The Effect of General Anesthesia On Test-Retest Reliability of Resting-State FMRI Metrics And Optimization of Scan Length

Faezeh Vedaei (Faezeh.Vedaei@jefferson.edu)
Thomas Jefferson University
Mahdi Alizadeh
Thomas Jefferson University
Victor M Romo
Thomas Jefferson University
Feroze B. Mohamed
Thomas Jefferson University
Chengyuan Wu
Thomas Jefferson University

Research Article

Keywords: resting-state, fMRI, anesthesia, test-retest reliability, scan length

Posted Date: October 26th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-992129/v1

Abstract

Resting-state functional magnetic resonance imaging (rs-fMRI) has been known as a powerful tool in neuroscience. However, exploring the test-retest reliability of the metrics derived from rs-fMRI BOLD signal is essential particularly in the studies of patients with neurological development. Two factors affecting reliability of rs-fMRI measurements including the effect of anesthesia and scan length have been estimated in this study. A total of 9 patients with drug-resistant epilepsy (DRE) of requiring interstitial thermal therapy (LITT) were scanned in two states of awake and under anesthesia. At each state, two rs-fMRI sessions were obtained that each one lasted 15 minutes, and the effect of scan length was evaluated. Voxel-wise rs-fMRI metrics including amplitude of low fractional frequency fluctuation (ALFF), amplitude of low fractional frequency fluctuation (fALFF), functional connectivity (FC), and regional homogeneity (ReHo) were measured. Intraclass correlation coefficient (ICC) was calculated to estimate the reliability between two sessions of scanning for both states. Overall, our finding revealed that reliability improves under anesthesia as well as by increasing the scanning length of the scanning sessions. Furthermore, we showed that the optimal scan length to achieve reliable rs-fMRI measurements was 3.1 – 7.5 minutes shorter in an anesthetized, compared to wakeful state.

Introduction

Resting-state functional Magnetic Resonance Imaging (rs-fMRI) estimates blood-oxygen-level-dependent (BOLD) signal fluctuations at low frequency (< 0.1 Hz) correspond to synchronized variations in spontaneous neuronal activity in the resting brain. The ability to measure brain functional connectivity noninvasively and with no instructed task has invaluable utility in clinical application and neuroscience research in animal and human populations. However, functional connectivity is not static, and the connectivity strengths may alter during a single session or between sessions resting-state scanning. Thus, despite its potential, rs-fMRI suffers from significant variabilities involved in the BOLD signal that may cause difficulties in replication of results across studies [1,2]. Several factors can contribute to the intra-subject variabilities and affect the reliability of estimated brain functional connectivity including head motion, physiological noise such as cardiac and respiratory effects, MRI acquisition parameters, and data analysis/standardization strategies [3–8].

General anesthesia has been used in animal and human studies to reduce the potential of motion artifacts and fluctuations in behavior in order to generate a consistent mental state and increase control of certain aspects of a subject’s physiology [9,10] . At the same time, it has been shown that general anesthesia including intravenous and volatile anesthetics suppress neuronal activity and affect fMRI BOLD response in both task-based and resting-state fMRI studies in a dose-dependent manner [11–16]. Studies performed primarily in animals and humans have examined the effect of intravenous and volatile anesthetic agents on BOLD signal variability and demonstrated that anesthetics reduce the overall variations of intrinsic BOLD fluctuations and dynamical complexity [14,17–19]. Hence, we speculate that

BOLD signal and resultant brain functional connectivity measurement under this condition will be more consistent and reliable within-subjects and between-subjects over repeating experiments.

Additionally, scan duration has been known as another influential factor that affects the stability of rs-fMRI BOLD response. Recently, several studies have answered this question regarding how much data is needed to be acquired to achieve stable and reliable resting-state BOLD response. They estimated optimal resting-state scan duration particularly using graph theory measurements [20,2120,21], and network-based functional connectivity [22,23]. Statistical voxel-wise test-retest reliability has been used in recent rs-fMRI studies to estimate the reproducibility of measured brain functional connectivity due to its ability to explicitly model measurement's variance. However, fewer studies have employed it on regional spontaneous brain activity measures including Absolute and fractional amplitude of low frequency fluctuation (ALFF and fALFF) and regional homogeneity (ReHo) in the same cohort [8,24,25].

While prior studies investigated either the effect of general anesthesia or scan length in rs-fMRI in separate experiments, our study is the first human study combining these two influential factors to estimate reliability of rs-fMRI measurements. The present study examined the impact of general anesthesia on test-retest reliability of voxel-wise rs-fMRI metrics including ALFF, fALFF, ReHo, and functional connectivity. Secondarily, we investigated the effect of resting-state scan length on reliability of the measurements. We aimed to determine standardized scan lengths of rs-fMRI which ensure robust and reliable BOLD responses in both conditions of wakefulness and under anesthesia.

Methods

-Participants

Candidates for this study included patients who presented to the Comprehensive Epilepsy Center at Thomas Jefferson University Hospitals with drug-resistant epilepsy (DRE) who have been deemed appropriate candidates for laser interstitial thermal therapy (LITT). These patients were selected because as part of their routine clinical course they required an awake MRI for preoperative surgical planning, followed by an MRI under general anesthesia during the LITT procedure, with approximately two weeks in-between. All patients underwent the standard informed consent process before being including in the study. A total of 9 patients (4 males and 5 females, age 28-60 years old) were enrolled. Each patient completed two rs-fMRI sequences during each scanning session (~ 15 min apart) (Figure 1). During the awake rs-fRMI scan, patients were instructed to keep eyes open and fix to a point with no motion during the scanning to avoid any motion artifact over respiratory challenges. To minimize head movement, straps and foam pads were used to fix the head comfortably during the scanning. The study was approved by the institutional review board (IRB) of Thomas Jefferson University Hospital.

-Data acquisition

Both MRI sessions were performed on a 3.0T Achieva Phillips scanner with an eight-channel head coil. fMRI images were acquired axially using a single-shot echo planar imaging (EPI) sequence in the same anatomical location prescribed for T1-weighted images. The T1-weighted imaging parameters used were: FOV=24.0cm, voxel size=1.0×1.0×1.0mm^3, matrix size=352×352, TR=7.5ms, TE=3.4ms and slice thickness=1mm. Functional MR imaging parameters were FOV=23.0cm, voxel size=3.5×3.5×3.5 mm^3, matrix size= 128×128, TR=2s, TE=25ms and number of averages=1. Patients were instructed to keep eyes open and fix to a point with no motion during the scanning to avoid any motion artifact over respiratory challenges.

-Anesthesia administration

The second MRI scan was acquired during the LITT procedure and under general anesthesia. All patients were evaluated by an anesthesiologist and underwent institutional standard of pre-anesthetic preparation. Every patient received a standard induction of intravenous propofol (130-300 mg) 15-20 minutes before scanning. After endotracheal intubation, sevoflurane anesthesia was administered through the endotracheal tube and maintained with 0.6-1.2 mean alveolar concentration (MAC) and 100% Fractional Inspiratory Oxygen (FIO2) during MRI acquisition. Mean arterial blood pressure was also maintained at 65-75 mmHg and end tidal carbon dioxide ($ETCO_2$) at 30-35 mmHg throughout the procedure. After MRI acquisition and the LITT procedure were completed, patients were reversed with Sugamadex 2-4 mg/kg, extubated as per routine, and were observed in a post-anesthetic care unit during their recovery.

-Data analysis

All rs-fMRI data were preprocessed using Data Processing & Analysis for Resting-State Brain Imaging (DPABI, V5.1_201201; http://rfmri.org/dpabi) [46] based on Statistical Parametric Mapping (SPM12; http://www.fil.ion. ucl.ac.uk/spm) running on MATLAB R2020b (The Math Works, Inc., Natick, MA, USA). The pre-processing steps are listed as follow: the first 10 volumes were discarded to get the steady state MRI signal and adaptation of participants in the scanning environment. The remaining volumes were corrected for slice timing and head motion using six rigid body motion parameters. Next, for each individual T1-weighted structural image was co-registered to the mean of the realigned EPI images. Then, all resting-state data were spatially normalized to the EPI template in Montreal Neurological Institute (MNI) space with a resampling voxel size of 3×3×3 mm. Due to the sensitivity of rs-fMRI measurements to micro head motions, the Friston 24-parameter model (the 24 parameters including 6 head motion parameters, 6 head motion parameters of the previous scan, and the 12 corresponding squared items) was applied to regress out the head motion effects from the realigned data [36]. Further, signal from white matter and cerebrospinal fluid were regressed out and filtered with a temporal band-pass of 0.01–0.08 Hz to reduce the effects of low-frequency drifts and high-frequency respiratory and cardiac noise.

All data processing steps were restricted within gray matter, for which a gray matter mask was created. Statistical parametric mapping 12 (https://www.fil.ion.ucl.ac.uk/spm/) was used to conduct segmentation of gray matter. Then, the probabilistic maps were binarized using *fslmaths* tools (cutoff = 0.2) and were averaged to generate the gray matter mask.

-mALFF/mfALFF calculation

The amplitude of low frequency fluctuation (ALFF) measures the intensity of regional spontaneous brain activity at rest; and the fractional ALFF (fALFF) is felt to have higher sensitivity and specificity [29] with less inclusion of artifacts from vascular signals. For each subject, before ALFF/fALFF calculation, spatial smoothing (Gaussian kernel of full-width half maximum, FWHM = 6 mm) was performed. The time courses of rs-fMRI signal were converted to the frequency range using a Fast Fourier Transform (FFT), and the square root of the power spectrum was measured and averaged across the 0.01-0.08 Hz domain. Voxel-wise fALFF was measured as the ratio of power in low-frequency band (0.01-0.08 Hz) to power of the entire frequency range (0-0.55 Hz). For standardization purpose, for each subject ALFF/fALFF of each voxel was divided by the global mean of ALFF/fALFF within the gray matter mask to generate the mALFF/mfALFFs maps.

-mReHo calculation

ReHo measurement was performed after band-pass filtering (0.01-0.08 Hz). This is accomplished on a voxel-based basis by calculating Kendall's coefficient of concordance (KCC) for a given time series that is assigned as the center voxel with those of its nearest 27 neighboring voxels (Eq. 1).

(1):

$$w = \frac{\sum(R_i) - n(\overline{R_i})^2}{\frac{1}{12}K^2(n^3 - n)}$$

In this formula *w* is the KCC (range from 0 to1) among given voxels; *K* is the number of neighboring voxels (here, *K* = 27, including the center voxel); $\bar{R}_i$ is the mean rank across nearest neighbors (27 voxels) at the *ith* time point; and *n* is the total number of time points [47,48]. To avoid any artificial coherence due to spatial smoothing, ReHo was calculated before spatial smoothing. For standardization purpose, ReHo value at each voxel was divided by the global mean of ReHo within the gray matter mask to obtain the mReHo maps. Spatial smoothing with an isotropic Gaussian kernel of 6 mm full-width half- maximum (FWHM) was performed after ReHo calculation.

-Voxel-wise functional connectivity (zFC)

For each subject, voxel-wise functional connectivity (FC) was measured by estimating Pearson's correlations between the time series of any pairs of brain voxels which resulted in individual z-functional connectivity matrix within the gray matter mask. Then, for a given voxel i, FC was measured using the equation as follows (Eq. 2):

(2):

$$FC\,(i) = \frac{1}{N_{voxels} - 1} \sum_{j \neq i} Z_{ij}\,, \qquad r_{ij} > r_0$$

where z_{ij} was the Fisher's Z-transformed version of correlation coefficient, r_{ij}, between voxel i and voxel j, and r_0 was a correlation threshold that was used to exclude weak correlations possibly arising from noises (r_0=0.2 in this study). r_{ij} was converted to z_{ij} using Fisher's Z-transformation. N_{voxels} was also defined as total number of voxels within gray matter mask (N_{voxels} = 544833) [49,50].

-Reliability of rs-fMRI metrics

Test-retest reliability has been shown to reflect the stability of a test measure under repeated experiments. While numerous statistical methods have been introduced to assess test-retest reliability such as Pearson's correlation, and Kendall coefficient of concordance, Intraclass Correlation Coefficient (ICC) has been shown to be the primary and most efficient approach in estimating functional connectivity reliability [24,51–53]. In this study, resting-state fMRI reliability of the metrics including of ALFF, fALFF, ReHo, and FC were estimated using ICC for awake and under anesthesia states according to the following equation (Eq. 3) [53]:

(3):

$$ICC = \frac{MS_b - MS_w}{MS_b + (K-1)MS_w}$$

Where, MS_b represents the between-subject mean square, MS_w represents the within-subject mean square at voxel level, and K is the number of sessions. For all the rs-fMRI metrics test-retest reliability was calculated by extracting the average of each metric from the gray matter mask in both state of awake and under anesthesia. It was classified as poor (ICC = 0-0.4), fair (ICC = 0.4-0.55), good (ICC = 0.55-0.75), and excellent (ICC = 0.75-1.0) [25,54]. For the rs-fMRI metrics, voxel-wise ICC maps were generated. Then, we extracted the average and standard deviation (SD) of the ICC measurements within the gray mask for both states of awake and under anesthesia.

-Effect of scan length

To investigate the effect of scan length on the rs-fMRI metrics, for each participant rs-fMRI data was truncated into 15 bins with i^{th} bin containing the first i minutes of data acquisition. Thus, 15 bins of data with a scan length ranging from 1 to 15 minutes were generated for each candidate for each state of awake and under anesthesia. Next, the same preprocessing steps as those applied on the entire scan mentioned above were conducted for each scan length. Further, voxel-wise rs-fMRI metrics including ALFF, fALFF, ReHo, FC, and related ICC maps were generated for each scan length.

In order to optimize the scan length for each state of awake and under anesthesia, the standardized non-linear logarithmic function curves (y = a + b ln(x)) were fitted to the ICC maps created for the rs-fMRI metrics corresponding to each scan length ranging from 1 to 15 minutes. The standardized scan length was defined as the time point where the measured derivative of the logarithmic function was at least 0.01 ((|dy/dx|) < 0.01) [23].

Results

-Test-retest reliability of rs-fMRI metrics in awake and under anesthesia states

We found increase of reliability under anesthesia state for the rs-fMRI metrics including fALFF, ReHo, and particularly for FC (0.65 vs. 0.83 for awake and under anesthesia respectively). Brain maps of ICC measurements are represented in Figure 3. Aldo, Table 1, and Figure 2 show the detail of the mean/SD values of ICC measurements comparing between awake and under anesthesia states. In addition, higher number of voxels with excellent ICC (0.75 < ICC < 1) were observed under anesthesia in fALFF and ReHo within gray matter mask (94.04% and 24.12% respectively). Whereas number of voxels with excellent ICC were higher in awake state for ALFF and FC metrics (5.17% and 18.20% respectively). Average of ICC for all the rs-fMRI metrics extracted within gray matter mask with excellent ICC also were higher under anesthesia than in awake state (Table 1).

-Optimization of scan length of rs-fMRI in awake and under anesthesia states

We found the best fit of non-linear logarithmic function for the ICC of rs-fMRI metrics in awake and under anesthesia (Figure 4). Next, we measured the time point where the logarithmic functions begin to reach a plateau ((|dy/dx|) < 0.01). The optimal scan lengths estimated in awake state were 18.7 minutes for ALFF, 14.7 minutes for fALFF, 20.3 minutes for FC, and 17.5 minutes for ReHo. We found lower optimal scan lengths under anesthesia for all the rs-fMRI metrics: 11.9 minutes for ALFF, 8.6 minutes for fALFF, 17.2 minutes for FC, and 10 minutes for ReHo (Table 2). Among quantitative maps, fALFF reached to the optimal values in the shortest scan duration. The scatter plots and logarithmic functions were fit to the ICC of rs-fMRI metrics and the related optimal scan lengths were shown in Figure 4.

Furthermore, by extracting the average and SD of rs-fMRI ICCs within the gray matter mask, we found improvement of reliability and decrease of variability as the function of scan length for both states of awake and under anesthesia (Figure 5).

Discussion

The application of general anesthesia during rs-fMRI studies unavoidably interferes with resting-state BOLD signal. Hereupon, understanding confounding effects of anesthetics is of the essence in MRI acquisition, design, and data analysis in the experiments need to get involved with anesthetics. This study provides the comprehensive evaluation of voxel-wise test-retest reliability of rs-fMRI metrics and its relationship with scan length in two states of awake and under volatile gaseous anesthesia. Voxel-wise ALFF, fALFF, ReHo, and FC were chosen to estimate the reliability of resting-state BOLD response since they have been demonstrated as robust measurements and stable to noise interference [25]. Reliable test-retest reliability of single subject metric is crucial to obtain imaging biomarkers that assist in detection and evaluation of developmental changes of neurological diseases [26]. Nevertheless, test-retest reliability has been shown to be regionally variable across the entire brain. Indeed, higher order cortical networks including default mode network, sensory/motor network, visual network, as well as dorsal attention, ventral attention, and frontoparietal control are known to have highest reproducibility [24]. Hence, this experiment was constrained within gray matter area through the entire data processing.

Two main observations were noted. First, our findings showed improvement of reliability of rs-fMRI metrics measurements under anesthesia compared to the awake state that was more prominent for the scans less than 12 minutes. Second, following non-linear logarithmic regression modeled to the rs-fMRI metrics, we showed that optimal scan length is shorter under anesthesia than in awake state. Although previous studies examined the effect of anesthesia and scan length as separate experiments on variability of resting-state BOLD signal to some extent [12,22,23,27,28], none has addressed this question by comparing test-retest reliability of rs-fMRI metrics obtained in awake and under anesthesia states and combining this factor with the impact of scan length.

To that end, we compared ICC maps obtained from rs-fMRI metrics including voxel-wise ALFF, fALFF, ReHo, and FC in awake and anesthetized states. In terms of the effect of anesthesia, our finding indicated that test-retest reliability enhances under anesthesia across the rs-fMRI metrics was more significant for FC; meanwhile, ALFF exhibited the least ICC difference between awake and under anesthesia states. Additionally, among good-to-excellent test-retest reliability of the rs-fMRI measurements, fALFF showed moderate values in both awake and anesthetized states that might be linked to the high sensitivity of this metric to detecting physiological signals compared to ALFF [24,29,30,41] . This finding is consistent with the prior literature reporting that ALFF have higher test-retest reliability than fALFF. As such, it is recommended to use fALFF, or at least combine it with ALFF, in rs-fMRI studies because of the high sensitivity of this measurement to reduce physiological noise [8].

-Effects of Anesthesia on Intracranial Physiology

Our results suggest that neurophysiology characteristic of the anesthetic agents may play a main role in reduction of BOLD signal variability under anesthesia which leads to enhancement of the test-retest reliability of rs-fMRI metrics compare to wakefulness state. It has been shown that both intravenous and inhaled anesthetic agents such as propofol and sevoflurane modulate γ-aminobutyric acid type A ($GABA_A$) receptors, which is the fastest inhibitory neurotransmitter receptor in the central neural system. Specially, propofol one of the potent modulators of $GABA_A$ receptors, enhances the gating of the receptors and thereby reduces neural excitability. The volatile anesthetics including sevoflurane enhances $GABA_A$ receptor function which leads to increasing channel opening and inhibition enhancement at both synaptic and extrasynaptic receptors. by reducing neural excitability [13,32]. Such reduction in the receptive field size under anesthesia has been revealed for somatosensory cortical and sub-cortical neurons which suppresses consciousness through actions that control sleep-wake states [11,16,33].

Our results were qualitatively comparable to the results from previous literature finding that anesthesia reduces temporal variability across midline cortical regions during loss of consciousness that was measured by standard deviation (SD) of BOLD response [14,18,19,34]. A recent animal study comparing BOLD signal variability in awake and anesthetized rats showed reduction of variability across much of the brain under anesthesia. They proposed that variability can be used as a robust signature of consciousness that distinguishes anesthesia-induced unconsciousness from the awake state [17].

Several factors may impact on reliability of resting-state BOLD response, and as such, test-retest reliability can vary fundamentally between datasets ranging from poor-to-excellent ICC [35]. We note that, since head motion during scanning can affect test-retest reliability that was intrinsically restricted under anesthesia, higher-order regression models including Friston 24-parameter model were employed at individual level to minimize the head motion artifacts in both states of awake and under anesthesia [5,8,36–38]. In addition, other variations related to non-neuronal physiology, such as white matter (WM) and cerebrospinal fluid (CSF), that represent nuisance signals were excluded as covariate factors in both states [6]. Therefore, it is unlikely that our results of ICC differences reflected motion artifacts and other physiological noises. Rather, in interpreting our finding, we affirm that improvement of test-retest reliability under anesthesia is associated with suppression of neural factor contributed to resting-state BOLD response variability. Although test-retest reliability of resting-state BOLD response improves under anesthesia, due to high reliability of rs-fMRI metrics including ALFF, ReHo, and FC, our results revealed that these rs-fMRI metrics as robust biomarkers can effectively be employed in rs-fMRI studies in either awake or anesthetized state.

-Effect of Scan length in Awake and Anesthetized States

Numerous studies have proven that scan length is one the key parameters in design of MRI acquisition that plays an important role in generating robust and stable results, particularly in the experiments candidates are anesthetized or have difficulty staying during scanning. They have reported the appropriate scan duration to achieve stable brain's function connectivity in wakeful state [7,21–23,35,39]. For instance, R.M. Birn et al. 2013 showed that test-retest reliability of functional connectivity can be significantly improved by increasing the length of imaging to 9-13 minutes or longer [22]. In line with this, a previous study using machine-learning classifier suggested that minimum of 15-25 minutes rs-fMRI imaging in a single-subject is required to obtain moderate reproducibility of quantitative functional connectivity [40]. Consistent with these, another study in the context of seed based functional connectivity within resting-state networks (RSN), examined the effect of scan duration using 9 distinct time points (3 to 27 minutes) on test-retest reliability and found improvement of ICC over time, until the time point where the plateau reached around 12-16 min and 8-12 min for intra-session and inter-session respectively [22]. Additionally, a recent study on Human Connectome Project (HCP) computed ICC for rs-fMRI graph metrics. Depending on the sample size and the number of time points (duration of scan), they found that for large sample size (for instance, 100 subjects) most of the global and regional graph metrics were reliable for minimum scan duration of 7 minutes. For small sample size of 40 subjects, they found most of the global graph metrics were reliable in long scan duration of 14 min. However, at regional level graph metrics were reliable in the areas located at the default mode network, visual and motor areas [20].

While previous studies estimated the optimal scan duration to obtain reliable resting-state function connectivity measures, our study is the first standardizing scan length of the rs-fMRI studies in anesthetized and wakefulness states using test-retest reliability of rs-fMRI metrics. In this experiment, rs-fMRI metrics were computed for 15 time points (with 1 min interval between) of 15-min scan length individually in both states of awake and under anesthesia. For each metric, ICC measurements over 15 time points were fitted to the logarithmic function using non-linear regression ($y = a + b \ln(x)$, where x defines as scan duration (min)). The optimized scan duration was determined at the time where the standardized function reached the plateau (($|dy/dx|$) < 0.01). Our results agree well with previous literature finding that test-retest reliability improves as scan duration increases [21]. Furthermore, we propose that resting-state scans under anesthesia require shorter scan length to achieve stable and reliable response compared to scans in awake state. According to the optimized scan length calculated for distinct rs-fMRI metrics, we suggest that optimal scan length range of 14.7-20.3 min and 8.6-17.2 min for rs-fMRI scans in awake and under anesthesia respectively are required to ensure stable and reproducible rs-fMRI measurements. Our findings agree well with the previous study suggesting that reliability of functional connectivity can be improved by increasing the imaging duration to 13 min and longer in awake state [22]. Moreover, we showed that general anesthesia affects significantly on reliability of the resting-state metrics which results in shorter optimal scan length compared to awake state. Additionally, among the quantitative maps, fALFF reaches to the optimal ICC in shorter scan duration (8.6 min under anesthesia and 14.7 min in awake) and FC in longer scan length (17.2 min under anesthesia and 20.3 min in awake) that can be useful in determination of the appropriate imaging biomarker in rs-fMRI studies with different scan lengths.

Our study for the first proposed a systematic approach standardizing rs-fMRI scan length for either wakefulness or anesthetized state that can predict the trend of reliability over any timepoint. However, prior analytical studies employed different strategies to determine the optimal scan duration that were elaborative and less effective [22,41,42]. For instance, a recent study investigated ICC of dynamic function connectivity (dFC) using different TRs and following sliding window approach and trade-off between large and small window sizes. They defined the appropriate time point at the window size where the highest reliability was arisen. They proposed that dFC ICC follows an inverted U curve which gives the optimal window size in the middle range of time [43].

There are some potential limitations involved in our study. First, this is important to keep in mind that resting-state BOLD response may be modulated by number of participants included in the study. It has been shown that poor reproducibility is associated with small sample size due to the low statistical power of ICC measurements. Indeed, sample size defines the number of degrees of freedom, which is a key element in determining the statistical power of ICC measurement at a group level. Prior literature reported the minimum number of 20 subjects that permits reliable rs-fMRI measurements [20,44,45]. As such, a trade-off between number of subjects and scan length is needed to achieve reliable outcomes. Hence, we suggest that smaller sample size (below 40 subjects) requires longer scan duration to obtain reproducible results as we reported in our experiment. Future research is needed to take this into account by investigating reliability with a larger sample size at different time points. Taking together, number of factors including sample size, scan duration, and effect of general anesthesia play the key roles in reliability of rs-fMRI metrics that can be optimized through an appropriate trade-off between these parameters.

Secondly, we note that our finding of improvement of reliability under anesthesia is consistent for the scan duration of 15 minutes that we acquired in our experiment. Based on non-linear logarithmic regression modeled to the ICC measurements, we showed that for all the rs-fMRI metrics at different time points, ICC is higher under anesthesia than in awake state. However, the trend of ICC for the measurements can change beyond 15 minutes. As such, ICC may drop under anesthesia than in awake at specific time point. This conclusion is convincing commensurate with the pathophysiology of anesthetics in which by increasing time the effect of the anesthetics may be diminished, or physiological noise may generate fluctuations and intervene the estimated fMRI BOLD signal.

Finally, we note that this study was focused on the scans acquired for a distinct repetition time (TR) or number of volumes per specific time. However, scan duration and number of acquired volumes are mutually dependent. On the other hand, different TR may impact BOLD signal due to different factors involved, such as time of recovery and in-flow effect. Thus, shorter TR (rapid imaging pulse sequences) would improve not only the efficiency of detecting connections, but also enhance the test-retest reliability of estimated rs-fMRI BOLD signal that can be the investigated by future research.

With increasing interest in reproducible findings using rs-fMRI, there has been growing number of studies evaluating reliability of rs-fMRI typically measured by the ICC. Recent findings from prior literature have pointed towards the low reliability of rs-fMRI BOLD response, along with the number of factors including study design and analysis decisions that play the main roles to boost the reproducibility of the measurements. Our results contribute to the body of research standardizing the reliability of rs-fMRI metrics through combining factors including application of anesthetic agents and scan duration. In the present study we comprehensively examined the reliability of rs-fMRI metrics within spontaneous low-frequency fluctuations (0.01-0.08 Hz) of brain activity from scans up to 15 minutes in duration. Our findings demonstrated improvement of reliability under anesthesia compared to awake state across the time points. Additionally, we revealed that shorter acquisitions are required to achieve comparable reliability under anesthesia. This systematic exploration of reliability of rs-fMRI measures is helpful in implication of anesthetic agents in clinical practice and longitudinal studies of patients need to get scanned under anesthesia.

Declarations

Authors' contributions:

F.A. designed the study, processed the acquired data and wrote the paper, M.A. helped with the statistics and data analysis, V.R. supervised and managed clinical and surgical aspects of the project and edited the paper, F. B. M. advised on the analysis and manuscript and edited the paper, C. W. was the principal study supervisor, supervised clinical and surgical aspects of the project and edited the paper. All authors reviewed the manuscript.

Acknowledgements:

We thank the support from the Department of Anesthesiology and Department of Neurosurgery of Thomas Jefferson University, Philadelphia, PA for helping on acquiring and analysis of data for this study.

Declaration of conflicting interests:

The authors declare no competing interests.

References

1. Malia, F. & Norton2, M. I. Mason1,§,*, John D. Van Horn1,†, Daniel M. Wegner3, Scott T. Grafton1,‡, and C. N. M. Wandering Minds: The Default Network and Stimulus-Independent Thought. *Science.* 393–395 (2007).
2. Rémi Patriata, E. K., Molloyb, Timothy, B., d Meierc, G. R., Kirke, Veena, A. & Naird, M. E. Meyeranda,f, Vivek Prabhakaranb,c,d,g,h, and R. M. B. The effect of resting condition on resting-state fMRI reliability and consistency: A comparison between resting with eyes open, closed, and fixated. Neuroimage 463–473(2013) doi:10.1016/j.neuroimage.2013.04.013.
3. Birn, R. M. The role of physiological noise in resting-state functional connectivity., **62**, 864–870 (2012).
4. Finn, E. S. *et al.* Can brain state be manipulated to emphasize individual differences in functional connectivity?, **160**, 140–151 (2017).
5. Cong, C. *et al.* and W. W. S. One-year test-retest reliability of intrinsic connectivity network fMRI in older adults., 1471–1483 https://doi.org/10.1016/j.neuroimage.2012.03.027 (2012).
6. Parkes, L., Fulcher, B., Yücel, M. & Fornito, A. An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI., **171**, 415–436 (2018).
7. Wang, Y. *et al.* Intra-session test-retest reliability of functional connectivity in infants., **239**, 118284 (2021).
8. Zuo, X. N. & Xing, X. X. Test-retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: A systems neuroscience perspective. *Neurosci. Biobehav. Rev*, **45**, 100–118 (2014).
9. Levine, A. T., Li, B., Barnes, P., Lomber, S. G. & Butler, B. E. Assessment of anesthesia on physiological stability and BOLD signal reliability during visual or acoustic stimulation in the cat. *J. Neurosci.*

Methods, **334**, 108603 (2020).

10. Wu, T. L. *et al.* Effects of isoflurane anesthesia on resting-state fMRI signals and functional connectivity within primary somatosensory cortex of monkeys. *Brain Behav*, **6**, 1–12 (2016).
11. Aksenov, D. P., Li, L., Miller, M. J., Iordanescu, G. & Wyrwicz, A. M. Effects of anesthesia on BOLD signal and neuronal activity in the somatosensory cortex. *J. Cereb. Blood Flow Metab*, **35**, 1819–1826 (2015).
12. Grandjean, J., Schroeter, A., Batata, I. & Rudin, M. Optimization of anesthesia protocol for resting-state fMRI in mice based on differential effects of anesthetics on functional connectivity patterns., **102**, 838–847 (2014).
13. Hemmings, H. C. *et al.* Emerging molecular mechanisms of general anesthetic action. *Trends Pharmacol. Sci*, **26**, 503–510 (2005).
14. Huang, Z. *et al.* Decoupled temporal variability and signal synchronization of spontaneous brain activity in loss of consciousness: An fMRI study in anesthesia., **124**, 693–703 (2016).
15. Jonckers, E. *et al.* Different anesthesia regimes modulate the functional connectivity outcome in mice. *Magn. Reson. Med*, **72**, 1103–1112 (2014).
16. Moody, O. A. *et al.* The Neural Circuits Underlying General Anesthesia and Sleep. *Anesth. Analg*, **132**, 1254–1264 (2021).
17. Baria, A. T. *et al.* Bold temporal variability differentiates wakefulness from anesthesia-induced unconsciousness. *J. Neurophysiol*, **119**, 834–848 (2018).
18. Bettinardi, R. G., Tort-Colet, N., Ruiz-Mejias, M., Sanchez-Vives, M. V. & Deco, G. Gradual emergence of spontaneous correlated brain activity during fading of general anesthesia in rats: Evidences from fMRI and local field potentials., **114**, 185–198 (2015).
19. Huang, Z. *et al.* Altered temporal variance and neural synchronization of spontaneous brain activity in anesthesia. *Hum. Brain Mapp*, **35**, 5368–5378 (2014).
20. Termenon, M., Jaillard, A., Delon-Martin, C. & Achard, S. Reliability of graph analysis of resting state fMRI using test-retest dataset from the Human Connectome Project., **142**, 172–187 (2016).
21. Wang, J. H. *et al.* Graph theoretical analysis of functional brain networks: Test-retest evaluation on short- and long-term resting-state functional MRI data.*PLoS One***6**, (2011).
22. Birn, R. M. *et al.* The effect of scan length on the reliability of resting-state fMRI connectivity estimates., **83**, 550–558 (2013).
23. Liao, X. H. *et al.* Functional brain hubs and their test-retest reliability: A multiband resting-state functional MRI study., **83**, 969–982 (2013).
24. Somandepalli, K. *et al.* Short-term test-retest reliability of resting state fMRI metrics in children with and without attention-deficit/hyperactivity disorder. *Dev. Cogn. Neurosci*, **15**, 83–93 (2015).
25. Li, Z., Kadivar, B. S. A., Pluta, B. S. J., Dunlop, B. S. J., Wang, Z. & Ph.D4, and PhD2, *. Test-retest Stability Analysis of Resting Brain Activity Revealed by BOLD fMRI. *J Magn Reson Imaging*, **36**, 344–354 (2012).

26. Shah, L. M., Cramer, J. A., Ferguson, M. A., Birn, R. M. & Anderson, J. S. Reliability and reproducibility of individual differences in functional connectivity acquired during task and resting state. *Brain Behav*, **6**, 1–15 (2016).

27. Becq, G. J. P. C. *et al.* Functional connectivity is preserved but reorganized across several anesthetic regimes. NeuroImage vol. 219(Elsevier Inc., 2020).

28. Elliott, M. L. *et al.* General functional connectivity: Shared features of resting-state and task fMRI drive reliable and heritable individual differences in functional brain networks., **189**, 516–532 (2019).

29. Zou, Q. H. *et al.* An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: Fractional ALFF. *J. Neurosci. Methods*, **172**, 137–141 (2008).

30. Fang, X. *et al.* Abnormal regional homogeneity (ReHo) and fractional amplitude of low frequency fluctuations (fALFF) in first-episode drug-naïve schizophrenia patients comorbid with depression. *Brain Imaging Behav*, https://doi.org/10.1007/s11682-021-00465-0 (2021).

31. Zoua, Q. H. *et al.* Xiang-Yu Longa, c,, and, E. An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: Fractional ALFF. J Neurosci Methods 172, 137–141(2008).

32. Franks, N. P. Molecular targets underlying general anaesthesia. *Br. J. Pharmacol*, **147**, 72–81 (2006).

33. Mashour, G. A. Top-down mechanisms of anesthetic-induced unconsciousness. *Front. Syst. Neurosci*, **8**, 1–10 (2014).

34. Pawela, C. P. *et al.* Resting-state functional connectivity of the rat brain. *Magn. Reson. Med*, **59**, 1021–1029 (2008).

35. Teeuw, J., Pol, H., Boomsma, H. E., Brouwer, R. M. & D. I. & Reliability modelling of resting-state functional connectivity., **231**, 117842 (2021).

36. Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. J. & Turner, R. Movement-related effects in fMRI time-series. *Magn. Reson. Med*, **35**, 346–355 (1996).

37. Ciric, R. *et al.* Mitigating head motion artifact in functional connectivity MRI. *Nat. Protoc*, **13**, 2801–2826 (2018).

38. Mahadevan, A. S., Tooley, U. A., Bertolero, M. A., Mackey, A. P. & Bassett, D. S. Evaluating the sensitivity of functional connectivity measures to motion artifact in resting-state fMRI data., **241**, 118408 (2021).

39. Mejia, A. F. *et al.* Improved estimation of subject-level functional connectivity using full and partial correlation with empirical Bayes shrinkage., **172**, 478–491 (2018).

40. Anderson, J. S., Ferguson, M. A. & Lopez-Larson, M. & Yurgelun-Todd, D. Reproducibility of single-subject functional connectivity measurements. *Am. J. Neuroradiol*, **32**, 548–555 (2011).

41. Andellini, M., Cannatà, V., Gazzellini, S., Bernardi, B. & Napolitano, A. Test-retest reliability of graph metrics of resting state MRI functional brain networks: A review. *J. Neurosci. Methods*, **253**, 183–192 (2015).

42. Choe, A. S. *et al.* Comparing test-retest reliability of dynamic functional connectivity methods., **158**, 155–175 (2017).
43. Zhang, C., Baum, S. A., Adduru, V. R., Biswal, B. B. & Michael, A. M. Test-retest reliability of dynamic functional connectivity in resting state fMRI., **183**, 907–918 (2018).
44. Thirion, B. *et al.* Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses., **35**, 105–120 (2007).
45. Button, K. S. *et al.* Power failure: Why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci*, **14**, 365–376 (2013).
46. Yan, C. G., Wang, X., Di, Zuo, X. N. & Zang, Y. F. DPABI: Data Processing & Analysis for (Resting-State) Brain Imaging. Neuroinformatics 14, 339–351(2016).
47. Jiang, L., Zuo, X. N. & Regional Homogeneity A Multimodal, Multiscale Neuroimaging Marker of the Human Connectome. *Neuroscientist*, **22**, 486–505 (2016).
48. Zang, Y., Jiang, T., Lu, Y., He, Y. & Tian, L. Regional homogeneity approach to fMRI data analysis., **22**, 394–400 (2004).
49. Wang, L. *et al.* Overlapping and segregated resting-state functional connectivity in patients with major depressive disorder with and without childhood neglect. *Hum. Brain Mapp*, **35**, 1154–1166 (2014).
50. Dai, Z. *et al.* Identifying and mapping connectivity patterns of brain network hubs in Alzheimer's disease. *Cereb. Cortex*, **25**, 3723–3742 (2015).
51. Gisev, N., Bell, J. S. & Chen, T. F. Interrater agreement and interrater reliability: Key concepts, approaches, and applications. *Res. Soc. Adm. Pharm*, **9**, 330–338 (2013).
52. Noble, S., Scheinost, D. & Constable, R. T. A decade of test-retest reliability of functional connectivity: A systematic review and meta-analysis., **203**, 116157 (2019).
53. Shrout, P. E. & Fleiss, J. L. Intraclass correlations: Uses in assessing rater reliability. *Psychol. Bull*, **86**, 420–428 (1979).
54. Morales, C. *et al.* Test–Retest Reliability of Functional Magnetic Resonance Imaging Activation for a Vergence Eye Movement Task. *Neurosci. Bull*, **36**, 506–518 (2020).

Tables

Table 1. Mean ± SD and number of voxels of ICC of rs-fMRI metrics including amplitude of low frequency fluctuation (ALFF), fractional amplitude of low frequency fluctuation (fALFF), functional connectivity (FC), and regional homogeneity (ReHo) in two different masks where $0 < ICC < 1$ and $0.75 < ICC < 1$ in two states of awake and under anesthesia.

		Number of Voxels		MEAN ± SD	
fMRI metrics	ICC range	Awake	Anesthesia	Awake	Anesthesia
ALFF	0 < ICC < 1	54833	54833	0.81 ± 0.15	0.80 ± 0.20
	0.75 <ICC < 1	35397	33566	0.86 ± 0.06	0.88 ± 0.06
fALFF	0 < ICC < 1	54833	54833	0.51 ± 0.51	0.59 ± 0.28
	0.75 <ICC < 1	9891	19193	0.82 ± 0.05	0.85 ± 0.06
FC	0 < ICC < 1	54833	54833	0.65 ± 0.20	0.83 ± 0.13
	0.75 <ICC < 1	21637	17699	0.83 ± 0.05	0.84 ± 0.06
ReHo	0 < ICC < 1	54833	54833	0.77 ± 0.15	0.80 ± 0.17
	0.75 <ICC < 1	28860	35790	0.84 ± 0.05	0.88 ± 0.06

Table 2. List of optimized scan lengths for rs-fMRI metrics at derivative logarithmic function $|dy/dx| = 0.01$ ($y = a + b \ln(x)$). ALFF: amplitude of low frequency fluctuation; fALFF: fractional amplitude of low frequency fluctuation; FC: functional connectivity; ReHo: regional homogeneity.

Resting-state fMRI metric	Optimized scan length in awake state (min)	Optimized scan length in anesthetized state (min)
ALFF	18.7	11.9
fALFF	14.7	8.6
Functional connectivity	20.3	17.2
ReHo	17.5	10

Figures

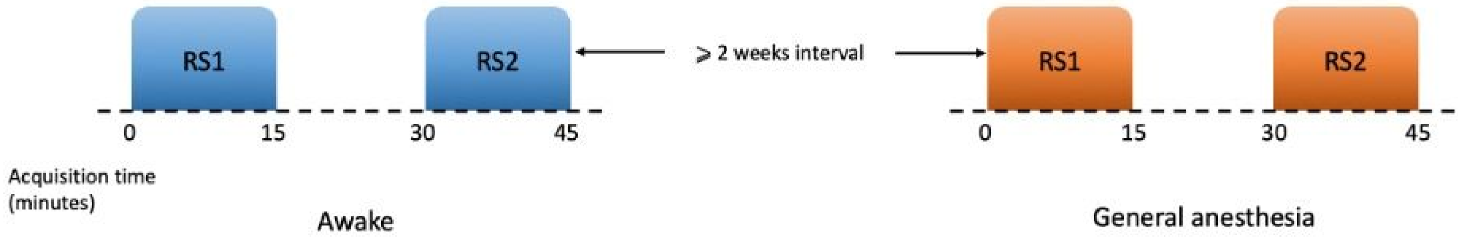

Figure 1

Schematic of resting-state paradigms. Each session of scan consisted of two sets of rs-fMRI which lasted 15 minutes, and at least 15 minutes gap between them. There was an interval of at least two weeks between two sessions of awake and under general anesthesia.

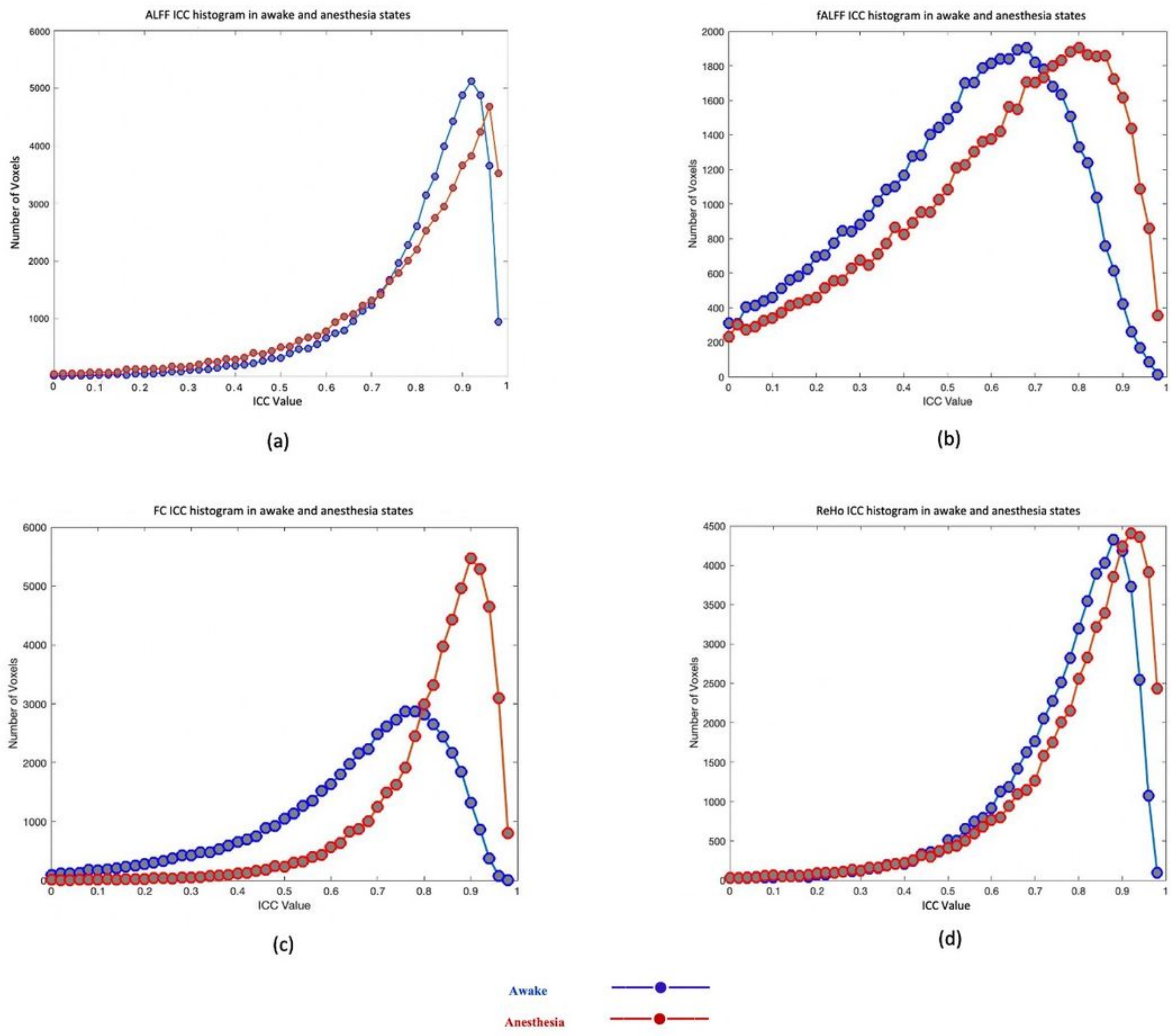

Figure 2

Test-retest reliability (ICC) histograms of rs-fMRI metrics including amplitude of low frequency fluctuation (ALFF) (a), fractional amplitude of low frequency fluctuation (fALFF) (b), functional connectivity (FC) (c), and regional homogeneity (ReHo) (d) in two states of awake and under anesthesia.

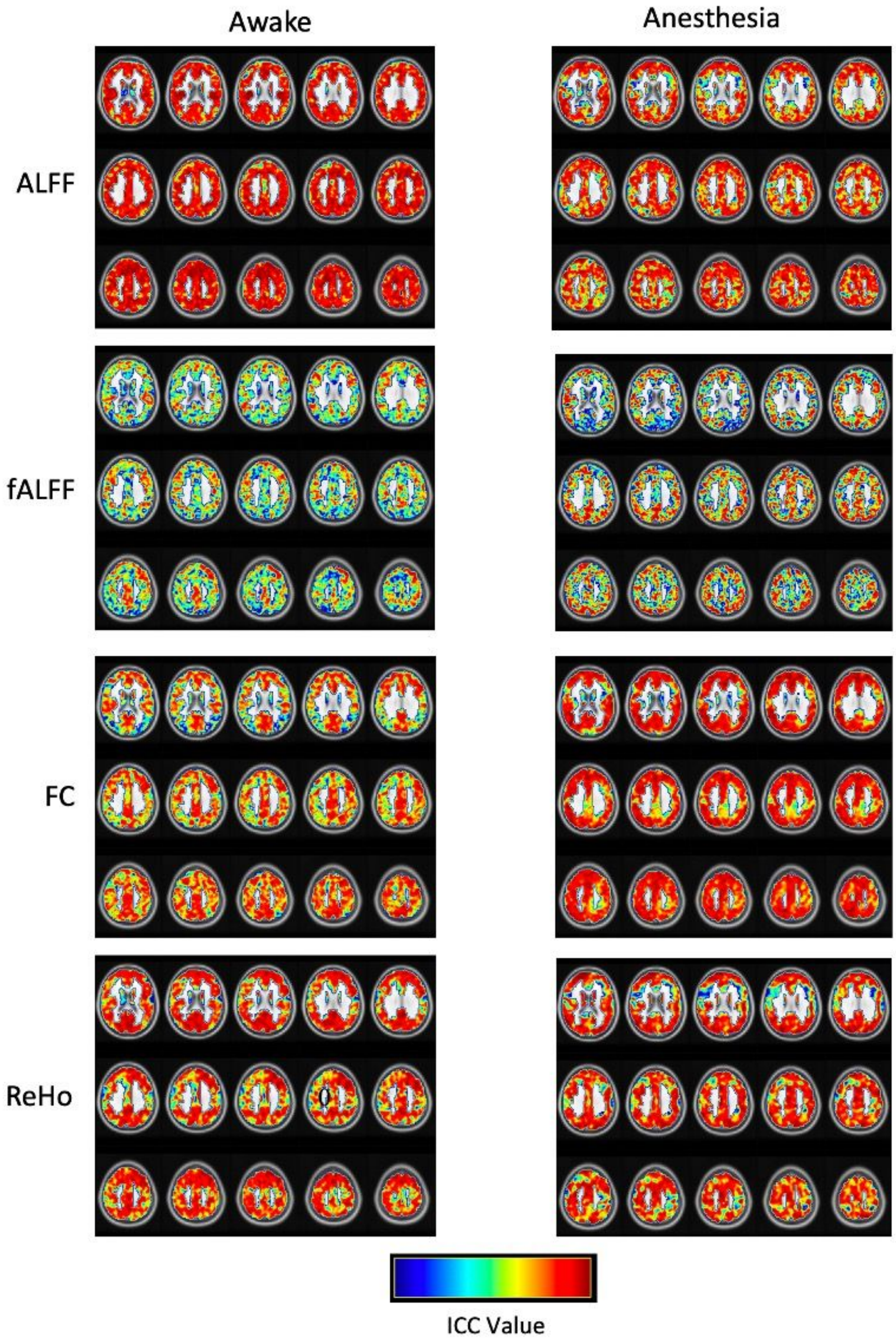

Figure 3

ICC maps of amplitude of low frequency fluctuation (ALFF), fractional amplitude of low frequency fluctuation (fALFF), functional connectivity (FC), and regional homogeneity (ReHo) in two states of awake and under anesthesia (0 < ICC < 1).

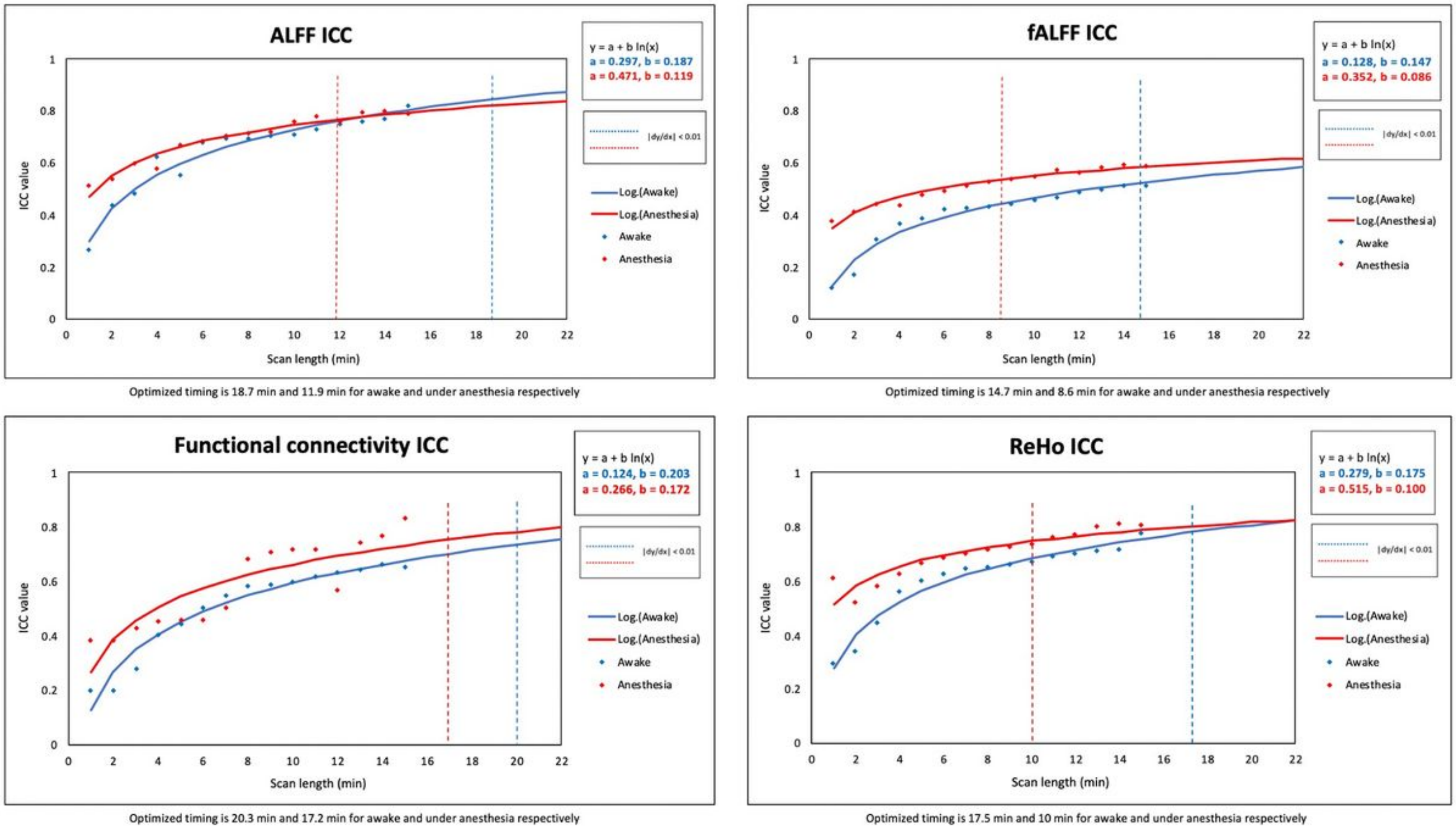

Figure 4

Scatter plots and standardized logarithmic fits of ICC of rs-fMRI metrics including amplitude of low frequency fluctuation (ALFF), fractional amplitude of low frequency fluctuation (fALFF), functional connectivity (FC), regional homogeneity (ReHo).

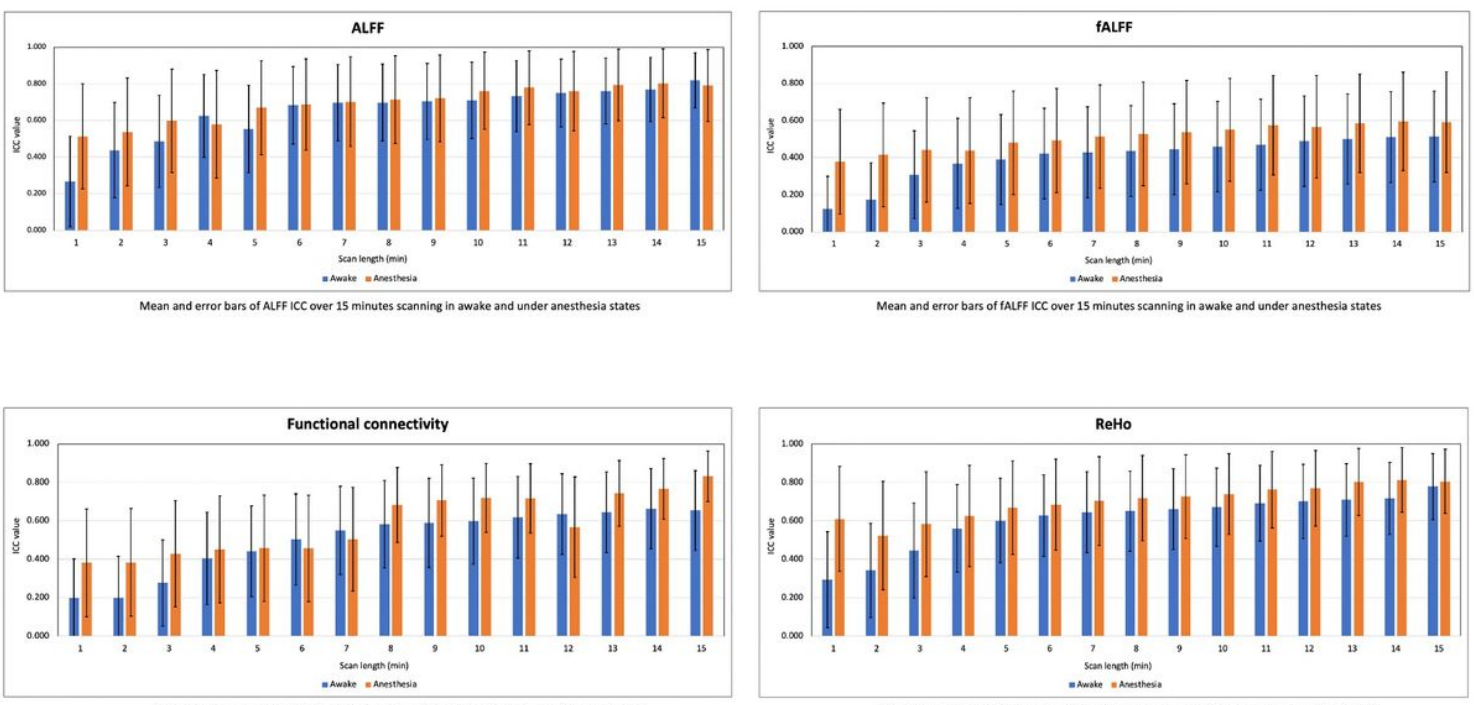

Figure 5

Means and error bars of ICC for amplitude of low frequency fluctuation (ALFF), fractional amplitude of low frequency fluctuation (fALFF), functional connectivity (FC), and regional homogeneity (ReHo) over 1 to 15 min scan lengths.