

Data-Driven and Predefined ROI-Based Quantification of Long-Term Resting-State fMRI Reproducibility

Xiaomu Song,¹ Lawrence P. Panych,² and Nan-Kuei Chen³

Abstract

Resting-state functional magnetic resonance imaging (fMRI) is a promising tool for neuroscience and clinical studies. However, there exist significant variations in strength and spatial extent of resting-state functional connectivity over repeated sessions in a single or multiple subjects with identical experimental conditions. Reproducibility studies have been conducted for resting-state fMRI where the reproducibility was usually evaluated in predefined regions-of-interest (ROIs). It was possible that reproducibility measures strongly depended on the ROI definition. In this work, this issue was investigated by comparing data-driven and predefined ROI-based quantification of reproducibility. In the data-driven analysis, the reproducibility was quantified using functionally connected voxels detected by a support vector machine (SVM)-based technique. In the predefined ROI-based analysis, all voxels in the predefined ROIs were included when estimating the reproducibility. Experimental results show that (1) a moderate to substantial within-subject reproducibility and a reasonable between-subject reproducibility can be obtained using functionally connected voxels identified by the SVM-based technique; (2) in the predefined ROI-based analysis, an increase in ROI size does not always result in higher reproducibility measures; (3) ROI pairs with high connectivity strength have a higher chance to exhibit high reproducibility; (4) ROI pairs with high reproducibility do not necessarily have high connectivity strength; (5) the reproducibility measured from the identified functionally connected voxels is generally higher than that measured from all voxels in predefined ROIs with typical sizes. The findings (2) and (5) suggest that conventional ROI-based analyses would underestimate the resting-state fMRI reproducibility.

Key words: functional network; intra-class correlation coefficient; long-term; reproducibility; resting-state fMRI; support vector machine

Introduction

RESTING-STATE FUNCTIONAL magnetic resonance imaging (rsfMRI) is a technology that measures baseline regional neuronal connectivity independent of an explicit task (Greicius et al., 2003). It is different from task-based fMRI that identifies and characterizes brain functional regions specific to different task stimuli. RsfMRI is a promising tool for neuroscience research and clinical studies, such as early diagnosis of neurodegenerative disease (e.g., Alzheimer's disease), characterization of disease progression, evaluation of drug effects and treatment efficacy, and neurosurgical planning (Broyd et al., 2009; Ferreira and Busatto, 2013; Fox and Greicius, 2010; Fox and Raichle, 2007; Kollndorfer et al., 2013; Lee et al., 2013; Sonuga-Barke and Castellanos, 2007; Vlioger et al., 2004).

Due to the nonstationary nature of fMRI, rsfMRI data exhibit considerable intra- and inter-subject variations. In other words, functional connectivity patterns identified from rsfMRI data collected from a subject might not be well repeatable from the same subject at a different time or from different subjects. Therefore, before rsfMRI can be used as a reliable imaging biomarker, it is important to evaluate its reproducibility within and between individual subjects.

Studies have been performed to investigate the rsfMRI reproducibility either in a short term (e.g., from a day up to several months) (Amann et al., 2009; Braun et al., 2012; Damoiseaux et al., 2006; Honey et al., 2009; Kristo et al., 2014; Meindl et al., 2010; Van de Ven et al., 2004; Wisner et al., 2013), or a long term (e.g., >1 year) (Chou et al., 2012; Fiecas et al., 2013; Guo et al., 2012; Liang et al., 2012; Shehzad et al., 2009; Wang et al., 2011). Most of

¹Department of Electrical Engineering, School of Engineering, Widener University, Chester, Pennsylvania.

²Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.

³Brain Imaging and Analysis Center, Duke University Medical Center, Durham, North Carolina.

these studies assessed rsfMRI reproducibility using all voxels in predefined anatomical or functional regions-of-interest (ROIs) (Braun et al., 2012; Chou et al., 2012; Fiecas et al., 2013; Liang et al., 2012; Shehzad et al., 2009; Wang et al., 2011). A major concern with the ROI-based fMRI reproducibility studies is that not all of the voxels within predefined ROIs are part of the network-of-interest and functionally connected to each other, potentially affecting the accuracy of the reproducibility measures.

In two recent studies, rsfMRI reproducibility was investigated using functionally connected voxels identified by a brain mapping technique (Guo et al., 2012; Wisner et al., 2013), such as seed-based analysis (Biswal et al., 1995; Fox and Raichle, 2007; Greicius et al., 2003) or independent component analysis (ICA) (Beckmann et al., 2005; Van de Ven et al., 2004). The seed-based analysis is hypothesis-driven and identifies a functional network-of-interest by measuring the correlation of each voxel's temporal profile with a seed region defined *a priori*. ICA is a data-driven method that explores the fMRI data structure to identify interest components without *a priori* model of brain activity.

In this work, rsfMRI reproducibility was investigated by comparing the reproducibility measures obtained from (1) only functionally connected voxels identified by a brain mapping technique in predefined ROIs, and (2) all voxels in the predefined ROIs in seven major functional networks in human brains based upon a set of rsfMRI data, which were collected from six subjects over 1.5 years with nine sessions for each subject (Chou et al., 2012). Functionally connected voxels in the resting-state networks were identified using a recently developed data-driven support vector machine (SVM)-based technique (Song and Chen, 2014). The calculation of reproducibility measures using functionally connected voxels is termed data-driven quantification of reproducibility in this study, in contrast to predefined ROI-based quantification of reproducibility where all voxels in each ROI are included when computing the reproducibility measures.

Besides a comparison between the reproducibility measures obtained from the data-driven and ROI-based analyses, the effects of ROI size on the reproducibility measures were examined for the ROI-based analysis. We also investigated whether there exists a connection between the overall functional connectivity strength and reproducibility.

Materials and Methods

fMRI experiments

Eight right-handed healthy subjects (seven males, one female, aged from 21 to 46 with an average of 30.33 ± 8.64 years) without any neurological diseases and/or psychiatric issues were recruited as participants in an fMRI study (Chou et al., 2012). A 1.5 Tesla GE MR scanner was used to collect data. Each subject was scanned in nine sessions. The inter-session gap ranged from 21 to 133 days, with an average of 54.3 ± 24.1 days. The entire experiment period for each subject ranged from 384 to 554 days, with an average of 463 ± 58.4 days. In each scan session, 24 echo planar imaging (EPI) axial-slices were acquired to cover the entire brain with 6 mm slice thickness (no gap). EPI parameters included a repetition time (TR) of 2.5 sec, an echo time (TE) of 50 msec, and a flip angle of 90° . The field of view was $24 \text{ cm} \times 24 \text{ cm}$, and the matrix size was 64×64 , correspond-

ing to an in-plane resolution of $3.75 \times 3.75 \text{ mm}^2$. A total of 114 volumes were acquired during each resting-state fMRI scan, and the first 6 volumes were discarded as dummy scans. Thus, 108 volumes from each scan session were used in the analysis.

Due to computer storage failure and other technical issues, not all acquired data were retrievable. There were nine sessions of data for six of the subjects, eight sessions for one subject, and only five sessions for one subject. In this study, the data from the six subjects with all nine sessions were used to evaluate the long-term resting-state fMRI reproducibility.

Preprocessing

The physiological fluctuations in white matter and cerebral spinal fluid were first regressed out from whole brain. Then, a band-pass filtering between 0.01 and 0.1 Hz was performed to remove scanner drift and retain low-frequency signals-of-interest below 0.1 Hz. The head motion effects were regressed out using the Friston's 24-parameter model followed by the Power's scrubbing procedure (Friston et al., 1996; Power et al., 2012). The motion corrected data were spatially normalized to the Montreal Neurological Institute (MNI) 2 mm template (Avants et al., 2009). All above preprocessing steps were performed using the C-PAC tool (Sikka et al., 2012). The spatially normalized data were spatially smoothed using a wavelet domain Bayesian noise removal method (Song et al., 2006), which transformed fMRI data into the wavelet domain using the Haar wavelet, and performed a multiscale Gaussian mixture model-based Bayesian wavelet shrinkage. This method can efficiently attenuate spatial noise while preserving signal details without over-smoothing the data.

Networks-of-interest

In this study, seven resting-state functional networks that have been identified in previous works were examined for the rsfMRI reproducibility (Allen et al., 2011; Greicius et al., 2003; Joel et al., 2011; Maneshi et al., 2012; Markett et al., 2014; Schmidt et al., 2013; Toro et al., 2008; Uddin et al., 2009; Watanabe et al., 2013; Woodward et al., 2011; Ystad et al., 2010), including the default mode network (DMN), fronto-parietal attention network (FPAN), visual network, auditory network, sensori-motor network (SMN), basal ganglia network (BGN), and frontal network. To detect each network using the SVM-based technique, a $6 \times 6 \times 6 \text{ mm}^3$ seed region reported in the previous work was used. Table 1 lists the MNI coordinates of the seed regions of the seven networks. Fifty-two ROIs from these networks were investigated where each ROI is defined as a cubic region centered at the MNI coordinates reported in previous studies and listed in Table 2 (Allen et al., 2011; Greicius et al., 2003; Joel et al., 2011; Maneshi et al., 2012; Markett et al., 2014; Schmidt et al., 2013; Toro et al., 2008; Uddin et al., 2009; Watanabe et al., 2013; Woodward et al., 2011; Ystad et al., 2010).

When the SVM-based network mapping was performed, an ROI mask of size $18 \times 18 \times 18 \text{ mm}^3$ was used where only the identified functionally connected voxels within each ROI mask were used to quantify reproducibility. When all voxels in the ROIs were used, four different ROI sizes were considered: $6 \times 6 \times 6$, $10 \times 10 \times 10$, $14 \times 14 \times 14$, and $18 \times 18 \times 18 \text{ mm}^3$. The purpose of choosing different ROI sizes was to examine the impact of ROI size on the

TABLE 1. SEED REGIONS AND THEIR MNI COORDINATES

Network	Seed region	MNI coordinates	References
DMN	PCC	-2, -54, 26	Greicius et al. (2003), Uddin et al. (2009)
FPAN	IPS	-23, -70, 46	Markett et al. (2014), Toro et al. (2008), Watanabe et al. (2013)
Visual	PVC	-2, -82, 4	Joel et al. (2011), Maneshi et al. (2012), Woodward et al. (2011)
Auditory	Primary auditory cortex	-48, -24, 9	Allen et al. (2011), Maneshi et al. (2012), Schmidt et al. (2013)
SMN	PMC	-38, -22, 60	Joel et al. (2011)
BGN	Putman	25, -1, 0	Allen et al. (2011), Ystad et al. (2010)
Frontal	IFG	50, 23, 2	Allen et al. (2011)

BGN, basal ganglia network; DMN, default mode network; FPAN, fronto-parietal attention network; IFG, inferior frontal gyrus; IPS, intraparietal sulcus; MNI, Montreal Neurological Institute; PCC, posterior cingulate cortex; PMC, primary motor cortex; PVC, primary visual cortex; SMN, sensori-motor network.

rsfMRI reproducibility measures. The 52 ROIs from the seven networks form a total of 1326 ROI pairs.

Network detection

In most existing seed- or ICA-based network mapping approaches, the functionally connected voxels were identified using a fixed threshold, and it remains unclear how an opti-

mal statistical threshold should be chosen for longitudinal fMRI data in the presence of significant inter-session and inter-subject signal variations (Achard and Bullmore, 2007; He et al., 2008; Wang et al., 2011). To more objectively assess the rsfMRI reproducibility, here we propose to use a recently developed SVM-based method to identify functionally connected voxels (Song and Chen, 2014), and

TABLE 2. PART OF ROIs IN THE SEVEN RESTING STATE FUNCTIONAL NETWORKS AND THEIR MNI COORDINATES

Network	ROI	MNI coordinates	Network	ROI	MNI coordinates
Default mode network	PCC	-2 -54 26	Sensori-motor network	LPMC	-38 -22 60
	VMPFC	2 56 0		RPMC	39 -23 60
	AMPFC	1 55 26		LVPMC	-34 -16 52
	LSFG	-14 36 59		RVPMC	36 -20 52
	RSFG	17 35 58		SMA	-1 -8 57
	LITG	-62 -33 -20		LS1	-36 -38 58
	RITG	66 -17 -19		RCD	32 -52 -28
	LPHG	-22 -26 -21		Basal ganglia network	LPM
RPHG	25 -26 -18	RPM	30 -14 -8		
LLPC	-47 -71 35	LSN	-12 -15 -18		
RLPC	54 -61 36	RSN	13 -17 -16		
Fronto-parietal network	LIPS	-23 -70 46	LSTN		-18 -18 -8
	RIPS	25 -62 53	RSTN		12 -18 -7
	LIPL	-42 -48 51	LP		-22 4 1
	RIPL	48 -41 54	RP		18 3 -3
	LDPFC	-48 21 38	LC	-33 -20 12	
	Frontal network	RDPFC	43 21 38	RC	33 -19 10
		MCC	0 -31 31	RIFG	50 23 2
		LPC	-9 -76 36	LIFG	-42 39 5
RPC		10 -73 39	RSMG	58 -36 36	
Visual network		BLG	1 -87 -2	LIPL	-58 -40 49
		BCG	1 -71 13	RMFG	49 22 25
		RLG	17 -55 -9	LMFG	-48 17 29
		LLG	-15 -56 -8	BSMG	-1 32 46
	Auditory network	PVC	-2 -82 4	LSTG	-51 -18 7
				RSTG	52 -15 7
				RMCC	2 -4 49

AMPFC, anterior medial prefrontal cortex; BCG, bi calcarine gyrus; BLG, bi lingual gyrus; BSMG, bi superior medial gyrus; LC, left claustrum; LDPFC, left dorsolateral prefrontal cortex; LIFG, left inferior frontal gyrus; LIPL, left inferior parietal lobule; LIPS, left intraparietal sulcus; LITG, left inferior temporal gyrus; LLG, left lingual gyrus; LLPC, left lateral parietal cortex; LMFG, left middle frontal gyrus; LP, left pallidus; LPC, left precuneus; LPHG, left parahippocampal gyrus; LPM, left putamen; LPMC, left primary motor cortex; LS1, left primary somatosensory cortex; LSFG, left superior frontal gyrus; LSN, left substantia nigra; LSTG, left superior temporal gyrus; LSTN, left subthalamic nucleus; LVPMC, left ventral premotor cortex; MCC, middle cingulate cortex; RC, right claustrum; RCD, right cerebellum dentate; RDPFC, right dorsolateral prefrontal cortex; RIFG, right inferior frontal gyrus; RIPL, right inferior parietal lobule; RIPS, right intraparietal sulcus; RITG, right inferior temporal gyrus; RLG, right lingual gyrus; RLPC, right lateral parietal cortex; RMCC, right middle cingulate cortex; RMFG, right middle frontal gyrus; ROI, region-of-interest; RP, right pallidus; RSTG, right superior temporal gyrus; RSTN, right subthalamic nucleus; RVPMC, right ventral premotor cortex; SMA, supplementary motor area; VMPFC, ventral medial prefrontal cortex.

then measure the within- and between-subject reproducibility from the functionally connected voxels.

Based upon the observation that voxels in a single functional network constitute less than 50% of all brain voxels (Laird et al., 2011), the SVM-based method treats the mapping of a resting-state network as an outlier detection process where functionally connected voxels in a network are considered as “outliers.” The one-class SVM (OCSVM) is used to provide an initial identification of the “outliers” (Schölkopf et al., 2001), and it needs an estimation of the ratio of functionally connected voxels to all voxels to initialize the identification. This ratio is network-, session-, and subject-specific and cannot be accurately estimated. It is usually estimated based on experiential knowledge or set with a relatively large value (close to 0.4–0.5) to guarantee a sufficient mapping sensitivity, which could lead to significant mis-detections. Thus, after the OCSVM processing, a spatial- and feature-domain prototype selection procedure is performed to identify voxels that are most probably correctly classified by OCSVM, and the identified voxels are used to train a two-class SVM (TCSVM) (Burges, 1998; Vapnik, 1998) to reclassify the input voxels and obtain refined network mapping results.

Each input voxel to OCSVM and TCSVM is represented by a feature vector consisting of multiple features computed using the voxel’s time course, its neighboring voxels, and a seed from a network-of-interest. The combination of the OCSVM and TCSVM implements a semi-supervised network mapping where *a priori* information from the network-of-interest is obtained from the seed. This method can better define the boundary between functionally connected and unconnected voxels in a feature space even in the presence of longitudinal fMRI data variations, and reliably identify functionally connected voxels either in an individual slice or in whole brain for an individual subject or at the group level.

Since the SVM-based detection method still uses a seed to obtain prior information of a network-of-interest, the method only detects the network-of-interest associated with the chosen seed instead of multiple networks at a time. In this study, different seeds were used to analyze whole brain fMRI data to identify functionally connected voxels in the seven resting-state networks at the group level. The reproducibility measures were computed using the same set of functionally connected voxels across multiple subjects and scan sessions.

Reproducibility study

An averaged time course profile was extracted from each ROI, by including either (1) only functionally connected voxels in the ROI detected by the SVM-based method, or (2) all voxels in the ROI. Pearson’s correlation coefficient (CC) between the average time courses of each pair of ROIs was calculated, and the Fisher r -to- z transformation was performed to normalize the CC values. The reproducibility of inter-regional correlation strength of all ROI pairs in the seven resting-state networks was investigated with intra-class correlation coefficient (ICC) across multiple scan sessions for each individual subject, and across subjects.

Reproducibility of individual subjects. The ICC was used to quantify the rsfMRI reproducibility for individual subjects across multiple sessions based on the inter-regional correlation between each pair of ROIs. There are different ICC mea-

asures, and ICC(2,1) defined in Equation (1) was used in our study (Shrout and Fleiss, 1979):

$$ICC(2, 1) = \frac{BMS - EMS}{BMS + (k + 1)EMS + k(JMS - EMS)/n} \quad (1)$$

where BMS is the between-target mean square; EMS is the error mean square; JMS is the between-session mean square, and k is the number of sessions for each target. The ICC value ranges between 0 and 1 (with negative ICC values set to 0). An ICC value greater than 0.6 indicates substantial or outstanding reproducibility, an ICC value between 0.4 and 0.6 implies moderate reproducibility, and an ICC value less than 0.4 indicates poor reproducibility (Chou et al., 2012; Landis and Koch, 1977). ICC(2,1) was used to measure the within-subject reproducibility with random session effects where the inter-regional correlation of each ROI pair was the target, and BMS was calculated for each subject across all ROI pairs using all nine data sessions.

Between-subject study. For the between-subject reproducibility study, ICC(3,1) defined in Equation (2) was used to quantify the reproducibility of inter-regional correlation of ROI pairs with random subject effects (Shrout and Fleiss, 1979):

$$ICC(3, 1) = \frac{BMS - EMS}{BMS + (k - 1)EMS} \quad (2)$$

where BMS and EMS are defined in the same ways as those in Equation (1), and BMS was computed for the inter-regional correlation of each ROI pair across all subjects using nine data sessions.

Coefficient of variation (CV) (Chou et al., 2012; Gasparovic et al., 2011; Soltysik et al., 2011), which is defined in Equation (3), was computed for the CC value of each ROI pair using the data of all subjects and sessions:

$$CV = \sigma/\mu \quad (3)$$

where μ and σ are the average and standard deviation of CC values. A small CV means a small inter-session and inter-subject variation and indicates acceptable reproducibility. A CV value less than 0.2 was suggested in previous studies to indicate acceptable reproducibility (Chou et al., 2012; Gasparovic et al., 2011), and this threshold was used in this work as well.

ROI size, inter-regional correlation, and reproducibility. In this study, we investigated how ROI size could affect the strength of inter-regional correlation and reproducibility. The 1326 ROI pairs were divided into three groups in terms of their CC values: $CC < 0.25$, $0.25 \leq CC < 0.5$, and $CC \geq 0.5$, and the number of ROI pairs falling into each group was counted for each ROI size, and compared over the four different ROI sizes. The numbers of ROI pairs showing substantial, moderate, and poor between-subject reproducibility obtained using the four ROI sizes were counted and compared as well. In addition, we examined whether there existed any association between the inter-regional correlation strength and reproducibility. The distributions of CC values of all 1326 ROI pairs against ICC values obtained from the four different ROI sizes were investigated, and the numbers of ROI pairs exhibiting substantial or moderate

reproducibility were counted in the three CC groups. Voxel-wise ICC maps in an individual slice that covered part of DMN were generated and compared with a correlation map calculated between the seed in posterior cingulate cortex (PCC) and all voxels in the same slice.

Temporal signal-to-noise ratio. The study of rsfMRI reproducibility requires reliable fMRI data. Temporal signal-to-noise-ratio (tSNR) is an important factor reflecting fMRI reliability (Bennett and Miller, 2010). The tSNR of an individual voxel is defined as the mean of the voxel's time course divided by its standard deviation. In our work, the tSNR of an ROI was calculated from the averaged time course, and the session tSNR of each data session was defined as the average tSNR of the 52 ROIs. The session tSNR was calculated for all data sessions of all subjects. A comparison was performed between the session tSNR values computed using functionally connected voxels and those calculated using all voxels in the predefined ROIs with the four different sizes. To evaluate the complete fMRI data, tSNR was calculated using motion corrected data without the scrubbing procedure. The obtained session tSNR values were evaluated across sessions and subjects.

Head motion effects. Increased subject head motion effects lead to decreased tSNR of fMRI data (Bennett and Miller, 2010; Kristo et al., 2014; Murphy et al., 2006). In addition, head motion effects may introduce artificial correlations between the ROI pairs, and affect the calculation of the reproducibility measures. In this work, the maximum and average root-mean-square (RMS) framewise displacements were calculated (Jenkinson et al., 2002). Correlation analyses were performed across all subjects and sessions to examine whether there exist significant correlations between the two motion effect measurements and the inter-regional correlation computed using the motion corrected data. A two-tailed *t*-test was used to evaluate the significance of the correlation values at the 0.05 significance level.

Results

Individual subjects

Figure 1 shows a comparison between the ICC(2,1) values computed using the correlation of functionally connected voxels (detected by the SVM-based method) and all voxels in the predefined ROIs with four different sizes. Seven symbols are used here to represent the seven networks, and for each symbol there are six instances corresponding to six subjects. The averages of the ICC(2,1) values over the seven networks for each subject are shown in Table 3. It was observed from Figure 1 and Table 3 that when functionally connected voxels were used, higher reproducibility was obtained for most networks of the six subjects compared to the predefined ROI-based analyses. A two-tailed *t*-test indicates that this increase is significant at the 2.6×10^{-4} , 4.9×10^{-4} , 6.0×10^{-4} , and 1.8×10^{-3} levels when the ROI size is $6 \times 6 \times 6$, $10 \times 10 \times 10$, $14 \times 14 \times 14$, and $18 \times 18 \times 18$ mm³, respectively. It was also noted that the fMRI reproducibility generally increases with the ROI size if all voxels in each ROI are used to calculate the average time course profiles. For example, when the ROI size was increased from $6 \times 6 \times 6$ to $10 \times 10 \times 10$ mm³, or from $10 \times 10 \times 10$ to $14 \times 14 \times 14$ mm³,

the increase in fMRI reproducibility is significant at the 3.3×10^{-5} or 5.1×10^{-3} level. When the ROI size was further increased to $18 \times 18 \times 18$ mm³, increases in the average ICC were still observed at the 2.1×10^{-3} significance level.

Table 3 shows that three subjects exhibit substantial reproducibility and the others have moderate reproducibility when only functionally connected voxels are included in quantifying the reproducibility. When all voxels in the $6 \times 6 \times 6$ mm³ ROIs are used, only one subject has moderate reproducibility, and the others have poor reproducibility. Even when using a larger ROI ($10 \times 10 \times 10$ or $14 \times 14 \times 14$ or $18 \times 18 \times 18$ mm³), only one subject exhibits substantial reproducibility.

Inter-subject study

Figure 2(a) shows the average inter-regional correlation of all ROI pairs computed using functionally connected voxels detected by the SVM-based method across all subjects and sessions, and there are 130 ROI pairs with $CC \geq 0.5$. Figures 2(b–e) illustrate the average correlation of all ROI pairs calculated using all voxels in the predefined ROIs with different sizes: (Fig. 2(b)) $6 \times 6 \times 6$ mm³ (13 ROI pairs with $CC \geq 0.5$), (Fig. 2(c)) $10 \times 10 \times 10$ mm³ (50 ROI pairs with $CC \geq 0.5$), (Fig. 2(d)) $14 \times 14 \times 14$ mm³ (84 ROI pairs with $CC \geq 0.5$), and (Fig. 2(e)) $18 \times 18 \times 18$ mm³ (116 ROI pairs with $CC \geq 0.5$). The use of only functionally connected voxels results in more ROI pairs showing $CC \geq 0.5$ compared to those using all voxels in the predefined ROIs even when the ROI size is increased up to $18 \times 18 \times 18$ mm³.

Figure 3(a) shows the ICC(3,1) values calculated using the inter-regional correlation computed from only functionally connected voxels. There are 37 ROI pairs showing substantial reproducibility, 311 pairs having moderate reproducibility, and 978 pairs exhibiting poor reproducibility. Figures 3(b–e) show the ICC(3,1) values of the inter-regional correlation computed using all voxels in the predefined ROIs with the four different sizes. The numbers of ROI pairs showing substantial, moderate, and poor reproducibility are listed in Table 4. It was found that when only functionally connected voxels are used, more ROI pairs exhibit substantial and moderate reproducibility compared to those using all voxels in each predefined ROI. Meanwhile, fewer ROI pairs have poor reproducibility.

Figure 4 illustrates the between-subject reproducibility using CV of correlation among all ROI pairs, where the dark red color represents ROI pairs with $CV < 0.2$ (acceptable reproducibility) and the light green indicates ROI pairs with $CV \geq 0.2$. Figure 4(a) was obtained using the correlation of functionally connected voxels in all ROI pairs across all subjects and sessions. There are 19 ROI pairs showing acceptable between-subject reproducibility. Figures 4(b–e) were obtained based on the inter-regional correlation computed using all voxels in each predefined ROI with the four sizes. The numbers of ROI pairs showing acceptable reproducibility are 2, 4, 7, and 13, respectively, for ROI sizes $6 \times 6 \times 6$, $10 \times 10 \times 10$, $14 \times 14 \times 14$, and $18 \times 18 \times 18$ mm³. More ROI pairs demonstrate a small CV value when only functionally connected voxels are included in the calculation. If all voxels in each predefined ROI are used to calculate the correlation, then a larger ROI size results in more ROI pairs showing acceptable reproducibility.

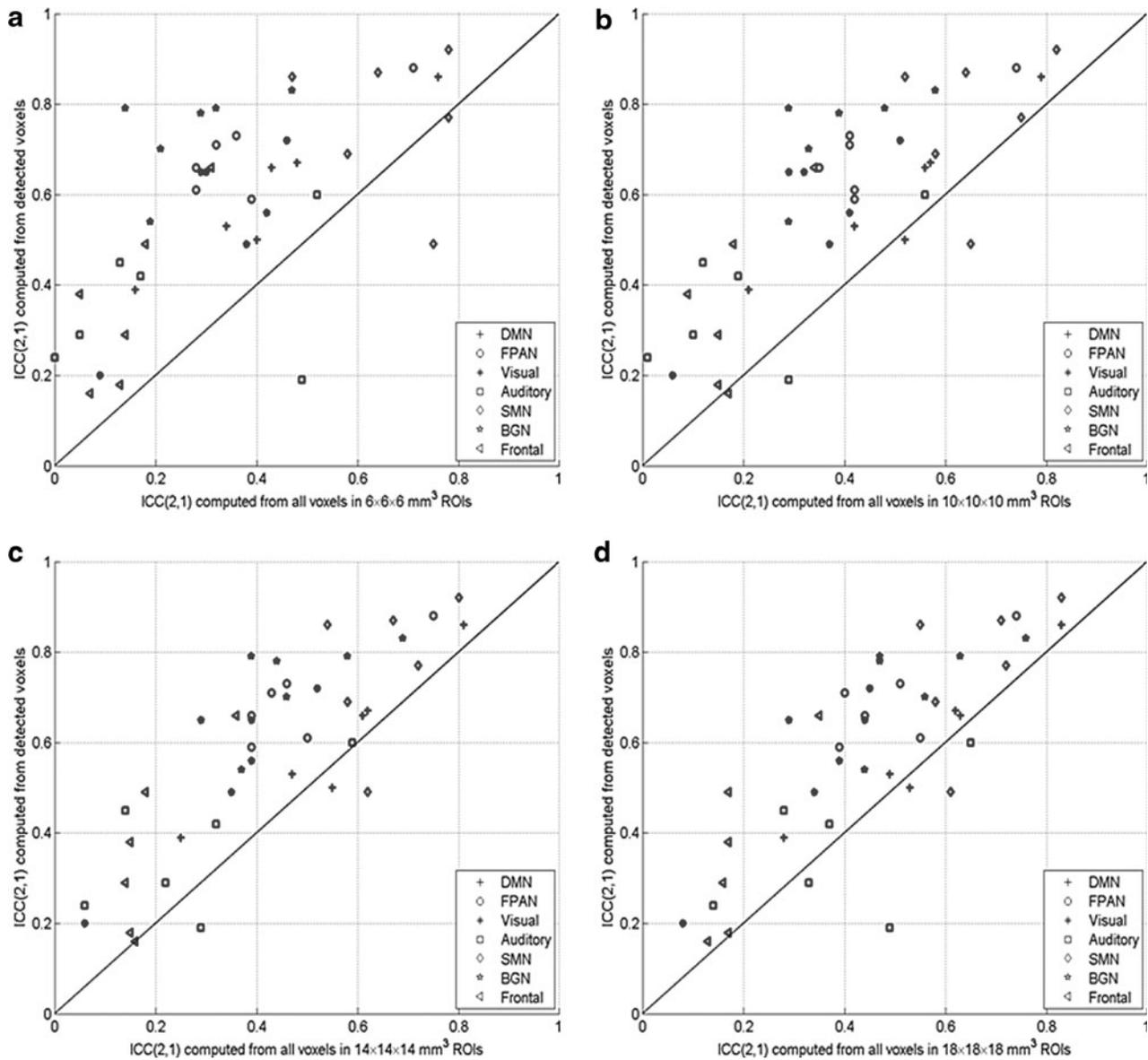


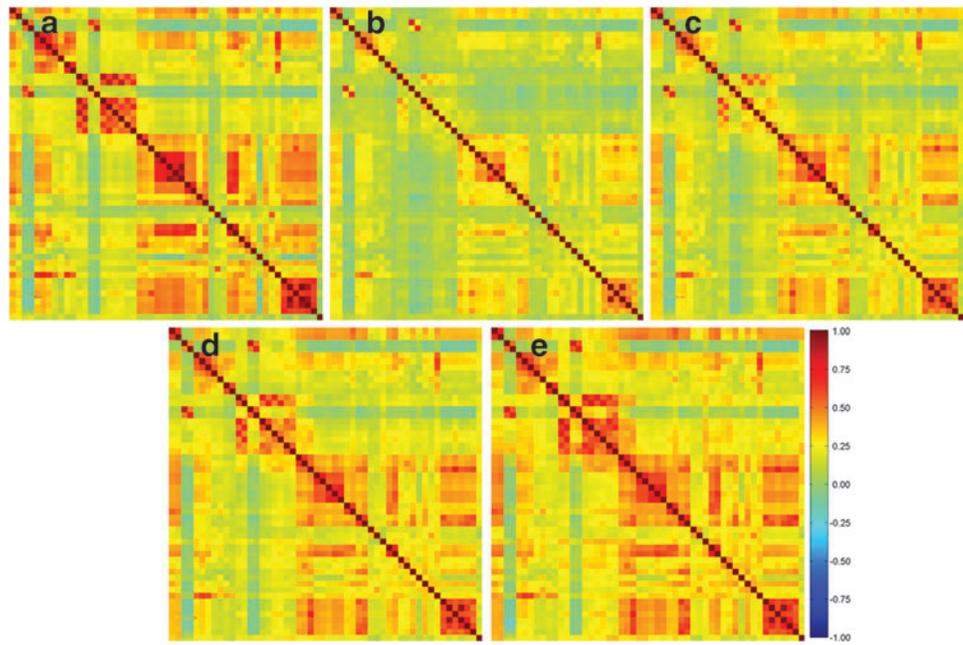
FIG. 1. A comparison of within-subject reproducibility of the seven resting-state networks in terms of the inter-regional correlation of 1326 region-of-interest (ROI) pairs computed using functionally connected voxels detected by the support vector machine (SVM)-based method and all voxels in each predefined ROI with a size of (a) $6 \times 6 \times 6 \text{ mm}^3$, (b) $10 \times 10 \times 10 \text{ mm}^3$, (c) $14 \times 14 \times 14 \text{ mm}^3$, and (d) $18 \times 18 \times 18 \text{ mm}^3$. ICC, intra-class correlation coefficient; DMN, default mode network; FPAN, fronto-parietal attention network; SMN, sensori-motor network; BGN, basal ganglia network.

TABLE 3. A COMPARISON OF AVERAGE WITHIN-SUBJECT REPRODUCIBILITY OF THE SIX SUBJECTS QUANTIFIED BY ICC(2,1) WHEN THE INTER-REGIONAL CORRELATION WAS CALCULATED USING FUNCTIONALLY CONNECTED VOXELS AND ALL VOXELS IN THE PREDEFINED ROIS OF DIFFERENT SIZES

	<i>Sub 1</i>	<i>Sub 2</i>	<i>Sub 3</i>	<i>Sub 4</i>	<i>Sub 5</i>	<i>Sub 6</i>
Connected voxels	0.6	0.78	0.42	0.48	0.56	0.65
ROI: $6 \times 6 \times 6 \text{ mm}^3$	0.3	0.57	0.25	0.32	0.3	0.35
ROI: $10 \times 10 \times 10 \text{ mm}^3$	0.36	0.62	0.29	0.36	0.34	0.4
ROI: $14 \times 14 \times 14 \text{ mm}^3$	0.42	0.65	0.31	0.38	0.36	0.43
ROI: $18 \times 18 \times 18 \text{ mm}^3$	0.45	0.66	0.36	0.41	0.39	0.47

ICC, intra-class correlation coefficient.

FIG. 2. Average correlation of 1326 ROI pairs across all subjects and sessions computed using (a) functionally connected voxels, and all voxels in each predefined ROI with a size of (b) $6 \times 6 \times 6 \text{ mm}^3$, (c) $10 \times 10 \times 10 \text{ mm}^3$, (d) $14 \times 14 \times 14 \text{ mm}^3$, and (e) $18 \times 18 \times 18 \text{ mm}^3$. The correlation coefficient (CC) values on the diagonal lines indicate the correlations of the ROIs to themselves and were not considered in this study. Color images available online at www.liebertpub.com/brain



ROI size, inter-regional correlation, and between-subject reproducibility

Table 5 shows the numbers of ROI pairs falling into the three inter-regional correlation groups when functionally connected voxels and all voxels in the predefined ROIs were used to calculate the correlation. In the ROI-based analysis, an increase in the ROI size leads to more ROI pairs having $CC \geq 0.25$. The average CC values of all ROI pairs are 0.16, 0.21, 0.26, and 0.3, respectively, when the ROI size

increases from $6 \times 6 \times 6$ to $18 \times 18 \times 18 \text{ mm}^3$. Meanwhile, the number of ROI pairs showing substantial reproducibility increases first then decreases and the turning point is $10 \times 10 \times 10 \text{ mm}^3$ as shown in Table 4. The number of ROI pairs having moderate reproducibility increases with the ROI size, while the number of those exhibiting poor reproducibility decreases with the ROI size when the ROI size is below $14 \times 14 \times 14 \text{ mm}^3$. The number of ROI pairs showing poor reproducibility remains the same when the ROI size is increased from $14 \times 14 \times 14$ to $18 \times 18 \times 18 \text{ mm}^3$, as

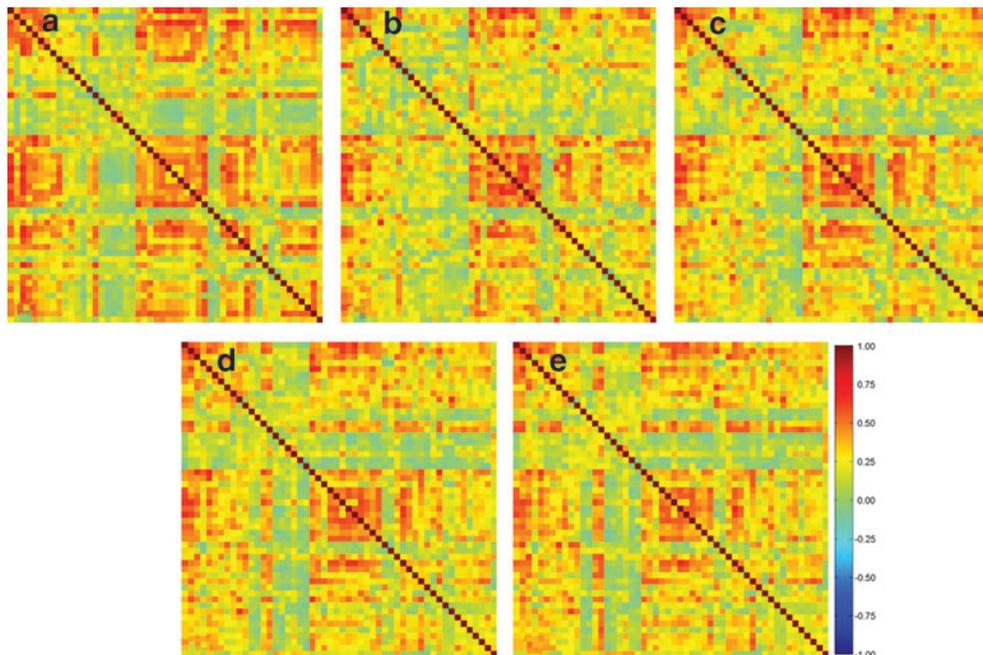


FIG. 3. (a–d) Intra-class correlation coefficient values [ICC(3,1)] calculated using the correlation of all ROI pairs obtained from the average time course of (a) functionally connected voxels, and all voxels in each predefined ROI with a size of (b) $6 \times 6 \times 6 \text{ mm}^3$, (c) $10 \times 10 \times 10 \text{ mm}^3$, (d) $14 \times 14 \times 14 \text{ mm}^3$, and (e) $18 \times 18 \times 18 \text{ mm}^3$. The ICC values on the diagonal lines were not considered in this work. Color images available online at www.liebertpub.com/brain

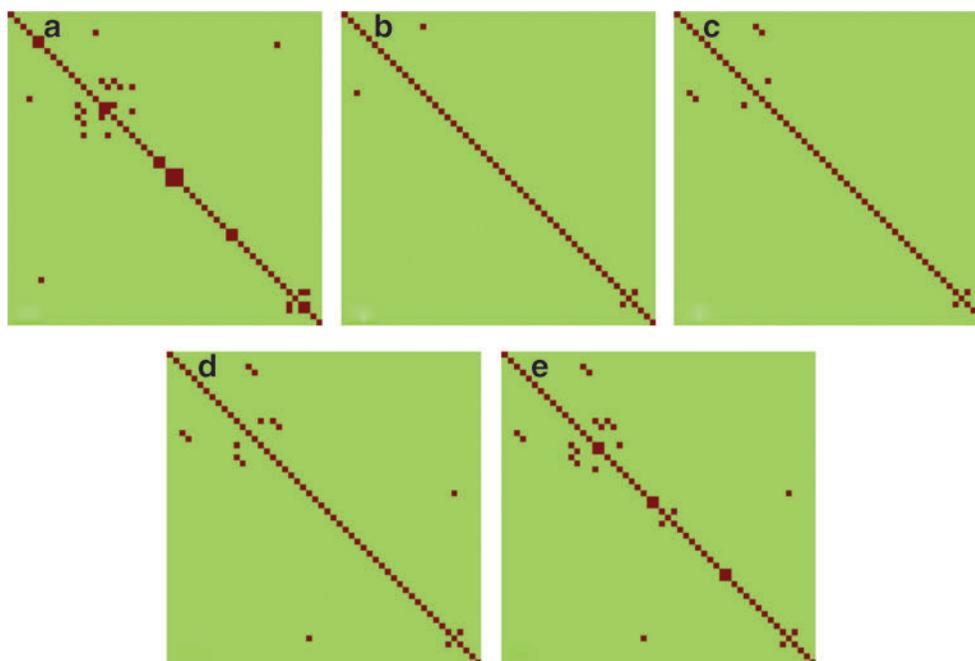


FIG. 4. Coefficient of variation (CV) maps of the correlation of all ROI pairs computed using the average time courses of (a) functionally connected voxels, and all voxels in each predefined ROI with a size of (b) $6 \times 6 \times 6 \text{ mm}^3$, (c) $10 \times 10 \times 10 \text{ mm}^3$, (d) $14 \times 14 \times 14 \text{ mm}^3$, and (e) $18 \times 18 \times 18 \text{ mm}^3$. The pairs in dark red are those with $CV < 0.2$, and the others with green color are ROI pairs with $CV \geq 0.2$. There are more ROI pairs showing small CV values in (a) than the others. The CV values on the diagonal lines were not considered in this study. Color images available online at www.liebertpub.com/brain

shown in Table 4. If we add the numbers of ROI pairs showing substantial and moderate reproducibility together, then the sum number increases first with the ROI size and then remains the same when the ROI size is greater than $14 \times 14 \times 14 \text{ mm}^3$.

Correlation strength and reproducibility

It was noted from Figures 2 and 3 that part of ROI pairs showing substantial or moderate reproducibility also have relatively high inter-regional correlations. Figure 5 shows the distribution of ICC(3,1) values against CC values of the 1326 ROI pairs. The horizontal lines in the ICC-CC plane divide the ROI pairs in terms of $CC < 0.25$, $0.25 \leq CC < 0.5$, and $CC \geq 0.5$, and the vertical lines separate the ROI pairs in term of poor, moderate, and substantial reproducibility. The numbers of ROI pairs showing moderate or substantial reproducibility were counted in terms of the three CC groups, and these are listed in Table 6. When functionally connected voxels identified by the SVM-based method

were used, about 64.94% of ROI pairs showing moderate or substantial reproducibility have $CC \geq 0.25$. In the ROI-based analysis, this ratio is 53.56%, 65.38%, 65.07%, and 66.54% when the ROI size is increased from $6 \times 6 \times 6$ to $18 \times 18 \times 18 \text{ mm}^3$. Therefore, greater connectivity strength is more likely to have higher reproducibility levels. But this does not necessarily mean that an ROI pair with high reproducibility also has high connectivity strength. It was observed from Figure 5 that some ROI pairs showing substantial reproducibility have low CC values, and some pairs with poor reproducibility have high CC values. In addition, there are more ROI pairs showing high CC values with poor reproducibility and low CC values with substantial reproducibility when functionally connected voxels are used, compared with those using all voxels in the predefined ROIs when the ROI size is below $18 \times 18 \times 18 \text{ mm}^3$. Table 7 lists 37 ROI-pairs showing substantial reproducibility identified using functionally connected voxels with their average CC values. There are 14 ROI pairs with $CC > 0.4$, and only 4 pairs with $CC > 0.5$. The anatomical distances of these ROI

TABLE 4. NUMBERS OF ROI PAIRS SHOWING SUBSTANTIAL, MODERATE, AND POOR BETWEEN-SUBJECT REPRODUCIBILITY WHEN THE INTER-REGIONAL CORRELATION WAS CALCULATED USING FUNCTIONALLY CONNECTED VOXELS AND ALL VOXELS IN THE PREDEFINED ROIS WITH FOUR DIFFERENT SIZES

	Substantial reproducibility	Moderate reproducibility	Poor reproducibility
Connected voxels	37	311	978
ROI: $6 \times 6 \times 6 \text{ mm}^3$	26	213	1087
ROI: $10 \times 10 \times 10 \text{ mm}^3$	28	232	1066
ROI: $14 \times 14 \times 14 \text{ mm}^3$	14	258	1054
ROI: $18 \times 18 \times 18 \text{ mm}^3$	11	261	1054

TABLE 5. NUMBERS OF ROI PAIRS COUNTED UNDER DIFFERENT LEVELS OF INTER-REGIONAL CORRELATION

	$CC \geq 0.5$	$0.25 \leq CC < 0.5$	$CC < 0.25$
Connected voxels	130	468	728
ROI: $6 \times 6 \times 6 \text{ mm}^3$	13	288	1025
ROI: $10 \times 10 \times 10 \text{ mm}^3$	50	413	863
ROI: $14 \times 14 \times 14 \text{ mm}^3$	84	556	686
ROI: $18 \times 18 \times 18 \text{ mm}^3$	116	713	497

CC, correlation coefficient.

pairs range from 27.58 to 108.9 mm. More specifically, the anatomical distances of 4 pairs are below 50 mm, 22 pairs have distances between 50 and 90 mm, and the remaining 11 pairs have distances above 90 mm. Among the 37 ROI pairs, 4 pairs are from DMN and 1 pair is in SMN.

Figure 6(a) shows an average correlation map overlaid on an individual slice covering part of DMN. It was obtained by averaging the CC values calculated between the seed in PCC and all voxels in this slice across all subjects and sessions. Figures 6(b–d) are voxel-wise ICC maps overlaid on the same slice as shown in Figure 6(a). These ICC maps were generated using CC values computed using the same PCC seed from different pairs of sessions of the six subjects. The ICC map in Figure 6(b) was obtained using the first two sessions that were separated by ~ 55 days, the map in Figure 6(c) was obtained using the first and fifth sessions with a time interval about 220 days, and Figure 6(d) was cal-

culated using the first and last sessions with an average interval of 463 days. It was observed that regions with a higher correlation to the seed do not necessarily demonstrate higher reproducibility. For example, a high average correlation can be observed at the PCC region in this slice, which does not exhibit high reproducibility in the same area.

Session temporal signal-to-noise ratio

Figure 7 shows a comparison between the session tSNR computed using the functionally connected voxels and those using all voxels in the predefined ROIs with the four sizes. Most of the session tSNR values calculated using functionally connected voxels identified by the SVM-based method are higher than those computed using all voxels in the predefined ROIs except when the ROI size is $18 \times 18 \times 18 \text{ mm}^3$. Figure 8 shows the average session tSNR values calculated across (Fig. 8(a)) sessions (within-subject average) and (Fig. 8(b)) subjects (between-subject average). The within- and between-subject average of session tSNR values computed from functionally connected voxels are higher than those obtained using all voxels in the predefined ROIs when the ROI size is less than $18 \times 18 \times 18 \text{ mm}^3$.

Head motion effects

Figure 9 shows the scatter plots of the average RMS framewise displacement versus the average inter-regional correlation computed using functionally connected voxels (Fig. 9(a)), and all voxels in the predefined ROIs of four

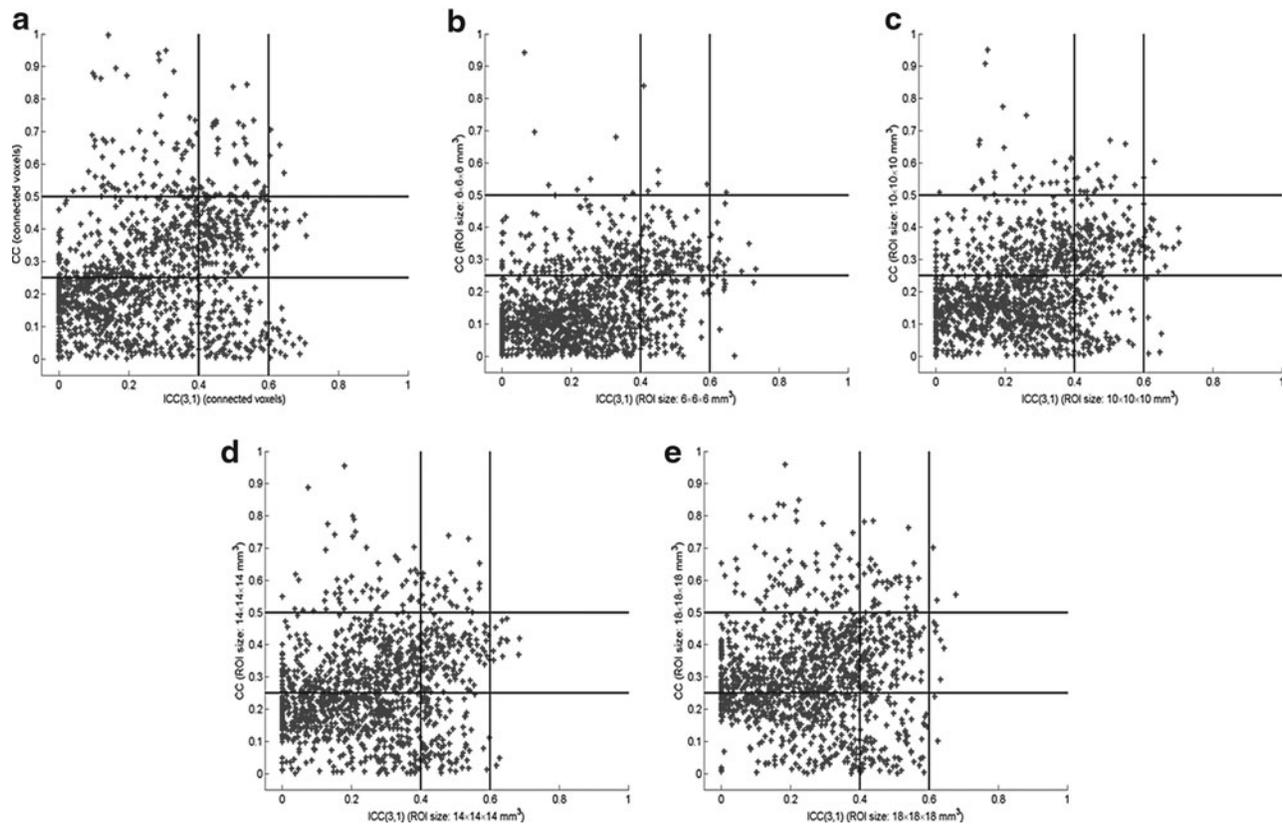


FIG. 5. Scatter plots of 1326 ROI pairs in the ICC-correlation coefficient (CC) domain. ICC(3,1) and CC values were computed using (a) functionally connected voxels, and all voxels in each predefined ROI with a size of (b) $6 \times 6 \times 6 \text{ mm}^3$, (c) $10 \times 10 \times 10 \text{ mm}^3$, (d) $14 \times 14 \times 14 \text{ mm}^3$, and (e) $18 \times 18 \times 18 \text{ mm}^3$.

TABLE 6. NUMBERS OF ROI PAIRS SHOWING SUBSTANTIAL OR MODERATE REPRODUCIBILITY COUNTED UNDER THE THREE DIFFERENT LEVELS OF INTER-REGIONAL CORRELATION

	<i>ICC > 0.4</i>		
	<i>CC ≥ 0.5</i>	<i>0.25 ≤ CC < 0.5</i>	<i>CC < 0.25</i>
Connected voxels	51	175	122
ROI: $6 \times 6 \times 6 \text{ mm}^3$	6	122	111
ROI: $10 \times 10 \times 10 \text{ mm}^3$	19	151	90
ROI: $14 \times 14 \times 14 \text{ mm}^3$	28	149	95
ROI: $18 \times 18 \times 18 \text{ mm}^3$	37	144	91

different sizes (Figs. 9(b–e)) over all subjects and sessions. No significant correlation between the motion effect and the inter-regional correlation was observed at the 0.05 significance level. In addition, the inter-regional correlation computed using the functionally connected voxels shows the least correlation with this motion measurement. Figure 10 shows the scatter plots of the maximum RMS framework displacement versus the average inter-regional correlation where Figures 10(a–e) follow the same order as that of Figure 9. The correlation analysis results indicate that there is no significant correlation between this motion measurement and the inter-regional correlation at the 0.05 level. The inter-regional correlation computed from functionally connected voxels shows the least correlation with the motion effect.

Discussion

The study of individual subjects indicates that substantial or moderate long-term within-subject reproducibility can be achieved for rsfMRI. If all voxels in the predefined ROIs are used, a larger ROI size provides higher within-subject reproducibility. With an increase in the ROI size, more voxels are used to calculate the average time course of each ROI, and thus the session-specific variations are expected to be further suppressed. It was observed that when only functionally connected voxels were used to calculate the inter-regional correlation between ROIs, increased within-subject reproducibility was obtained as compared with those obtained from all voxels in the predefined ROIs. This finding is not surprising because in analyses that include all voxels in the predefined ROIs, the voxels that do not belong to the same network could lead to higher inter-session variation. It was noted that the average ICC values obtained from the largest

ROI are still lower than those calculated from functionally connected voxels identified by the SVM-based method. The results suggest that the use of only functionally connected voxels in a functional network is preferable. Similarly, the inter-subject study results indicate that the use of functionally connected voxels results in more ROI pairs with acceptable between-subject reproducibility than the use of all voxels in the predefined ROIs. Our results suggest that the within- and between-subject reproducibility of the rsfMRI experiments are likely underestimated in previous studies where the inter-regional correlation was calculated from all voxels in predefined ROIs.

The numbers of functionally connected voxels detected by the SVM-based method (even after excluding voxels outside the $18 \times 18 \times 18 \text{ mm}^3$ ROI masks) are quite different across the ROIs, ranging from 12 to 679 with an average of 176. Note that the corresponding within-subject reproducibility measure is still higher than that using all 343 voxels in the $14 \times 14 \times 14 \text{ mm}^3$ ROIs, and all 729 voxels in the $18 \times 18 \times 18 \text{ mm}^3$ ROIs. This is expected because the functionally connected voxels carry more power from signals-of-interest. In the inter-subject study, the use of functionally connected voxels results in more ROI pairs showing substantial and moderate reproducibility compared to the analyses using all voxels in the predefined ROIs. The data-driven SVM method can ensure that the identified voxels are relevant to the network-of-interest, resulting in reliable estimations of connectivity strength. On the other hand, the predefined ROIs may include voxels that are not part of the network-of-interest, leading to unreliable estimations of connectivity strength.

The four ROI sizes used in this work are close to those used in previous reproducibility studies. For instance, an ROI size of $12 \times 12 \times 12 \text{ mm}^3$ (cubic) was used in the work of Chou and associates (2012), and a size of 5 mm (sphere) was considered in another two studies (Shehzad et al., 2009; Wang et al., 2011). Varying-size ROIs defined based on the anatomical automatic labeling atlas (AAL) (Tzourio-Mazoyer et al., 2002), Harvard-Oxford atlas (Kennedy et al., 1998), or functional ROIs (Shirer et al., 2012) have also been widely used (Braun et al., 2012; Chou et al., 2012; Fiecas et al., 2013; Liang et al., 2011; Wang et al., 2011). Although the effects of ROI size on rsfMRI reproducibility measures were rarely investigated in detail in previous studies, some information can still be identified from existing results. For example, from the work of Wang and associates (2011) it was found that the reproducibility measures computed from the AAL template is in general greater than those computed

TABLE 7. ROI PAIRS SHOWING SUBSTANTIAL REPRODUCIBILITY AND THEIR INTER-REGIONAL CORRELATION

<i>ROI pairs</i>	<i>ICC</i>	<i>CC</i>									
RPHG-RITG	0.6	0.26	AMPFC-RPHG	0.6	0.11	LPC-RPHG	0.67	-0.04	LIPL-RSTG	0.63	0.41
LSN-RITG	0.63	0.07	PCC-RPHG	0.65	0.21	RMFG-RPHG	0.69	-0.02	LIPL-LLG	0.64	0.42
LLG-RITG	0.65	0.46	RLPC-RPHG	0.63	0.09	SMA-RPHG	0.63	-0.09	MCC-BLG	0.64	0.16
RLG-RITG	0.62	0.46	BLG-RPHG	0.69	-0.06	LS1-RPHG	0.6	-0.11	LIPL-BCG	0.61	0.42
BLG-RITG	0.7	0.44	BCG-RPHG	0.61	-0.09	RVPMC-RPHG	0.64	-0.12	RPC-BCG	0.63	0.66
BCG-RITG	0.65	0.41	PVC-RPHG	0.66	-0.08	RSTG-VMPFC	0.71	0.38	RPC-PVC	0.6	0.62
PVC-RITG	0.69	0.43	LIPS-RPHG	0.63	-0.02	LSTG-VMPFC	0.64	0.41	LPMC-LIFG	0.62	0.34
RIPS-RITG	0.64	0.57	RIPS-RPHG	0.65	0.01	LSN-PCC	0.61	0.06	LS1-LIFG	0.62	0.35
RIPL-RITG	0.65	0.41	RPC-RPHG	0.7	-0.05	RMFG-LSN	0.62	-0.03	LVPMC-LIFG	0.62	0.38
									RCD-RPMC	0.61	0.71

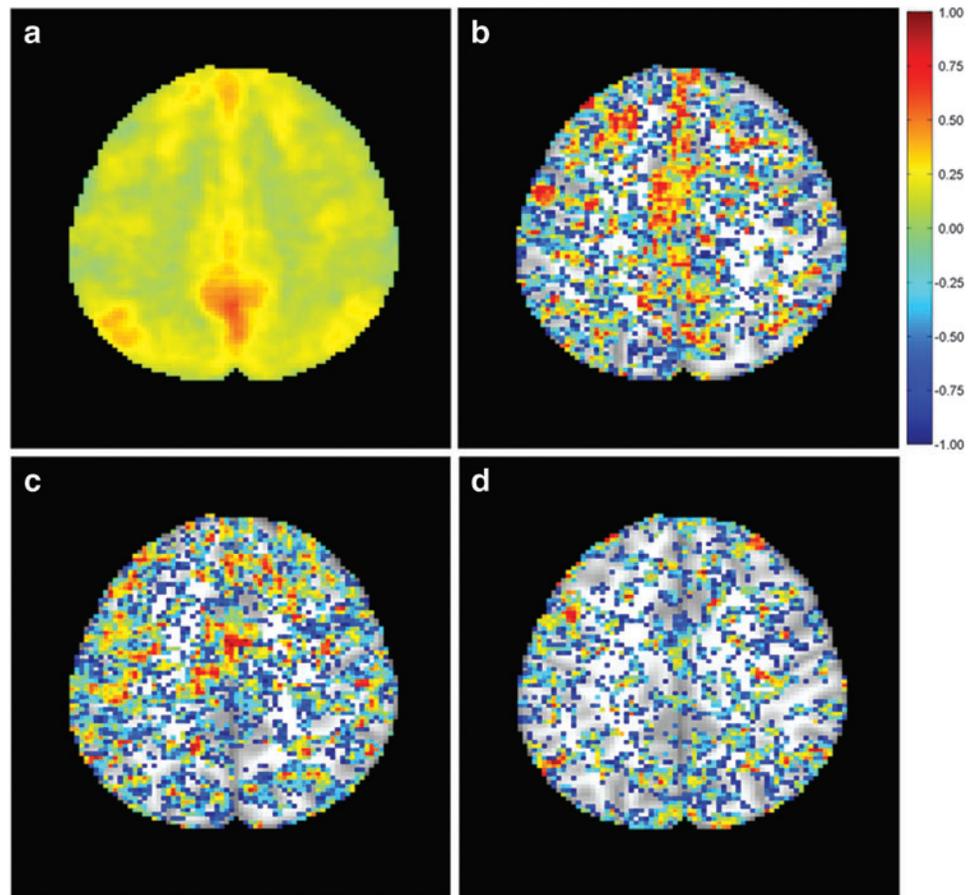


FIG. 6. (a) Average correlation between a seed in posterior cingulate cortex and voxels in a slice covering part of default mode network (DMN). (b) ICC map overlaid on the same slice computed from the first two sessions of the six subjects. (c) ICC map computed using the first and fifth sessions of the six subjects. (d) ICC map calculated using the first and ninth sessions of the six subjects. Color images available online at www.liebertpub.com/brain

from 5 mm sphere functional ROIs. Considering the sizes of ROIs defined from AAL are usually greater than 5 mm, this observation is consistent to our findings. An ROI size of $18 \times 18 \times 18 \text{ mm}^3$ is not typical in most ROI-based studies, and the major reason of investigating it was to examine how an increase of the ROI size would affect the reproducibility measures. In addition, the sizes of some ROIs defined from AAL or other anatomical templates are close to $18 \times 18 \times 18 \text{ mm}^3$ although they have different shapes and their anatomical locations do not exactly match. This work may provide an initial guidance about possible effects of using large ROIs in rsfMRI reproducibility studies.

An interesting finding of this study is that the relationship between the ROI size and reproducibility is not linear. An increase in ROI size would result in more ROI pairs showing higher between-subject reproducibility. However, when the ROI size is further increased, more irrelevant voxels are included, and in this case the increase in between-voxel variations may be more pronounced than the decrease in across-session and across-subject fluctuations. As a result, the number of ROI pairs with substantial reproducibility starts to decrease after the ROI size is increased beyond a certain point (e.g., $>10 \times 10 \times 10 \text{ mm}^3$). Our findings suggest that a large ROI size might lead to underestimated reproducibility measures in the predefined ROI-based reproducibility studies. In addition, a large ROI size reduces the spatial specificity of network connectivity, making the reproducibility study results less reliable.

In the predefined ROI-based analysis, the turning point of the ROI size for the number of ROI pairs showing substantial

reproducibility is $10 \times 10 \times 10 \text{ mm}^3$, which contains 125 voxels. Compared to the number of voxels in the ROIs of the other three sizes, this number is the closest one to the average number (176) of functionally connected voxels identified by the SVM-based method. But it does not mean that the use of this ROI size can achieve the same level of reproducibility as the use of identified functionally connected voxels.

RsfMRI reproducibility is influenced by multiple factors, such as the time interval between scan sessions, short-term or long-term, motion correction, and existence of physiological noise and other artifacts. The experimental results from this study indicate that reproducibility is partially connected to connectivity strength of ROI pairs. Greater connectivity strength has a higher chance to exhibit moderate or substantial reproducibility. This is consistent to the observations in previous reproducibility studies (Patriat et al., 2013; Shehzad et al., 2009; Wang et al., 2011). On the other hand, the study of ROI pairs in the seven networks indicates that higher connectivity strength does not necessarily correspond to higher reproducibility. This is verified by the distribution of CC values in the ICC-CC plane shown in Figure 5, and the list of specific ROI pairs shown in Table 7. This finding implies relatively large variations of inter-regional correlations across subjects, which is also reflected by the CV values obtained in the inter-subject study.

The anatomical distance of each ROI pair, which is defined as the Euclidean distance between the centroids of two ROIs, was computed for all 37 ROI pairs showing substantial reproducibility listed in Table 7. The distances range

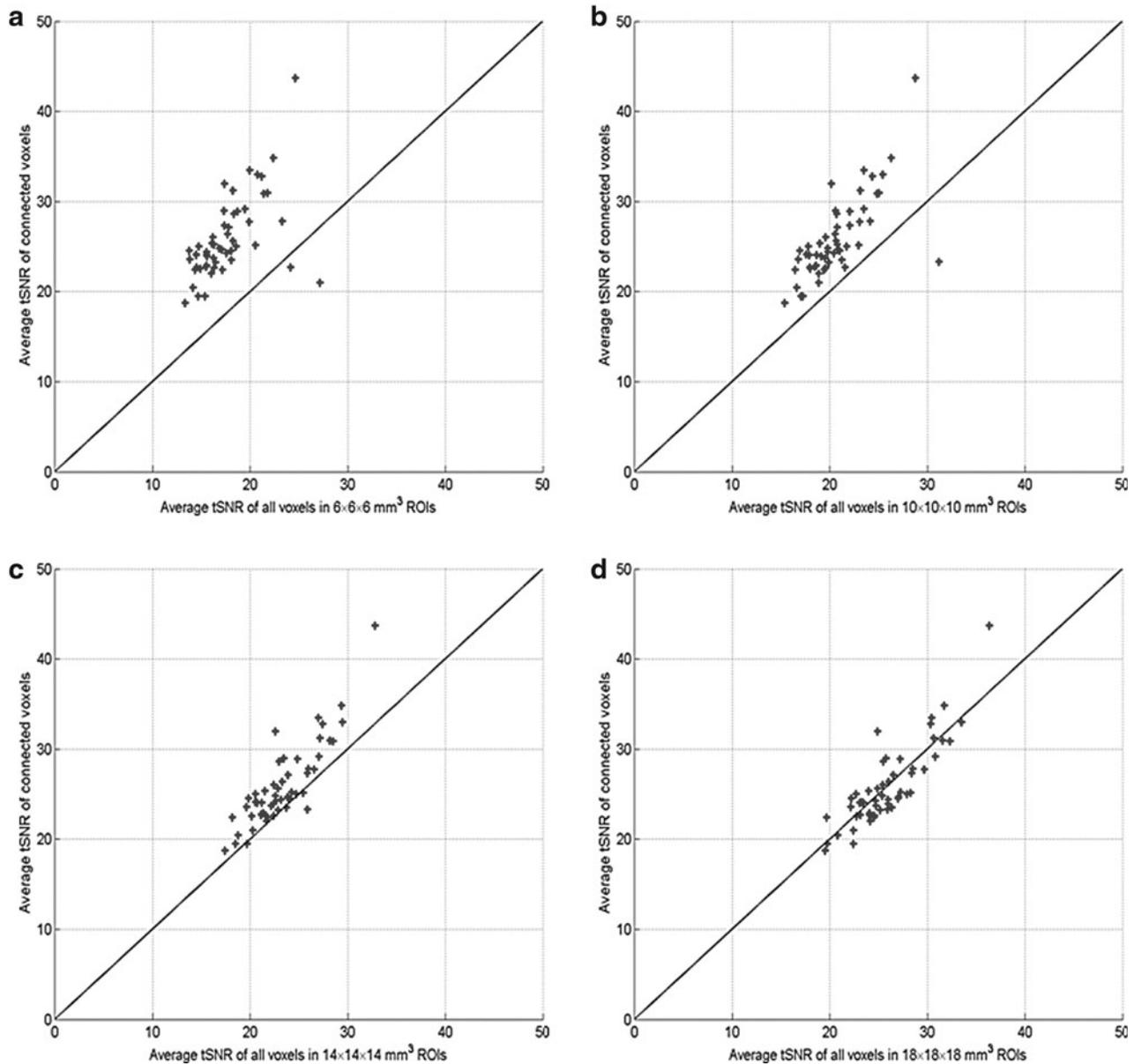


FIG. 7. A comparison of temporal signal-to-noise-ratio (tSNR) of all 54 sessions from the six subjects computed using functionally connected voxels and all voxels in the predefined ROIs with a size of (a) $6 \times 6 \times 6 \text{ mm}^3$, (b) $10 \times 10 \times 10 \text{ mm}^3$, (c) $14 \times 14 \times 14 \text{ mm}^3$, and (d) $18 \times 18 \times 18 \text{ mm}^3$.

from 27.58 to 108.9 mm with a median of 71.41 mm. Among these ROI pairs, 10.81% of them have anatomical distances below 50 mm, 59.46% have distances between 50 and 90 mm, and the remaining 29.73% have distances above 90 mm. This indicates that a majority of ROI pairs showing substantial reproducibility have moderate inter-regional anatomical distances. In addition, there are more ROI pairs in DMN than other investigated networks showing substantial reproducibility. This implies that brain regions from DMN could be considered with a high priority in rsfMRI reproducibility studies.

It was confirmed in this study that ICC values are apparently influenced by the time interval between sessions. With an increased time interval, the overall area showing high reproducibility decreases, as shown in Figure 6. This in-

dicates that a long-term study exhibits relatively low reproducibility compared with a short-term one. In addition, regions showing higher short-term reproducibility do not necessarily overlap with regions showing higher middle- or long-term reproducibility. Moreover, even at a similar time interval, the regions showing higher reproducibility in a certain session pair only partially overlap with regions showing higher reproducibility in another session pair (e.g., sessions 1, 2 vs. sessions 2, 3, or sessions 1, 3 vs. sessions 3, 5, etc.).

The session tSNR is an indicator of the rsfMRI reliability, and a higher session tSNR value implies an increased reliability of fMRI data. The experimental results show that the session tSNR computed from functionally connected voxels is greater than those calculated using all voxels in the predefined ROIs when the ROI size is not very large.

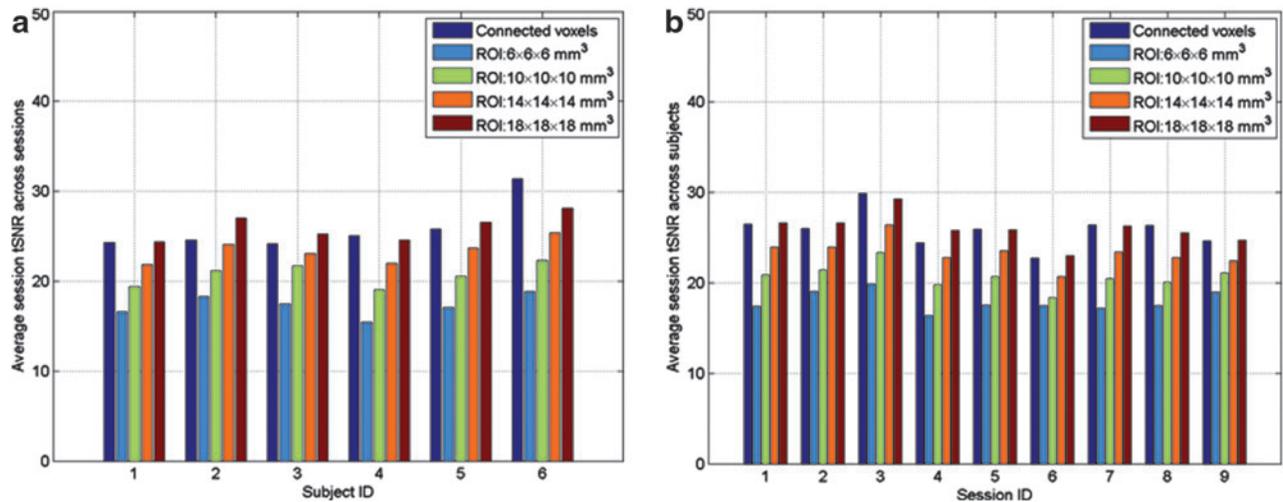


FIG. 8. The average session tSNR calculated across: (a) nine sessions for each of the six subjects, and (b) six subjects for each of the nine sessions. The session tSNR values were calculated using functionally connected voxels detected by the SVM-based method and all voxels in the predefined ROIs with the four different sizes. Color images available online at www.liebertpub.com/brain

This indicates that the temporal variation of rsfMRI can be better represented by functionally connected voxels than all voxels in the predefined ROIs. In the analyses using the predefined ROIs, a larger ROI size leads to a higher tSNR in each session. But as mentioned before, a large ROI size

may lose the specificity of functional connectivity and its application could be limited.

The fMRI datasets used in this work are the same as those used in a previous reproducibility study (Chou et al., 2012). However, the approaches used here are somewhat different

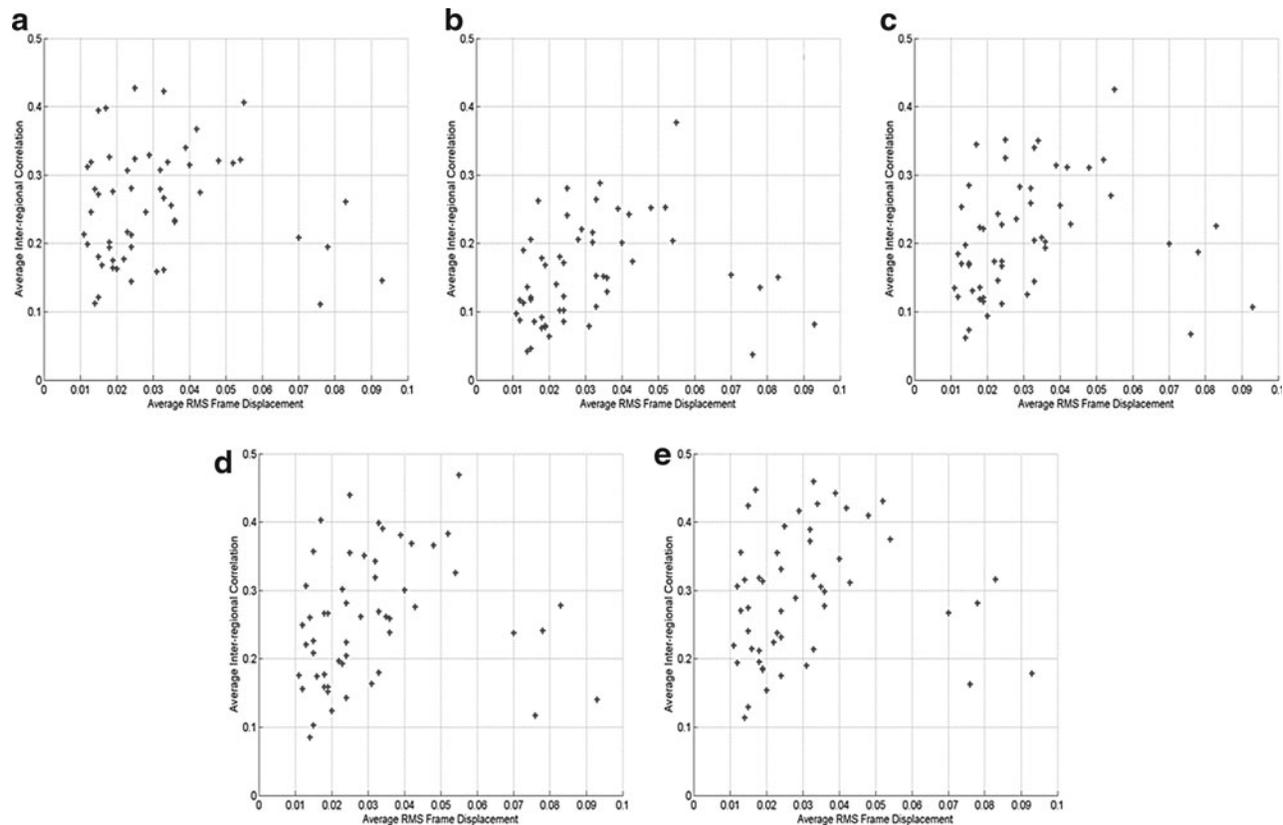


FIG. 9. Scatter plots of the average inter-regional correlation versus the average root-mean-square (RMS) frame displacement for all subjects and sessions. The correlation was computed using (a) functionally connected voxels ($r = -0.016$, $p = 0.908$), and all voxels in each predefined ROI with a size of (b) $6 \times 6 \times 6 \text{ mm}^3$ ($r = 0.172$, $p = 0.214$), (c) $10 \times 10 \times 10 \text{ mm}^3$ ($r = 0.154$, $p = 0.267$), (d) $14 \times 14 \times 14 \text{ mm}^3$ ($r = 0.151$, $p = 0.276$), and (e) $18 \times 18 \times 18 \text{ mm}^3$ ($r = 0.135$, $p = 0.332$).

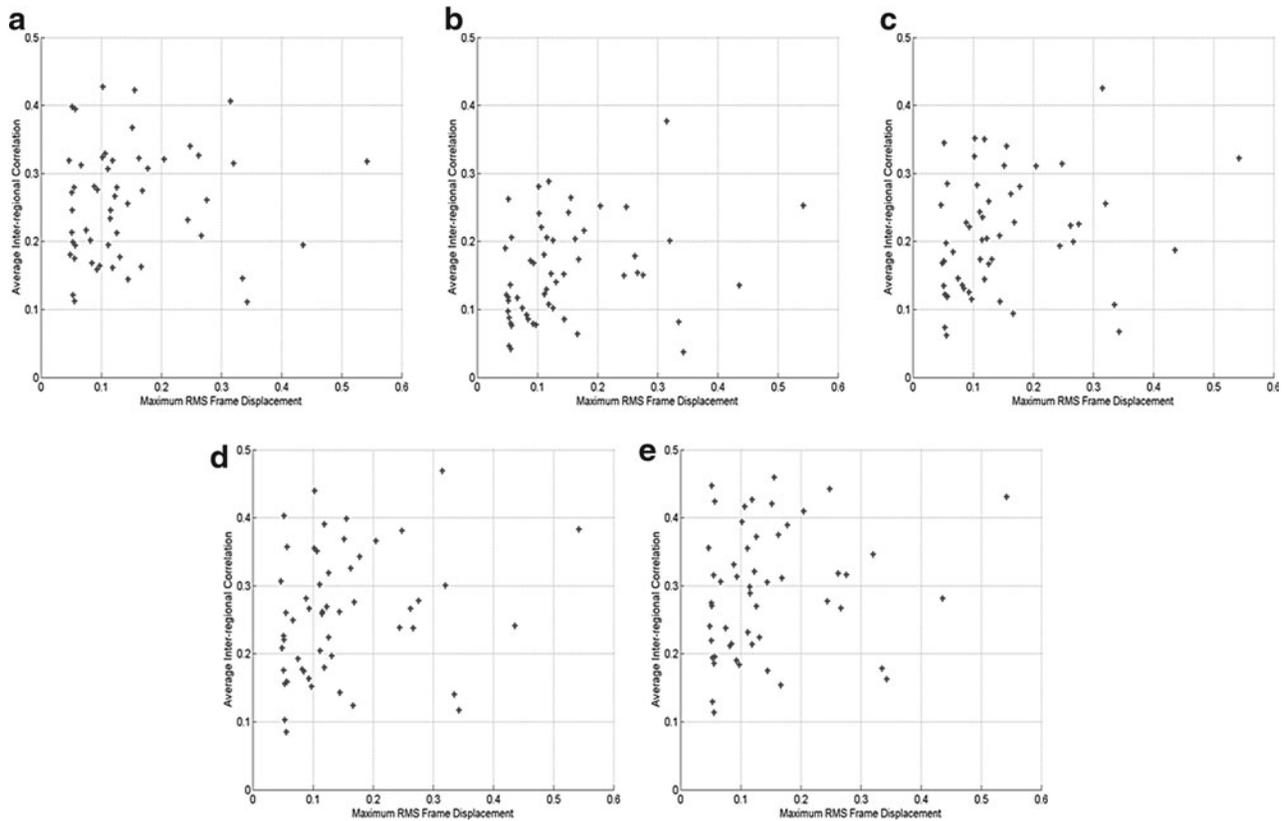


FIG. 10. Scatter plots of the average inter-regional correlation versus the maximum RMS frame displacement for all subjects and sessions. The correlations was computed using (a) functionally connected voxels ($r=0.075$, $p=0.591$), and all voxels in each predefined ROI with a size of (b) $6 \times 6 \times 6 \text{ mm}^3$ ($r=0.262$, $p=0.056$), (c) $10 \times 10 \times 10 \text{ mm}^3$ ($r=0.229$, $p=0.096$), (d) $14 \times 14 \times 14 \text{ mm}^3$ ($r=0.221$, $p=0.108$), and (e) $18 \times 18 \times 18 \text{ mm}^3$ ($r=0.208$, $p=0.131$).

from the previous one. In the work of Chou and associates (2012), the ROI size of $12 \times 12 \times 12 \text{ mm}^3$ was used for the predefined ROI-based reproducibility study of DMN where 16 nodes were investigated, and 116 ROIs were defined using the AAL template to study the reproducibility of connectivity strength of 6670 ROI pairs. In addition, in the preprocessing, the removal of physiological noises in white matter and cerebral spinal fluid and the scrubbing in motion correction were not performed, and the spatial smoothing methods were also different. It was found in our work that the physiological noise may bring additional inter-regional correlation between ROIs and these preprocessing steps resulted in significant differences in the computed connectivity strength and reproducibility measures compared to those without the physiological noise removal and scrubbing. Therefore, the numerical results in our study are not directly comparable to the previous work using the same datasets. However, the relatively high within-subject reproducibility and low between-subject reproducibility were identified in both studies.

This study has a few limitations. First, the number of subjects is not very large, and thus the findings need to be confirmed in future studies with a larger subject population. Second, the seed positions in the seven networks were obtained from previous works. Recent studies show that changes in seed positions may potentially lead to different network mapping results (Ma et al., 2007; Yan et al., 2013; Yang et al., 2013), and how this would affect the repro-

ducibility measures needs more investigation. Third, some factors that may contribute to inter-session variations were not quantified, such as residual movement artifacts after motion correction, registration inaccuracy after spatial normalization, head position changes across scan sessions, and brain plasticity. Finally, over the 1.5 years' period of data collection, the changes of the MRI scanner's physical properties and the upgrade of data acquisition software may potentially introduce additional variations to the acquired data. Our study was conducted with a research scanner where several other longitudinal MRI projects were performed at the same time. To minimize possible effects from this factor, daily quality assurance examinations were performed to ensure that the hardware performance was consistent. In addition, structural and functional MRI data acquired immediately before and immediately after system maintenance/software upgrade were compared, to ensure that the MRI quality is minimally affected.

Conclusion

In this work, the long-term rsfMRI reproducibility was investigated based on connectivity strength of selected ROI pairs in seven major functional networks. The rsfMRI reproducibility measures were calculated from (1) functionally connected voxels identified by a brain mapping technique and (2) all voxels in the predefined ROIs with four different

sizes. Experimental fMRI data collected from six subjects over 1.5 years' time were used in the study, and a newly developed SVM-based data-driven method was used to identify functionally connected voxels in seven major networks. Experimental results indicate that substantial or moderate long-term within-subject reproducibility and acceptable between-subject reproducibility can be obtained under resting-state when only functionally connected voxels were used to calculate the reproducibility measures. On the other hand, in the ROI-based study where all voxels in each predefined ROI were included in the analysis, an increase in ROI size leads to increased tSNR and inter-regional correlation, but the number of ROI pairs showing substantial reproducibility increases initially and then decreases with the ROI size. In addition, the reproducibility measured from only functionally connected voxels was generally higher than that measured from all voxels in the predefined ROIs when using typical ROI sizes. It was also observed that ROI pairs showing high connectivity strength are more likely to have high reproducibility. The experimental results suggest that (1) the conventional predefined ROI-based analyses could underestimate the reproducibility of rsfMRI, particularly when the ROI size is small or very large, and (2) rsfMRI reproducibility should be assessed by analyzing only identified functionally connected voxels that more precisely reflect functional connectivity in the networks-of-interest.

Acknowledgments

This research was partially supported by NIH R01-NS037992 grant (to L.P.P.) and NIH R01-NS074045 grant (to N.-K.C.).

Author Disclosure Statement

No competing financial interests exist.

References

- Achard S, Bullmore E. 2007. Efficiency and cost of economical brain functional networks. *PLoS Comput Biol* 3:e17.
- Allen E, Erhardt E, Damaraju E, et al. 2011. A baseline for the multivariate comparison of resting-state networks. *Front Syst Neurosci* 5:1–23.
- Amann M, Hirsch J, Gass A. 2009. A serial functional connectivity MRI study in healthy individuals assessing the variability of connectivity measures: reduced interhemispheric connectivity in the motor network during continuous performance. *Magn Reson Imaging* 27:1347–1359.
- Avants BB, Tustison N, Song G. 2009. Advanced Normalization Tools (ANTS). *Insight J* 2:1–35.
- Beckmann C, DeLuca M, Devlin J, Smith S. 2005. Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc B Lond B Biol Sci* 360:1001–1013.
- Bennett C, Miller M. 2010. How reliable are the results from functional magnetic resonance imaging. *Ann N Y Acad Sci* 1191:133–155.
- Biswal B, Yetkin F, Haughton V, et al. 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med* 34:537–541.
- Braun U, Plichta M, Christine Esslinger C, et al. 2012. Test-retest reliability of resting-state connectivity network characteristics using fMRI and graph theoretical measures. *Neuroimage* 59:1404–1412.
- Broyd S, Demanuele C, Debener S, Helps S, James C, Sonuga-Barke E. 2009. Default-mode brain dysfunction in mental disorders: a systematic review. *Neurosci Biobehav Rev* 33:279–296.
- Burges C. 1998. A tutorial on support vector machines for pattern recognition. *Data Min Knowl Discov* 2:121–167.
- Chou Y, Panych L, Dickey C, Petrella J, Chen N. 2012. Investigation of long-term reproducibility of intrinsic connectivity network mapping: a resting-state fMRI study. *AJNR Am J Neuroradiol* 33:833–838.
- Damoiseaux J, Rombouts S, Barkhof F, et al. 2006. Consistent resting-state networks across healthy subjects. *Proc Natl Acad Sci U S A* 103:13848–13853.
- Ferreira L, Busatto G. 2013. Resting-state functional connectivity in normal brain aging. *Neurosci Biobehav Rev* 37:384–400.
- Fiecas M, Ombao H, van Lunen D, et al. 2013. Quantifying temporal correlations: a test-retest evaluation of functional connectivity in resting-state fMRI. *Neuroimage* 65:231–241.
- Fox M, Greicius M. 2010. Clinical applications of resting state functional connectivity. *Front Syst Neurosci* 4:1–13.
- Fox M, Raichle M. 2007. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 8:700–711.
- Friston KJ, Williams S, Howard R, Frackowiak RS, Turner R. 1996. Movement-related effects in fMRI time-series. *Magn Reson Med* 35:346–355.
- Gasparovic C, Bedrick E, Mayer A, et al. 2011. Test-retest reliability and reproducibility of short-echo-time spectroscopic imaging of human brain at 3T. *Magn Reson Med* 66:324–332.
- Greicius M, Krasnow B, Reiss A, Menon V. 2003. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc Natl Acad Sci U S A* 100:253–258.
- Guo C, Kurth F, Zhou J, Mayer E, Eickhoff S, et al. 2012. One-year test-retest reliability of intrinsic connectivity network fMRI in older adults. *Neuroimage* 61:1471–1483.
- He Y, Chen Z, Evans A. 2008. Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer's disease. *J Neurosci* 28:4756–4766.
- Honey C, Sporns O, Cammoun L, et al. 2009. Predicting human resting-state functional connectivity from structural connectivity. *Proc Natl Acad Sci U S A* 106:2035–2040.
- Jenkinson M, Bannister P, Brady J, Smith S. 2002. Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17:825–841.
- Joel S, Caffo B, Van Zijl P, Pekar J. 2011. On the relationship between seed-based and ICA-based measures of functional connectivity. *Magn Reson Med* 66:644–657.
- Kennedy DN, Lange N, Makris N, Bates J, Meyer J, Caviness VS Jr. 1998. Gyri of the human neocortex: an MRI-based analysis of volume and variance. *Cereb Cortex* 8:372–384.
- Kollndorfer K, Fischmeister F, Kasprian G, Prayer D, Schöpf V. 2013. A systematic investigation of the invariance of resting-state network patterns: is resting-state fMRI ready for pre-surgical planning? *Front Hum Neurosci* 7:95.
- Kristo G, Rutten G, Raemaekers M, Gelder B, Rombouts S, Ramsey N. 2014. Task and task-free fMRI reproducibility comparison for motor network identification. *Hum Brain Mapp* 35:340–352.
- Laird A, Fox PM, Eickhoff S, et al. 2011. Behavioral interpretations of intrinsic connectivity networks. *J Cogn Neurosci* 12:4022–4037.

- Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33:159–174.
- Lee M, Smyser C, Shimony J. 2013. Resting-State fMRI: a review of methods and clinical applications. *AJNR Am J Neuroradiol* 34:1866–1872.
- Liang X, Wang J, Yan C, et al. 2012. Effects of different correlation metrics and preprocessing factors on small-world brain functional networks: a resting-state functional MRI study. *PLoS One* 7:e32766.
- Ma L, Wang B, Chen X, Xiong J. 2007. Detecting functional connectivity in the resting brain: a comparison between ICA and CCA. *Magn Reson Imaging* 25:47–56.
- Maneshi M, Moeller F, Fahoum F, et al. 2012. Resting-state connectivity of the sustained attention network correlates with disease duration in idiopathic generalized epilepsy. *PLoS One* 7:e50359.
- Markett S, Reuter M, Montag C, et al. 2014. Assessing the function of the fronto-parietal attention network: insights from resting-state fMRI and the attentional network test. *Hum Brain Mapp* 35:1700–1709.
- Meindl T, Teipel S, Elmouden R, et al. 2010. Test-retest reproducibility of the default-mode network in healthy individuals. *Hum Brain Mapp* 31:237–246.
- Murphy K, Bodurka J, Bandettini PA. 2006. How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. *Neuroimage* 34:565–574.
- Patriat R, Molloy EK, Meier TB, Kirk GR, Nair VA, Meyerand ME, Prabhakaran V, Birn RM. 2013. The effect of resting condition on resting-state fMRI reliability and consistency: a comparison between resting with eyes open, closed, and fixated. *Neuroimage* 78:463–473.
- Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59:2142–2154.
- Schmidt S, Akrofi K, Carpenter-Thompson J, Husain F. 2013. Default mode, dorsal attention and auditory resting state networks exhibit differential functional connectivity in tinnitus and hearing loss. *PLoS One* 8:e76488.
- Schölkopf B, Platt J, Shawe-Taylor J, Smola A, Williamson R. 2001. Estimating the support of high-dimensional distribution. *Neural Comput* 13:1443–1471.
- Shehzad Z, Kelly A, Reiss P, et al. 2009. The resting brain: unconstrained yet reliable. *Cereb Cortex* 19:2209–2229.
- Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD. 2012. Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cereb Cortex* 22:158–165.
- Shrout P, Fleiss J. 1979. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 86:420–428.
- Sikka S, Cheung B, Khanuja R, et al. 2012. Towards Automated Analysis of Connectomes: The Configurable Pipeline for the Analysis of Connectomes (C-PAC). 5th INCF Congress of Neuroinformatics: 117.
- Soltysik D, Thomasson D, Rajan S, et al. 2011. Head-repositioning does not reduce the reproducibility of fMRI activation in a block-design motor task. *Neuroimage* 56:1329–1337.
- Song X, Chen N. 2014. A SVM-based quantitative fMRI method for resting state functional network detection. *Magn Reson Imaging* 32:819–831.
- Song X, Murphy M, Wyrwicz A. Spatiotemporal Denoising and Clustering of fMRI Data. In Proceedings of the IEEE International Conference on Image Processing, Atlanta, GA, 2006, p. 2857.
- Sonuga-Barke E, Castellanos F. 2007. Spontaneous attentional fluctuations in impaired states and pathological conditions: a neurobiological hypothesis. *Neurosci Biobehav Rev* 31:977–986.
- Toro R, Fox PT, Paus T. 2008. Functional coactivation map of the human brain. *Cereb Cortex* 18:2553–2559.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289.
- Uddin L, Kelly A, Biswal B, et al. 2009. Functional connectivity of default mode network components: correlation, anticorrelation, and causality. *Hum Brain Mapp* 30:625–637.
- Van de Ven V, Formisano E, Prvulovic D, et al. 2004. Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. *Hum Brain Mapp* 22:165–178.
- Vapnik V. 1998. *Statistical Learning Theory*. Wiley-Interscience: New York.
- Vlietier J, Majoie B, Leenstra S, Den Heeten J. 2004. Functional magnetic resonance imaging for neurosurgical planning in neurooncology. *Eur Radiol* 14:1143–1153.
- Wang JH, Zuo XN, Gohel S, Milham M, Biswal B, He Y. 2011. Graph theoretical analysis of functional brain networks: test-retest evaluation on short- and long-term resting-state functional MRI data. *PLoS One* 6:e21976.
- Watanabe T, Hirose S, Wada H, et al. 2013. A pairwise maximum entropy model accurately describes resting-state human brain networks. *Nat Commun* 4:1–10.
- Wisner K, Atluri G, Lim K, MacDonald A III. 2013. Neurometrics of intrinsic connectivity networks at rest using fMRI: retest reliability and cross-validation using a meta-level method. *Neuroimage* 76:236–251.
- Woodward ND, Rogers B, Heckers B. 2011. Functional resting-state networks are differentially affected in schizophrenia. *Schizophr Res* 130:86–93.
- Yan F, Wu C, Cheng S, Lim K, Hsu Y, Liu H. 2013. Resting-state functional magnetic resonance imaging analysis with seed definition constrained by regional homogeneity. *Brain Connect* 3:438–449.
- Yang L, Lin F, Zhou Y, Xu J, Yu C, Pan W, Lei H. 2013. Iterative cross-correlation analysis of resting state functional magnetic resonance imaging data. *PLoS One* 8:e58653.
- Ystad M, Eichele T, Lundervold A, Lundervold A. 2010. Subcortical functional connectivity and verbal episodic memory in healthy elderly—a resting state fMRI study. *Neuroimage* 52:379–388.

Address correspondence to:

Xiaomu Song
 Department of Electrical Engineering
 School of Engineering
 Widener University
 Kirkbride Hall
 Room 369
 One University Place
 Chester, PA 19013

E-mail: xmsong@widener.edu