored using the intra-class correlation coefficient (ICC). I2C2 is an extent of ICC to verify the reliability in high-dimensional data as imaging studies. 13 subjects of test-retest resting-state fMRI are used to investigate reliability. I2C2 of four ROIs are also computed (Caudate, Cingulate, Cuneus and Precentral regions). Functional connectivity is found to have moderate reliability ranging 0.6244 to 0.6941. 95% confidence interval of I2C2 is calculated by nonparametric bootstrap in which CI of Caudate region I2C2 has the shortest length.

Keywords: Reliability; fMRI; Bootstrap; Correlation; Classical measurement error

INTRODUCTION

Functional MRI is a noninvasive neuroimaging method with the high spatial resolution to diseases diagnosis and treatment [1]. fMRI investigates the function of the brain using Blood Oxygen Level Dependent (BOLD) signal. The differences between the oxyhemoglobin and deoxyhemoglobin make BOLD signals [2]. BOLD signals were analyzed in the different types of understanding. Detecting activated region during a specific task and stimulus is one of the most common fMRI studies [3]. Predicting brain activities and classifying healthy and patient people is another important issue [4]. Recently, understanding brain connectivity has received increased interests [5, 6]. In all kinds of studies, it is very important to verify if the results of researchers are reliable or not. Reliability is the key to create a method and instrument. Reliability in an experiment means the results should have the least variability. In this situation, a research is regarded as trustworthy and can lead to true conclusions [7]. Recently, reliability of brain imaging methods has gained more interest. This interesting issue focuses on instruments and methods in brain studies. For example, reproducibility of Functional MR imaging results has been studied by 1.5-T MR systems from different companies. Reliability of value was compared by Mann-Whitney Test [8]. Reproducibility of MRI measurement in a longitudinal study was investigated by ICC method [9]. In the test-retest experiment on stroke patients and healthy volunteers, the reliability of results was studied by ROC curves [10]. Reliability of fMRI data was evaluated by I2C2. I2C2 is a kind of reliability method that generalizes ICC to brain imaging and easy to calculate even in imaging studies with high-dimensional data [11].

MATERIALS AND METHODS

For computing the reliability of functional connectivity, the study with two replication was focused on. $Y_{ij}(\mathrm{v},\mathrm{t})$ is defined as BOLD time series for subject i=1,…, I, scanning session j=1,…, J at time t, t=1,…,T and voxel v, v=1,..,V. So, Pearson correlation is used for calculating the functional connectivity of each subject and session for each voxel-pairs [5].

$$W_{ij}(\mathrm{v},\mathrm{v}') = \frac{\sum_{t=1}^{T}\{\mathrm{Y}_{ij}(\mathrm{v},\mathrm{t}) - \bar{\mathrm{Y}}_{ij}(\mathrm{v},.)\}\{\mathrm{Y}_{ij}(\mathrm{v}',\mathrm{t}) - \bar{\mathrm{Y}}_{ij}(\mathrm{v}',.)\}}{\sqrt{\sum_{t=1}^{T}\{\mathrm{Y}_{ij}(\mathrm{v},\mathrm{t}) - \bar{\mathrm{Y}}_{ij}(\mathrm{v},.)\}^2 \sum_{t=1}^{T}\{\mathrm{Y}_{ij}(\mathrm{v}',\mathrm{t}) - \bar{\mathrm{Y}}_{ij}(\mathrm{v}',.)\}^2}}$$

$\bar{\mathrm{Y}}_{ij}(\mathrm{v},.)$ and $\bar{\mathrm{Y}}_{ij}(\mathrm{v}',.)$ are average of time series for pair-voxel v and v' over time. In classical measurement error model, it is assumed that $X_i(\mathrm{v},\mathrm{v}')$ is unobserved correlation coefficient. Also $W_{ij}(\mathrm{v},\mathrm{v}')$ denote as observed value of $X_i(\mathrm{v},\mathrm{v}')$ from each replication and $U_{ij}(\mathrm{v},\mathrm{v}')$ is measurement error [12].

$$W_{ij}(\mathrm{v},\mathrm{v}') = X_i(\mathrm{v},\mathrm{v}') + U_{ij}(\mathrm{v},\mathrm{v}')$$

We assume $X_i(\mathrm{v},\mathrm{v}')$ are independent between subjects and $U_{ij}(\mathrm{v},\mathrm{v}')$ are independent between subjects and replications. Also $X_i(\mathrm{v},\mathrm{v}')$ and $U_{ij}(\mathrm{v},\mathrm{v}')$ are mutually independent. The .variance of $W_{ij}(\mathrm{v},\mathrm{v}')$ is called by between-subject variance and variance of $U_{ij}(\mathrm{v},\mathrm{v}')$ by within-subject variance. In this paper, The I2C2 describes the reliability of correlation measurement over time. Image intra-class correlation coefficient is defined as below [11]:

$$ICC = \frac{\sigma^2_{between} - \sigma^2_{within}}{\sigma^2_{between} + \sigma^2_{within}}$$

Based on replication data, the variances calculated as following [11]:

$$\sigma^2_{within} = \frac{1}{\sum_{i=1}^{I}(J_i - 1)} \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{v\in V}\{\mathrm{W}_{ij}(\mathrm{v},\mathrm{v}') - \bar{\mathrm{W}}_{i.}(\mathrm{v},\mathrm{v}')\}^2$$

$$\sigma^2_{between} = \frac{1}{\sum_{i=1}^{I} J_i - 1} \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{v\in V}\{\mathrm{W}_{ij}(\mathrm{v},\mathrm{v}') - \bar{\mathrm{W}}_{..}(\mathrm{v},\mathrm{v}')\}^2$$

Within-subject variance, $\bar{\mathrm{W}}_{i.}(\mathrm{v},\mathrm{v}')$ is the average of correlation coefficient of each pair-voxels over replication. Also, $\bar{\mathrm{W}}_{..}(\mathrm{v},\mathrm{v}')$ is the average of correlation coefficient of each pair-voxels over subjects and replication in between-subject variance.The range of ICC value is -1 to 1 and has been classified as poor (<0.5), moderate (0.5 to 0.75), good (0.75 to 0.9) and excellent (>0.9) [13]. According to I2C2 formula, the value of I2C2 tends to be 1 in less within-subject variability and shows excellent reliability.This paper investigated the reliability of functional connectivity in resting-state fMRI data using the I2C2 method. To illustrate this point, a resting state data is analyzed with two scanning session from 13 healthy subjects (7 males, 6 females) who participated in "ADHD-200 Global Competition" as the control subject [14]. Correlation map would be calculated for voxels of four different ROIs (Caudate, Cingulate, Cuneus, and Precentral) in all subjects and two replications. Structural and functional preprocessing of "ADHD-200" data have been done by The Athena Pipeline [14]. This data source is available in www.intrc.org. Using WFU_pickatlas toolbox, Mask of specific ROIs is prepared and is co-registered by SPM [15, 16]. All other processes have been performed in MATLAB software. To generate a 95% confidence interval for I2C2, nonparametric bootstrap is applied. The sample size for each resample is equal to the sample size of the original data set. The same process is repeated for each sample and finally I2C2 is calculated for each of them [13].

RESULTS

Bootstrap sampling were done with 100, 250 replication. Fig.1a shows the results for 100 samples and Fig.2b displays the results of 250 samples. Estimated I2C2 of all four regions that have approximately normal distribution are shown in figure 1.

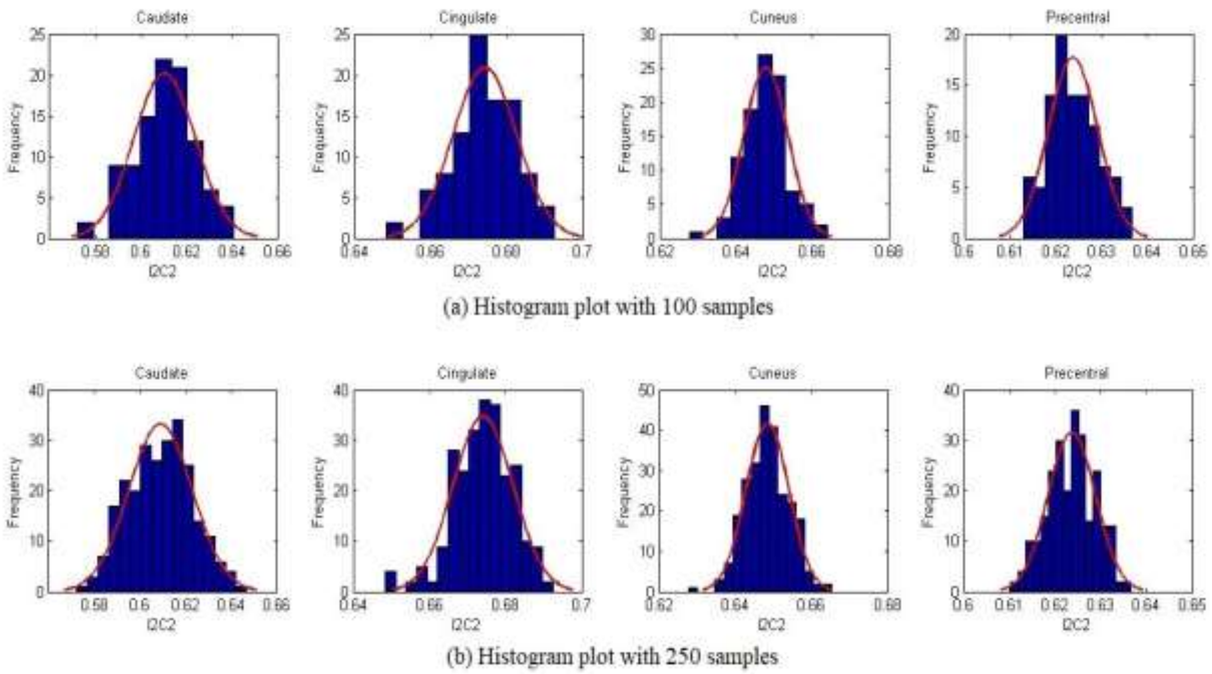

Figure 1. Histogram plots of bootstrap sampling for Caudate, Cingulate, Cuneus and Precentral ROIs.

I2C2 was calculated for four different RIOs. I2C2 values are 0.6244 for Caudate, 0.6941 for Cingulate, 0.6631 for Cuneus and 0.6388 for Precentral. According to normality assumption for estimated I2C2 distribution, 95% confidence interval for true I2C2 of each ROI had been reported in Table.1. According to Table.1, confidence interval with 100 samples is (0.6218, 0.6270) for caudate, (0.6925, 0.6957) for Cingulate, for (0.6620, 0.6642) Cuneus and (0.6377, 0.6399) for Precentral. The results of 250 samples have narrower confidence interval than 100 samples for all regions. Confidence interval is (0.6227, 0.6261) for caudate, (0.6931, 0.6951) for Cingulate, (0.6624, 0.6638) for Cuneus and (0.6381, 0.6395) for Precentral.

Table 1. True ICC and 95% bootstrap confidence interval with 100 and 250 resampling for Caudate, Cingulate, Cuneus and Precentral ROIs.

ROI	True Value	No of bootstrap	Lower 95% CI	Upper 95% CI
Caudate	0.6244	N=100	0.6218	0.6270
		N=250	0.6227	0.6261
Cingulate	0.6941	N=100	0.6925	0.6957
		N=250	0.6931	0.6951
Cuneus	0.6631	N=100	0.6620	0.6642
		N=250	0.6624	0.6638
Precentral	0.6388	N=100	0.6377	0.6399
		N=250	0.6381	0.6395

DISCUSION

In the present study, extension of the intra-class correlation coefficient (I2C2) is used to evaluate the reliability of functional connectivity. In rs-fMRI with two replications of 13 healthy subjects, it was found that the reliability of results is at the moderate level . Calculated I2C2 for Caudate is (0.6244), for Cingulate (0.6941), for Cuneus (0.6631) and for Precentral is (0.6388); consequently Caudate and Cingulate had the least and most reliability. Shou et al applied the I2C2 to investigate the reliability of connectivity map in test-retest rs-fMRI of 20 healthy subjects. In four regions, I2C2 values were calculated with ranging approximately 0.20 to 0.37. The results showed that correlation maps had large variability between sessions and estimated values are not reliable [5]. In comparison with this study's results, it can be claimed that the more participating the cause, the more the variability. As a result, in this situation reliability has an inverse relation with the number of subjects in a study. In a language imaging study with two replications in different days, the reliability of results was studied. There were 21 subjects (12 healthy participants, 9 stroke patients) in the experiment with the same language imaging. Using area under receiver operating characteristic (ROC) curves, Chen and Small found out that the healthy subjects are more reliable than stroke patients [10].

Also, Manoach et al. worked on the reliability of results in patients with schizophrenia. The results showed that overall reliability in healthy participants is higher than in participants with brain injury [17]. Generally speaking, studies of healthy people are more reliable than patient subjects.

CONCLUSION

Reliability and reproducibility are needed to prove the results of any kind of researches as correct and trustworthy and image studies are not an exception. Using I2C2 with an extension of intra class correlation coefficient, the reliability of functional connectivity is investigated in functional MRI data. The method has been used for scanning with several sessions, different numbers of participants, healthy or patient subjects, patients with different diseases, etc. All these situations can affect the results of imaging study, so care should be exercised about the scenario determined for increasing the reliability. fMRI data have large variability in replication of the experiment, so the reliability of results is an effective factor to reach the true conclusion.

ACKNOWLEDGMENTS

This project was financially supported by grant No.12413 from the Department of Biostatistics, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences. This paper was extracted from the MSc thesis of the first author.

"The authors declare no conflict of interest"

REFERENCES

1.Lazar N. The statistical analysis of functional MRI data: Springer Science & Business Media; 2008.

2.Zhang L, Guindani M, Vannucci M. Bayesian Models for fMRI Data Analysis. Wiley interdisciplinary reviews Computational statistics. 2015;7(1):21-41.

3.Poldrack RA, Mumford JA, Nichols TE. Handbook of functional MRI data analysis: Cambridge University Press; 2011.

4.Cheng W, Ji X, Zhang J, Feng J. Individual classification of ADHD patients by integrating multiscale neuroimaging markers and advanced pattern recognition techniques. Frontiers in systems neuroscience. 2012;6.

5.Shou H, Eloyan A, Nebel MB, Mejia A, Pekar JJ, Mostofsky S, et al. Shrinkage prediction of seed-voxel brain connectivity using resting state fMRI. NeuroImage. 2014;102:938-44.

6.Borumandnia N, Majd HA, Zayeri F, Baghestani AR, Tabatabaee M, Faeghi F. Human brain functional connectivity in resting-state fMRI data across the range of weeks. MIDDLE EAST JOURNAL OF FAMILY MEDICINE.7(10):148.

7.Bennett CM, Miller MB. How reliable are the results from functional magnetic resonance imaging? Annals of the New York Academy of Sciences. 2010;1191(1):133-55.

8.Vlieger E-J, Lavini C, Majoie CB, Den Heeten GJ. Reproducibility of functional MR imaging results using two different MR systems. American journal of neuroradiology. 2003;24(4):652-7.

9.McGuire SA, Wijtenburg SA, Sherman PM, Rowland LM, Ryan M, Sladky JH, et al. Reproducibility of quantitative structural and physiological MRI measurements. Brain and behavior. 2017;7(9).
10.Chen EE, Small SL. Test–retest reliability in fMRI of language: group and task effects. Brain and language. 2007;102(2):176-85.
11.Shou H, Eloyan A, Lee S, Zipunnikov V, Crainiceanu A, Nebel M, et al. Quantifying the reliability of image replication studies: The image intra-class correlation coefficient (I2C2). Cognitive, affective & behavioral neuroscience. 2013;13(4):714.
12.Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. Measurement error in nonlinear models: a modern perspective: CRC press; 2006.
13.Wagstaff DA, Elek E, Kulis S, Marsiglia F. Using a nonparametric bootstrap to obtain a confidence interval for Pearson'sr with cluster randomized data: a case study. The journal of primary prevention. 2009;30(5):497-512.
14.Bellec P, Chu C, Chouinard-Decorte F, Benhajali Y, Margulies DS, Craddock RC. The Neuro Bureau ADHD-200 Preprocessed repository. Neuroimage. 2017;144:275-86.
15.Maldjian JA, Laurienti PJ, Burdette JH. Precentral gyrus discrepancy in electronic versions of the Talairach atlas. Neuroimage. 2004;21(1):450-5.
16.Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage. 2003;19(3):1233-9.
17.Manoach DS, Halpern EF, Kramer TS, Chang Y, Goff DC, Rauch SL, et al. Test-retest reliability of a functional MRI working memory paradigm in normal and schizophrenic subjects. American Journal of Psychiatry. 2001;158(6):955-8.