OBJECTIVE: The aim of this study was to analyze the usefulness of preoperative language functional magnetic resonance imaging (fMRI), by correlating fMRI data with intraoperative cortical stimulation results for patients with brain tumors.

METHODS: Naming and verb generation tasks were used, separately or in combination, for 14 right-handed patients with tumors in the left hemisphere. fMRI data obtained were analyzed with SPM software, with two standard analysis thresholds ($P < 0.005$ and then $P < 0.05$). The fMRI data were then registered in a frameless stereotactic neuronavigational device and correlated with direct brain mapping results. We used a statistical model with the fMRI information as a predictor, spatially correlating each intraoperatively mapped cortical site with fMRI data integrated in the neuronavigational system (site-by-site correlation). Eight patients were also studied with language fMRI postoperatively, with the same acquisition protocol.

RESULTS: We observed high variability in signal extents and locations among patients with both tasks. The activated areas were located mainly in the left hemisphere in the middle and inferior frontal gyri (F2 and F3), the superior and middle temporal gyri (T1 and T2), and the supramarginal and angular gyri. A total of 426 cortical sites were tested for each task among the 14 patients. In frontal and temporoparietal areas, poor sensitivity of the fMRI technique was observed for the naming and verb generation tasks (22 and 36%, respectively) with $P < 0.005$ as the analysis threshold. Although not perfect, the specificity of the fMRI technique was good in all conditions (97% for the naming task and 98% for the verb generation task). Better correlation (sensitivity, 59%; specificity, 97%) was achieved by combining the two fMRI tasks. Variation of the analysis threshold to $P < 0.05$ increased the sensitivity to 66% while decreasing the specificity to 91%. Postoperative fMRI data (for the cortical brain areas studied intraoperatively) were in accordance with brain mapping results for six of eight patients. Complete agreement between pre- and postoperative fMRI studies and direct brain mapping results was observed for only three of eight patients.

CONCLUSION: With the paradigms and analysis thresholds used in this study, language fMRI data obtained with naming or verb generation tasks, before and after surgery, were imperfectly correlated with intraoperative brain mapping results. A better correlation could be obtained by combining the fMRI tasks. The overall results of this study demonstrated that language fMRI could not be used to make critical surgical decisions in the absence of direct brain mapping. Other acquisition protocols are required for evaluation of the potential role of language fMRI in the accurate detection of essential cortical language areas.

KEY WORDS: Brain tumors, Cortical mapping, Functional magnetic resonance imaging, Image-guided surgery
 Patients

Fourteen patients (nine male and five female patients; age range, 14–68 yr; average age, 43 yr) with brain tumors in presumed language areas were studied with language fMRI data integrated with infrared light-based frameless stereotactic devices and intraoperative cortical stimulation findings in this study. Intraoperative cortical stimulation has become a standard method for localizing language areas during neurosurgical procedures and is considered the standard method for cortical brain mapping. To better understand the usefulness of language fMRI among neurosurgical patients, we prospectively analyzed the relationship between language fMRI data and intraoperative cortical stimulation findings in this study.

 Patients and Methods

Functional magnetic resonance imaging (fMRI) is now used in different areas of neurosurgery, such as surgical planning for brain tumor surgery (1, 24, 25, 30, 33) or epilepsy or pain surgery (12, 25). Several authors have emphasized the good spatial correlation between intraoperative cortical stimulation and motor fMRI findings (13, 25, 30, 33), sometimes using fMRI data integrated with infrared light-based frameless stereotactic devices (24, 25, 27, 30). Language fMRI has been successfully used to test language dominance among volunteers and patients (3, 9). More specific correlations between language fMRI data and direct cortical mapping findings have rarely been studied (5, 6, 15, 22, 27), however, and uncertainties persist regarding the ability of language fMRI to detect cortical language areas among neurosurgical patients (27).

Intraoperative cortical stimulation has become a standard method for localizing language areas during neurosurgical procedures (16, 18, 20) and is considered the standard method for cortical brain mapping. To better understand the usefulness of language fMRI among neurosurgical patients, we prospectively analyzed the relationship between language fMRI data and intraoperative cortical stimulation findings in this study.

 Patients

Fourteen patients (nine male and five female patients; age range, 14–68 yr; average age, 43 yr) with brain tumors in presumed language areas were studied with language fMRI data integrated with infrared light-based frameless stereotactic devices (StealthStation; Sofamor-Danek, Broomfield, CO) and cortical brain mapping, between September 1998 and April 2001. All subjects were native (n = 13) or high-proficiency (n = 1) French speakers. These patients represented a subgroup of 26 patients who underwent language fMRI at our institution during this period; 12 patients were excluded from the final analysis because they did not meet all of the requirements of the study (the patients were not surgically treated, were operated on without brain mapping, or exhibited modification of their language performance between the fMRI procedure and surgery). Before the fMRI procedure (and after the operation), all patients underwent neuropsychological and language examinations, to confirm their ability to perform the required tasks. This testing included evaluations of written and oral comprehension and denomination, language fluency, reading, computation, dictation, repetition, copying, and object handling. Aphasic patients and those with more than 10% errors in the naming tasks were excluded from the study. Although all patients spontaneously stated that they were right-handed, the degree of handedness of the subjects was assessed with the Edinburgh Handedness Inventory (19): the subjects were asked what hand they usually preferred to perform various daily acts, and handedness results ranged from 100 for completely right-handed subjects to −100 for completely left-handed subjects. In our study, the handedness indices ranged between 100 and 60, which indicated a clear right predominance for all subjects. All of the patients and their families gave their informed consent for language area study with fMRI.

fMRI Activation Tasks Chosen

Two different tasks were chosen, i.e., a naming task and a verb generation task. These tasks were chosen for their ability to activate both receptive and expressive language areas and for their ability to activate usually large cerebral regions. Thirty images representing concrete nouns (such as daily-life objects or animals) were used during the fMRI procedures, and the same images were used intraoperatively for direct brain mapping. We preferred to use whispered rather than silent responses during the fMRI experiments, to confirm the patients’ ability to perform the task. Patients were positioned in the head coil of a 1.5-T Magnetom Vision magnetic resonance imaging (MRI) scanner (Siemens, Erlangen, Germany). Optimization of the magnetic field was performed with the automatic map-shim procedure, to reach a gradient tolerance of 0.001 mT/m. fMRI data were obtained with a gradient echo/echo-planar imaging single-shot sequence (TE, 60 ms; flip angle, 90 degrees; slice number, 10; matrix size, 64 × 64; field of view, 200 mm; slice thickness, 5 mm; distance factor, 0.5 mm). The 10 slices were positioned parallel to the anterior commissure-posterior commissure axis from the base of the brain to the vertex. We used an automated shim procedure to improve the magnetic field homogeneity. A staff member was always present near the patient during the acquisitions, to control the procedure, to encourage the patient to perform the task as well as possible, and to ensure that the patient followed the start and stop signals. During the procedure, the patient alternated periods of rest and periods of activation. Each period (rest or activation) lasted 30 seconds, with 10 slices being acquired every 3 seconds. Alternating rest and activation periods were repeated six times, with each procedure beginning with a period of rest. Each period was controlled vocally via the headphones. During the 30 seconds of the naming task, the patients were asked to name five objects shown by special MRI-design glasses (Resonance Technology Inc., Northridge, CA) at the rate of one image every 5 seconds, after which a white screen was shown for 1 second. For the verb generation task, the subjects were asked to find verbs in relation to the objects presented. During the rest periods, the patients were asked to relax, with eyes opened. For all paradigms, the control task was presented first. Patients’ verbal responses to stimuli were controlled and noted during the procedure. No patient was included in this study if more than 10% errors were made during the fMRI naming and verb generation procedures.

The fMRI data were analyzed with SPM 96 (7) and SPM 99 (when it became available) software, on a Sun SPARC workstation (Sun Microsystems, Inc., Mountain View, CA). The first three images of each run were discarded, to allow signal stabilization, and the remaining slices were realigned to correct for the subjects’ movements during scanning, using the
first volume of images as a reference. Detection of the activated voxels was performed on a pixel-by-pixel basis. The initial analysis threshold used for all patients was \( P < 0.005 \), and the cluster extent was more than 25 pixels. A second analysis threshold \( (P < 0.05) \) was also used, to determine the effects of the analysis threshold on the final results. An analysis combining fMRI verb generation and naming tasks (adding procedure) was also performed, and results were correlated with brain mapping findings. If large movements of the head (>5 mm) occurred, motion artifacts were not corrected and the data were excluded from analysis. We used the general linear model implemented in SPM, where conditions (rest or activation) represent independent variables, after global normalization to cancel differences among scans.

Neuronavigation Registration Procedures

The neuronavigational data were acquired from a three-dimensional (3D) data set obtained with a 3D magnetization prepared rapid acquisition gradient echo sequence (TR, 15 ms; TE, 7 ms; flip angle, 12 degrees; 128 partitions; field of view, 300 mm; matrix size, 256 \( \times \) 256; slab thickness, 150 mm; voxel size, 1 \( \times \) 1 \( \times \) 1 mm\(^3\); number of acquisitions, 1; time of acquisition, 10 min). The center of the 3D block was positioned according to the axis joining the anterior commissure and the posterior commissure. These 3D sequences lasted 13 minutes each. The fMRI views were reformatted for integration into the anatomic images with ANALYZE software (Mayo Clinic, Rochester, MN), ImMerge software (InnovMetric Software, Sainte-Foy, Quebec, Canada), and the graphic tools of the neuronavigational system. After views were overlaid on axial anatomic slices, detected functional areas of activation were edited in color with a standard computer printer, to be directly used by the neurosurgeon for patient counseling and surgical planning. The patients were surgically treated within 1 to 7 days after preoperative functional brain mapping.

Cortical Mapping and Correlation Procedures

A strict frameless stereotactic technique was used during surgery, to minimize the technical problems of the neuronavigational system (procedures demonstrating >3 mm of inaccuracy at the beginning of the operation were rejected, and correlation of patient versus preoperative MRI data was performed until a satisfactory level of spatial accuracy was obtained). All of the spatial correlations between fMRI data and cortical mapping results were performed just after opening of the dura and before any tumor removal, which could have been a source of spatial mislocalization because of brain shift.

Patients underwent surgery with the conscious surgery technique (18). After standard premedication, a three-point headholder was applied, the sites of the pins and the incision were well infiltrated with bupivacaine, and 0.05 to 0.1 mg of fentanyl and 40 mg of propofol were administered intravenously. The patients remained sedated (drowsy) with a continuous infusion of propofol during the skin incision and craniotomy, until speech mapping was required. Intraoperative cortical stimulation was used to localize areas of functional cortex after determination of the afterdischarge threshold with electrocorticography. The cortex was directly stimulated with the bipolar electrode of an Ojemann cortical stimulator (1-mm electrodes separated by 5 mm; Radionics, Burlington, MA). The current amplitude was progressively increased by 1 mA, beginning at 2 mA. We used a standard procedure of stimulation with biphasic square-wave pulses of 1 millisecond at 60 Hz, with a maximal train duration of 4 seconds. When a functional site was identified, it was marked with a sterile marker (0.25 cm\(^2\)) and then another area (5 mm distant) was tested. We studied the entire area exposed during craniotomy with cortical and subcortical mapping. Our policy during tumor removal was to spare the functional areas identified with this testing, by resecting tumor tissue no more than 1 cm from eloquent cortex (distance of the resection margin from the nearest functional site) and preserving the subcortical fibers.

Conditions of Correlation

The exact locations of functional sites were compared by using 3D reconstructions of the brain surface with integrated functional data (Fig. 1). Strict spatial conditions of correlation were defined as follows. fMRI and direct cortical mapping data were said to be correlated if they were within the same 1-cm range (matching criteria), according to the high accuracy level provided by the neuronavigational system. Deep fMRI activations (intrasulcal activations) were considered to match cortical mapping results if they were within the same 1-cm area and in the same sulcus. If direct brain mapping within a sulcus was not possible for any reason, then fMRI data were not included in the correlation and were considered to be lost for strict correlation. Although this can be a matter of debate, the essential language sites used for correlation were associated with 1) speech arrest, 2) typical anomia, or 3) hesitations, paraphrasias, neologisms, or consistent delays in response. Language locations were tested at least three times. Precentral gyrus sites producing speech arrest, which were associated with the articulatory final process of speech, were also included in the correlation. To improve the clarity of intraoperative views, cortical sites producing no language impairment were not systematically noted.

Statistical Analyses

Data analysis was performed by an independent statistician. Spatial naming and verb generation tasks were compared with spatial direct stimulation data separately and then in combination. Specificity and sensitivity were calculated with this method. To assess the correlation between fMRI and cortical mapping results, we compared a series of models to predict the positions of language areas on the basis of the fMRI information. A benchmark model that randomly predicted the language areas was used to judge the improvement provided by the fMRI information. Predictive accuracy was assessed by using a loss function that penalized predictions on the basis of the distance to the true language areas identified with cortical mapping. As a simplification, a single language area was predicted for each brain with
the model and the distances from that single point to each of the actual language areas were added, with the overall loss being obtained by addition of the losses for each brain. A naive model using the centroid of the largest scan patch for prediction reduced the loss to 36.5% of that incurred with the benchmark model. The subsequent three models attempted to take into account the possible statistical correlations between brains to improve predictions. To reduce dimensionality in the modeling approach, the data vectors were binary 0,1 vectors reflecting the row-column positions of the cortical mapping language points, as well as those of the scan areas, in a standardized grid. The size of the vectors was equal to the number of rows plus the number of columns in the grid. The models produced probability predictions for the binary 0,1 vectors of language points, and the highest-probability row and column constituted the single prediction. For simplicity, the predictor variable was obtained by adding the verb test and noun test scan vectors. For the predictive loss computation, we adopted a cross-validation strategy, which consisted of stacking all brains in a large vector and sequentially treating each brain as a missing value for comparison with the actual value.

Postoperative fMRI

Eight patients underwent postoperative control language fMRI studies, performed with the same procedures as preoperative fMRI studies (analysis threshold, $P < 0.005$). Because these control fMRI studies were performed only for research purposes, not all patients accepted the examination. The aim of these control fMRI studies was to compare pre- and postoperative fMRI data, for assessment of the value of the fMRI technique with the cortical brain mapping data.

RESULTS

Histopathological and Patient Findings

Patient and histopathological data are summarized in Table 1. Lesions were in the left hemisphere in all cases. Gliomas ($n = 11$) originated from the frontal (posterior part of F2 and F3) or temporoparietal (T1, T2, and angular gyrus) areas. One patient with partial seizures (consisting of paraphasias and speech arrest) and a low-grade astrocytoma located in the left supplementary motor area was also studied. Seven gliomas were of low histopathological grade (World Health Organization Grade II), and four were of high histopathological grade (World Health Organization Grade III or IV). Three patients with meningiomas were also studied. No patient in this series exhibited a significant language impairment precluding direct brain mapping. Functional language sites observed intraoperatively were always respected with a margin of at least 1 cm. fMRI language sites not correlated with brain mapping data were not always respected. Two patients (Patients 2 and 8) exhibited slight language deficits a few days after surgery. These transient language deficits were attributable to the proximity of the area of resection to cortical language areas. The language areas were preserved for both patients. Patient 14 had a tumor in the left supplementary motor area, which was resected. He developed a typical supplementary motor area syndrome with mutism, which resolved completely in a few weeks. None of the patients in this series exhibited postoperative language deficits, according to the neuropsychological tests performed 4 to 8 weeks after surgery.

Language fMRI Findings

High variability in the signal extent and localization was observed among patients for both tasks and both thresholds. For both tasks, the activated areas were located mainly in the mid-
dle and inferior frontal gyri (F2 and F3), the superior and middle temporal gyri (T1 and T2), and the supramarginal and angular gyri in the left hemisphere but also in the superior parietal gyrus, thalamus, and insula. The verb generation task activated the frontal areas more constantly. To a lesser extent, there were some activations in the same regions in the right hemisphere. Visual cortical areas were also bilaterally activated. No studies needed to be discarded because of excess head motion. Intratumoral activations were observed in six cases.

**fMRI-Cortical Stimulation Correlations**
Sites of correlation were analyzed separately. Because of technical considerations (position of the craniotomy and difficulty of access to some sites of activation), not all activated areas detected by fMRI could be correlated with intraoperative brain mapping data. None of the activations observed within tumors were correlated with direct brain mapping data, and none of those patients exhibited permanent postoperative language deficits. These intratumoral activations were not used for validation. In this study, the largest currents that did not evoke afterdischarges ranged from 3 to 10 mA.

**fMRI Naming Task**
Among the 14 patients, a total of 426 sites were tested. These 426 sites represented the sum of the cortical sites tested intraoperatively and correlated with fMRI data. Twenty-two sites of positive cortical stimulation were observed (nine sites of

### TABLE 1. Clinical and radiological data for the patients with brain tumors

<table>
<thead>
<tr>
<th>Patient no.</th>
<th>Age (yr)/sex</th>
<th>Diagnosis</th>
<th>Gyrus location, left side</th>
<th>Number/location of positive stimulation sites</th>
<th>Location of main areas of fMRI activation</th>
<th>Postoperative fMRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>55/F</td>
<td>Astrocytoma, Grade III</td>
<td>T2</td>
<td>1/precentral gyrus</td>
<td>Precentral and F3 gyri</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>33/F</td>
<td>Astrocytoma, Grade II</td>
<td>F2</td>
<td>1/precentral gyrus</td>
<td>Precentral, F3, and T1 gyri</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>35/M</td>
<td>Oligodendroglioma</td>
<td>Posterior part of F2 and F3</td>
<td>1/precentral gyrus</td>
<td>Precentral, F3, and T1 gyri</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>64/M</td>
<td>Meningioma</td>
<td>Angular gyrus</td>
<td>1/angular gyrus</td>
<td>Precentral, F2, F3, and T1 gyri</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>37/M</td>
<td>Meningioma</td>
<td>Pterional</td>
<td>None</td>
<td>Precentral, F3, T1, and T2 gyri</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>40/M</td>
<td>Oligoastrocytoma</td>
<td>T2</td>
<td>1/F3</td>
<td>Precentral, F2, F3, and T1 gyri</td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td>31/M</td>
<td>Oligodendroglioma</td>
<td>Posterior part of T2</td>
<td>None</td>
<td>F3, T1, and T2 gyri</td>
<td>+</td>
</tr>
<tr>
<td>8</td>
<td>52/F</td>
<td>Meningioma</td>
<td>T1</td>
<td>2/precentral and T2 gyri</td>
<td>Precentral, F3, and T1 gyri</td>
<td>+</td>
</tr>
<tr>
<td>9</td>
<td>68/F</td>
<td>Glioblastoma</td>
<td>Posterior part of T2</td>
<td>1/T2</td>
<td>Precentral, F3, and T2 gyri</td>
<td>+</td>
</tr>
<tr>
<td>10</td>
<td>14/F</td>
<td>Astrocytoma, Grade II</td>
<td>Temporal lobe</td>
<td>1/T2</td>
<td>Precentral, F2, T1, and T2 gyri</td>
<td>+</td>
</tr>
<tr>
<td>11</td>
<td>24/M</td>
<td>Astrocytoma, Grade II</td>
<td>Insula and F3</td>
<td>5/F3, F2, and precentral gyr</td>
<td>Precentral, F3, and T1 gyri</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>64/M</td>
<td>Glioblastoma</td>
<td>Posterior part of T2</td>
<td>5/T1 and T2</td>
<td>Pre- and postcentral, F3, and angular gyr</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>53/M</td>
<td>Glioblastoma</td>
<td>T2</td>
<td>2/T1</td>
<td>F3 and T1 gyri</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>37/M</td>
<td>Astrocytoma, Grade II</td>
<td>F1 and SMA</td>
<td>1/F2</td>
<td>Precentral, F3, and T1 gyri</td>
<td></td>
</tr>
</tbody>
</table>

*a F1, superior frontal gyrus; F2, middle frontal gyrus; F3, inferior frontal gyrus; T1, superior temporal gyrus; T2, middle temporal gyrus; SMA, supplementary motor area; fMRI, functional magnetic resonance imaging; +, postoperative fMRI performed.*

**Language Functional Magnetic Resonance Imaging**

**Neurosurgery**

**Volume 52 | Number 6 | June 2003 | 1339**
speech arrest, nine sites of anomia, and four sites of hesitations, paraphasia, or delayed response). Five of those sites were concordant with fMRI signals (analysis threshold, $P < 0.005$), whereas 17 positive cortical sites were not associated with a fMRI signal (sensitivity, 22%) (Table 2). Negative cortical sites corresponding to fMRI signals were observed in 12 cases (specificity, 97%; predictive value, 29%). A less-restrictive analysis threshold ($P < 0.05$) increased sensitivity to 45% and decreased specificity to 95% (predictive value, 33%).

**fMRI Verb Generation Task**

The 426 cortical sites tested intraoperatively were also correlated with the verb generation fMRI data. Of the 22 positive cortical sites, 8 were concordant with fMRI signals (analysis threshold, $P < 0.005$), whereas 14 positive cortical sites were not associated with a fMRI signal (sensitivity, 36%) (Fig. 2; Table 3). Negative cortical sites corresponding to fMRI signals were observed in eight cases (specificity, 98%; predictive value, 50%). A less-restrictive analysis threshold ($P < 0.05$) increased sensitivity to 54% and decreased specificity to 95% (predictive value, 37%).

**fMRI Verb Generation and Naming Tasks Combined**

The 426 cortical sites tested intraoperatively were also correlated by combining data from the verb generation and naming fMRI tasks. Of the 22 positive cortical sites, 13 were concordant with fMRI signals, whereas 9 positive cortical sites were not associated with a fMRI signal (sensitivity, 59%) (Table 4). Negative cortical sites corresponding to fMRI signals were observed in 12 cases (specificity, 97%; predictive value, 52%). A less-restrictive analysis threshold ($P < 0.05$) increased sensitivity to 66% and decreased specificity to 91% (predictive value, 50%).

In summary, results were not satisfactory, in terms of correlation, with the naming and verb generation tasks used separately. Better correlation was observed when the two fMRI tasks were combined.

### TABLE 2. Functional magnetic resonance imaging and intraoperative cortical stimulation correlation using the naming task

<table>
<thead>
<tr>
<th></th>
<th>Positive cortical stimulation (n = 22)</th>
<th>Negative cortical stimulation (n = 404)</th>
</tr>
</thead>
<tbody>
<tr>
<td>fMRI-positive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>signal (n = 17)</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>fMRI-negative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>signal (n = 409)</td>
<td>17</td>
<td>392</td>
</tr>
</tbody>
</table>

$fMRI$, functional magnetic resonance imaging.

**Postoperative fMRI**

Patients 3 to 10 underwent postoperative fMRI, 1 to 2 months after surgery. For each patient, the thresholds of analysis for fMRI signals used in the preoperative fMRI studies ($P < 0.005$) were chosen. Again, the activated areas were located mainly in the left hemisphere (and partially in the right) in the middle and inferior frontal gyri (F2 and F3), the superior and middle temporal gyri (T1 and T2), and the supramarginal and angular gyri but also in the superior parietal gyrus, thalamus, and insula. Visual cortical areas were also bilaterally activated. Areas of activation were located in the same regions as in the preoperative studies but not constantly. For two patients (Patients 6 and 7), two areas of activation that were preoperatively observed close to the tumor area and were not correlated with brain mapping findings were not present postoperatively (Fig. 3). For Patient 9, an area not activated preoperatively corresponded to a typical area of anomia during brain mapping. This area, which was located just above the tumor and was respected during removal, was activated postoperatively (Fig. 4). For Patient 8, one intraoperatively detected area (in the temporal lobe) was not detected in the postoperative fMRI studies. Finally, for Patient 5, one activated area in the frontal region that was detected in both pre- and postoperative fMRI studies was not observed during careful cortical mapping. Therefore, postoperative fMRI data (in the cortical areas studied intraoperatively) agreed with brain mapping data for six of eight patients (Patients 3, 4, 6, 7, 9, and 10). Complete agreement between pre- and postoperative fMRI results and direct brain mapping data was observed for only three of eight patients (Patients 3, 4, and 10).

**DISCUSSION**

**fMRI-Cortical Stimulation Correlations**

The possibility of detecting not only the laterality of language but also the exact locations of essential cortical language sites is of major interest for patients with brain tumors or those undergoing epilepsy surgery. To be safely used for surgical planning or intraoperatively (integrated with a neuro-navigation system, for example), fMRI must be able 1) to accurately detect language areas and 2) to ensure that, when no activation is present, no essential language sites will be identified with stimulation. Validation of the fMRI technique in the language field is fundamental. Neurosurgeons have long used cortical and subcortical stimulations in both epilepsy surgery (20) and brain tumor surgery (16, 18), to preserve language functions. This technique is safe and accurate and has proven its usefulness for the detection of essential language sites among a large number of patients (16, 18, 20). No other brain mapping technique has been used so extensively in this field. Direct cortical mapping has been validated by clinical results, i.e., if observed essential language sites are respected, then the risk of permanent aphasia is low. From a neurosurgical point of view, fMRI should yield similar information, to be useful in
surgical planning. It could seem natural to attempt to correlate direct brain mapping with fMRI by using similar tasks; however, the conceptual and technical differences between the two techniques seem important. First, by its nature, intraoperative brain mapping can sample only a limited brain territory (i.e., the exposed brain area). Furthermore, the validation of intrasulcal activations is often problematic; this could be considered a major drawback of this mode of correlation. Finally, activation studies can identify sites of neural activity associated with a particular function but their ability to demonstrate essential areas can be questioned. Intraoperative brain mapping detects essential language areas (associated with speech arrest or anomia), which is probably different from fMRI identification of cortical areas involved in language (even using the same tasks). Determination of whether fMRI data are clinically relevant would require a large, prospective, statistically significant study, testing not the correlation between direct cortical mapping and fMRI data but the correlation between preoperative fMRI data and postoperative language deficits.

How should we correlate fMRI data and direct brain stimulation findings in practice? We chose site-by-site correlation with the fMRI data observed in 3D reconstructions of the brain. 3D brain reconstructions allow easy comparisons of fMRI data and intraoperative views which can also be guided with the neuronavigational system probe. For both specific activation tasks used in this study, we observed an imperfect correlation between language fMRI data and direct cortical mapping results. Better results could be obtained by using combined fMRI tasks. However, correlations were too inconstant to allow the use of language fMRI for surgical planning without direct

<table>
<thead>
<tr>
<th>TABLE 3. Functional magnetic resonance imaging and intraoperative cortical stimulation correlation using the verb generation task*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No. of sites (n = 14 patients, 426 sites)</strong></td>
</tr>
<tr>
<td>Positive cortical stimulation (n = 22)</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>fMRI-positive signal (n = 16)</td>
</tr>
<tr>
<td>fMRI-negative signal (n = 410)</td>
</tr>
</tbody>
</table>

* fMRI, functional magnetic resonance imaging.

<table>
<thead>
<tr>
<th>TABLE 4. Functional magnetic resonance imaging and intraoperative cortical stimulation correlation using the verb generation task and the naming task*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No. of sites (n = 14 patients, 426 sites)</strong></td>
</tr>
<tr>
<td>Positive cortical stimulation (n = 22)</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>fMRI-positive signal (n = 25)</td>
</tr>
<tr>
<td>fMRI-negative signal (n = 401)</td>
</tr>
</tbody>
</table>

* fMRI, functional magnetic resonance imaging.
brain mapping as a control method. The reasons for this discrepancy are probably multifactorial. The choice of activation paradigms is of paramount importance in fMRI and can be questioned first. A naming task was chosen because it was thought to be very close to the intraoperative cortical stimulation procedure. Picture naming is also a standard method for diagnosing speech/language impairments (8). In fMRI, however, this task has been said to produce poorly lateralized activity, with important visual cortex activation (6). Some authors stated that some brain areas, including the dorsolateral frontal and lateral temporal cortices, were more frequently activated with a verb generation task than with a naming task (21). In this study, we observed activation of common brain areas but also activation of areas specific to the two tasks. The naming task exhibited poor localization value for both frontal and temporoparietal areas. The apparent high specificity of the fMRI technique in this study (97 and 98% for the naming and verb generation tasks, respectively) must also be interpreted cautiously; these results partially depend on the large number (n = 404) of negative cortical areas observed with electrostimulation (95% of the cortical sites tested). Because cortical language data are extremely localized (measured in small areas of 1–2 cm², with sharp boundaries) (18), it is not surprising that only a few language sites were observed.

The correlation of cortical stimulation results with language fMRI data has been less frequently studied than the correlation with motor fMRI data. Although some authors claimed that there was a good correlation between language fMRI data and direct brain mapping results (5, 6, 15, 22, 26), others demonstrated more contrasting results (27). Using different activation tasks involving picture naming and listening to recordings of spoken words, Ruge et al. (26) demonstrated a high degree of correlation between fMRI and direct brain mapping data for three patients. Also studying three patients, Lurito et al. (15) demonstrated that the correlation between the two techniques in receptive language areas was strong but not perfect, with a parietal language area that was detected with direct brain mapping not being detected with fMRI. Fitzgerald et al. (6), using five different activation tasks (word reading, visual and auditory verb generation, and listening to single words and text) and correlating their results with electrostimulation data for 11 patients, observed that fMRI data were very sensitive for detecting language areas but were poorly specific. Pouratian et al. (22), studying 10 patients with arteriovenous malformations and correlating fMRI data with intraoperative brain mapping findings, observed a specificity of 66% and a sensitivity of 96% for the language fMRI technique. Mueller et al. (16), testing two patients with tumors in frontal areas, observed a good correspondence between fMRI findings (areas activated with a word generation task) and electrostimulation results. Finally, Rutten et al. (27), treating eight patients by using fMRI data integrated into a surgical guidance system, observed an overall specificity of 61%, a sensitivity of 100% for seven patients, and an overall predictive value for the fMRI technique of 51%. Those authors stressed the importance of fMRI data variations among subjects, analysis thresholds, and tasks, which could influence the correspondence between cortical mapping and fMRI findings. In our opinion, all of these results, although generally promising, do not allow neurosurgeons to use lan-

**FIGURE 3.** Pre- and postoperative fMRI data comparisons. A 40-year-old man with a mixed glioma in the left T2 gyrus (Patient 6), with seizures, was tested with fMRI pre- and postoperatively (same acquisition procedures, same task, and same analysis threshold). No language deficits were observed pre- or postoperatively. A, axial (upper) and coronal (lower), anatomic, T1-weighted, MRI scans, showing the tumor located in the left T2 gyrus. B, preoperative fMRI findings (combined verb generation and naming tasks), showing one language-activated area immediately posterior to the tumor and one in the insula, just anterosuperior to the lesion (white arrows). C, intraoperative view, showing a language area (L) corresponding to the activated area observed in the insula (covered by the dura). However, no cortical language area (N) was observed either posterior to the tumor, as expected (white arrow), or elsewhere in the cortex before tumor removal. Dotted white line, tumor area. D, postresection view. E, 2-month postoperative, axial (upper) and coronal (lower), anatomic, T1-weighted, MRI scans. F, 2-month postoperative fMRI findings. No activated area posterior to the tumor can be observed.
language fMRI to make critical surgical decisions in the absence of direct brain mapping. A solution may lie in the combined use of different fMRI tasks, which could increase the sensitivity and specificity for the detection of essential language areas (6, 11, 27, 28). However, this approach could prolong the language fMRI acquisition procedure, which is already time-consuming (1 h in our study).

**Methodological Issues**

Among the methodological issues raised by the correlation between cortical brain mapping and language fMRI, the significance thresholds chosen to generate language activation maps must first be discussed. Individual fMRI activation signals are dependent on the analysis thresholds used to analyze raw fMRI data, among other factors. Variations in the significance thresholds can lead to changes in the spatial extent and number of activated cortical areas (Fig. 5). The choice of the best analysis thresholds for language fMRI signals remains an unsolved issue. Some authors previously noted that some patients must be treated with more restrictive thresholds than others (6). Our analysis thresholds ($P < 0.005$ and $P < 0.05$) were chosen arbitrarily. Similar thresholds were chosen by some other authors for similar language studies (21). It must also be noted that fMRI data for language function are far less reliably and easily recorded than are those for motor function. In contrast to fMRI motor data, the signal changes in language testing are very small, leading to extremely sensitive threshold variations.

Although many studies on fMRI signals have addressed their usefulness in clinical practice, several issues regarding the real spatial accuracy of fMRI are still open to debate. Indeed, the possibility of detecting variations in fMRI signals associated with activation tasks in large extraparenchymal veins that are not directly involved in the activation process has been named the “vein effect.” Some authors have claimed that this vein effect could be a serious limitation to accurate detection of neuronal areas implicated in activation. The anatomic proximity of intrasulcal cortical activated areas and some draining veins has influenced the debate (31). In this study and when validating motor fMRI data (25), we did not always observe draining veins in areas of signal validation, even when we validated intrasulcal located fMRI signals. It must also be noted that the blood oxygen level-dependent (BOLD) effect could be modified in tissue surrounding the tumor area (especially with high-grade astrocytomas). Some authors have suggested that disruption of normal vascular responses in the periphery of brain tumors could potentially enhance, rather than reduce, vascular responses (17). Many factors could be involved in the discrepancy between our pre- and postoperative language fMRI data. Among them, the stability of the BOLD effect in these pathological conditions can be questioned. Furthermore, the susceptibility of the fMRI signal (especially in nonprimary areas) to the attention of the patient during the fMRI procedure (32), the age of the subject (10), deficits of the patient (1, 25), and tumor growth (4) are all factors that potentially influence results. Recent advances in fMRI demonstrated a direct link between neural activity and BOLD contrast signal (14), although some questions regarding the exact nature of the neural activity detected with the BOLD technique remained (2). Although these issues regarding fMRI signals do not automatically have clinical importance, neurosurgeons must bear in mind the potential limitations and uncertainty of this technique for brain mapping.
To our knowledge, a comparison of pre- and postoperative language fMRI studies and direct brain mapping has never been performed. However, it was demonstrated for motor fMRI that an absent or barely detectable preoperative signal could become detectable postoperatively (23, 29), posing the problem of fMRI sensitivity. For two of our patients, two preoperatively detected areas that were not correlated with direct brain mapping findings corresponded to nonactivated areas in postoperative fMRI studies (Fig. 3). In contrast, in one of our postoperative fMRI studies, we detected a language area that was not present preoperatively, which corresponded to a typical zone of anomia (Fig. 4). In these three apparently contradictory cases, the postoperative studies (performed after the tumor and mass effect had been removed) seemed more reliable than the preoperative studies. This was contradicted by the results for Patients 5 and 8, which demonstrated that postoperative fMRI data did not always agree with brain mapping data. Although we want to remain prudent in drawing conclusions regarding these individual cases, these results at least call into question the value of naming or verb generation fMRI tasks (and their reproducibility) among patients with brain tumors.

We have considered that, for neurosurgical applications, precise gyrus localization should be the rule. The use of a neuronavigational system with integrated fMRI data is important, because of its accuracy in defining brain anatomic features (gyri) and its facilitation of correlations. However, we demonstrated in a previous work (25) that the registration of fMRI data in anatomic slices or with a frameless stereotactic neuronavigational device could remain a potential source of functional mislocalizations. To be used for neurosurgical purposes, fMRI echo-planar images (T2-weighted), which are anatomically imprecise, must be integrated into anatomic slices. Different pixel sizes for fMRI and anatomic slices, anatomic distortion of echo-planar images, and differences in slice thicknesses between fMRI and the neuronavigational system are all problems to be solved in merging images. The spatial accuracy of integrated motor fMRI data in a neuronavigational system can be validated by combining intraoperative brain mapping, somatosensory evoked potential monitoring, and the use of standard anatomic landmarks (25). Validation is far more difficult for language fMRI data, because of the relatively unpredictable patterns of language anatomic features. We cannot assume that distortion factors in image merging have not influenced our overall results. It was also demonstrated that stimulation maps generated with naming or verb generation tasks could differ (18), which represents another factor that could be a source of the discrepancy between our fMRI activation tasks and stimulation mapping. The specific correlation between the fMRI verb generation task and direct mapping should have been performed by using the verb generation task for brain mapping. Finally, three factors significantly limit the use of language fMRI in neurosurgical practice. First, fMRI is very sensitive to cortical changes but not to subcortical areas such as white matter tracts, which can be tested only with direct brain mapping. Therefore, the functional information provided by fMRI is probably incomplete. The second limitation is that the performance of language fMRI, the analysis of data, and the final use of the data in clinical practice require the cooperation of several specialists (neurosurgeons, radiologists, and computer engineers). The data analysis is time-consuming and not straightforward. The cost-effectiveness of this procedure must be questioned, perhaps eventually limiting it to selected patients. Finally, fMRI is restricted to patients with no language deficits; we did not include patients with significant language deficits (which we defined as more than 10% errors during picture naming) in this study because of the serious analysis bias produced by language deficits in the fMRI technique. The behavior of the language fMRI-activated areas among patients with language deficits is unknown, which remains a serious limitation for the preoperative evaluation of some patients.

**CONCLUSIONS**

With the activation paradigms used in this study, we observed no evidence that language fMRI could be used to make critical surgical decisions in the absence of direct brain mapping. Although fMRI has been successfully used to evaluate language hemispheric dominance, more studies are needed to evaluate the potential role of language fMRI in neurosurgical practice. The combination of multiple different fMRI tasks could optimize both the sensitivity and the specificity of the language fMRI technique, allowing better determination of language areas. The fact that fMRI and intraoperative brain
mapping are probably not exclusive of each other for language detection must also be accepted.

REFERENCES


tests and verb generation tests and duplicated these tests during surgery. In some of the patients, a postoperative fMRI study was used for comparison with preoperative localization and stimulation mapping. This article is important in that it remains to be seen whether fMRI can be used to predict consistently and accurately the preoperative localization of not only cerebral dominance (i.e., Broca’s area) but also the localization of key language functions (e.g., naming, reading). The real issue has to do with being able to use this information to predict the localization of these functions accurately, thus obviating the need for stimulation mapping. In this analysis, the authors demonstrate that there was a fair amount of variability in the signal extent and localization for both tasks and that there was activation all over the left hemisphere in these right-handed patients, as well as in some sites on the right side. As expected, verb generation was seen more commonly in frontal areas, which is probably the most often used preparative test used to find language dominance at present. Interestingly, activation seen within tumor tissue could not be verified at the time of surgery; yet, after undergoing resection, none of these patients had postoperative deficits. In essence, naming was tested intraoperatively in 426 sites in these patients, and 22 sites were found to be positive for stimulation, but only 5 sites corresponded to positive fMRI activation. Seventeen sites were not associated with fMRI activation, however. In some circumstances, sites that tested negative for stimulation mapping corresponded to positive fMRI sites (12 cases). With regard to verb generation, in the 22 positive cortical sites, stimulation corresponded to activation in only 8 sites but did not in 14 sites. Negative cortical sites corresponded to positive fMRI sites in eight cases. When the two tasks were combined, there was positive correlation with fMRI in 13 of the 22 positive cortical sites. Thus, the authors conclude that there is better correlation when both naming and verb generation tests are used, but the correlation still is not sufficient to replace direct stimulation mapping during surgery. Postoperative fMRI agreed with the brain mapping data in six of eight patients, which is reassuring, but complete agreement between pre- and postoperative fMRI and stimulation mapping was seen in only three of eight patients.

The take-home message of this study is that fMRI cannot reliably predict language dominance or laterality, nor can it reliably predict the exact location of naming and reading sites. Therefore, fMRI cannot be substituted for actual stimulation mapping to make surgical decisions regarding excision of tissue. This message is important, and this study is one of the few available that specifically correlates intraoperative stimulation mapping with fMRI testing. Once again, this study confirms that fMRI by itself cannot be used to decide which cortical regions can be resected. The authors also point out that fMRI cannot detect subcortical functional tissue, which can indeed be found with the use of direct subcortical stimulation mapping. This is another reason to remember that stimulation mapping is the “gold standard” for cortical and subcortical localization of function and that fMRI using any language task cannot be substituted for it at present. The authors provide an excellent study on this presently controversial subject.

Mitchel S. Berger
San Francisco, California

This article by Roux et al. compares perioperative fMRI and intraoperative surgical mapping in a series of 14 patients. The authors found a total of 22 sites that elicited a language deficit on the basis of stimulation mapping. Despite combining fMRI object naming and verb generation tasks, an fMRI site concordant to the stimulation error sites was detected for only 13 of 22 sites.

This article addresses some of the important methodological differences between activation-based fMRI and inactivation-based stimulation mapping techniques in the examination of language function. There is general agreement that fMRI cannot replace stimulation mapping on the basis of the current state of technology. Clearly, much development is needed to standardize tasks and procedures in fMRI.

Methodologically, the authors have used object naming as a stimulation task in comparison with both verb generation and object naming as fMRI tasks. It has been shown that the stimulation maps that these tasks generate differ. To better compare the authors’ results, it would have been helpful to apply both tasks with the use of both stimulation mapping and fMRI techniques. In addition, during the fMRI testing, the investigators had the patients whisper rather than silently perform the tasks “to confirm the patients’ ability to perform the task.” Most centers use silent speech generation to avoid movement artifact. In addition, language disruption sites within dominant face motor cortex were included in the 22 stimulation-positive sites. How many of the 22 sites were in fact motor cortex? fMRI analysis performed with the use of silent speech tasks generally does not activate motor cortex; the authors’ methods are thus inherently biased against the accuracy of fMRI. A schematic representation of the topographical location of the 22 positive stimulation sites, together with the location of the fMRI sites, would allow better direct comparison of the results.

Studies such as this one are important to the various centers working to learn how to refine fMRI for operative planning. On the basis of the “negative” experience of this group and the experiences of other investigators with fMRI language mapping relative to stimulation mapping, hypotheses can be tested directly to investigate the different results obtained with inactivation-based stimulation mapping and activation-based fMRI mapping of language function.

Guy M. McKhann II
Joy Hirsch
New York, New York
brain areas by performing MRI. Other reports have described attempts to find a similar correlation with regard to motor function. In several of these studies, there was certainly no precise relationship found, meaning that distances of up to 1 cm for the location of corresponding areas determined by using the two stimulation methods were considered acceptable. The fact remains that, to date, insufficient information has been found concerning the precise reliability of fMRI data with regard to the congruence of motor or language areas determined electrophysiologically. The task of verifying fMRI findings electrophysiologically is not made easier by the necessity of using two electrodes on the brain surface, which renders electrophysiological stimulation methods precise only to within 0.5 to 1 cm.

Also obvious is that the surgeon would like to have a higher degree of resolution in using fMRI data when it comes time to make a decision regarding the resection of a tumor border. In my experience, the tumor border visible on the brain surface is not infrequently closer than 1 cm to the area where fMRI has located some important function that is delineated on the brain surface with the aid of fused navigation MRI and fMRI. A safety zone of 1 cm would therefore lead to leaving behind tumor parts.

This article touches on an important point—namely, the reliability and the topographical precision of electrical stimulation. We do know that not all sites where speech arrest is induced are totally indispensable sites for language generation. Here is the old problem with electric stimulation mapping: is every site where speech arrest can be induced responsible for speech generation? Obviously, the process of verifying fMRI methodology with electrical stimulation technology is tedious, and I guess that this study will not be the last one in which speech sites detected with electrostimulation are not detected by fMRI and vice versa.

The authors have performed an important job in pointing out that the significance of an enhanced cortical area in an fMRI examination still is not always precisely defined. In other words, an uncritical belief in the beautiful imagery of MRI scans needs to be retested continually. I think that these authors address a problem of major clinical interest by highlighting the need for a precise workup of the correlation between fMRI findings and intraoperative electrophysiological stimulation mapping findings.

Johannes Schramm
Bonn, Germany