Long-Term Test-Retest Reliability of Resting-State Networks in Healthy Elderly Subjects and Patients with Amnestic Mild Cognitive Impairment

Janusch Blautzika,∗, Daniel Keesera, Albert Bermana, Marco Paolinia, Valerie Kirschc, Sophia Muellera, Ute Coatesa, Maximilian Reiserb, Stefan J. Teipeld and Thomas Meinid

a Institute for Clinical Radiology, Ludwig-Maximilians-University Munich, Munich, Germany
b Department of Psychiatry and Psychotherapy, Ludwig-Maximilians-University, Munich, Germany
c Department of Neurology, Ludwig-Maximilians-University, Munich, Germany
d Department of Psychiatry, University of Rostock, and DZNE, German Center for Neurodegenerative Disorders, Rostock, Germany

Accepted 1 December 2012

Abstract. The investigation of cerebral resting-state networks (RSNs) by functional magnetic resonance imaging (fMRI) is a promising tool for the early diagnosis and follow-up of neuropsychiatric and neurodegenerative disorders like Alzheimer’s disease (AD). In this context, the determination of inter-session reliability of these networks is crucial. However, data on network reliability in healthy elderly subjects is rare and does not exist for patients with amnestic mild cognitive impairment (aMCI), a prodromal stage of AD. Therefore, the aim of this study was to investigate the long-term test-retest reliability of RSNs in both groups. Twelve healthy controls (HC) and 13 aMCI patients underwent resting-state (fMRI) and neuropsychological testing (CERAD test battery) twice, at baseline and after 13–16 months. Resting-state fMRI data was decomposed into independent components using independent component analysis. Inter-session test-retest reliability of the resulting RSNs was determined by calculating voxel-wise intra-class correlation coefficients. Overall test-retest reliability of corresponding RSNs was moderate to high in both groups, but significantly higher in the HC group compared to the aMCI group (p < 0.001), while the cognitive performance within the CERAD test battery remained stable over time in either group. Most reliable RSNs derived from the HC group and were associated with sensory and motor as well as higher order cognitive and the default-mode function. Particularly low reliability was found in basal frontal regions, which are known to be prone to susceptibility-induced noise. We conclude that stable RSNs may represent healthy aging, whereas decreased RSN reliability may indicate progressive neuro-functional alterations before the actual manifestation of clinical symptoms.

Keywords: Default-mode network, functional magnetic resonance imaging, independent component analysis, intra-class correlation coefficient, mild cognitive impairment, resting-state networks, test-retest reliability

INTRODUCTION

Resting-state functional magnetic resonance imaging (fMRI) has evolved as a powerful tool for mapping functional connectivity in large-scale neural
networks. These resting-state networks (RSNs) represent coherent patterns of spontaneous low-frequency (0.01–0.1 Hz) fluctuations in the blood-oxygen level dependent (BOLD) signal and strongly overlap with cortical regions known to be involved in primary and higher cognitive functions [1–7]. Over the past years, RSNs have become an important field of research since a growing number of cross-sectional studies have reported alterations in these intrinsic connectivity systems in various neuropsychiatric and neurodegenerative disorders [8–11]. For instance, significant disruptions of a connectivity pattern referred to as the default-mode network (DMN) have been described for pathologic conditions such as Alzheimer’s disease (AD) [12–14] and amnestic mild cognitive impairment (aMCI) [15, 16], which is believed to represent a prodromal form of AD [17]. Specific changes in the connectivity pattern of the DMN have been shown to allow the differentiation between healthy aging, aMCI, and AD [12, 16] and between different stages of AD [18] suggesting the possibility of applying that network as a non-invasive, readily available and radiation exposure free imaging marker for the detection of incipient AD or its follow-up [13, 18].

However, one important prerequisite for the application of resting-state fMRI for scientific and clinical purposes is the knowledge of the test-retest reliability of RSNs. Short-term test-retest reliability of RSNs has been already demonstrated to be moderate to high [1, 19, 20]. Longer-term inter-session studies addressing this question are still rare in healthy young [21] and elderly [22, 23] subjects and, to the best of our knowledge, have not yet been addressed in aMCI patients.

The present study investigates the long-term (>1 year) test-retest reliability of RSNs in healthy elderly subjects and aMCI patients in order to further establish the applicability of these imaging markers as a diagnostic tool to monitor subjects at increased risk of developing dementia.

METHODS
Participants

We prospectively studied 12 right-handed healthy elderly controls (HC) (4 female; mean age = 67.8 ± 7.3 y, range = 59–83 y; education = 12.8 ± 3.4 y) and 13 right-handed subjects with aMCI (6 female; mean age = 72.8 ± 7.3 y, range = 60–88 y; education = 11.2 ± 2.3 y). The participants were measured at baseline and on a follow-up examination after 12 to 16 months. All participants gave their written informed consent. Prior to the MRI examination, subjects were carefully screened for contraindications to the MRI exam (e.g., pacemaker) both verbally and by means of a questionnaire.

The study was carried out in accordance with the Declaration of Helsinki and approved by the local ethics committee.

All subjects underwent longitudinal clinical examination, laboratory testing, structural MRI, and neuropsychological testing at two sessions (baseline and follow-up, respectively).

Both sessions were separated by approximately 13–16 months (HC = 477 ± 175 days, range = 352–979 days; aMCI = 411 ± 47 days, range = 315–484 days). Amnestic MCI patients were recruited from the university’s Memory Clinic and met the Mayo clinic criteria [24] for aMCI. Healthy subjects were recruited among the spouses of patients; they had no subjective memory complaints and scored within 1 SD of the age and education adjusted means of the Mini-Mental-Status Examination (MMSE) [25], all items of the Consortium to Establish a Registry for Alzheimer’s Disease (CERAD) cognitive battery [26], the Clock-drawing-test [27], and the trail-making test [28].

White matter hyperintensities as shown on T2w FLAIR imaging exceeding 10 mm in diameter or more than 3 in number were defined as exclusion criteria to avoid participation of subjects with significant cerebrovascular lesions.

Statistical analysis of neuropsychological data

At both sessions, neuropsychological testing included the CERAD battery subtests (MMSE, verbal fluency, Boston naming test, word list learning, word list recall, word list recognition, constructional praxis, and recall of constructional praxis) and the Clock drawing test. Inter-session changes in neuropsychological data were analyzed using paired t-tests and inter-group differences were analyzed with independent t-tests as implemented in SPSS version 20.0. To adjust for multiple comparisons (9 neuropsychological tests), the significance threshold of \( p < 0.05 \) was Bonferroni-corrected resulting in \( p_{corr} < 0.0056 \) [29].

Magnetic resonance imaging

Imaging was performed using a 3.0 Tesla Magnetom (TRIO, Siemens, Erlangen, Germany) with...
For anatomical reference, a sagittal high-resolution magnetization-prepared rapid gradient-echo sequence (MPRAGE) was acquired with the following imaging parameters: field of view (FoV): 250 mm; voxel size: 0.8 × 0.8 × 0.8 mm³; time of repetition (TR): 1400 ms; time of echo (TE): 7.61 ms; flip angle (FA): 20°; number of slices: 160.

For detection of white matter lesions, a 2-dimensional T2-weighted fluid attenuation inversion recovery (FLAIR) sequence was acquired in axial orientation (FoV: 230 mm; voxel size: 0.9 × 0.9 × 5.0 mm³; TR: 9000 ms; TE: 117 ms; FA: 180°; number of slices: 28).

For functional imaging, the subjects were instructed to close their eyes and to think of nothing in particular, and not to fall asleep. During the examination, light in the scanner room was dimmed. There were no additional task-fMRI scans preceding the resting-state fMRI scan. Functional data were recorded by means of a BOLD sensitive echo-planar gradient-echo sequence in axial orientation (120 volumes; FoV: 192 mm; voxel size: 3 × 3 × 4 mm³; slice gap: 0.4 mm; imaging matrix: 64 × 64; TR: 3000 ms; TE: 30 ms; FA: 80°; number of slices: 36). Before starting imaging, 3D-field shimming was performed using automated shimming algorithms implemented in the scanner.

Statistical analyses of functional data


Data preprocessing

The initial five volumes of every functional run were discarded to account for T1 saturation effects. Pre-statistics processing was done in four steps: (1) Individual high-resolution T1-weighted images were processed using AFNI; (2) head motion correction was done using MCFLIRT (Motion Correction using FMRIB’s Linear Image Registration Tool) [30]; (3) the skull was removed using BET (Brain Extraction Tool) [31] followed by spatial smoothing using a 5-mm FWHM Gaussian kernel with high-pass temporal filtering (Gaussian-weighted least squares straight line fitting with sigma = 100 s); (4) registration of functional data to the corresponding individual high-resolution T1-weighted images and transformation to the MNI152 standard brain space was performed applying FLIRT (FMRIB’s Linear Image Registration) version 5.5 [30].

Independent component analysis

Functional data was analyzed using MELODIC (Multivariate Exploratory Linear Optimized Decomposition into Independent Components) software, version 3.10 in combination with a dual regression analysis as implemented in FSL 4.1.7 [32]. Dual regression has been recently validated to generate highly reliable and reproducible independent components (ICs) over the short- and long-term run [21].

In a main approach, time-courses of all participants (25) and occasions (2), for a total of 50 scans, were spatially concatenated in the MNI152 standard space to allow for the estimation of a covariance matrix, which was used for reduction of individual fMRI data sets via Probabilistic Independent Component Analysis. Reduced individual data sets of all participants and scans were further processed with TC-GICA (Temporal Concatenation Group ICA) to calculate multi-subject multi-session group-level components (hereinafter referred to as 1 TC-GICA approach). We allowed the MELODIC 3.10 software to automatically estimate the number of ICs.

In a complementary approach, group-level components were calculated for each group and each session separately (resulting in four different TC-GICA; hereinafter referred to as 4 TC-GICA approach). For these analyses, the number of ICs was fixed to the number automatically estimated in the 1 TC-GICA approach.

To be considered as RSNs, connectivity patterns were required to extent over functionally relevant brain areas equivalently to RSNs detected in previous studies [1, 14, 21, 33], and to contain signals of ultra-slow frequencies between 0.01–0.1 Hz. This frequency range has been shown to be characteristic for RSNs using BOLD fMRI [2, 3, 5–7, 34]. Connectivity patterns not meeting these criteria were regarded as artifacts.

Test-retest approach

All TC-GICA derived group-level ICs containing RSNs were reconstructed into individual ICs for each participant and session applying dual regression [14, 21, 35]. To evaluate the test-retest reliability of RSNs, voxel-wise intra-class correlation coefficients (ICCs) were calculated for corresponding connectivity patterns based on individual-level ICs [36].
Table 1 Results of the neuropsychological testing

<table>
<thead>
<tr>
<th></th>
<th>HC baseline</th>
<th>HC follow-up</th>
<th>aMCI baseline</th>
<th>aMCI follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>29.17 ± 0.72</td>
<td>28.67 ± 1.30</td>
<td>26.46 ± 1.2</td>
<td>26.46 ± 1.33</td>
</tr>
<tr>
<td>Verbal fluency</td>
<td>24.92 ± 7.01</td>
<td>23.5 ± 5.95</td>
<td>20.38 ± 4.31</td>
<td>17.54 ± 4.91</td>
</tr>
<tr>
<td>Boston naming</td>
<td>14.42 ± 1.44</td>
<td>14.02 ± 0.29</td>
<td>14.08 ± 1.55</td>
<td>13.46 ± 1.9</td>
</tr>
<tr>
<td>Word list learning</td>
<td>22.25 ± 3.60</td>
<td>22.08 ± 4.56</td>
<td>18.0 ± 4.5</td>
<td>16.69 ± 3.73</td>
</tr>
<tr>
<td>Word list recall</td>
<td>8.0 ± 1.76</td>
<td>7.75 ± 2.09</td>
<td>4.62 ± 2.02</td>
<td>5.08 ± 2.47</td>
</tr>
<tr>
<td>Word list recognition</td>
<td>9.67 ± 0.65</td>
<td>9.83 ± 0.58</td>
<td>8.0 ± 2.8</td>
<td>7.92 ± 2.56</td>
</tr>
<tr>
<td>Constructional praxis</td>
<td>10.92 ± 0.29</td>
<td>10.75 ± 0.45</td>
<td>10.46 ± 0.88</td>
<td>10.38 ± 0.87</td>
</tr>
<tr>
<td>Constructional praxis recall</td>
<td>10.5 ± 1.10</td>
<td>10.25 ± 1.36</td>
<td>7.62 ± 2.61</td>
<td>7.77 ± 3.03</td>
</tr>
<tr>
<td>Clock drawing test</td>
<td>1.00 ± 0.0</td>
<td>1.00 ± 0.0</td>
<td>2.31 ± 0.63</td>
<td>2.46 ± 0.78</td>
</tr>
</tbody>
</table>

The presented numbers in Table 1 depict the mean value ± standard deviation of the CERAD battery subtests and the clock drawing test. HC, healthy controls; aMCI, patients with amnestic mild cognitive impairment; MMSE, Mini-Mental State Examination.

Table 2 Inter-session and inter-group changes in cognitive performance

<table>
<thead>
<tr>
<th></th>
<th>HC baseline versus follow-up</th>
<th>aMCI baseline versus follow-up</th>
<th>HC versus aMCI baseline</th>
<th>HC versus aMCI follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>0.082</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Verbal fluency</td>
<td>0.349</td>
<td>0.044</td>
<td>0.062</td>
<td>0.012</td>
</tr>
<tr>
<td>Boston naming</td>
<td>0.275</td>
<td>0.025</td>
<td>0.577</td>
<td>0.015</td>
</tr>
<tr>
<td>Word list learning</td>
<td>0.890</td>
<td>0.262</td>
<td>0.016</td>
<td>0.004</td>
</tr>
<tr>
<td>Word list recall</td>
<td>0.660</td>
<td>0.351</td>
<td>0.000</td>
<td>0.008</td>
</tr>
<tr>
<td>Word list recognition</td>
<td>0.551</td>
<td>0.897</td>
<td>0.056</td>
<td>0.012</td>
</tr>
<tr>
<td>Constructional praxis</td>
<td>0.339</td>
<td>0.837</td>
<td>0.100</td>
<td>0.206</td>
</tr>
<tr>
<td>Constructional praxis recall</td>
<td>0.586</td>
<td>0.855</td>
<td>0.003</td>
<td>0.016</td>
</tr>
<tr>
<td>Clock drawing test</td>
<td>4.549</td>
<td>0.000</td>
<td>2.31 ± 0.63</td>
<td>2.46 ± 0.78</td>
</tr>
</tbody>
</table>

Numbers in Table 2 show the results of the paired and independent t-tests (p-values). Significant differences (p_corr < 0.0056, Bonferroni-corrected for multiple comparisons) are italicized. HC, healthy controls; aMCI, patients with amnestic mild cognitive impairment; MMSE, Mini-Mental State Examination. Italic values represent significant differences.

Measurement of head motion

To account for head motion, which has been shown to systematically affect fMRI derived connectivity measures, the mean absolute displacement of each functional brain volume as compared to the previous volume was calculated [37] for each subject.

Correlation analyses

In order to detect possible associations between cognitive performance and resting-state functional connectivity, we correlated the neuropsychological test scores with the degree of functional connectivity of individual RSNs by calculating Pearson product-moment correlation coefficients. The degree of functional connectivity was determined as the mean z-value of an individual RSN, which was extracted from the underlying IC by setting a threshold at z ≥ 3.0. To account for multiple testing, the significance threshold of p < 0.05 was Bonferroni-corrected for each analysis [29]. Correlation analyses were performed with SPSS Statistics version 20.0.

RESULTS

HC and aMCI patients did not significantly differ in mean age (t-test, p = 0.07), gender distribution (χ² = 0.11; Fisher’s exact test, p = 0.53), years of education (t-test, p = 0.18), and the time period between both examinations (t-test, p = 0.25). No significant cerebrovascular lesions were observed in either group.

Neuropsychological testing

At baseline, aMCI patients showed significantly (p_corr < 0.0056) lower performance in the MMSE, word list recall, recall of constructional praxis, and the clock drawing test than HC (Tables 1 and 2, column 3), consistent with the characterization of this group as aMCI in several cognitive domains.

There were no significant inter-session (baseline versus follow-up) differences in the CERAD battery subtests and the Clock drawing test in either group at a significance level of p_corr < 0.0056 (Tables 1 and 2, columns 1 & 2).

Measurement of head motion

To account for head motion, which has been shown to systematically affect fMRI derived connectivity measures, the mean absolute displacement of each functional brain volume as compared to the previous volume was calculated [37] for each subject.

Correlation analyses

In order to detect possible associations between cognitive performance and resting-state functional connectivity, we correlated the neuropsychological test scores with the degree of functional connectivity of individual RSNs by calculating Pearson product-moment correlation coefficients. The degree of functional connectivity was determined as the mean z-value of an individual RSN, which was extracted from the underlying IC by setting a threshold at z ≥ 3.0. To account for multiple testing, the significance threshold of p < 0.05 was Bonferroni-corrected for each analysis [29]. Correlation analyses were performed with SPSS Statistics version 20.0.

RESULTS

HC and aMCI patients did not significantly differ in mean age (t-test, p = 0.07), gender distribution (χ² = 0.11; Fisher’s exact test, p = 0.53), years of education (t-test, p = 0.18), and the time period between both examinations (t-test, p = 0.25). No significant cerebrovascular lesions were observed in either group.

Neuropsychological testing

At baseline, aMCI patients showed significantly (p_corr < 0.0056) lower performance in the MMSE, word list recall, recall of constructional praxis, and the clock drawing test than HC (Tables 1 and 2, column 3), consistent with the characterization of this group as aMCI in several cognitive domains.

There were no significant inter-session (baseline versus follow-up) differences in the CERAD battery subtests and the Clock drawing test in either group at a significance level of p_corr < 0.0056 (Tables 1 and 2, columns 1 & 2).
During the course of the study, 4 aMCI patients showed overall cognitive decline in MMSE scores without meeting the criteria for dementia. None of the HC converted to MCI or dementia; none of the MCI converted to AD. No patient was excluded from the study.

Resting-state connectivity patterns

Concatenating the time-courses of all subjects and sessions (1 TC-GICA approach), the MELODIC software generated 24 ICs (Fig. 1). Fourteen maps were considered as RSNs. These networks were highly consistent with connectivity patterns reported in previous studies [1, 14, 21, 33]. Ten component maps were assumed to be associated with non-neural noise, such as head motion (IC 10), white matter (IC 14 & 15), and large vessel or CSF pulsation (IC 1, 3, 4, 7, 12, 20, and 23).

IC 2 and 11 were characterized by a primarily lateralized fronto-parietal connectivity pattern, which is closely functionally integrated in a wide range of cognitive processes [33, 38, 39]. Connectivity patterns represented in IC 5, 13, and 17 were recently described as subsystems of the DMN [21], which among others is discussed to support internally directed mental activity [33, 40, 41]. In that context, IC 13 was most similar to the classic network described by Raichle and colleagues [42]; this pattern includes the precuneus, the posterior and anterior cingulate cortex, the ventromedial prefrontal cortex, the parietal regions, and the thalamus bilaterally. IC 5 showed mainly the middle temporal gyrus extending into the parieto-temporo-occipital junction with additional involvement of the posterior cingulate and the precuneus cortex as well as the inferior and superior frontal gyrus bilaterally. IC 17 mainly involved the ventromedial prefrontal cortex, the anterior and the posterior cingulate cortex, and the precuneus cortex. IC 6 predominantly included the superior temporal lobe representing the auditory cortex [1]. IC 8 combined the posterior parieto-temporal cortex extending into the lateral occipital cortex with parietal areas. Both IC 9 and 21 involved pre- and postcentral gyri forming a sensorimotor network.

Fig. 1. Component maps for all 25 subjects as generated by the TC-GICA sorted by the decreasing percentage of variance explained by each component. Sagittal, coronal, and axial images are displayed in radiological convention (x-, y-, and z-coordinates of each slice in the MNI152 standard space are given in parenthesis). TC-GICA, temporal concatenation group independent component analysis.
The aMCI group (mean = 0.38, SD = 0.09; session 1: 

& 0.08) than for those deriving from 0.23 to 0.65 (Table 3). Modal ICCs were signif-

canently higher for RSNs deriving from the HC group 

and the parieto-temporal junction; this pattern strongly overlaps with 

a network that has been proposed to play a major role in executive control and working memory function [1, 33, 44, 45]. IC 18 and 22 included areas within 

the visual cortex. IC 19 encompassed regions pre-
odominantly located in the orbitofrontal cortex [21]. IC 24 combined the superior parietal cortex extend-
ing into the occipito-temporal cortex and precentral 

regions. This set of interconnected brain regions has 

been described previously as the dorsal attention net-

work [46].

Within the 24 group-level ICs that were generated in separate TC-GICAs for each group and session, i.e., in the 4 TC-GICA approach, we found 10 RSNs that were present in both groups at both sessions and there-

fore could be subjected to the ICC analysis. These RSNs were equivalent to the connectivity patterns rep-

resented in IC 2, 8, 9, 11, 13, 18, 19, 21, 22, and 24 of the 1 TC-GICA approach (maps demonstrating the connectivity patterns derived from the 4 TC-GICA approach are not shown here).

Reliability analyses

In case of the 1 TC-GICA approach, the overall inter-session test-retest reliability of connectivity pat-
terns, which was expressed by the first most frequent 

ICC in each IC (“modal ICC”, see [21] for further explanation), was moderate to high with values ranging from 0.23 to 0.65 (Table 3). Modal ICCs were signif-

icantly higher for RSNs deriving from the HC group and included 

the left- and right-hemispheric fronto-parietal network (IC 2 & 11), the auditory cortex (IC 6), one sensorimotor 

network (IC 9), all three sub-networks of the DMN (IC 5, 13 & 17), a network associated with executive control and working memory function (IC 16), a visual network (IC 18), and the dorsal attention network (IC 24). The only RSN in the aMCI group that expressed a 

network (IC 18), and the dorsal attention network (IC 24). A relatively high modal ICC was also observed in the “classic” sub-network of the DMN (IC 13, modal ICC = 0.48) in the HC 

group.

The modal ICCs derived from the 4 TC-GICA approach were significantly lower than those derived from the 1 TC-GICA approach (p < 0.001).

Maps showing the voxel-wise ICCs, i.e., the degree of the long-term test-retest reliability of a specific voxel, demonstrated that the most reliable voxels within an IC were generally localized in the core regions of the original RSNs (Figs. 2–5).

Measurements of head motion

No significant inter-session differences in head dis-
placement existed in the HC (0.025 ± 0.009 mm versus 0.027 ± 0.011 mm, p = 0.52) and in the aMCI group (0.027 ± 0.006 mm vs. 0.025 ± 0.006 mm, p = 0.33).

Also, no significant inter-group differences in head displacement were present at either session (session 1: p = 0.54; session 2: p = 0.58). These results demon-
strate that motion artifacts are unlikely to be the leading

<table>
<thead>
<tr>
<th>IC</th>
<th>HC</th>
<th>aMCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 TC-GICA</td>
<td>4 TC-GICA</td>
<td>1 TC-GICA</td>
</tr>
<tr>
<td>2</td>
<td>0.63</td>
<td>0.53</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
<td>*</td>
</tr>
<tr>
<td>6</td>
<td>0.5</td>
<td>*</td>
</tr>
<tr>
<td>8</td>
<td>0.48</td>
<td>0.4</td>
</tr>
<tr>
<td>9</td>
<td>0.55</td>
<td>0.53</td>
</tr>
<tr>
<td>11</td>
<td>0.63</td>
<td>0.55</td>
</tr>
<tr>
<td>13</td>
<td>0.55</td>
<td>0.46</td>
</tr>
<tr>
<td>16</td>
<td>0.5</td>
<td>*</td>
</tr>
<tr>
<td>17</td>
<td>0.65</td>
<td>*</td>
</tr>
<tr>
<td>18</td>
<td>0.5</td>
<td>0.45</td>
</tr>
<tr>
<td>19</td>
<td>0.33</td>
<td>0.2</td>
</tr>
<tr>
<td>21</td>
<td>0.43</td>
<td>0.36</td>
</tr>
<tr>
<td>22</td>
<td>0.48</td>
<td>0.33</td>
</tr>
<tr>
<td>24</td>
<td>0.55</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 3

Inter-session test-retest reliability

(mean = 0.18, SD = 0.09, p < 0.001) (Table 3). Com-
parable to the 1 TC-GICA approach, notably reliable RSNs with an ICC ≥ 0.5 were present only in the HC 
group and included both fronto-parietal networks (IC 2 & 11), one sensorimotor network (IC 9), and the dor-
sal attention network (IC 24). A relatively high modal ICC was also observed in the “classic” sub-network of the DMN (IC 13, modal ICC = 0.48) in the HC 
group.

The most reliable RSNs (defined as modal ICC ≥ 0.5) derived mainly from the HC group and included 

the left- and right-hemispheric fronto-parietal network (IC 2 & 11), the auditory cortex (IC 6), one sensorimotor 

network (IC 9), all three sub-networks of the DMN (IC 5, 13 & 17), a network associated with executive control and working memory function (IC 16), a visual network (IC 18), and the dorsal attention network (IC 24). The only RSN in the aMCI group that expressed a 

network (IC 18), and the dorsal attention network (IC 24). A relatively high modal ICC was also observed in the “classic” sub-network of the DMN (IC 13, modal ICC = 0.48) in the HC 
group.

The modal ICCs derived from the 4 TC-GICA approach were significantly lower than those derived from the 1 TC-GICA approach (p < 0.001).

Maps showing the voxel-wise ICCs, i.e., the degree of the long-term test-retest reliability of a specific voxel, demonstrated that the most reliable voxels within an IC were generally localized in the core regions of the original RSNs (Figs. 2–5).

Measurements of head motion

No significant inter-session differences in head dis-
placement existed in the HC (0.025 ± 0.009 mm versus 0.027 ± 0.011 mm, p = 0.52) and in the aMCI group (0.027 ± 0.006 mm vs. 0.025 ± 0.006 mm, p = 0.33).

Also, no significant inter-group differences in head displacement were present at either session (session 1: p = 0.54; session 2: p = 0.58). These results demon-
strate that motion artifacts are unlikely to be the leading
cause of the observed difference in long-term test-retest reliability of RSNs between both groups.

**Correlation analyses**

Considering the problem of multiple comparisons (9 neuropsychological tests and 14 (10) RSNs, resulting in a Bonferroni corrected $p_{corr} < 0.000397$ in the 1 TC-GICA approach and $p_{corr} < 0.000556$ in the 4 TC-GICA approach), there were no significant associations between the neuropsychological test scores and the degree of functional connectivity of RSNs. The correlation coefficients ranged from $-0.29$ to $+0.4$ (1 TC-GICA approach; $p > 0.004$; Table 4) and $-0.29$ and $+0.34$ (4 TC-GICA approach; $p > 0.023$; data not shown here), respectively.

There were also no significant inter-session differences in the degree of functional connectivity, determined as the mean $z$-value of a connectivity pattern, between corresponding RSNs in either group in both TC-GICA approaches ($p > 0.05$).

**DISCUSSION**

In this study, we investigated the long-term (>1 year) test-retest reliability of RSNs in HC and patients with aMCI in order to further establish the applicability of these imaging markers as a diagnostic tool to monitor subjects at increased risk of developing dementia. We found moderate to high reliability for most RSNs in the HC group. Reliability of RSNs was significantly decreased in the aMCI group, likely reflecting the decreased intra-subject stability of intrinsic functional connectivity patterns, especially as group differences in major confounders like head motion could be ruled out.

Two different approaches were performed to decompose the resting-state fMRI data set into independent components: Initially, the voxel time-courses of all participants and sessions were combined in one Temporal Concatenation Group ICA (TC-GICA; “1 TC-GICA approach”). In a supplemental approach, we calculated a separate TC-GICA for each group and session.
Fig. 3. Voxel-wise intra-class correlation coefficient (ICC) maps thresholded at ICC ≥0.8 demonstrating the reliability of corresponding group-level components based upon individual-level independent components (ICs) derived from patients with aMCI. For this analysis, fMRI data of all healthy elderly and aMCI subjects and sessions were combined in one TC-GICA (1 TC-GICA approach). Sagittal, coronal, and axial reliability maps are displayed in radiological convention (x-, y-, and z-coordinates of each slice in the MNI152 standard space are equivalent to those in Fig. 1). TC-GICA, temporal concatenation group independent component analysis.

When applying the 1 TC-GICA approach, we observed that the overall inter-session test-retest reliability of most RSNs was moderate to high in the HC group, but significantly decreased in aMCI subjects (p < 0.001); the most reliable connectivity patterns derived from the HC group and included brain regions associated with auditory (IC 6), sensorimotor (IC 9), visual (IC 18), and higher order cognitive functions (IC 2 & 11, IC 16, 24) as well as the default-mode function (IC 5, 13 & 17). Particularly low reliability was found in a network including inferior frontal regions (IC 19), which are known to display a low signal-to-noise ratio due to susceptibility artifacts [47].

The high consistency of most of the networks is compatible with results from prior studies investigating the test-retest reliability of RSNs in young healthy subjects across a short-term period, even though these studies used different methods for data decomposition and correlation analyses. For instance, Danoiseaux and colleagues reported a high consistency for RSNs comparable to our “classic” sub-network of the DMN (IC 13, their Fig. 3/map B), the fronto-parietal networks (IC 2 & 11, their Fig. 3/map C & D), the visual network (IC 22, their Fig. 3/maps A & E), and the dorsal attention network (IC 24, their Fig. 3/map H) within a period of up to 14 days [1]. Chen and colleagues found a remarkably high inter-session (12–24 days) consistency for RSNs similar to our sensorimotor network (IC 9, their Fig. 1/IC 10), the “classic” sub-network of the DMN (IC 13, their Fig. 1/IC 11), both fronto-parietal networks (IC 2 & 11, their Fig. 1/IC 16 & 14), and for the dorsal attention network (IC 24, their Fig. 1/IC 13) [48].

Moreover, our results for HC subjects are in line with the findings presented by Zuo and colleagues, who examined the one-year test-retest reliability of RSNs in young healthy subjects [21]. Using a comparable voxel-wise ICC approach based upon individual-level ICs derived from a multi-subject multi-session TC-GICA, they demonstrated that the most reliable connectivity patterns include RSNs associated with sensory, motor and higher order cognitive function as
well as the “default-mode” networks; the RSNs represented in their study (their ICs 4 & 15, 9, 13, 7 & 19, 8, 10, 12 & 14) are very similar to those found in this work (the corresponding ICs are 18 & 22, 24, 9 & 21, 11 & 2, 16, 13, 17 & 5) and the modal ICCs reported by Zuo and colleagues for these specific connectivity patterns in young healthy subjects (0.45–0.65) are similar to the numbers determined in our study for healthy elderly subjects (modal ICC: 0.43–0.65).

Most of the RSNs found in the 1 TC-GICA approach were also present when performing the supplemental 4 TC-GICA approach, suggesting that these RSNs are reproducible across approaches. As expected, the long-term test-retest reliability of RSNs was significantly lower using the supplemental analysis ($p<0.001$).

However, the results were similar with regard to the fact that the test-retest reliability of RSNs was significantly higher in the HC than in the aMCI group and that the most consistent RSNs, which derived only from the HC group, included connectivity patterns extending over brain regions associated with sensorimotor (IC 9), higher order cognitive (IC 2, 11 & 24), and the default-mode function (IC 13).

Combining the findings mentioned above, it seems that RSNs associated with sensory, motor, and higher order cognitive function and those supporting internal processes as the DMN sub-networks generally exhibit very stable functional connectivity patterns across short-term and longer-term periods during healthy conditions, i.e., also during healthy aging.

Previous studies have demonstrated that progressive alterations in RSN connectivity constitute a common feature during pathologic conditions associated with declining cognitive performance, e.g., in case of the DMN during the conversion from healthy aging to MCI and from MCI to dementia [12, 13, 15] as well as during the progression of AD [18, 49]. The significantly lower long-term test-retest reliability of RSNs in patients with aMCI observed in this study may therefore represent disease-associated neural alterations that in general affect resting-state functional connectivity in these subjects over time. However, no significant inter-session differences in the degree of RSN connectivity were detectable in this group; this result may be caused by very subtle alterations of functional connectivity comprising both connectivity increases and...
Fig. 5. Voxel-wise intra-class correlation coefficient (ICC) maps thresholded at ICC ≥0.8 demonstrating the reliability of corresponding group-level components based upon individual-level independent components (ICs) derived from patients with aMCI. For this analysis, fMRI data of each group (healthy elderly and patients with aMCI respectively) and session underwent a separate TC-GICA (1 TC-GICA approach). Sagittal, coronal, and axial reliability maps are displayed in radiological convention (x-, y-, and z-coordinates of each slice in the MNI152 standard space are equivalent to those in Fig. 1). TC-GICA, temporal concatenation group independent component analysis.

According to the neuropsychological tests, there were no significant inter-session changes in cognitive performance in either group. The aMCI subjects’ cognitive stability and concurrent reduction of long-term RSN reliability imply that alterations in resting-state functional connectivity may precede cognitive alterations, e.g., due to an initial recruitment of compensatory mechanisms on the cognitive level commonly referred to as the “cognitive reserve” [50, 51]. Compensatory mechanisms may also account for the lack of association between the degree of RSN connectivity and the neuropsychological data observed in this study.

Also, the age factor may have contributed to the observed overall reduction of RSN reliability in aMCI patients since subjects of this group tended to be older than the control group (72.8 ± 7.3 versus 67.8 ± 7.3, p = 0.07). In that context, previous studies have shown that aging per se modifies the connectivity of RSNs, e.g., of the DMN [52] as well as of networks associated with attention and somatosensory functions [53]. It is therefore possible that the higher average age of aMCI patients, although not significant, has contributed to an accelerated long-term alteration of neural connectivity resulting in lower test-retest reliability of RSNs in that group.

In the present work, we demonstrate by means of the group-level ICC maps that the highest degree of inter-session test-retest reliability usually affects those voxels, which originally are located within a specific network, i.e., those voxels with the highest connectivity strength within an IC. This result is compatible with previous findings by Shehzad and colleagues who demonstrated in a seed-based analysis that voxel-wise test-retest reliability of resting-state fMRI measures is generally more robust in case of voxels with highly correlated signal time-courses [54].

There are some limitations of the present study that have to be considered. First, only 12 HC and 13 aMCI subjects were included in the study. The small (and possibly unrepresentative) sample may restrict the reliability of the data. Further studies with larger groups of participants are needed for the validation of the
Table 4

|                        | IC 2 | IC 3 | IC 4 | IC 5 | IC 6 | IC 7 | IC 8 | IC 9 | IC 10 | IC 11 | IC 12 | IC 13 | IC 14 | IC 15 | IC 16 | IC 17 | IC 18 | IC 19 | IC 20 | IC 21 | IC 22 | IC 23 | IC 24 |
|------------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Verbal fluency         | 0.23 | 0.13 | 0.09 | 0.21 | 0.34 | 0.22 | 0.32 | 0.25 | 0.29 | 0.32 | 0.14 | 0.13 | 0.23 | 0.14 | 0.13 | 0.23 | 0.17 |
| Boston naming          | 0.10 | 0.11 | -0.01 | 0.17 | 0.96 | 0.24 | 0.23 | 0.11 | 0.05 | 0.04 | 0.05 | 0.12 | 0.08 |
| MMSE                   | 0.24 | -0.04 | 0.09 | 0.19 | 0.36 | 0.23 | 0.24 | 0.33 | 0.28 | 0.27 | 0.12 | 0.24 | 0.23 | 0.16 |
| Word list learning     | 0.091 | 0.738 | 0.518 | 0.193 | 0.010 | 0.114 | 0.088 | 0.020 | 0.053 | 0.056 | 0.408 | 0.099 | 0.102 | 0.270 |
| Constructional praxis  | 0.31 | 0.07 | 0.19 | 0.40 | 0.39 | 0.24 | 0.24 | 0.39 | 0.36 | 0.22 | 0.09 | 0.16 | 0.27 | 0.18 |
| Word list recall       | 0.24 | -0.22 | 0.25 | 0.29 | 0.26 | 0.15 | 0.16 | 0.26 | 0.25 | 0.09 | 0.05 | 0.07 | 0.05 | -0.03 |
| Constructional praxis  | 0.095 | 0.017 | 0.079 | 0.044 | 0.068 | 0.307 | 0.279 | 0.063 | 0.076 | 0.549 | 0.721 | 0.610 | 0.716 | 0.824 |
| Word list recognition  | 0.16 | -0.16 | 0.02 | 0.32 | 0.14 | 0.05 | 0.10 | 0.23 | 0.14 | 0.08 | 0.00 | 0.01 | 0.04 | -0.01 |
| Recall of constructional praxis | 0.21 | -0.06 | 0.10 | 0.28 | 0.22 | 0.13 | 0.15 | 0.26 | 0.25 | 0.14 | 0.11 | 0.12 | 0.15 | 0.05 |
| Clock drawing test     | 0.141 | 0.094 | 0.497 | 0.406 | 0.117 | 0.366 | 0.306 | 0.064 | 0.078 | 0.315 | 0.436 | 0.380 | 0.313 | 0.728 |

Table 4 demonstrates the results of the correlation analyses determining the association between the degree of a RSN's functional connectivity and the neuropsychological test scores (CERAD test battery). Adjusted for multiple comparisons resulting in a Bonferroni corrected $p_{corr} < 0.000397$ no significant associations exist. Results are shown for RSNs derived from the 1 TC-GICA approach. IC, independent component; $\rho$, Pearson’s correlation coefficient; TC-GICA, temporal concatenation group independent component analysis.
presented results. Second, it has been shown that fluctuations in the low-frequency domain of the BOLD signal where resting-state fluctuations are detected may comprise aliasing non-BOLD noise from respiratory or cardiovascular function, which may not be completely removed by ICA approaches [55]. These physiological functions, which were not monitored during the acquisition of the MRI data, may have influenced the detected RSNs and consequently the reliability measurements. Moreover, the presented reliability results may also be influenced by parameters as the technological equipment, e.g., variation in the MR scanner mode, or biological functions as fluctuations in vigilance. The impact of these parameters on RSN connectivity and consequently test-retest reliability measurements should be investigated in future studies.

Our results suggest that RSNs are reliable over the long-term in both healthy elderly and patients with aMCI. The degree of reliability, however, seems to be different between both groups with generally more stable RSNs in healthy elderly indicating that healthy aging may be associated with stable resting-state functional connectivity. Correspondingly, the overall lower reliability of RSNs in the aMCI group may represent a disease associated progression of neuro-functional alterations over time in these patients that precede changes in cognitive performance; an additional age-associated impact on the presented results in this study cannot be excluded and should be clarified in future studies.

Before being able to apply resting-state connectivity patterns in clinical routine, it will be essential, among others, to better understand the relationship between functional connectivity and cognitive performance including the impact of compensatory cognitive mechanisms; future studies may therefore systematically examine the degree of network disruptions as a function of cognitive performance and they may also illuminate the role of specific network regions in the process of cognitive alterations.

DISCLOSURE STATEMENT

Authors’ disclosures available online (http://www.j-alz.com/disclosures/view.php?id=1600).

REFERENCES


deactivations: Change with age and dementia of the Alzheimer type. Proc Natl Acad Sci USA 100, 14504-14509.

