

# Impact of cognitive performance on the reproducibility of fMRI activation in schizophrenia

Olivier Maïza, MD; Bernard Mazoyer, MD, PhD; Pierre-Yves Hervé, PhD;  
Annick Razafimandimby, PhD; Sonia Dollfus, MD, PhD; Nathalie Tzourio-Mazoyer, MD

Maïza, Mazoyer, Hervé, Razafimandimby, Dollfus, Tzourio-Mazoyer — Centre d'imagerie–neurosciences et applications aux pathologies, Universités de Caen et Paris Descartes; Maïza, Mazoyer, Dollfus — Centre Hospitalier Universitaire de Caen, Caen; Mazoyer — Institut universitaire de France, Paris, France

Previously published at [www.jpn.ca](http://www.jpn.ca)

**Background:** Longitudinal functional magnetic resonance imaging (fMRI) studies in patients with schizophrenia allow exploration of the course of the illness and brain activity after therapy. A crucial question, however, is whether fMRI findings are reliable, because they can be affected by performance deficits in patients with schizophrenia. Our aim was to evaluate the reproducibility of fMRI activations in highly integrated language areas in patients with schizophrenia, taking into account task performance. **Methods:** Ten patients with schizophrenia and 10 matched healthy controls were scanned twice, 21 months apart, while performing a story comprehension task. The reproducibility of the activations in each participant was evaluated globally by the percentage of spatial overlap between the 2 sessions and locally by a voxel-wise computation of the between-session relative standard deviation. We performed between-group comparisons both with and without the inclusion of comprehension scores (measuring task performance) as a covariate. **Results:** On average, patients with schizophrenia had significantly lower comprehension scores than controls (4.5/12 v. 7.8/12,  $p = 0.002$ ). The mean spatial overlap between fMRI sessions was 30.6% in the patient group and 47.0% in the control group ( $p = 0.017$ ). Locally, the lower reproducibility in patients was most prominent in the left posterior middle temporal gyrus, inferior frontal gyrus and medial prefrontal cortex ( $p < 0.001$  uncorrected for multiple comparisons). Comprehension scores were positively correlated with both reproducibility measures in patients (overlap:  $r = 0.82$ ,  $p = 0.004$ ; relative standard deviation: several significant clusters at  $p < 0.001$ ). When we included the comprehension scores as a covariate, most of the local between-group differences in reproducibility were removed, and the difference in overlap was not significant. **Limitations:** Owing to the small sample size, we could not investigate the impact of clinical subtypes and different types of medications on reproducibility. **Conclusion:** Our findings suggest that the greater variability in activation in patients with schizophrenia compared with controls concerns high-level areas and is mainly attributable to deficient task performance. Consequently, cognitive performance must be carefully controlled when longitudinal fMRI studies are undertaken.

## Introduction

Functional magnetic resonance imaging (fMRI) is the method of choice for evaluating neurofunctional correlates of symptoms and cognitive impairments in patients with schizophrenia. Most studies have had a cross-sectional design, but a few fMRI studies have adopted a longitudinal design, with 2 types of objectives. Some authors have attempted to identify functional biomarkers that are stable through the course of the illness and that may represent trait markers.<sup>1-3</sup> Others

have assessed the effects of cognitive therapies<sup>4,5</sup> or medications<sup>6-13</sup> on brain activity during cognitive tasks. However, these studies have had conflicting findings because of several methodologic limitations.<sup>14</sup> In particular, they did not address the question of whether fMRI activation was reproducible in patients, which is crucial for establishing the validity of a longitudinal experimental paradigm.<sup>15</sup>

In healthy participants, the test-retest reliability of fMRI activation has been investigated in the context of sensory<sup>16-22</sup> and motor tasks<sup>21-29</sup> as well as for various cognitive tasks addressing

**Correspondence to:** Dr. S. Dollfus, Centre d'imagerie–neurosciences et application aux pathologies, Unité Mixte de Recherche, 6232 Centre national de la recherche scientifique, commissariat à l'énergie atomique, Universités de Caen et Paris Descartes, Centre Hospitalier Universitaire, Caen, France; [dollfus-s@chu-caen.fr](mailto:dollfus-s@chu-caen.fr)

*J Psychiatry Neurosci* 2010;35(6):378-89.

Submitted Aug. 26, 2009; Revised Feb. 23, 2010; Accepted May 3, 2010.

DOI: 10.1503/jpn.090103

memory,<sup>30–32</sup> executive functions<sup>33–35</sup> and language.<sup>36–38</sup> Classically, the reproducibility of fMRI activations has been assessed by considering indices such as the percentage of spatial overlap between individual statistically thresholded activation maps at different time points.<sup>19</sup> Reproducibility, as assessed by the percentage of spatial overlap, appears to be higher for motor tasks (54%–84%)<sup>22,26</sup> than for sensory tasks (49%–64%).<sup>16,20</sup> Memory tasks yielded less reproducible activations (with spatial overlap ranging from 36% to 41%).<sup>31,32</sup> For language tasks, similar to the greater reproducibility of motor compared with sensory tasks, word generation elicited more reproducible activations (spatial overlap 75%) than story comprehension (spatial overlap 50%).<sup>38</sup>

However, this type of reproducibility index can be affected by a dependence on the arbitrary statistical threshold that is applied to define activated areas. Despite the good reliability of estimated blood oxygen level-dependent (BOLD) signals, the difference between the task and reference may be slightly above the statistical threshold in one session and slightly beneath threshold in another session.<sup>39,40</sup> Several authors have circumvented this problem by computing the percentage of overlap at different thresholds.<sup>16,20,22,28,41</sup> Another way of alleviating this bias is to threshold the activation maps so as to keep the number of activated voxels constant across sessions.<sup>42</sup> Otherwise, reproducibility indices that are independent of a statistical threshold can be used. In particular, the intraclass correlation coefficient (ICC)<sup>18,26,34,41,43</sup> and coefficient of variation (CV)<sup>24,44</sup> are both computed on a region-of-interest or voxel-wise basis and thus allow local assessment of reproducibility. Because of methodologic differences and marked functional differences between the various cerebral regions investigated, it is difficult to compare the results obtained with these indices across studies.

Despite the interest in using longitudinal fMRI studies in clinical populations, few studies have assessed the reproducibility of activations in patients<sup>41,45–48</sup> or populations at high risk of illness.<sup>49,50</sup> Compared with healthy controls, participants at high risk of Alzheimer disease<sup>50</sup> or schizophrenia<sup>49</sup> had similar reproducibility of activations, whereas patients who had experienced a stroke had less reproducible activations.<sup>47</sup>

Only 1 study has investigated the test-retest reliability of measuring fMRI activation within a working memory paradigm in patients with schizophrenia.<sup>51</sup> Relying on ICC to assess reproducibility, Manaöch and colleagues<sup>51</sup> showed that patients with schizophrenia exhibited less consistent activations across sessions in areas involved in high-order cognitive processes compared with healthy participants. This raises the question of whether cognitive performance has an impact on the reproducibility of activation. Indeed, some patients performed poorly, but the authors suggested that poor performance was not the cause of low reproducibility, and they concluded that the variability in activation might be intrinsic to schizophrenia.<sup>51</sup> Nevertheless, given the potential of longitudinal fMRI studies in schizophrenia, the effect of performance on reproducibility deserves precise evaluation. Furthermore, ICC, in contrast to CV, assumes that the groups being compared have similar between-subject variability.<sup>52</sup> However, between-subject variability is higher among patients with schizophrenia than among healthy participants,

particularly for working memory tasks.<sup>53</sup> Thus, CV appears to be a more relevant measure than ICC for comparing the reproducibility of fMRI activations between patients with schizophrenia and healthy participants.

In the present study, we assessed the effects of performance on the reproducibility of fMRI activations in patients with schizophrenia and healthy participants. We used 2 different reproducibility indices, allowing both global and local assessment of the reproducibility of fMRI activations. We assessed the effect of task performance on both of these indices.

## Methods

### Participants

We selected participants from the 21 pairs of patients with schizophrenia and matched healthy controls who participated in a previous cross-sectional fMRI study.<sup>54</sup> Ten pairs underwent a second fMRI session and were included in the present study. Eleven of the patients were not included: 3 did not give their consent, 3 were admitted to hospital between the 2 sessions, 4 did not stay in contact with their practitioner and in one case, the data from the second session were not usable owing to an excessive intersession difference of positioning in the scanner, which prevented accurate intersession coregistration.

All participants were free of auditory deficits, neurologic disorders and cerebral abnormality, as assessed by a brain  $T_1$ -weighted MRI. All participants reported that French was their primary language and used their right hand for writing. They all had a verbal intelligence quotient (IQ) above 80 (Wechsler Adult Intelligence Scale<sup>55</sup>). Patients and healthy participants (controls) were matched one-to-one for sex, age at the first fMRI session, level of education and handedness (Edinburgh Inventory score<sup>56</sup>). Schizophrenia was diagnosed by clinicians according to the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (DSM-IV) criteria.<sup>57</sup> The patients were stabilized out-patients with no hospital admissions or psychotic exacerbations during the follow-up period. At each session, their clinical state was evaluated by use of the Positive and Negative Syndrome Scale (PANSS).<sup>58</sup> Control participants were free of psychotic disorders and substance dependence (including alcohol) as assessed by the Structured Clinical Interview of the DSM III-revised (SCID).<sup>59</sup>

The local ethics committee (Comité de Protection des Personnes Nord-Ouest, France) approved this study, and the participants provided informed, written consent. Consent was obtained from the schizophrenia patients after an interview with a practitioner. All were outpatients who were capable of understanding the information that was given to them.

### Experimental paradigm

All participants underwent 2 fMRI sessions, separated by an average of 21 months. The intersession interval was significantly longer in patients than in controls (patients: mean 17 mo, standard deviation [SD] 5 mo; controls mean 26 [SD 5] mo;  $t_{18} = 4.79$ ,  $p < 0.001$ , 2-sample  $t$ -test).

At each session, each participant was invited to listen to a

factual story in French that described a sports competition involving intricate social interactions between the characters (Appendix 1, available online at [www.jpn.ca](http://www.jpn.ca)). The French story condition (hereinafter the story condition) was contrasted with listening to the same story in Tamil, a language to which no participant had been exposed to before the experiment. Phonologically, Tamil is relatively similar to French, in the sense that it does not involve tonal patterns, and words and phrases can be segmented by the participants even though the language sounds foreign to them.<sup>60</sup> The Tamil condition was chosen to enable the subtraction of activations elicited by phonologic processing of auditory stimuli. The detection of activity in high-order language integration areas was enhanced upon contrasting the 2 conditions (story v. Tamil).

For each condition, presentation of the auditory stimuli followed a block design with 9 alternating 30-second duration blocks of either the Tamil condition (5 blocks) or the story condition (4 blocks), starting and ending with the Tamil condition. Participants were in a dark room and were instructed to listen attentively to the stimuli while keeping their eyes closed.

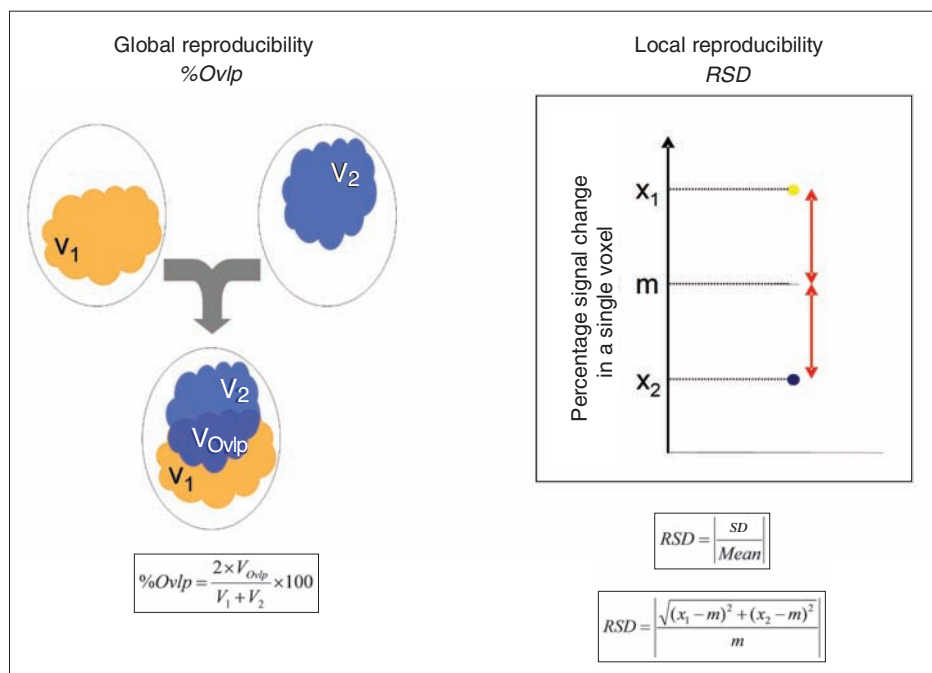
Shortly after scanning, task performance was assessed with a 12-item questionnaire, designed to investigate both the overall understanding of the story and the comprehension of selected details. Oral answers to each item were scored 1, 0.5 or 0 for an exact, partial, or inexact or no response, respectively.

*Neuroimaging data acquisition and preprocessing*

At both sessions, all participants underwent both anatomic

MRI and fMRI scanning with a General Electric Signa 1.5 T magnetic resonance imager. Both scans covered the same field of view (240 mm). We acquired anatomic data with a high-resolution,  $T_1$ -weighted MRI scan using a fast, 3-dimensional, spoiled gradient echo sequence that used a spectral selective inversion recovery pulse (matrix size  $256 \times 256 \times 124$ , slice thickness 1.5 mm). The fMRI data were acquired with an echo-planar imaging BOLD sequence (repetition time 6 s, echo time 60 ms, flip angle  $90^\circ$ , matrix size  $64 \times 64 \times 32$ , 50 volumes, slice thickness 3.8 mm).

The data preprocessing was specifically designed for this longitudinal study and has been previously described.<sup>3</sup> We performed the preprocessing using statistical parametric mapping software (SPM99; Wellcome Department of Imaging Neuroscience). To achieve optimal coregistration of fMRI data from both sessions in the same stereotaxic space, we first corrected the fMRI data for differences in time acquisition between slices and then for head motion; the data were then coregistered with the contemporary anatomic MRI data. We coregistered the anatomic MRI data between sessions, and the coregistration parameters were applied to the contemporary fMRI data. We estimated spatial normalization parameters for the Montreal Neurological Institute  $T_1$ -weighted template using the anatomic MRI data from session 1, and we applied these parameters to the corresponding fMRI data from both sessions. These spatially normalized fMRI data were then smoothed with a Gaussian kernel of 8 mm and a high-pass filter (0.0102 Hz cut-off).



**Fig. 1:** Schematic representation of the reproducibility indices.  $m$  = average of percent signal change across session 1 and session 2;  $RSD$  = relative standard deviation;  $SD$  = standard deviation;  $\%Ovlp$  = percentage of overlap;  $V_1$  = volume of activation at session 1;  $V_2$  = volume of activation at session 2;  $V_{Ovlp}$  = volume of voxels activated at session 1 and session 2;  $X_1$ : percent signal change at session 1;  $X_2$ : percent signal change at session 2.

Statistical analyses

Task performance

To assess the possible effects of illness and repetition on task performance, we performed a 3-way mixed-model analysis of variance (ANOVA) with the fixed effects of group (patients v. controls) and session (session 1 v. session 2), and subject as random effects. We evaluated the reproducibility of task performance across sessions separately for each group with Pearson correlation analyses.

fMRI data

We entered the preprocessed fMRI data for each participant from both sessions into a SPM99 regression model for computation of individual contrast maps (story v. Tamil) and the corresponding *t* maps. We subsequently used both of these maps to assess local and global reproducibility.

Global reproducibility

We assessed global reproducibility of the spatial distribution of activation by computing the percentage of spatial overlap for each participant. This value corresponded to the ratio, expressed in percent, of the number of voxels jointly activated at both sessions ( $V_{\text{ovlp}}$ ) to the number of voxels activated at session 1 ( $V_1$ ) or session 2 ( $V_2$ ).<sup>19</sup> It was computed as follows: percentage of spatial overlap =  $[(2 \times V_{\text{ovlp}}) / (V_1 + V_2)] \times 100$  (Fig. 1).

The percentage of spatial overlap was computed by thresholding the *t* maps at a constant activation volume ( $V_1 = V_2$ ) to alleviate the bias of an arbitrary statistical threshold. The con-

stant volume was set at 5000 voxels because, in both groups, it maximized the average percentage of spatial overlap and represented the best compromise between the inclusion of irrelevant voxels at high activation volumes (false positives) and the exclusion of relevant voxels (false negatives) at low activation volumes (Fig. 2). We computed the average *t* values of the 5000 most activated voxels of each individual *t* map. We assessed the stability across sessions of these average *t* values by computing correlation coefficients in controls and patients with schizophrenia.

For comparability with other studies, we also computed a traditional "fixed threshold" percentage of spatial overlap with the following significance thresholds:  $p < 0.001$ ,  $p < 0.01$  and  $p < 0.05$  uncorrected for multiple comparisons (corresponding to *t* value thresholds of 3.15, 2.35, 1.66, respectively).

We compared global reproducibility, which was assessed by the percentage of spatial overlap, between patients and controls with a Welch 2-sample *t* test. To assess the potential effect of task performance on reproducibility, we entered the percentage of spatial overlap data as a dependent variable in a linear model, with group, comprehension score and the interaction between the 2 as explanatory variables; thus, one regression slope was estimated per group. We centred the comprehension scores on the mean of the controls in our statistical model; this allowed us to make comparisons between the 2 groups at the same performance level, corresponding to that of the controls.

To assess the possible influence of other factors on reproducibility, we also analyzed whether the percentage of spatial overlap was correlated to age, verbal IQ, intersession duration,

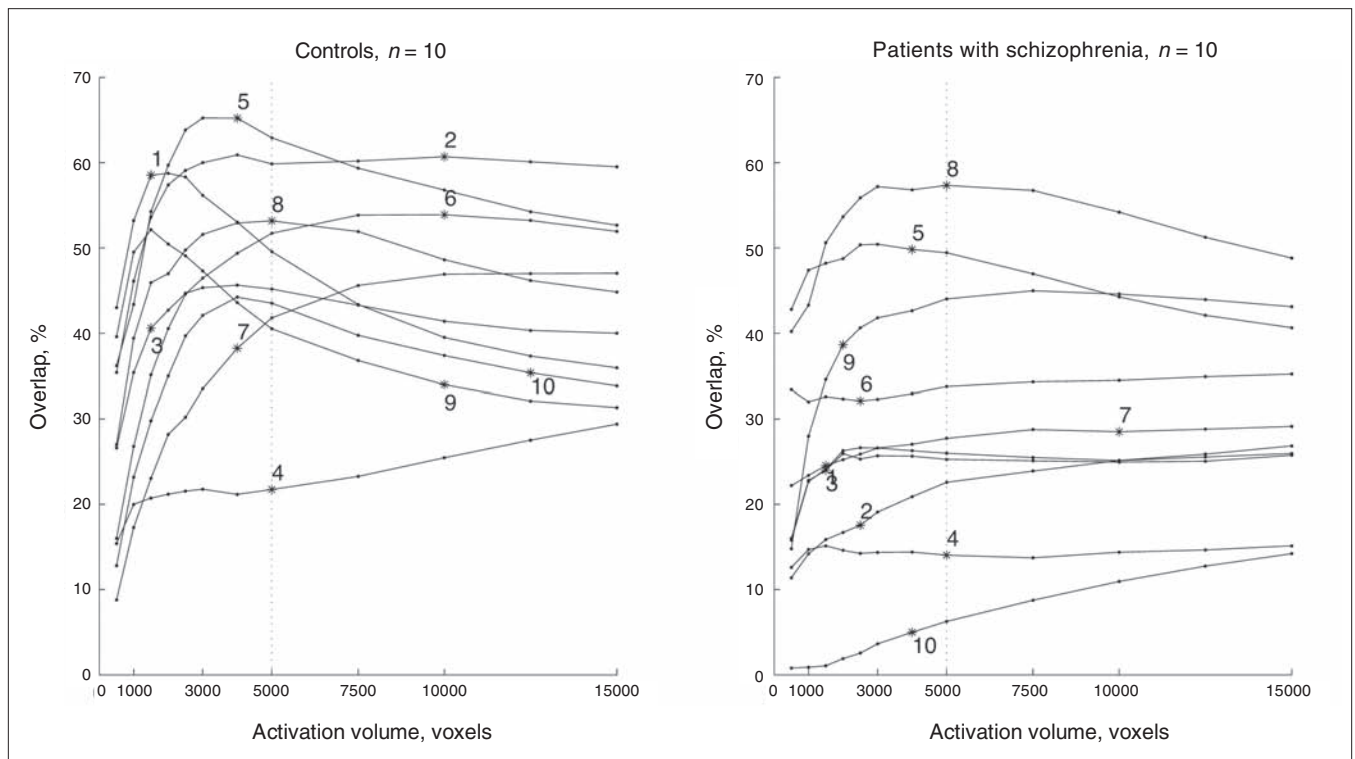


Fig. 2: Percentage of spatial overlap between the activation maps from 2 sessions as a function of a given activation volume threshold in patients with schizophrenia and controls.

absolute value of the difference between session comprehension scores, duration of illness, doses of antipsychotic medication and positive or negative PANSS scores. We performed correlation analyses separately for patients and controls.

To display the functional network used to evaluate global reproducibility, individual *t* maps of the 5000 most activated voxels were binarized, and we computed probability maps of activation for each group at each session.

### Local reproducibility

To assess local reproducibility on a voxel-wise basis, we computed individual relative standard deviation (RSD) maps across sessions from the 2 contrast maps of each participant (session 1 and session 2). Relative standard deviation is the absolute value of the CV, expressed as a percentage. We defined CV as the ratio of the SD to the mean. For each participant, and at each voxel, we computed the SD between the 2 sessions, divided it by the absolute value of the average of the 2 sessions, and then multiplied this value by 100.

To avoid extremely large RSD values, attributable to a very small average of the 2 sessions,<sup>61</sup> the maps were thresholded at an RSD value of 500%; the maps were then smoothed with a Gaussian kernel of 8 mm.

To identify regions with different patterns of reproducibility between patients and healthy participants, we compared the RSD maps between the groups (2-sample *t* tests, SPM99). To restrict this analysis to the regions belonging to the network involved in this task, we masked the individual RSD

images with a binary image representing the network of areas activated during session 1 or 2.<sup>34,49</sup> This mask was computed by performing a group analysis (1-sample *t* test on the contrast maps of the 20 participants for session 1 and 2). We applied a liberal threshold of  $p = 0.01$ , uncorrected for multiple comparisons, to the resulting map, which was then binarized. To map the effects of comprehension scores on reproducibility in each group, we computed the correlation between RSD and comprehension scores. Finally, to control for the effect of performance on the difference in RSD between groups, we compared the patient and control groups with a SPM99 multiple regression model, which included the comprehension scores as a covariate, centred over the mean of the healthy participants. The resulting analysis of covariance design was equivalent to the one used for the percentage of spatial overlap (i.e., the difference between groups was estimated at the performance level of the controls, with separate slopes in the 2 groups).

All analyses that dealt with the effect of the comprehension score on reproducibility included the scores from session 1. We repeated these analyses with the scores from session 2.

All voxel-wise statistical analyses were performed with SPM99. The significance threshold was set at  $p < 0.001$ , uncorrected for multiple comparisons, unless otherwise specified. Other analyses were performed with R (<http://cran.r-project.org>).

## Results

### Participants

Tables 1 and 2 show the clinical and demographic data of the participants. At both sessions, 3 patients were classified as paranoid, 2 as undifferentiated and 5 as experiencing residual symptoms. At both sessions, 7 patients were taking an atypical antipsychotic, 2 were taking a typical antipsychotic and 1 was taking both medicines. One patient changed antipsychotic medicines between the 2 sessions (haloperidol followed by flupenthixol).

For the patients with schizophrenia, the PANSS scores and the antipsychotic medication doses in chlorpromazine equivalents were not significantly different ( $p > 0.05$ ) between sessions and were highly correlated across sessions (Table 2). There was

**Table 1: Sociodemographic characteristics of the study groups**

Characteristic	Group; mean (SD) [range]*	
	Controls, <i>n</i> = 10	Schizophrenia patients, <i>n</i> = 10
No. of men	8	8
≥ 12 years of education, no. of participants	4	4
Age at session 1, yr	34.2 (10.1) [25–50]	34.2 (9.1) [24–48]
Edinburgh Inventory Score <sup>57</sup>	94 (12) [67–100]	87 (10) [78–100]
Verbal IQ	100 (8) [89–115]	98 (12) [82–110]
Illness duration, yr	NA	12.1 (8.6) [4–28]

IQ = intelligence quotient; NA = not applicable; SD = standard deviation.  
\*Unless otherwise indicated.

**Table 2: Clinical characteristics and comprehension scores for the schizophrenia patients and controls at each sessions**

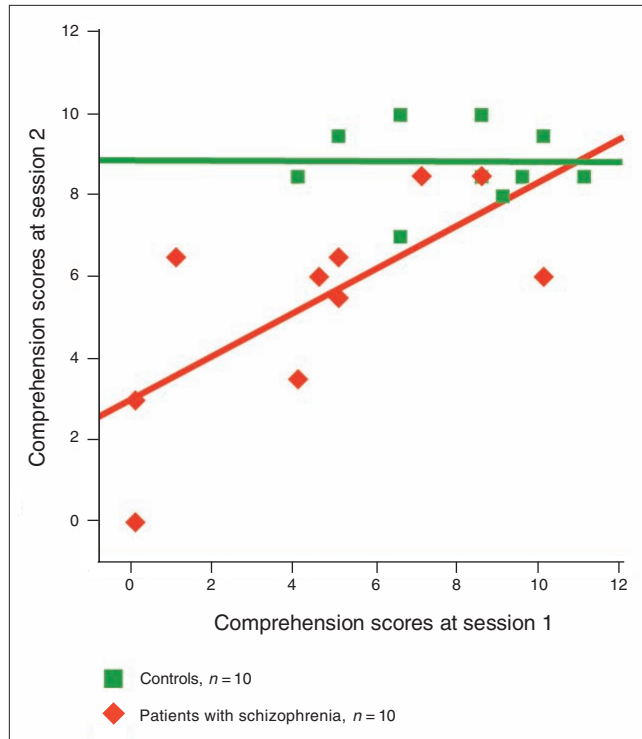
Group; measure	Session; mean (SD) [range]				<i>t</i> <sub>0</sub> value	<i>p</i> value	<i>r</i> value	<i>p</i> value
	Session 1		Session 2					
Schizophrenia patients, <i>n</i> = 10								
Total PANSS	52.2	(8.3) [41–66]	48.9	(8.1) [39–63]	1.3	0.22	0.52	0.12
Positive PANSS	11.8	(3.5) [7–18]	10.1	(2.9) [7–14]	2.15	0.06	0.71	0.021
Negative PANSS	13.8	(5.4) [8–27]	14.4	(4.8) [8–22]	0.38	0.71	0.89	< 0.001
Medication, CPZ	371.5	(228.7) [40–775]	354.7	(153.4) [125–550]	0.74	0.48	0.92	< 0.001
Comprehension scores	4.5	(3.4) [0–10]	5.4	(2.6) [0–8.5]	—	—	0.87	0.001
Controls, <i>n</i> = 10								
Comprehension scores	7.8	(2.3) [4–11]	8.8	(0.9) [7–10]	—	—	–0.015	0.97

CPZ = chlorpromazine equivalents; PANSS = Positive and Negative Syndrome Scale; <sup>58</sup>SD = standard deviation.

a trend toward lower positive PANSS scores at the second session ( $p = 0.06$ ), but the difference between session 1 and 2 was low (mean 1.7, SD 2.5) for a scale ranging from 7 to 49.

*Task performance*

The ANOVA of comprehension scores (Table 2) showed no group  $\times$  session interaction ( $F_{1,18} = 0.002, p = 0.96$ ). Controls performed significantly better than patients in both sessions ( $F_{1,18} = 12.32, p = 0.002$ ). The slight increase in scores between sessions was not significant ( $F_{1,18} = 2.82, p = 0.11$ ).



**Fig. 3:** Comprehension scores at both sessions in patients with schizophrenia and controls. Solid lines represent the regression lines (controls: scores at session 2 =  $-0.006 \times$  scores at session 1 + 8.85,  $R^2 = 0.0002$ ; schizophrenia patients: scores at session 2 =  $0.53 \times$  scores at session 1 + 3,  $R^2 = 0.49$ ).

Patient comprehension scores were strongly correlated across sessions ( $r = 0.70, p = 0.024$ ), but control scores were not correlated across sessions ( $r = -0.015, p = 0.97$ ). However, the range (7–10) of comprehension scores from session 2 was very restricted in this group (Fig. 3).

*Global reproducibility*

Patients had a significantly lower percentage of spatial overlap than controls (30.6% [SD 15.8%] v. 47% [SD 11.7%];  $t_{16.5} = 2.64, p = 0.017$ ; Table 3). This difference was not significant when we controlled for comprehension scores ( $t_{16} = 0.63, p = 0.53$ ). The percentage of spatial overlap was highly correlated with the comprehension scores in patients ( $r = 0.82, p = 0.004$ ) but not in controls ( $r = 0.12, p = 0.73$ ; Fig. 4). No other variable was significantly correlated with the percentage of spatial overlap in either group. The same analyses were performed with the percentage of spatial overlap computed at fixed statistical thresholds; this yielded the same results (Table 3). We found a significant correlation between the  $t$  values corresponding to the 5000 most activated voxels at the 2 sessions (controls:  $r = 0.65, p = 0.005$ ; schizophrenia patients:  $r = 0.42, p = 0.025$ ), showing that the threshold selection was reproducible.

Figure 5 shows the most consistently activated brain areas (i.e., areas activated in most patients or controls for each session). Activation patterns were similar for both sessions. In controls, the most consistently activated voxels were located along the left superior temporal sulcus, right anterior superior temporal sulcus, pars triangularis of the left inferior frontal gyrus and left medial prefrontal cortex. In patients, the most consistently activated areas were restricted to the bilateral anterior and middle parts of the superior temporal sulcus.

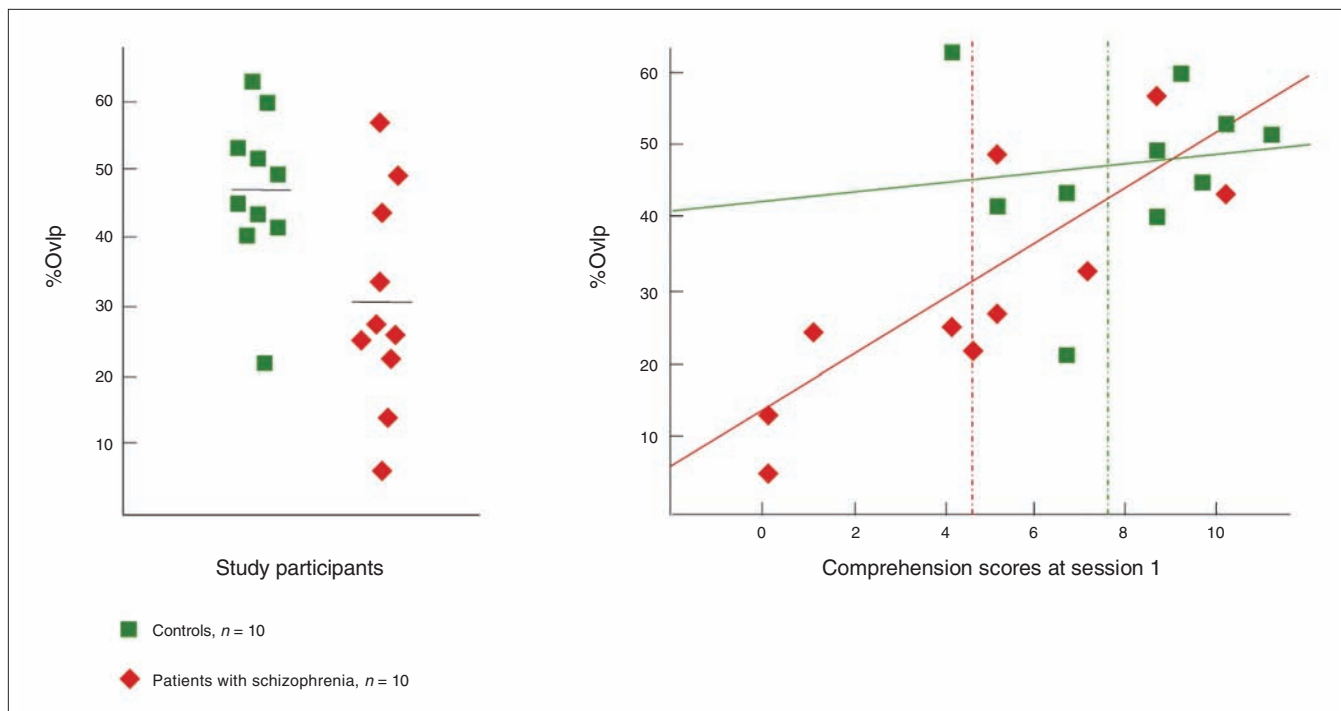
*Local reproducibility*

Figure 6A shows the reproducibility patterns in both groups. Controls had high reproducibility (i.e., low RSD) in the left cortical areas, particularly along the left superior temporal sulcus but also in the frontal inferior gyrus and medial prefrontal cortex. In contrast, the right hemispheric areas were

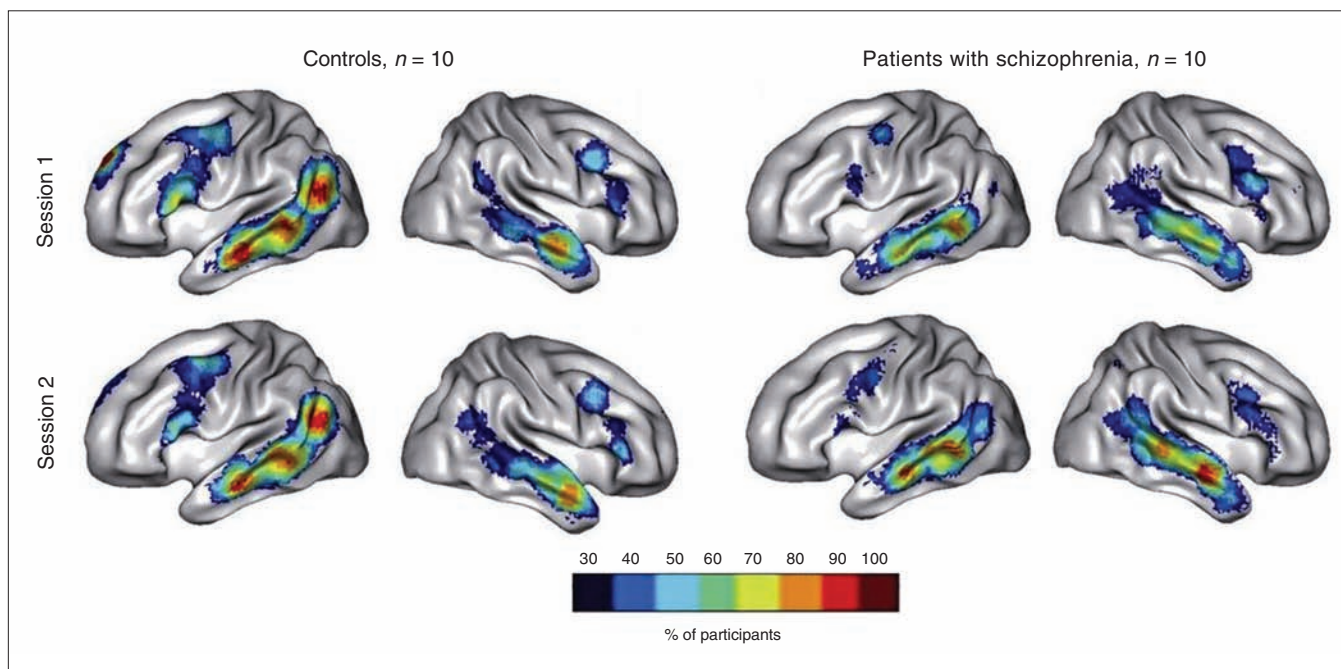
**Table 3: Percentage of overlap in schizophrenia patients and controls**

Measure; session	Group; mean (SD) [range]	
	Controls, n = 10	Schizophrenia patients, n = 10
Percentage of spatial overlap		
Fixed threshold of $p = 0.001^*$ ( $t = 3.15$ )	46.99 (11.13) [23.2–60.98]	24.11 (18.28) [0.51–57.07]
Fixed threshold of $p = 0.01^*$ ( $t = 2.35$ )	45.25 (10.07) [31.07–59.09]	26.35 (16.05) [2.89–51.82]
Fixed threshold of $p = 0.05^*$ ( $t = 1.66$ )	43.76 (10.58) [26.97–58.91]	27.60 (13.08) [5.00–43.96]
Fixed volume of 5000 voxels, %	47.00 (11.58) [21.74–62.92]	30.65 (15.81) [6.26–57.34]
Average $t$ value corresponding to 5000 voxels		
Session 1	3.34 (0.85) [2.47–4.91]	2.26 (0.74) [0.90–3.31]
Session 2	3.23 (1.07) [2.05–5.09]	2.69 (0.49) [2.09–3.86]

SD = standard deviation.  
\*Uncorrected for multiple comparisons.



**Fig. 4:** Global reproducibility in patients with schizophrenia and controls, and correlation between global reproducibility and task performance. **(Left)** The percentage of spatial overlap (%Ovlp) was significantly lower in patients with schizophrenia than in controls ( $t_{16.5} = 2.64, p = 0.017$ ). **(Right)** Percentage of overlap was not correlated with comprehension scores in session 1 for controls ( $r = 0.12, p = 0.73$ ) but was positively correlated with comprehension scores for patients with schizophrenia ( $r = 0.82, p = 0.004$ ). Solid lines represent the regression lines (controls: percentage overlap =  $0.64 \times \text{score} + 41.93, R^2 = 0.02$ ; schizophrenia patients: percentage overlap =  $3.79 \times \text{score} + 13.60, R^2 = 0.68$ ); dotted lines represent the mean comprehension scores in session 1.



**Fig. 5:** Activation probability maps in patients with schizophrenia and controls for each session. For each voxel, the colour scale indicates the percentage of participants in whom this voxel was activated among the 5000 most activated voxels. The maps were superimposed onto an inflated reconstruction of the Montreal Neurological Institute template.

more variable, with the exception of the anterior middle temporal gyrus. In patients, the areas with greatest reproducibility (i.e., areas exhibiting the lowest RSD) were located bilaterally in the middle part of the superior temporal sulcus.

Patients showed significantly less reproducibility than controls in 3 left hemispheric regions: the posterior part of the middle temporal gyrus up to the angular gyrus, the medial part of the superior frontal gyrus, and the pars triangularis of the inferior frontal gyrus (Table 4, Fig. 6B). In contrast, we did not detect a single area in patients that showed higher reproducibility than in controls.

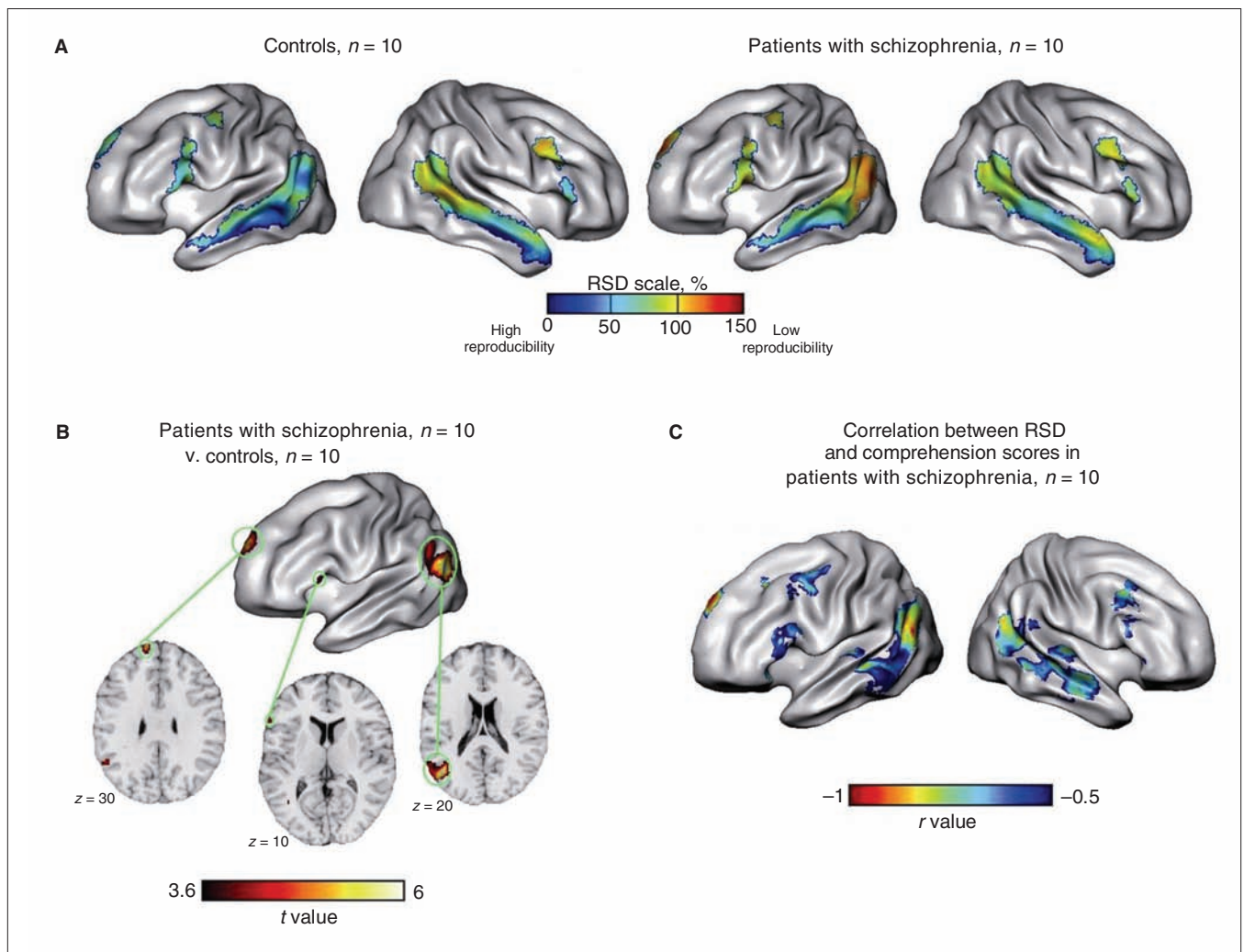
When we included the comprehension scores as a covariate, all differences between groups were removed, except at the cluster located in the posterior part of the left middle temporal gyrus. That cluster had significantly less reproducibility

in patients than in controls; however, after accounting for the comprehension scores, the extent was reduced from 434 to 27 voxels (Table 4).

In patients, the RSD maps were significantly negatively correlated with the comprehension scores in 3 left hemispheric clusters: 2 were located in the posterior part of the middle temporal gyrus and the angular gyrus, and 1 was located in the medial superior frontal gyrus (Table 4, Fig. 6C). These clusters overlapped with 2 of the areas that showed lower reproducibility in patients compared with controls in the between-group comparison analysis.

### Discussion

The main finding of our study was that performance during



**Fig. 6:** Local reproducibility in patients with schizophrenia and controls, and correlation between local reproducibility and task performance. (A) Average (within group) relative standard deviation (RSD) maps show local reproducibility of activation. Low RSD values correspond to high reproducibility. (B) Cortical areas that exhibited higher activation reproducibility in controls than in patients with schizophrenia: (top) SPM  $t$  map and (bottom) selected axial slices of the 2-sample  $t$  test that compared RSD maps of controls and patients were superimposed onto the Montreal Neurological Institute template. (C) Cortical areas that showed significant negative correlations in patients between RSD values and task performance score (session 1). The SPM99  $t$  map of the negative correlation between the local RSD and the performance score was converted into an  $r$  map (Pearson coefficient). The threshold of  $p = 0.001$  corresponds to an  $r$  value of 0.84.



a story comprehension task was the only relevant factor that influenced the reproducibility of fMRI activation in patients with schizophrenia compared with healthy participants. Indeed, without taking into account the effect of task performance, patients with schizophrenia exhibited lower global and local reproducibility compared with controls. Three left cortical areas showed lower reproducibility in patients compared with controls: the posterior part of the middle temporal gyrus up to the angular gyrus, the inferior frontal gyrus and the medial prefrontal cortex. Both percentage of spatial overlap and RSD were strongly correlated with task performance in patients. Accordingly, the inclusion of task performance as a covariate removed the intergroup differences in spatial overlap at the global level and in RSD at the local level, except in a small cluster located at the posterior part of middle temporal gyrus.

In this study, we controlled the various factors that might affect the reproducibility of fMRI results. Preprocessing of functional data was specifically designed to reduce signal misregistration by applying the same normalization parameters onto coregistered fMRI data from both sessions. Hence, misregistration between the 2 sessions was not a factor that influenced cross-session reproducibility. Furthermore, we minimized a potential learning effect due to the repetition of the same paradigm by using a long intersession interval of 21 months. Indeed, the differences in intersession task performances were neither significant nor correlated with the percentage of spatial overlap in controls and patients. Finally, the percentage of spatial overlap and RSD computations were not affected by dependence on a statistical threshold.

Our results in healthy participants are in line with previous studies that used a similar functional paradigm. With both types of overlap (fixed volume and fixed threshold), we found a mean percentage of spatial overlap of 47%, which was within the range of the percentages (35%–50%)<sup>37,38</sup> reported in 2 studies contrasting speech comprehension with reversed speech. However, such comparisons should be considered with caution given the differences in methods and tasks. Indeed, to date, there is no consensus about the most

appropriate method to study reproducibility or the cut-off points that should be used to make a qualitative judgment about the degree of reproducibility in fMRI.

At the local level, controls exhibited higher reproducibility in the left cortical regions compared with the right regions. This is consistent with the greater involvement of the left hemisphere in language processing in healthy right-handed individuals.<sup>62</sup> Highly activated regions are typically more reliable than less activated regions.<sup>61</sup>

Our main hypothesis was that deficient task performance, reflecting cognitive impairments associated with the illness, could explain the lower reproducibility of activation in patients with schizophrenia compared with controls. That is, the reduced reproducibility of fMRI activation in patients would be correlated to a deficiency in task performance and not a systematic phenomenon due to illness. Our results are consistent with this hypothesis. Indeed, patients had lower global reproducibility (percentage of spatial overlap of 30.6%) than controls (47%), but the local analysis with RSD maps showed that this lower reproducibility did not concern the whole network recruited by the story comprehension task. It was restricted to areas located beyond the language-related perisylvian cortex, in areas that underpinned higher-order processes involved in narrative comprehension.<sup>63</sup>

The left inferior frontal gyrus and posterior middle temporal gyrus are thought to be involved in the resolution of textual ambiguities and the construction of a representation of the global meaning of text.<sup>64,65</sup> Furthermore, medial prefrontal cortex involvement in the context of a narrative may reflect social cognition processes and especially “theory of mind,”<sup>66,67</sup> as well as more general inferential processes involved when listening to logically connected sentences.<sup>68</sup> These processes are known to be impaired in schizophrenia.<sup>69,70</sup> In contrast, the anterior and middle superior temporal sulcus are involved in lexicosemantic processing,<sup>65</sup> and activations in these areas exhibited similar reproducibility in patients and healthy participants. This strengthens the hypothesis that lower reproducibility in schizophrenia patients would preferentially concern integrative areas that sustain high-level cognitive processes. Consistently, in a

**Table 4: Areas of lower reproducibility in schizophrenia patients ( $n = 10$ ) compared with controls ( $n = 10$ ) with and without comprehension scores as a covariate and areas exhibiting a significant\* positive correlation between reproducibility and comprehension scores in schizophrenia patients**

Comparison; brain region	Estimated Brodmann area	MNI coordinate			Z value	Cluster size
		x	y	z		
Areas of lower reproducibility† in schizophrenia patients v. controls						
Left middle temporal gyrus extending to the angular gyrus	39	-40	-64	18	4.28	434
Left medial superior frontal gyrus	9	-12	64	30	3.97	61
Left pars triangularis of the inferior frontal gyrus	44	-58	26	10	3.54	13
Areas of lower reproducibility† in patients v. controls including comprehension scores as covariate						
Left middle temporal gyrus	39	-40	-64	18	3.45	27
Areas of significant positive correlation between reproducibility† and comprehension scores in patients						
Left middle temporal gyrus extending to the angular gyrus	39	-52	-64	20	4.47	65
Left middle temporal gyrus extending to the angular gyrus	39	-32	-60	24	3.54	16
Left medial superior frontal gyrus	9	-8	50	38	4.14	63

MNI = Montreal Neurological Institute.

\* $p < 0.001$ , uncorrected for multiple comparisons.

†The local index of reproducibility was relative standard deviation (RDS); low RSD values correspond to high reproducibility.

less complex task that preferentially targeted regions involved in the early phases of language processing, patients showed reproducibility similar to that observed in healthy participants.<sup>71</sup> This is in line with results observed in patients with schizophrenia compared with controls during a working memory task; a similar reproducibility was observed in regions associated with motor function, whereas a lower reproducibility was observed in patients in regions associated with higher-level cognitive functions.<sup>51</sup>

Comprehension scores give an indirect and global assessment of performance during the story comprehension task. These scores were highly correlated across sessions, and patients with schizophrenia had lower scores than controls. These scores appeared to be the main factor correlated to low reproducibility in patients. Indeed, comprehension scores in session 1 accounted for 67% of the spatial overlap variance in patients.

At the local level, the between-group differences in reproducibility were no longer significant after the scores were used as covariates. Locally, this performance effect only influenced areas of greater RSD in patients (left medial prefrontal cortex, left posterior superior temporal sulcus). When we included the comprehension scores as a covariate, only a few voxels in the posterior part of the middle temporal gyrus remained significantly more variable in patients than in controls. This residual difference between patients and controls could be explained by the fact that comprehension scores are neither a comprehensive nor direct assessment of the participants' behaviour during the task. There is a recall component in the comprehension scores because there is necessarily a delay between scanning and the comprehension test. It is likely that a more precise measure would have explained more of the differences. Indeed, the area that was still significantly less reproducible in patients was very close to the peak of the effect of comprehension scores on reproducibility of activations. Nevertheless, we cannot exclude that some other factors linked to the disease and its evolution might have contributed to this residual difference in reproducibility. Our patients, however, were reasonably stable, as assessed by the symptom scales.

All analyses dealing with the effect of comprehension scores were also conducted with scores from session 2, and the same pattern of results was found. Activation probability maps showed that most patients with schizophrenia did not recruit all of the cortical regions that healthy participants recruited during the task. One can speculate that the cerebral activity in these regions would be less correlated with the functional paradigm that led to poor signal strength and, thus, poor reproducibility.

In this study, we did not detect any factor, other than performance level, that significantly influenced the reproducibility of activation. In particular, clinical variables, including duration of illness, positive and negative PANSS scores and dose of antipsychotic medication, were not correlated with the percentage of spatial overlap in patients. All patients were clinically stable, with no significant changes in symptoms or medication doses, which were correlated strongly across sessions. The patients included in the present study had been ill for an average of 12 years; the effects of the illness course on

the brain may be less pronounced at this stage of the disease than during the first few years.<sup>72</sup>

### Limitations

Several limitations affected the present study. The intersession interval was shorter in patients than in controls. However, this was not correlated with the percentage of spatial overlap. The patients were heterogeneous with regard to clinical subtype and antipsychotic medication. Because of the small sample sizes, we could not investigate the influence of these factors on reproducibility. However, medication dose expressed in chlorpromazine equivalents was not correlated with spatial overlap. Finally, our results only concerned language-related tasks, and they should not be directly generalized to other cognitive domains.

### Conclusion

Our findings showed that the lower reproducibility of fMRI activation elicited by a language paradigm in patients with schizophrenia compared with healthy participants concerned only a part of the network of active areas (mostly the "integrative" areas), and the lower reproducibility was mainly explained by deficient task performance in patients. Therefore, when designing a longitudinal fMRI study in patients with schizophrenia, task performance should be carefully controlled.

The fMRI paradigm should yield reliable activations and be sensitive to illness-related changes. Ideally, there should not be any ambiguity regarding the cause of the observed difference, change in task performance or behaviour or evolution of the illness. If a modification in the terms of task performance occurs between 2 sessions, unless the effects of task performance and illness can be disentangled statistically, the results should be interpreted as neural correlates of improved or diminished performance at the task, in a context of illness. This would be the case when assessing the effects of a therapeutic intervention on brain function. If no change in behaviour is observed, and provided the paradigm (in this context, a "probe") yields reliable activations in stable patients, the neuroimaging observations can be interpreted as reflecting the evolution of the illness. In the latter case, the performance level of the patients should be high enough to guarantee reproducibility. Finally, there is nothing intrinsically associated with the disease that would preclude longitudinal fMRI studies in this population.

**Acknowledgements:** We thank G. Perchey, N. Delcroix and F. Lambertson for their assistance. This work was funded by the French Health Ministry in a Programme Hospitalier de Recherche Clinique and by the French Research Ministry.

**Competing interests:** None declared.

**Contributors:** Drs. Dollfus and Tzourio-Mazoyer designed the study. Drs. Maïza, Dollfus, Razafimandimby and Tzourio-Mazoyer acquired the data, which Drs. Maïza, Mazoyer and Hervé analyzed. Drs. Maïza and Hervé wrote the article, which Drs. Maïza, Mazoyer, Razafimandimby, Dollfus and Tzourio-Mazoyer reviewed. All authors approved the article that was submitted for publication.

## References

- Bertolino A, Blasi G, Caforio G, et al. Functional lateralization of the sensorimotor cortex in patients with schizophrenia: effects of treatment with olanzapine. *Biol Psychiatry* 2004;56:190-7.
- Mendrek A, Laurens KR, Kiehl KA, et al. Changes in distributed neural circuitry function in patients with first-episode schizophrenia. *Br J Psychiatry* 2004;185:205-14.
- Razafimandimby A, Maıza O, Herve PY, et al. Stability of functional language lateralization over time in schizophrenia patients. *Schizophr Res* 2007;94:197-206.
- Wykes T, Brammer M, Mellers J, et al. Effects on the brain of a psychological treatment: cognitive remediation therapy: functional magnetic resonance imaging in schizophrenia. *Br J Psychiatry* 2002;181:144-52.
- Wexler BE, Anderson M, Fulbright RK, et al. Preliminary evidence of improved verbal working memory performance and normalization of task-related frontal lobe activation in schizophrenia following cognitive exercises. *Am J Psychiatry* 2000;157:1694-7.
- Schlagenhauf F, Wustenberg T, Schmack K, et al. Switching schizophrenia patients from typical neuroleptics to olanzapine: effects on BOLD response during attention and working memory. *Eur Neuropsychopharmacol* 2008;18:589-99.
- Reske M, Kellermann T, Habel U, et al. Stability of emotional dysfunctions? A long-term fMRI study in first-episode schizophrenia. *J Psychiatr Res* 2007;41:918-27.
- Meisenzahl EM, Scheuerecker J, Zipse M, et al. Effects of treatment with atypical neuroleptic quetiapine on working memory function: a functional MRI follow-up investigation. *Eur Arch Psychiatry Clin Neurosci* 2006;256:522-31.
- Snitz BE, MacDonald A III, Cohen JD, et al. Lateral and medial hypofrontality in first-episode schizophrenia: functional activity in a medication-naïve state and effects of short-term atypical antipsychotic treatment. *Am J Psychiatry* 2005;162:2322-9.
- Yurgelun-Todd DA, Coyle JT, Gruber SA, et al. Functional magnetic resonance imaging studies of schizophrenic patients during word production: effects of D-cycloserine. *Psychiatry Res* 2005;138:23-31.
- Nahas Z, George MS, Horner MD, et al. Augmenting atypical antipsychotics with a cognitive enhancer (donepezil) improves regional brain activity in schizophrenia patients: a pilot double-blind placebo controlled BOLD fMRI study. *Neurocase* 2003;9:274-82.
- Honey GD, Bullmore ET, Soni W, et al. Differences in frontal cortical activation by a working memory task after substitution of risperidone for typical antipsychotic drugs in patients with schizophrenia. *Proc Natl Acad Sci U S A* 1999;96:13432-7.
- Stip E, Fahim C, Mancini-Marie A, et al. Restoration of frontal activation during a treatment with quetiapine: an fMRI study of blunted affect in schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 2005;29:21-6.
- Davis CE, Jeste DV, Eyler LT. Review of longitudinal functional neuroimaging studies of drug treatments in patients with schizophrenia. *Schizophr Res* 2005;78:45-60.
- Cho RY, Ford JM, Krystal JH, et al. Functional neuroimaging and electrophysiology biomarkers for clinical trials for cognition in schizophrenia. *Schizophr Bull* 2005;31:865-9.
- Miki A, Raz J, van Erp TG, et al. Reproducibility of visual activation in functional MR imaging and effects of postprocessing. *AJNR Am J Neuroradiol* 2000;21:910-5.
- Kiehl KA, Liddle PF. Reproducibility of the hemodynamic response to auditory oddball stimuli: a six-week test-retest study. *Hum Brain Mapp* 2003;18:42-52.
- Specht K, Willmes K, Shah NJ, et al. Assessment of reliability in functional imaging studies. *J Magn Reson Imaging* 2003;17:463-71.
- Rombouts SA, Barkhof F, Hoogenraad FG, et al. Test-retest analysis with functional MR of the activated area in the human visual cortex. *AJNR Am J Neuroradiol* 1997;18:1317-22.
- Rombouts SA, Barkhof F, Hoogenraad FG, et al. Within-subject reproducibility of visual activation patterns with functional magnetic resonance imaging using multislice echo planar imaging. *Magn Reson Imaging* 1998;16:105-13.
- Waldvogel D, van Gelderen P, Immisch I, et al. The variability of serial fMRI data: correlation between a visual and a motor task. *Neuroreport* 2000;11:3843-7.
- Yetkin FZ, McAuliffe TL, Cox R, et al. Test-retest precision of functional MR in sensory and motor task activation. *AJNR Am J Neuroradiol* 1996;17:95-8.
- Yoo SS, Wei X, Dickey CC, et al. Long-term reproducibility analysis of fMRI using hand motor task. *Int J Neurosci* 2005;115:55-77.
- Marshall I, Simonotto E, Deary IJ, et al. Repeatability of motor and working-memory tasks in healthy older volunteers: assessment at functional MR imaging. *Radiology* 2004;233:868-77.
- Loubinoux I, Carel C, Alary F, et al. Within-session and between-session reproducibility of cerebral sensorimotor activation: a test-retest effect evidenced with functional magnetic resonance imaging. *J Cereb Blood Flow Metab* 2001;21:592-607.
- Raemaekers M, Vink M, Zandbelt B, et al. Test-retest reliability of fMRI activation during prosaccades and antisaccades. *Neuroimage* 2007;36:532-42.
- Swallow KM, Braver TS, Snyder AZ, et al. Reliability of functional localization using fMRI. *Neuroimage* 2003;20:1561-77.
- Havel P, Braun B, Rau S, et al. Reproducibility of activation in four motor paradigms. *J Neurol* 2006;253:471-6.
- Liu JZ, Zhang L, Brown RW, et al. Reproducibility of fMRI at 1.5 T in a strictly controlled motor task. *Magn Reson Med* 2004;52:751-60.
- Miller MB, Van Horn JD, Wolford GL, et al. Extensive individual differences in brain activations associated with episodic retrieval are reliable over time. *J Cogn Neurosci* 2002;14:1200-14.
- Wagner K, Frings L, Quiske A, et al. The reliability of fMRI activations in the medial temporal lobes in a verbal episodic memory task. *Neuroimage* 2005;28:122-31.
- Machielsen WC, Rombouts SA, Barkhof F, et al. FMRI of visual encoding: reproducibility of activation. *Hum Brain Mapp* 2000;9:156-64.
- Wei X, Yoo SS, Dickey CC, et al. Functional MRI of auditory verbal working memory: long-term reproducibility analysis. *Neuroimage* 2004;21:1000-8.
- Aron AR, Gluck MA, Poldrack RA. Long-term test-retest reliability of functional MRI in a classification learning task. *Neuroimage* 2006;29:1000-6.
- Neumann J, Lohmann G, Zysset S, et al. Within-subject variability of BOLD response dynamics. *Neuroimage* 2003;19:784-96.
- Rutten GJ, Ramsey NF, van Rijen PC, et al. Reproducibility of fMRI-determined language lateralization in individual subjects. *Brain Lang* 2002;80:421-37.
- Harrington GS, Buonocore MH, Farias ST. Intrasubject reproducibility of functional MR imaging activation in language tasks. *AJNR Am J Neuroradiol* 2006;27:938-44.
- Maldjian JA, Laurienti PJ, Driskill L, et al. Multiple reproducibility indices for evaluation of cognitive functional MR imaging paradigms. *AJNR Am J Neuroradiol* 2002;23:1030-7.
- McGonigle DJ, Howseman AM, Athwal BS, et al. Variability in fMRI: an examination of intersession differences. *Neuroimage* 2000;11:708-34.
- Smith SM, Beckmann CF, Ramnani N, et al. Variability in fMRI: a re-examination of inter-session differences. *Hum Brain Mapp* 2005;24:248-57.
- Fernandez G, Specht K, Weis S, et al. Intrasubject reproducibility of presurgical language lateralization and mapping using fMRI. *Neurology* 2003;60:969-75.
- Knecht S, Jansen A, Frank A, et al. How atypical is atypical language dominance? *Neuroimage* 2003;18:917-27.
- Wei X, Yoo SS, Dickey CC, et al. Functional MRI of auditory verbal working memory: long-term reproducibility analysis. *Neuroimage* 2004;21:1000-8.

44. Tjandra T, Brooks JC, Figueiredo P, et al. Quantitative assessment of the reproducibility of functional activation measured with BOLD and MR perfusion imaging: implications for clinical trial design. *Neuroimage* 2005;27:393-401.
45. Brannen JH, Badie B, Moritz CH, et al. Reliability of functional MR imaging with word-generation tasks for mapping Broca's area. *AJNR Am J Neuroradiol* 2001;22:1711-8.
46. Kurland J, Naeser MA, Baker EH, et al. Test-retest reliability of fMRI during nonverbal semantic decisions in moderate-severe nonfluent aphasia patients. *Behav Neurol* 2004;15:87-97.
47. Chen EE, Small SL. Test-retest reliability in fMRI of language: group and task effects. *Brain Lang* 2007;102:176-85.
48. Kimberley TJ, Birkholz DD, Hancock RA, et al. Reliability of fMRI during a continuous motor task: assessment of analysis techniques. *J Neuroimaging* 2008;18:18-27.
49. Whalley HC, Gountouna VE, Hall J, et al. fMRI changes over time and reproducibility in unmedicated subjects at high genetic risk of schizophrenia. *Psychol Med* 2009;39:1189-99.
50. Clement F, Belleville S. Test-retest reliability of fMRI verbal episodic memory paradigms in healthy older adults and in persons with mild cognitive impairment. *Hum Brain Mapp* 2009;30:4033-47.
51. Manoach DS, Halpern EF, Kramer TS, et al. Test-retest reliability of a functional MRI working memory paradigm in normal and schizophrenic subjects. *Am J Psychiatry* 2001;158:955-8.
52. de Vet HC, Terwee CB, Knol DL, et al. When to use agreement versus reliability measures. *J Clin Epidemiol* 2006;59:1033-9.
53. Manoach DS, Gollub RL, Benson ES, et al. Schizophrenic subjects show aberrant fMRI activation of dorsolateral prefrontal cortex and basal ganglia during working memory performance. *Biol Psychiatry* 2000;48:99-109.
54. Dollfus S, Razafimandimby A, Delamillieure P, et al. Atypical hemispheric specialization for language in right-handed schizophrenia patients. *Biol Psychiatry* 2005;57:1020-8.
55. Wechsler D. *Manuel de l'Echelle d'intelligence de Wechsler pour Adultes* — 3e édition. Paris: ECPA; 2000.
56. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971;9:97-113.
57. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders. 4th ed.* Washington (DC): The Association; 1994.
58. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 1987;13:261-76.
59. Spitzer RL, Williams JB, Gibbon M, et al. The Structured Clinical Interview for DSM-III-R (SCID). I: History, rationale, and description. *Arch Gen Psychiatry* 1992;49:624-9.
60. Mazoyer B, Tzourio N, Frak V, et al. The cortical representation of speech. *J Cogn Neurosci* 1993;5:467-79.
61. Caceres A, Hall DL, Zelaya FO, et al. Measuring fMRI reliability with the intra-class correlation coefficient. *Neuroimage* 2009;45:758-68.
62. Springer JA, Binder JR, Hammeke TA, et al. Language dominance in neurologically normal and epilepsy subjects: a functional MRI study. *Brain* 1999;122:2033-46.
63. Mar RA. The neuropsychology of narrative: story comprehension, story production and their interrelation. *Neuropsychologia* 2004;42:1414-34.
64. Virtue S, Haberman J, Clancy Z, et al. Neural activity of inferences during story comprehension. *Brain Res* 2006;1084:104-14.
65. Vigneau M, Bouchoux V, Herve PY, et al. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage* 2006;30:1414-32.
66. Xu J, Kemeny S, Park G, et al. Language in context: emergent features of word, sentence, and narrative comprehension. *Neuroimage* 2005;25:1002-15.
67. Fletcher PC, Happe F, Frith U, et al. Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. *Cognition* 1995;57:109-28.
68. Ferstl EC, von Cramon DY. What does the frontomedian cortex contribute to language processing: Coherence or theory of mind? *Neuroimage* 2002;17:1599-612.
69. Corcoran R. Inductive reasoning and the understanding of intention in schizophrenia. *Cogn Neuropsychiatry* 2003;8:223-35.
70. Sprong M, Schothorst P, Vos E, et al. Theory of mind in schizophrenia: meta-analysis. *Br J Psychiatry* 2007;191:5-13.
71. Maiza O, Razafimandimby A, Herve PY, et al. Reproducibility of fMRI activations in patients with schizophrenia. *Schizophr Bull* 2009;35S1:162.
72. Gur RE, Cowell P, Turetsky BI, et al. A follow-up magnetic resonance imaging study of schizophrenia. Relationship of neuroanatomical changes to clinical and neurobehavioral measures. *Arch Gen Psychiatry* 1998;55:145-52.


# ZELDOX<sup>®</sup>


## COVERED BY FORMULARY ACROSS CANADA<sup>†‡</sup>

Now in Ontario as a General Benefit!

† Except PEI


‡ ZELDOX is listed as a General Benefit on the Ontario Drug Benefit Formulary and by the New Brunswick Prescription Drug Formulary. ZELDOX is listed as a full benefit to both lists of the Quebec formulary. ZELDOX has been added to the Alberta Health and Wellness Drug Benefit List as a full benefit. ZELDOX is covered by the Saskatchewan Drug Plan & Extended Benefits. ZELDOX has been added to Part 1 of the Manitoba Drug Benefit and Interchangeability List. ZELDOX is listed as a Limited Coverage Drug pursuant to the BC Pharmacare Special Authority Program. ZELDOX is also covered by the Nova Scotia Pharmacare, and by the Newfoundland and Labrador Prescription Drug Program with Special Authorization. ZELDOX has been included as a limited use benefit on the Non-Insured Health Benefits (NIHB) Drug Benefit List.





Working together for a healthier world<sup>™</sup>

Member



TM Pfizer Inc., used under license  
ZELDOX<sup>®</sup> Pfizer Products Inc., Pfizer Canada Inc., licensee  
© 2010 Pfizer Canada Inc., Kirkland, Quebec. H9J 2M5