Assessments of variability in cortical and subcortical measurements and within-network connectivity of the brain using test-retest data*

Hosna Tavakoli, Reza Rostami, Mohammad-Reza Nazem-Zadeh

Abstract*— The evaluation and diagnosis of structural changes in brain caused by disease or treatment over time has become one of the important applications of medical imaging methods, in particular MRI, and it is growing. It is critical to evaluate the reliability of the changes in measurements observed in an individual patient for any clinical decision-making. In this paper, we calculated the repeatability coefficient (RC) as a measure of uncertainty for MRI measurements of subcortical volumes and cortical thickness, and within-network connectivity using test-retest data of 20 healthy subjects. We also evaluated changes in 13 patients who received 20 sessions of transcranial magnetic stimulation as a treatment. The most reliable measure seems to be in the thickness of the left occipital with RC% of 3.5 and the least reliable measure is the brain connectivity within visual network using Yeo's atlas with RC% of 29.4. The most sensitive measure to the percentage of true changes in treated patients is the connectivity within subcortical network of AAL with 76.9%.

Clinical Relevance— **The results of this study can be useful for evaluating changes in the gray matter structures or functional connectivity of the brain due to a neurological disease such as Alzheimer's or Parkinson's. Also, the obtained results can be used to evaluate the changes caused by any intervention or treatment that may have any positive or negative effect on the brain.**

I. INTRODUCTION

Magnetic resonance imaging (MRI) has been increasingly investigated as a biomarker for the detection of structural and functional changes in the brain [1, 2]. The validity of the observed changes in the MRI measurement at individual level is critical for the clinical decision-making systems and processes. This implies that the changes must be beyond an estimated level of uncertainty in identified variables in imaging, image analysis, or natural biological, physical, and physiological occurrences so that they can be considered true changes.

Uncertainty in the MRI measures has been investigated in various studies using different metrics. Previous studies have dealt with the concept of uncertainty and the potential applications of uncertainty analysis in MRI. The uncertainty analysis of longitudinal imaging has been used to assess the validity of measurements for the assessment of radiation-induced neurotoxicity in patients with low-grade or benign tumors undergoing partial brain radiation therapy [3]. Another application to be cited is the estimation of the uncertainty of interhemispheric change to identify the side of epilepsy in patients with temporal lobe epilepsy in a cross-sectional study [4].

Analysis of changes in medical imaging measurements at individual level is important in the context of individualized treatments. Whether a patient has a true change cannot be determined from the average change observed in a group of patients. Even if a group of patients has a statistically significant mean change, some people in the group may not have real changes. To evaluate an individual change, it is necessary to determine how much a change can be considered as a real change. [5].

Test-retest studies in MRI tried to address the reliability and reproducibility of this method in order to prove the benefit of using MRI. It is shown that T1- and DTI-derived tissue metrics including cortical and subcortical gray matter and white matter volume exhibited insignificant mean differences both across relocation and within-site repeat [6]. Some studies demonstrate the reliability of task-based fMRI to illustrate the strength of this modality [7, 8].

In longitudinal imaging studies, this range of uncertainty is defined as the repeatability coefficient (RC). Several RC studies have estimated diffusion indices. The wide range of RC indicates that many factors can affect RC, including imaging acquisition, data preprocessing, segmentation methods, and characteristics of the structures being studied. Estimated RC is necessary to determine how reliable a longitudinal individual change is for potential use in clinical decision-making setups.

In this study, we estimated RC for three MRI measures of different brain regions and networks to determine which measure is more reliable. We also evaluate the changes in these MRI measures in 13 patients who went under 20 sessions of transcranial magnetic stimulation (TMS) treatment by comparing them to RC ranges for each measure.

II. MATERIAL AND METHOD

A. *Subjects and image acquisition*

20 healthy subjects (10 female and 10 male, age = 31.4 ± 8.3) were recruited and underwent two sessions of MRI scanning in one day. The time interval between test and retest scans was 30 minutes. Each session was consisting of structural and resting-state functional MRI (rs-fMRI) conducted on a 3T MRI system (Prisma, Siemens, Erlangen, Germany). Each session included a T1-weighted image with following protocol: TR=1.9 s, TE=2.26 ms, FOV =250 mm, Matrix =256 × 256, Sagittal plane, Slice thickness=1 mm, Resolution= 1 x 1 x 1 mm, 176 slices. The resting-state scan lasted 396 s with the following parameters: TR=1.2 s, TE=2 ms, FOV =250 mm, Matrix =256 × 256, Sagittal plane, Slice thickness=1 mm, Resolution= 3 x 3 x 3 mm, 330 slices. We asked participants to close their eyes during the resting state.The ethic approval was provided by Iran University of Medical Science, Tehran, Iran.

---

*Research supported by The Tehran university of Medical sciences.

Hosna Tavakoli is with the Computational and Artificial Intelligence Department, Institute of Cognitive Science Studies, Tehran, Iran (e-mail: h.tavakoli868@gmail.com). Reza Rostami is with the Department of Psychology, University of Tehran, Iran (e-mail: rrostami@ut.ac.ir).

Mohammad Reza Nazem-Zadeh (the corresponding author) is with Research Center for Molecular and Cellular Imaging, Advanced Medical Technologies and Equipment Institute (AMTEI), Tehran University of Medical Sciences, Tehran, Iran (e-mail: mnazemza@tums.ac.ir). He is also with the Department of Neuroscience, Monash University, Melbourne, VIC, Australia (e-mail: Mohamad.Nazem-Zadeh@monash.edu).

B. *MRI measures*

Structural MRI measures included the subcortical volume (mm^3) of thalamus, caudate, putamen, pallidum, hippocampus, and amygdala; and the cortical thickness (mm) of fusiform, inferior parietal lobule (IPL), occipital, frontal pole, temporal pole, and insula for both right and left hemispheres, as the most important and widely used brain regions in MRI studies, calculated using Freesurfer (v6.0.0) software.

Another measure studied was the functional connectivity within the brain networks obtained for 6 functional networks based on the yeo article and also 7 networks based on the AAL atlas. Time series were extracted from resting-state fMRI data after preprocessing which included slice timing correction, reorienting, brain extraction, and registration. Each network included a certain number of nodes connected. The weight of the connection between two nodes was determined by the correlation between nodes' time series. A connectivity matrix was estimated for each brain network with dimension n by n, where n is the number of nodes in that network. The within-network connectivity was then calculated by averaging the correlation values in this matrix. All steps were performed using DPARSF advanced edition (version 5.1).

C. *Repeatability coefficient*

To determine the level of uncertainty, we estimated the RC values for each MRI measure as follows:

Let I_{ik} be the index observed value for the i^{th} subject and k^{th} replication, $i = 1, 2, \dots, n$, $k = 1, 2, \dots, K$ (in our test and retest dataset, $n = 20$, $k = 2$) as:

$$I_{ik}/\mu_i = 1 + \varepsilon_{ik} \qquad (1)$$

which relates I_{ik} to its true value μ_i for each subject through a residual relative error ε_{ik} with the within-subject variance $\sigma_w^2 = var\ (\varepsilon_{ik})$ in a normalized ANOVA model. The within- and between-subject means of squares (WMS and BMS) with χ^2 distributions of n (K − 1) and n-1 degrees of freedom are:

$$WMS = \frac{1}{n\,(k-1)} \sum_{i=1}^{n} \sum_{k=1}^{K} [\frac{I_{ik} - \bar{I}_t}{\bar{I}_t}]^2 \qquad (2)$$

$$BMS = \frac{K}{n-1} \sum_{i=1}^{n} [\frac{\bar{I}_t - \bar{I}}{\bar{I}}]^2 \qquad (3)$$

where $\bar{I}_t$ is the mean over replications for i^{th} subject, and $\bar{I}$ is the mean over all observations. The within-subject standard deviation can be estimated by $\hat{\sigma}_w^2 = WMS$. Rewriting the Equation (2) for K = 2:

$$\hat{\sigma}_w^2 = 1/n \sum_{i=1}^{n} 1/2\ [\frac{I_{it} - I_{ir}}{(I_{it} + I_{ir})/2}]^2 \qquad (4)$$

where t and r denote test and retest, respectively. The RC is given by $RC = 2.77\ \sigma_w$, which defines the 95% Confidence Interval (CI) of the normalized measurements to determine whether a change in an individual patient is a true change. The 95% Confidence Interval (CI) of the estimated RC is given by:

$$RC_L = 2.77 \sqrt{\frac{n.WMS}{\chi_n^2(0.975)}}$$

$$, RC_U = 2.77 \sqrt{\frac{n.WMS}{\chi_n^2(0.025)}}\ , \widehat{RC}\ \epsilon (RC_L, RC_U) \qquad (5)$$

Assuming there is no change in a structure between test and retest due to disease progression or the treatment, any change has to be due to random and/or systematic errors that could have originated from the variabilities caused by imaging device, image acquisition, patient re-positioning, image processing and analysis, and/or subject-specific natural biological, physical, and physiological variations [5].

D. *True individual longitudinal changes*

We estimated whether the change from pre to post treatment in each MRI measure in each structure and network at the individual level was a true change. Patients underwent MRI sessions before and after TMS (20 sessions, 6 weeks). We calculated a percentage change ($\Delta I_t\%$) in subcortical volume, cortical thickness, and within-network connectivity of from baseline to post-TMS scan. For more reliable results, we consider the interval (-RC_L, RC_L) to evaluate the changes. Three scenarios may occur: First, the calculated changes for each measurement fall within this interval. In this case, it is confident to say that no change has been observed in this measure due to the treatment. The second scenario is when the $\Delta I_t\%$ is above the range considered a positive change. In the third scenario, the change is below the interval, which indicates negative changes in this measurement due to treatment.

III. RESULTS

A. *RC ranges of MRI measures*

Using test-retest data from 20 healthy subjects, we estimated the uncertainty in three MRI measures. In subcortical volumes, the smallest RC% belonged to the caudate with 8% and the largest RC% is 17.5% belonged to the left amygdala. For cortical thickness, region with the largest RC% was left temporal pole with 16.8% and the smallest RC% is 3.5% for left occipital. RC% in within-network connectivity for AAL networks was between 12.5% to 27.1%, where the lowest value in RC% was for visual network and the highest one was for sensorimotor network. In Yeo networks, the RC% values were between (9.7%, 29.4%) which the minimum and maximum estimated RC% respectively belonging to frontoparietal and visual networks.

B. *Evaluation of longitudinal percentage changes in MRI measures in individual patients*

Tables 1 and 2 show the percentage of the patients who had changed beyond the uncertainty ranges in subcortical volume, cortical thickness, and within-network connectivity among all structures and networks compared to $(-RC_L, RC_L)$. The number of patients with changes out of uncertainty ranges (above or below), is considered as patients with true changes. Figure 1 and 2 is showing the $\Delta I_t\%$ of each patient for subcortical volumes and cortical thickness. The smallest percentage was observed in the volume of right hippocampus and in within connectivity of default-mode and ventral attention of Yeo networks with 0%. The largest number of true changes was 10 out of 13 patients observed in within-connectivity of subcortical network.

IV. DISCUSSION

To determine the quality of a biomarker, the repeatability and reliability are often examined in the first steps. In the last decade, MRI has become a common biomarker to investigate physiological changes and treatment-induced changes in psychiatric and neurobiological diseases [9, 10]. Meanwhile, less attention has been paid to its reliability and repeatability.

Test-retest studies in recent years have addressed this issue to an acceptable extent. However, there is still a lack of evidence in this field, especially since with the advancement of technology and improvement of analysis methods; different aspects of MRI have been used. In this study, we determined the uncertainty for three of the most widely used MRI measures, i.e., subcortical volume, cortical thickness, and within-network connectivity in brain networks by estimating the repeatability coefficient.

TABLE I. PERCENTAGE OF TRUE CHANGES IN STRUCTURAL MRI MEASURES

Subcortical regions	% true changes in volumes	**Cortical regions**	% true changes in thickness
L-Thalamus	23.1	**L-Fusiform**	23.1
L-Caudate	15.4	**L- inferior parietal lobule**	30.8
L-Putamen	38.5	**L-Occiptal**	38.5
L-Pallidum	30.8	**L-Frontalpole**	7.7
L-Hippocampus	7.7	**L-Temporalpole**	7.7
L-Amygdala	7.7	**L-Insula**	23.1
R-Thalamus	7.7	**R-Fusiform**	7.7
R-Caudate	23.1	**R-IPL**	15.4
R-Putamen	15.4	**R-Occiptal**	23.1
R-Pallidum	15.4	**R-Frontalpole**	30.8
R-Hippocampus	0	**R-Temporalpole**	23.1
R-Amygdala	23.1	**R-Insula**	38.5

L-: Left R-: Right

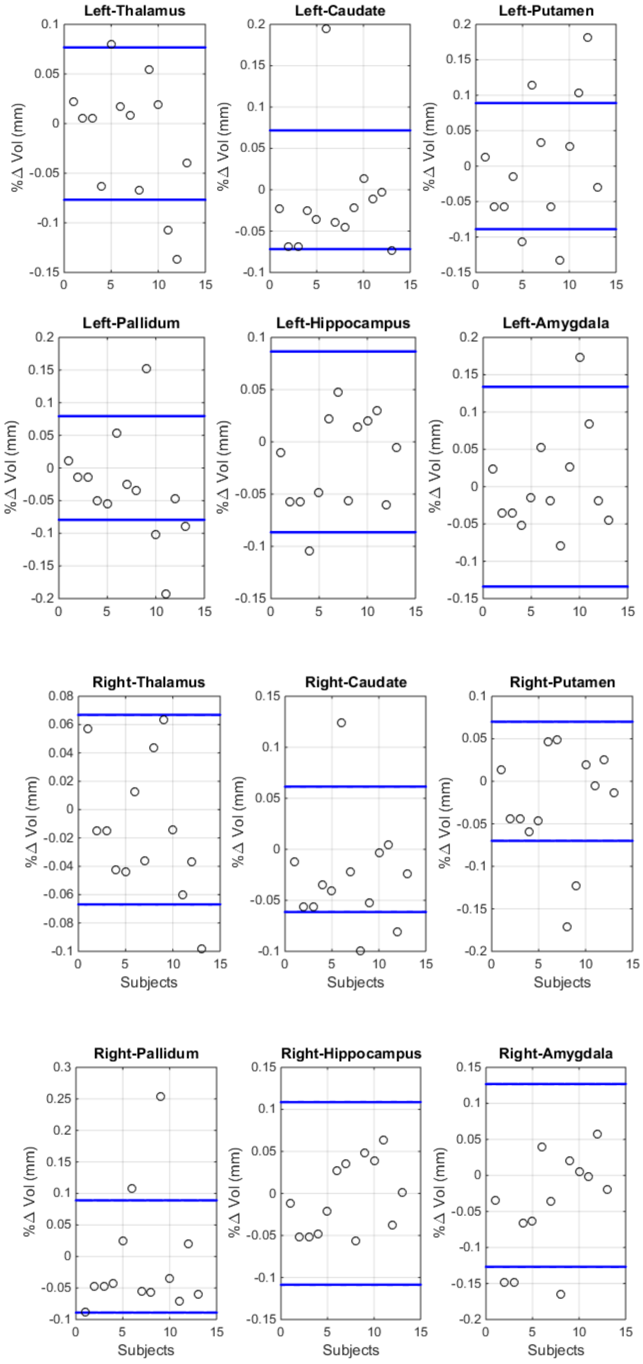

Figure 1. Percentage of subcortical volume changes for 13 patients (circles) in the RC interval (blue lines).

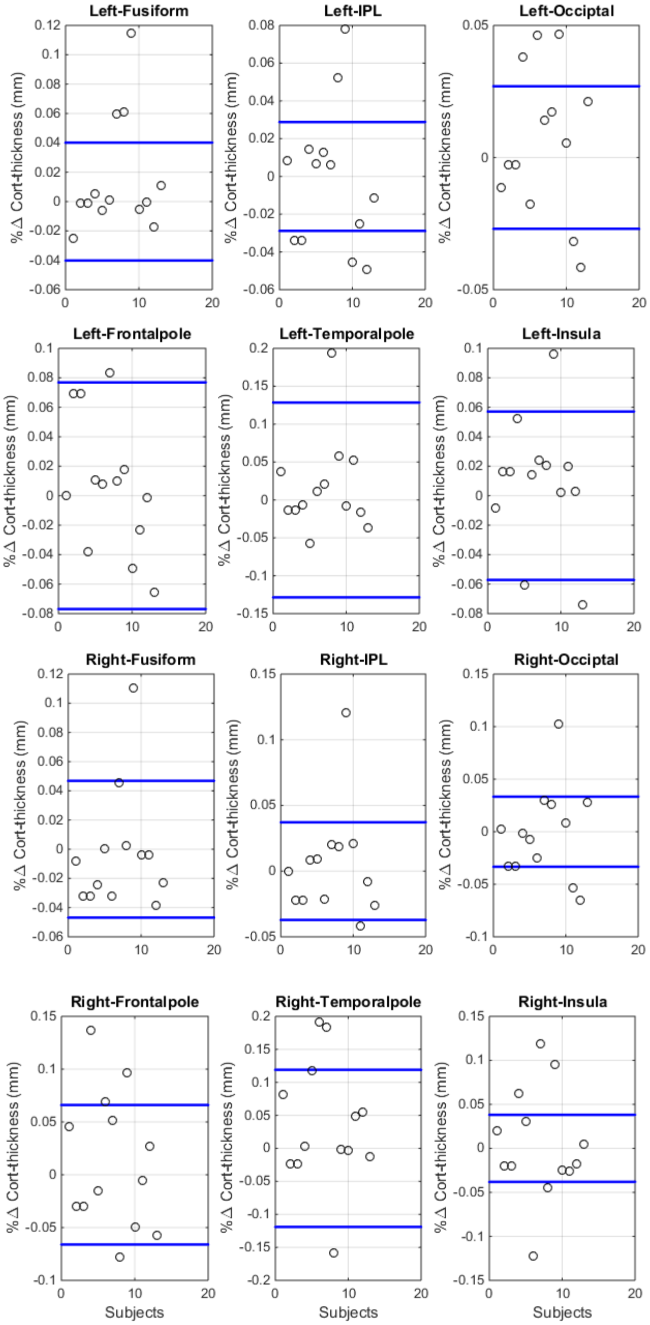

Figure 2. Percentage of cortical thickness changes for 13 patients (circles) in the RC interval (blue lines).

V. CONCLUSION

The results obtained above can be used to determine whether a change in brain structure or networks in an individual is a real change or not. Also, repeating the steps performed in this study for other MRI criteria will be a step forward in this field.

TABLE II. PERCENTAGE OF TRUE CHANGES IN NETWORK MEASURES

Yeo Networks	% true changes in within connectivity	**AAL Networks**	% true changes in within connectivity
Default-mode	0	**Default-mode**	7.7
Visual	7.7	**Auditory**	61.5
Dorsal attention	7.7	**Visual**	61.5
Fronto-parietal	15.4	**Attention**	38.5
Sensorimotor	15.4	**Sensorimotor**	53.8
Limbic	7.7	**Subcortical**	76.9
Ventral attention	0		

REFERENCES

1. Starmans, N.L., et al., *Twenty-four hour blood pressure variability and the prevalence and the progression of cerebral white matter hyperintensities.* Journal of Cerebral Blood Flow & Metabolism, 2023. **0**(0): p. 0271678X221149937.
2. Moraga-Amaro, R., et al., *Real-world data of non-invasive stimulation of the human insula-prefrontal cortices using deep TMS to treat anxiety for occupational stress and generalized anxiety disorder.* Psychiatry Research, 2023. **320**: p. 115036.
3. Nazem-Zadeh, M.R., et al., *Uncertainty in assessment of radiation-induced diffusion index changes in individual patients.* Phys Med Biol, 2013. **58**(12): p. 4277-96.
4. Nazem-Zadeh, M.-R., et al., *Lateralization of temporal lobe epilepsy using a novel uncertainty analysis of MR diffusion in hippocampus, cingulum, and fornix, and hippocampal volume and FLAIR intensity.* Journal of the Neurological Sciences, 2014. **342**(1-2): p. 152-161.
5. Barnhart, H.X. and D.P. Barboriak, *Applications of the repeatability of quantitative imaging biomarkers: a review of statistical analysis of repeat data sets.* Transl Oncol, 2009. **2**(4): p. 231-5.
6. Melzer, T.R., et al., *Test-retest reliability and sample size estimates after MRI scanner relocation.* Neuroimage, 2020. **211**: p. 116608.
7. Wuthrich, F., et al., *Test-retest reliability of a finger-tapping fMRI task in a healthy population.* Eur J Neurosci, 2023. **57**(1): p. 78-90.
8. Ibinson, J.W., et al., *Comparison of test-retest reliability of BOLD and pCASL fMRI in a two-center study.* BMC Med Imaging, 2022. **22**(1): p. 62.
9. Zhang, Y., et al., *Short-term Medication Effects on Brain Functional Activity and Network Architecture in First-Episode psychosis: a longitudinal fMRI study*. Brain Imaging and Behavior, 2023.
10. Tang, R., et al., *Longitudinal association of executive function and structural network controllability in the aging brain.* GeroScience, 2023. **45**(2): p. 837-849.