

Is Functional MR Imaging Assessment of Hemispheric Language Dominance as Good as the Wada Test?: A Meta-Analysis¹

R. Joshua Dym, MD
Judah Burns, MD
Katherine Freeman, DrPH
Michael L. Lipton, MD, PhD

Purpose:

To perform a systematic review and meta-analysis to quantitatively assess functional magnetic resonance (MR) imaging lateralization of language function in comparison with the Wada test.

Materials and Methods:

This study was determined to be exempt from review by the institutional review board. A systematic review and meta-analysis were performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. A structured Medline search was conducted to identify all studies that compared functional MR imaging with the Wada test for determining hemispheric language dominance prior to brain surgery. Studies meeting predetermined inclusion criteria were selected independently by two radiologists who also assessed their quality using the Quality Assessment of Diagnostic Accuracy Studies tool. Language dominance was classified as typical (left hemispheric language dominance) or atypical (right hemispheric language dominance or bilateral language representation) for each patient. A meta-analysis was then performed by using a bivariate random-effects model to derive estimates of sensitivity and specificity, with Wada as the standard of reference. Subgroup analyses were also performed to compare the different functional MR imaging techniques utilized by the studies.

Results:

Twenty-three studies, comprising 442 patients, met inclusion criteria. The sensitivity and specificity of functional MR imaging for atypical language dominance (compared with the Wada test) were 83.5% (95% confidence interval: 80.2%, 86.7%) and 88.1% (95% confidence interval: 87.0%, 89.2%), respectively.

Conclusion:

Functional MR imaging provides an excellent, noninvasive alternative for language lateralization and should be considered for the initial preoperative assessment of hemispheric language dominance. Further research may help determine which functional MR methods are most accurate for specific patient populations.

© RSNA, 2011

¹From the Departments of Radiology (R.J.D., J.B., M.L.L.), Epidemiology and Population Health (K.F.), Psychiatry and Behavioral Sciences (M.L.L.), The Dominick P. Purpura Department of Neuroscience (M.L.L.), and The Gruss Magnetic Resonance Research Center (M.L.L.), Albert Einstein College of Medicine and Montefiore Medical Center, 1300 Morris Park Ave, Bronx, NY 10461. Received July 12, 2010; revision requested September 4; revision received March 4, 2011; accepted May 10; final version accepted May 20.

Address correspondence to M.L.L. (e-mail: michael.lipton@einstein.yu.edu).

When neurosurgery for excision of an epileptogenic focus or brain mass is planned, it is essential to first determine its relationship with functionally important (“eloquent”) brain regions, as resection of such an area could result in devastating permanent neurologic deficits (1,2). Localization of cortex involved in language function is a particular concern prior to performance of a temporal lobectomy for treatment of epilepsy (3–5). Language localization can be determined with several methods and can guide surgical planning. In some cases, the risk of postoperative deficits may outweigh the benefits of surgery, while in others, a more conservative surgical approach may be designed to limit the risk of impairment (6,7).

The current standard of reference for preoperative language lateralization is the intracarotid amobarbital procedure, commonly known as the Wada test. Since the Wada test is invasive, several alternatives have been investigated, among which functional magnetic resonance (MR) imaging has been the most studied (8,9). As a noninvasive tool, functional MR imaging is considerably safer than Wada testing and other invasive means for determining language lateralization, such as electrocortical stimulation. Furthermore, because functional MR imaging does not require anesthesia, the results are free from the effects of sedation and the test may be easily repeated or extended if necessary (4,8,10). Another advantage of functional MR is its ability to provide more precise anatomic localization within a hemisphere, thus allowing for tailoring of surgery to avoid resection of eloquent cortex (6).

Advances in Knowledge

- Functional MR imaging is both sensitive and specific for the detection of atypical language dominance.
- The sensitivity and specificity of functional MR imaging in the assessment of language dominance may vary depending on the technique of implementation and interpretation.

Numerous previous studies have compared functional MR imaging with the Wada test in a variety of clinical and laboratory settings (4,11–32). Similarly, several review articles have sought to establish the utility of functional MR imaging as a technique suitable for widespread clinical use as an adjunct or alternative to Wada testing (8–10,32). The small size of these previously published studies, most including fewer than 20 patients, limits their statistical power and generalizability. Thus, despite the number of studies supporting the effectiveness of functional MR imaging for language localization, there remains no clear consensus regarding the true reliability of functional MR in this role.

We hypothesize that functional MR imaging is a highly accurate means for lateralizing language function in patients prior to brain surgery. The purpose of this study was to perform a systematic review and meta-analysis to quantitatively assess functional MR imaging lateralization of language function in comparison with the Wada test.

Materials and Methods

Study Design

The study performed involved review and analysis of the published medical literature. Since no patient-identifying data were involved, this study was determined to be exempt from review by the institutional review board. No industry support was provided for this study. The systematic review, meta-analysis, and manuscript preparation were conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (33).

Implications for Patient Care

- Functional MR imaging should be considered for the initial preoperative assessment of hemispheric language dominance.
- Consideration should be given to the potential for result variability depending on the specific functional MR method used.

A structured MEDLINE (Ovid) search was performed by two radiologists (R.J.D., J.B.) with 8 and 10 years of experience, respectively, in conducting a literature search. The search was designed to identify all full-length English-language articles published between January 1966 and June 2008 that compared functional MR imaging with the Wada test for determination of hemispheric language dominance prior to brain surgery. The following keywords were used to generate a list of potentially useful studies: “fMRI” OR “functional MRI” OR “functional MR imaging” OR “functional magnetic resonance imaging” AND “Wada” OR “IAT” (intracarotid amobarbital test) OR “IAP” (intracarotid amobarbital procedure) OR “amytal” OR “amobarbital.”

To select studies for the meta-analysis, the following inclusion criteria were applied: (a) examination of the same patients with both functional MR imaging and the Wada test; (b) preoperative examination of at least four patients with both methods; and (c) reporting of data on the agreement of the two methods in individual patients, not only as a whole. For relevant articles that did not provide a breakdown of their data as to the agreement of the two tests, attempts were made to contact the authors to obtain this information (23,34,35). A response to this request for clarification was received for one of three articles (F.G.W.,

Published online before print

10.1148/radiol.11101344 Content codes: **MR** **NR**

Radiology 2011; 261:446–455

Abbreviations:

CI = confidence interval
QUADAS = Quality Assessment of Diagnostic Accuracy Studies

Author contributions:

Guarantors of integrity of entire study, R.J.D., M.L.L.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; literature research, R.J.D., J.B., M.L.L.; statistical analysis, R.J.D., K.F., M.L.L.; and manuscript editing, R.J.D., J.B., M.L.L.

Potential conflicts of interest are listed at the end of this article.

written communication, March 27, 2008). In cases where aspects of the study design were unclear, attempts were made to contact the authors for clarification (17,31). A response to this request for clarification was received for one of two articles (C.K., written communication, June 24, 2009). For each study that met the inclusion criteria, only data regarding patients who met the above criteria were included. Specifically, any patients who did not have diagnostic results from both preoperative Wada and functional MR testing were excluded.

The abstracts of all of the articles yielded by the Medline search were reviewed independently by two radiologists (R.J.D., J.B.) to determine which articles met criteria for inclusion. Any discrepancies between the two compilations of eligible studies were reviewed by the two radiologists in consultation with a senior radiologist (M.L.L.) with more than 20 years of experience in conducting a literature search, and final decisions regarding inclusion of articles was decided by consensus. References of the included articles, as well as relevant review articles, were also reviewed (R.J.D., J.B.) in an attempt to identify additional eligible articles.

Methodological quality of the selected studies was assessed by using a version of the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool, as modified by the Cochrane Collaboration (36,37). Two radiologists independently reviewed the studies to evaluate their fulfillment of 11 different quality criteria and graded them accordingly (R.J.D., J.B.); differences were resolved by consensus.

Individual articles were reviewed for study and patient characteristics (R.J.D.), which were then compiled in spreadsheet format for analysis. Recorded study features included functional MR imaging design (block vs event related), functional MR imaging task paradigm (semantic decision vs word generation), evaluated territory (global vs regional [ie, frontal or temporal-parietal]), functional MR interpretation method (volume of activation vs magnitude of signal change), technique for calculating functional MR lateralization, and number of patients meeting meta-analysis inclusion

criteria. The recorded patient characteristics included reason for examination, side of any identified lesion or epileptic focus, and handedness.

To determine hemispheric language dominance, most authors utilized a lateralization index, which is based on the number of activated voxels in each region or hemisphere and/or the magnitude of signal change during activation. This calculation yields a value ranging from +1 (or +100) for patients with pure left hemispheric lateralization to -1 (or -100) for patients with all of their language activity within the right hemisphere. A threshold of 0.2 (or 20) was generally used, with values above that level indicative of typical left hemispheric dominance for language and values below -0.2 (or -20) denoting right hemispheric language dominance. Intermediate values were defined as bilateral or mixed dominance. For studies that provided numerical lateralization results but did not assign a hemispheric lateralization, we used the ± 0.2 (or ± 20) threshold for determining dominance, as this was the most commonly used threshold in the remainder of the studies. Results of Wada and functional MR imaging were classified as being either typical (left hemispheric language dominance) or atypical (right hemispheric

language dominance or bilateral language representation) for each patient case.

Statistical Analysis

A meta-analysis was performed by using data from all studies fulfilling the search criteria to determine the overall sensitivity and specificity of functional MR for determination of hemispheric language dominance. Sensitivity was defined as the ability of functional MR to depict atypical language representation, as determined by the standard of reference Wada test. Specificity was defined as the ability of functional MR to depict Wada-established typical language representation.

To perform the meta-analyses, a random-effects (linear mixed-effects) model was used to derive summary estimates of sensitivity and specificity, which incorporated the possible correlation between these measures, along with 95% confidence intervals (CIs) (38). This method corrects for zero frequencies within each study's screening table by using a value of 0.5. For supporting analyses, we derived separate estimates of sensitivity and specificity by pooling values (proportions) from individual studies derived by summing the weighted estimates (ie, sensitivity) across all studies and dividing by one less than

Figure 1

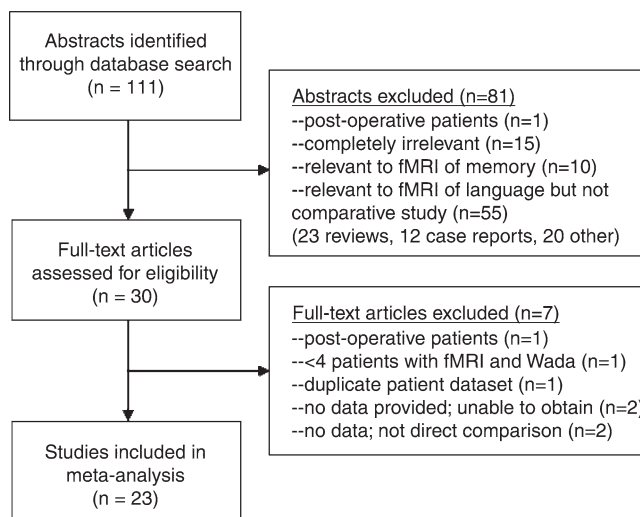


Figure 1: Flowchart of search results. *fMRI* = functional MR imaging.

Table 1

Methods of the Included Studies

Study	Year	Journal	Task Paradigm	Evaluated Territory	Interpretation Method	Lateralization Calculation
Binder et al (4)	1996	Neurology	Semantic decision	Global	Volume of activation	Lateralization index
Hertz-Pannier et al (11)	1997	Neurology	Word generation	Frontal	Volume of activation/ magnitude of activation*	Lateralization index
Bahn et al (12)	1997	AJR	Word generation	Global	Volume of activation	Voxel count
Yetkin et al (13)	1998	AJNR	Word generation	Frontal	Volume of activation	Lateralization index
Benson et al (14)	1999	Neurology	Word generation	Global	Magnitude of activation	Lateralization index
Lehéricy et al (15)	2000	Neurology	Word generation, story listening [†]	Frontal, temporal/parietal	Volume of activation	Lateralization index
Carpentier et al (16)	2001	Epilepsia	Semantic decision	Global	Volume of activation	Lateralization index
Gao et al (17)	2001	Chin Med J	Other [‡]	Not specified	Not specified	Not specified
Baciu et al (18)	2001	Epilept Disord	Semantic decision	Global	Volume of activation	Voxel count
Liégeois et al (19)	2002	Neuroimage	Word generation	Frontal	Volume of activation, magnitude of activation [§]	Lateralization index (volume of activation)
Gaillard et al (20)	2002	Neurology	Word generation	Global	Volume of activation	Lateralization index
Rutten et al (21)	2002	Neuroimage	Word generation	Global, frontal, temporal/ parietal	Volume of activation	Lateralization index
Spreer et al (22)	2002	Neuroradiology	Semantic decision	Global, frontal, temporal/ parietal	Volume of activation	Lateralization index
Woermann et al (23)	2003	Neurology	Word generation	Global	Volume of activation	Visual
Adcock et al (24)	2003	Neuroimage	Word generation	Global (volume of activation); frontal (magnitude of activation)	Volume of activation, magnitude of activation [§]	Lateralization index
Sabbah et al (25)	2003	Neuroimage	Word generation	Global	Volume of activation	Lateralization index
Gaillard et al (26)	2004	Neurology	Word generation	Global	Volume of activation	Visual
Deblaere et al (27)	2004	Neuroradiology	Word generation	Global, frontal, temporal/ parietal	Volume of activation	Lateralization index
Baciu et al (28)	2005	Neuroradiology	Semantic decision	Global	Volume of activation, magnitude of activation [§]	Lateralization index (volume of activation)
Benke et al (29)	2006	Epilepsia	Semantic decision	Global, frontal, temporal/ parietal	Volume of activation	Visual
Chlebus et al (30)	2007	Exp Brain Res	Word generation	Global	Volume of activation	Lateralization index
Kesavadas et al (31)	2007	Pediatr Radiol	Word generation	Global	Volume of activation	Visual
Szafarski et al (32)	2008	Epilepsy Behav	Word generation, semantic decision [§]	Global	Volume of activation	Lateralization index

* Combination of both techniques.

[†] Story listening task used for temporal lobe evaluation.

[‡] Language activities such as reciting poems.

[§] Each performed independently; results of more commonly used method included in overall meta-analysis.

^{||} Alternate technique was used for magnitude interpretation method.

the sum of the sample sizes—using raw data and using data corrected for zero frequencies; 95% CIs were derived by using the formula by Agresti and Coull (39). In addition, we compared results from these models with those from the bivariate random-effects model but var-

ied values of the correction factor from 0.5 to 0.2 by increments of 0.05. Finally, we also performed models without the one study with zero frequencies for calculating specificity. Weighting factors associated with design features were deemed unnecessary due to the similarity

of the design and methods of the individual studies. Prevalence rates were computed for each study, and an unweighted pooled estimate was derived. Forest and funnel (estimate vs reciprocal of standard error) plots were created for sensitivity and specificity for the set

Figure 2

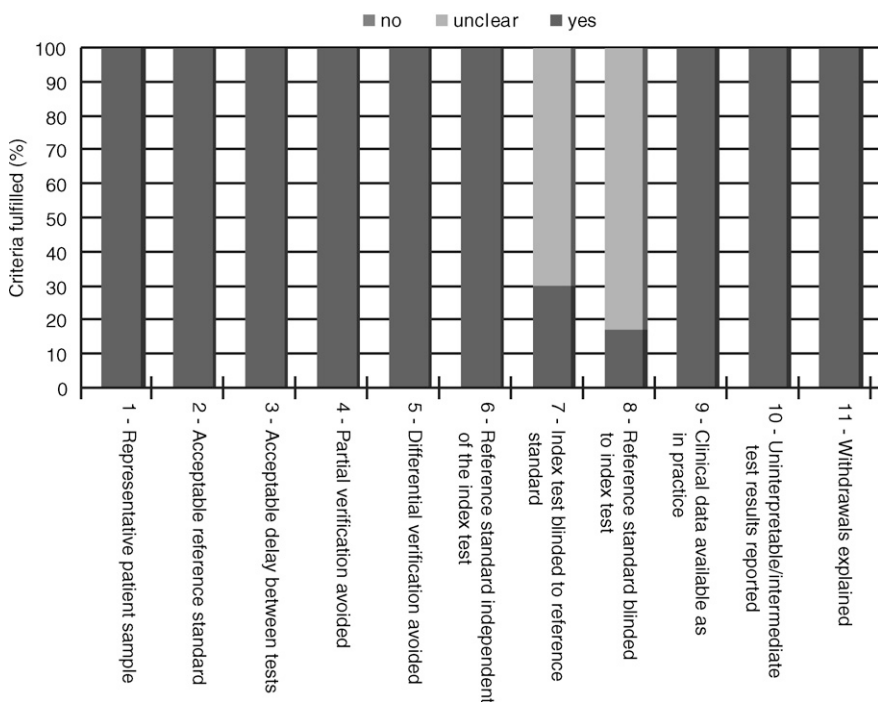


Figure 2: Bar chart for quality assessment by using QUADAS criteria.

of all studies. SAS software (version 9.2; SAS Institute, Cary, NC) was used to perform statistical analysis.

Subgroup analysis was also performed to determine whether particular techniques for the performance of functional MR imaging and analysis more closely correlate with the Wada test results. Patients included in the meta-analysis were subdivided into groups on the basis of the functional MR task paradigm, the evaluated territory, functional MR interpretation method, lateralization calculation technique (lateralization index or voxel count vs visual), the side of epileptic focus and/or lesion, and handedness. For studies using more than one type of the above-mentioned functional MR methods, including in the same patient, the results of each technique were treated as independent samples and were recorded in their respective subgroups. Paired analyses could not be performed because many of the studies did not provide full patient-level data. As such, our results are more conservative than they would have been had we been able to

use that method. Results of uncommonly used techniques, such as unique functional MR task paradigms, were not recorded as separate subgroups due to their low utilization across the entire series. For the overall meta-analysis, to maintain the highest possible uniformity, only the results of the technique more commonly used across the entire data set were included. Results for the overall meta-analysis were also calculated separately, with certain patient groups excluded, such as patients who only underwent a unilateral Wada testing and patients who did not have epilepsy.

Results

Systematic Review

Our MEDLINE search yielded 111 abstracts, of which 23 studies (21%) met criteria for inclusion in this meta-analysis (4,11–32) (Fig 1). These studies included a total of 442 eligible patients who preoperatively underwent both a Wada test and functional MR imaging.

A list of the included studies and their respective characteristics are summarized in Table 1.

All studies utilized a block design for functional MR imaging task administration. Among the 23 included studies, 16 (70%) used some form of a word generation task paradigm, while seven (30%) used a semantic decision task paradigm (Table 1). One study (4%) used both types of task paradigms and therefore was included in both subgroups; another study (4%) did not use either of these task paradigms and was not included in either subgroup. Several studies used unique task paradigms in addition to or in place of the commonly used word generation or semantic decision tasks.

Data analysis methods varied with regard to whether activation was assessed in each hemisphere (global) or within individual brain regions (regional, ie, frontal or temporal-parietal) (Table 1). Among the studies using a global technique, some assessed the hemispheres as a whole, while others either used regions of interest in multiple lobes in each hemisphere or combined the results of their regional analyses. A global method of analysis was utilized by 18 (78%) studies. This includes four studies that also provided results for frontal and temporal-parietal regional assessments. One study evaluated the frontal and temporal lobes independently but did not provide a combined result, and three studies only assessed activation within the frontal lobes. Thus, eight of 23 (35%) studies provided results of regional analysis.

The studies also used different methods of functional MR imaging interpretation for determining language lateralization (Table 1). In the majority of studies (20 of 23, 87%), lateralization of language dominance was determined on the basis of the volume of the activated brain, that is, by comparing the number of activated voxels within regions of interest in each hemisphere. Three of those studies also independently analyzed the functional MR images by using a second method for determining lateralization, in which they assessed the magnitude of functional MR signal change between periods of rest and task in

each hemisphere. One additional study used this magnitude change technique exclusively. In total, four of 23 (17%) studies used the magnitude change method.

Although a few studies in our meta-analysis used visual assessment ($n = 4$, 17%) or a comparison of the number of activated voxels in each hemisphere ($n = 2$, 9%) as their technique for determining lateralization, most studies ($n = 16$, 70%) utilized a formula to calculate a lateralization index (Table 1).

Of the 11 utilized QUADAS quality criteria, nine were fulfilled by 100% of the studies (Fig 2). Authors of most of the studies did not clearly specify whether the interpreters of the Wada test and/or functional MR imaging were blinded to the results of the other test nor did most comment as to the order of test performance. No studies were excluded due to poor quality, and no quality-based subgroup analysis was deemed necessary.

Meta-Analysis Results

Table 2 provides a breakdown of patients' Wada and functional MR imaging language lateralization results, as well as overall concordance between the two tests for each of the studies. Most studies demonstrated a high degree of concordance between functional MR and Wada testing. In 21 of 23 (91%) studies, there was at least 80% concordance between the two tests.

The results of the meta-analysis are listed in Table 3, and Forest plots are provided for sensitivity and specificity of the studies (Figs 3, 4). Overall, the sensitivity of functional MR imaging in depicting atypical language representation was 83.5% (95% CI: 80.2%, 86.7%) and specificity was 88.1% (95% CI: 87.0%, 89.2%). Meta-analysis results with exclusion of patients who had received only unilateral Wada tests were not significantly different. With the meta-analysis limited to only studies with epileptic patients, specificity was significantly higher (90.7% vs 88.1%).

Results from the subgroup analysis are also summarized in Table 3. Functional MR imaging with use of word generation tasks had significantly greater specificity than that with use of semantic

Table 2

Study	Study Results*				Wada/functional MR Concordance [†]
	WA/MA	WA/MT	WT/MA	WT/MT	
Binder et al (4)	4	0	0	18	22/22 (100)
Hertz-Pannier et al (11)	1	0	0	4	5/5 (100)
Bahn et al (12)	2	0	0	5	7/7 (100)
Yetkin et al (13)	1	0	0	12	13/13 (100)
Benson et al (14)	2	0	0	9	11/11 (100)
Lehéricy et al (15) [‡]	1	0	1	8	9/10 (90)
Carpentier et al (16)	2	0	2	6	8/10 (80)
Gao et al (17)	1	0	0	3	4/4 (100)
Baciu et al (18)	1	0	0	7	8/8 (100)
Liégeois et al (19)	3	0	0	1	4/4 (100)
Gaillard et al (20)	2	2	1	13	15/18 (83.3)
Rutten et al (21)	5	2	1	10	15/18 (83.3)
Spreer et al (22)	5	0	3	13	18/21 (85.7)
Woermann et al (23)	23	3	5	63	86/94 (91.5)
Adcock et al (24) [§]	3	1	1	14	17/19 (89.5)
Sabbah et al (25)	8	0	0	12	20/20 (100)
Gaillard et al (26)	2	0	0	11	13/13 (100)
Deblaere et al (27)	2	0	1	14	16/17 (94.1)
Baciu et al (28) [§]	2	2	3	6	8/13 (61.5)
Benke et al (29)	11	3	15	39	50/68 (73.5)
Chlebus et al (30)	3	0	2	10	13/15 (86.7)
Kesavadas et al (31)	5	0	0	0	5/5 (100)
Szafarski et al (32)	5	0	2	20	25/27 (92.6)

* WA = Wada atypical result, WT = Wada typical result, MA = functional MR atypical result, MT = functional MR typical result.
[†] Numbers in parentheses are percentages.
[‡] Results for frontal lobe evaluation with word generation task.
[§] Results for volume lateralization index method of interpretation.
^{||} Results for word generation task.

decision tasks (95.6% vs 69.5%), and there was a trend toward greater sensitivity of word generation tasks as well. Functional MR imaging with use of a global evaluation had significantly greater sensitivity than a regional temporal-parietal evaluation (86.6% vs 62.6%). When compared with a visual technique for designation of lateralization, the subgroup including the more precise quantitative methods of lateralization index or voxel count was found to have significantly higher sensitivity (86.1% vs 58.5%) and significantly higher specificity (90.8% vs 56.1%). Functional MR imaging was also shown to be significantly more specific for right-handed patients than for left-handed or ambidextrous patients (88.8% vs 74.6%). There was no

statistically significant difference between subgroups differing by interpretation method (volume vs magnitude activation) or location of brain lesion.

The results of the supporting analyses were consistent with the primary analysis for the full data set, as well as the subgroup analyses. Funnel plots for sensitivity and specificity are provided (Figs 5, 6); however, they are ambiguous as to any impression of publication bias because the x-axis is truncated at 1.0 (100% sensitivity or specificity), which precludes plot symmetry.

Discussion

Our review of the relevant literature revealed numerous studies that evaluated

Table 3

Meta-Analysis Results

Group	No. of Patients	Sensitivity*	Specificity*
Overall meta-analysis	442	83.5 (80.2, 86.7)	88.1 (87.0, 89.2)
Exclude: unilateral Wada	409	79.8 (76.1, 83.5)	87.8 (86.5, 89.0)
Exclude: nonepileptics	431	84.8 (81.4, 88.1)	90.7 (89.6, 91.9)
Task paradigm subgroups			
Word generation	296	88.9 (84.5, 93.4)	95.6 (93.9, 97.3)
Semantic decision	168	78.7 (68.5, 88.9)	69.5 (66.5, 72.5)
Evaluated territory subgroups			
Global	406	86.6 (83.1, 90.2)	92.2 (91.0, 93.4)
Frontal	156	75.9 (65.9, 85.9)	91.8 (88.7, 94.9)
Temporal/Parietal	132	62.6 (49.5, 75.6)	87.4 (83.7, 91.1)
Interpretation method subgroups			
Volume activation	422	85.8 (82.3, 89.2)	91.4 (90.2, 92.6)
Magnitude activation	47	88.0 (65.2, 100)	86.3 (75.8, 96.7)
Lateralization calculation subgroups			
Lateralization index or voxel count	260	86.1 (80.7, 91.5)	90.8 (88.9, 92.6)
Visual technique	182	58.5 (50.1, 67.0)	56.1 (53.0, 59.3)
Lesion location subgroups			
Right hemisphere	112	73.1 (60.6, 85.5)	88.4 (84.8, 92.0)
Left hemisphere or bilateral	175	85.6 (79.5, 91.7)	81.9 (78.9, 84.9)
Handedness subgroups			
Right	206	74.5 (66.1, 82.9)	88.8 (86.7, 90.9)
Left or ambidextrous	52	74.1 (64.8, 83.5)	74.6 (65.6, 83.5)

* Data are percentages. Data in parentheses are 95% CIs.

the reliability of functional MR imaging for language lateralization, relative to the Wada test. Individually, the results of these studies fail to provide a generalizable conclusion due to variability in study design and small sample sizes, limiting statistical power. Our meta-analysis suggests strongly that functional MR imaging is both sensitive and specific for the detection of atypical language representation, supporting its use for preoperative language assessment.

Medina et al (40) evaluated the role of functional MR imaging in the assessment of language dominance using a Bayesian analysis of 13 studies, with the Wada test and electrocortical stimulation as reference standards, to determine likelihood ratios for functional MR in determining language dominance. They concluded that functional MR increases the posttest probability of language dominance in various patient populations. Our meta-analysis addresses the more specific question of whether the results of functional MR

imaging and the Wada test correspond to an extent that would allow functional MR to serve as a substitute for Wada testing in the presurgical determination of language lateralization. This current analytic approach may have more direct practical application and generalizability to routine practice. We also examined the utility of different functional MR methods using our subgroup analysis. In addition, due in part to differences in inclusion criteria, our meta-analysis was more comprehensive: it included a greater number of studies, with results from nearly twice as many individual patients.

One challenge in interpreting the existing functional MR literature regarding determination of language dominance arises from the variety of functional MR techniques used in the numerous studies. In addition to our overall assessment of functional MR results obtained by combining the results of all of the varied techniques, we also performed subgroup analyses to assess whether particular

methods were more reliable than others. Subgroup analyses yielded several notable differences between the various functional MR methods. Regarding the results for the different functional MR tasks, a word generation task was significantly more specific than a semantic decision task. The lower specificity of a semantic decision task may be attributed to its greater likelihood of activating areas of cortex that are involved in other functions such as memory, which are only indirectly related to language and may not be absolutely essential for language function (41). It should be noted that although individual studies used several variations of word generation and semantic decision tasks, these studies were consolidated into two groups (word generation or semantic decision) for the subgroup analyses, due to sample size limitations.

Data were also grouped on the basis of which brain region was evaluated. Functional MR with use of a global evaluation appeared to have higher sensitivity than that with use of a regional frontal lobe evaluation, which in turn appeared more sensitive than an evaluation limited to the temporal-parietal lobes; however, only the difference between the global and temporal-parietal evaluation subgroups was statistically significant. With regard to specificity, the global and frontal evaluation subgroups were similar and both trended slightly higher than the temporal-parietal subgroup.

The higher sensitivity of the global evaluation subgroup is expected since the greater area evaluated allows more likely detection of atypical localization of language function. The lower sensitivity and apparently lower specificity of functional MR for the temporal-parietal subgroup may be due in part to the fact that most studies used a word generation task, which yields more robust frontal as compared with temporal lobe activation. A focused temporal lobe evaluation might be better served by the use of a semantic decision task, since many semantic processes can be localized to the temporal lobes (42). Overall, more studies focused their evaluation on the frontal lobe as compared with

the temporal lobe. While frontal and temporal lobe language dominance are generally concordant, patients with epilepsy may occasionally demonstrate dissociation between frontal and temporal dominance (43). Because temporal lobe dominance has more direct clinical relevance for preoperative planning prior to epilepsy surgery, further studies focusing on methods for temporal lobe evaluation would be useful.

In the calculation of lateralization, the significantly higher sensitivity and specificity of the lateralization index-voxel number subgroup supports the use of such strict quantitative methods of lateralization over a visual technique. Functional MR imaging was shown to be significantly more specific for right-handed patients than for left-handed or ambidextrous patients; this may be related to their higher rate of typical language representation. The lack of statistically significant differences between the subgroups, which differed by the interpretation method, is not particularly instructive, since very few patients (47 of 442) were analyzed with the magnitude interpretation method, and the resultant wide confidence intervals limit the usefulness of the results.

Two outlying studies (28,29) demonstrated relatively lower concordance of 61.5% and 73.5%, respectively, for functional MR imaging and Wada test. This may be related to the fact that both of those studies utilized a semantic decision task and one (29) used a visual technique for lateralization; both of these methods demonstrated lower sensitivity and specificity in this meta-analysis. The study with 61.5% concordance (28) only included 13 patients, so perhaps this is a statistical aberration.

There were several limitations to our study. As in all meta-analyses, important variation in methods exists among the included studies. By performing subgroup analyses, we aimed to limit the effect of variation between the studies and determine which method would in fact provide the best results. Subgroup analyses identified differences in the sensitivity and specificity of functional MR imaging for most of the different techniques. However, these distinctions do not all

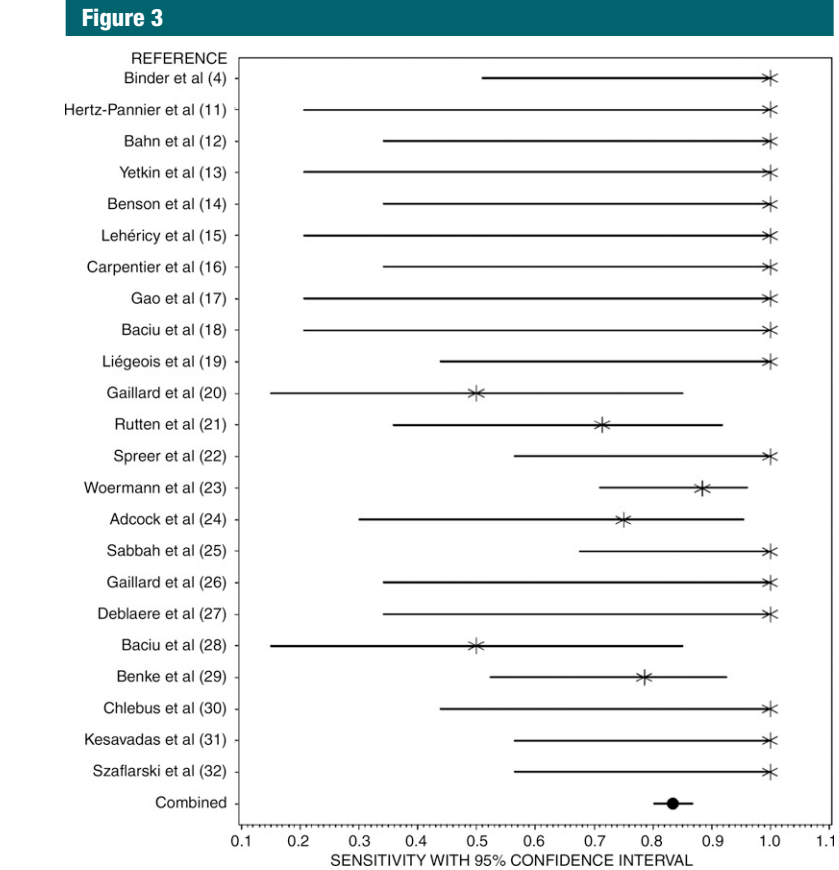


Figure 3: Forest plots of studies included in the meta-analysis show individual and combined estimates with 95% CIs for diagnostic sensitivity of functional MR imaging in determination of hemispheric language dominance. * = mean sensitivity for each study, ● = combined sensitivity for all 23 studies, horizontal lines = 95% CIs of sensitivity.

reach statistical significance, in some cases due to the smaller sizes of the subgroups, which result in wide confidence intervals.

While the subgroups were more homogeneous, the overall meta-analysis was fairly heterogeneous as it incorporated the results of various different methods of testing and data analysis. For studies included in the meta-analysis that used more than one technique, we strove to improve uniformity and relevance by using data generated with the most commonly utilized techniques. Nevertheless, the lack of complete uniformity due to the inclusion of varied techniques could diminish the power of the study. The fact that we demonstrated functional MR imaging to be fairly sensitive and specific despite our inclusion

of several different techniques underscores the strength of our findings.

Another potential limitation, inherent to any meta-analysis, is the possibility of publication bias. Since many of the studies in this meta-analysis cluster near the upper limit of 100% sensitivity and/or specificity, the funnel plots in this case are not clearly interpretable in evaluating for the presence of publication bias. Our meta-analysis is also limited in that valid positive and negative predictive values could not be calculated because the prevalence of atypical language representation varied widely between the studies.

In conclusion, functional MR imaging provides an accurate, noninvasive alternative to the Wada test and should be considered for the initial preoperative

Figure 5

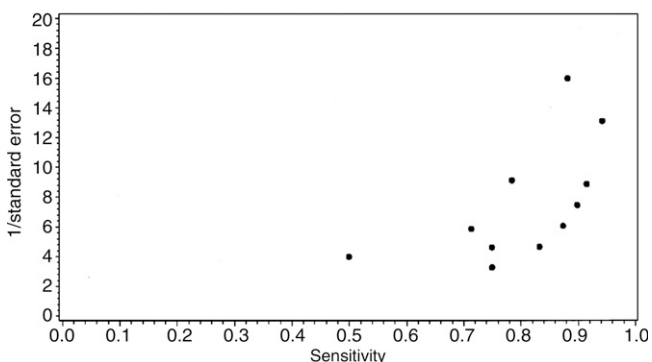


Figure 5: Funnel plot for corrected sensitivity.

Figure 6

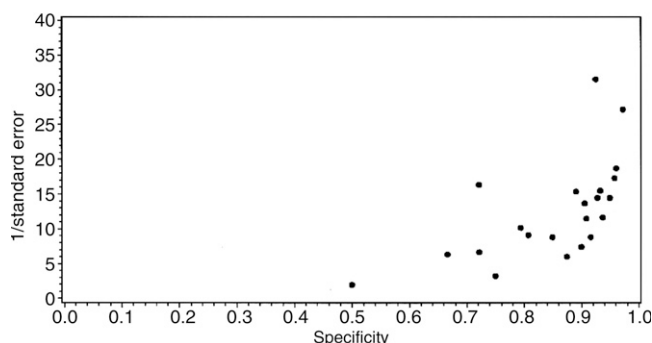


Figure 6: Funnel plot for corrected specificity.

Figure 4

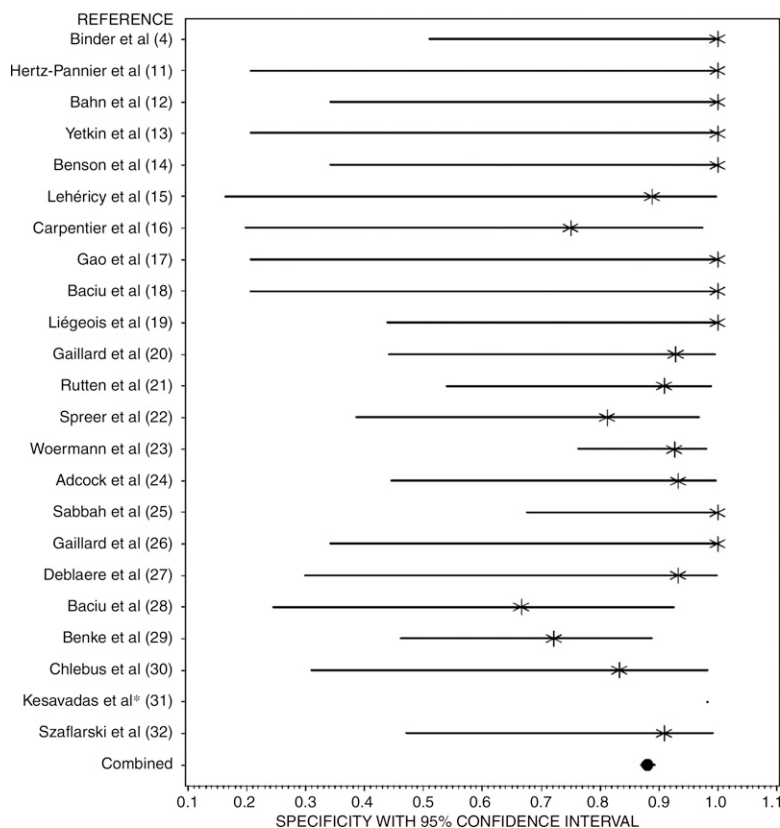


Figure 4: Forest plots of studies included in the meta-analysis show individual and combined estimates with 95% CIs for diagnostic specificity of functional MR imaging in determination of hemispheric language dominance. * = mean specificity for each study, ● = combined specificity for all 23 studies, horizontal lines = 95% CIs of specificity. * = No patients with typical language dominance; specificity is therefore undefined.

the overall effectiveness of functional MR imaging in clinical practice, even beyond the level of accuracy that we have demonstrated in this meta-analysis.

Disclosures of Potential Conflicts of Interest:

R.J.D. No potential conflicts of interest to disclose. **J.B.** No potential conflicts of interest to disclose. **K.F.** No potential conflicts of interest to disclose. **M.L.L.** No potential conflicts of interest to disclose.

References

1. Scoville WB, Milner B. Loss of recent memory after bilateral hippocampal lesions. *J Neurool Neurosurg Psychiatry* 1957;20(1):11-21.
2. Binder JR, Sabsevitz DS, Swanson SJ, Hammeke TA, Raghavan M, Mueller WM. Use of preoperative functional MRI to predict verbal memory decline after temporal lobe epilepsy surgery. *Epilepsia* 2008;49(8):1377-1394.
3. Devinsky O, Perrine K, Llinas R, Luciano DJ, Dogali M. Anterior temporal language areas in patients with early onset of temporal lobe epilepsy. *Ann Neurol* 1993;34(5):727-732.
4. Binder JR, Swanson SJ, Hammeke TA, et al. Determination of language dominance using functional MRI: a comparison with the Wada test. *Neurology* 1996;46(4):978-984.
5. Powell HW, Duncan JS. Functional magnetic resonance imaging for assessment of language and memory in clinical practice. *Curr Opin Neurol* 2005;18(2):161-166.
6. Baxendale SA, Thompson PJ, Duncan JS. Evidence-based practice: a reevaluation of the intracarotid amobarbital procedure (Wada test). *Arch Neurol* 2008;65(6):841-845.
7. Akanuma N, Koutroumanidis M, Adachi N, Alarcón G, Binnie CD. Presurgical assessment of memory-related brain structures: the Wada test and functional neuroimaging. *Seizure* 2003;12(6):346-358.

assessment of hemispheric language dominance. A large prospective study comparing several different functional MR techniques with the Wada test, with

postoperative outcomes analysis, would help to validate our results and determine the optimal techniques for this examination. This would likely improve

8. Abou-Khalil B. An update on determination of language dominance in screening for epilepsy surgery: the Wada test and newer noninvasive alternatives. *Epilepsia* 2007;48(3):442–455.
9. Baxendale S. The Wada test. *Curr Opin Neurol* 2009;22(2):185–189.
10. Klöppel S, Büchel C. Alternatives to the Wada test: a critical view of functional magnetic resonance imaging in preoperative use. *Curr Opin Neurol* 2005;18(4):418–423.
11. Hertz-Pannier L, Gaillard WD, Mott SH, et al. Noninvasive assessment of language dominance in children and adolescents with functional MRI: a preliminary study. *Neurology* 1997;48(4):1003–1012.
12. Bahn MM, Lin W, Silbergeld DL, et al. Localization of language cortices by functional MR imaging compared with intracarotid amobarbital hemispheric sedation. *AJR Am J Roentgenol* 1997;169(2):575–579.
13. Yetkin FZ, Swanson S, Fischer M, et al. Functional MR of frontal lobe activation: comparison with Wada language results. *AJNR Am J Neuroradiol* 1998;19(6):1095–1098.
14. Benson RR, FitzGerald DB, LeSueur LL, et al. Language dominance determined by whole brain functional MRI in patients with brain lesions. *Neurology* 1999;52(4):798–809.
15. Lehéricy S, Cohen L, Bazin B, et al. Functional MR evaluation of temporal and frontal language dominance compared with the Wada test. *Neurology* 2000;54(8):1625–1633.
16. Carpentier A, Pugh KR, Westerveld M, et al. Functional MRI of language processing: dependence on input modality and temporal lobe epilepsy. *Epilepsia* 2001;42(10):1241–1254.
17. Gao X, Jiang C, Lu C, Shen T. Determination of the dominant language hemisphere by functional MRI in patients with temporal lobe epilepsy. *Chin Med J (Engl)* 2001;114(7):711–713.
18. Baciú M, Kahane P, Minotti L, et al. Functional MRI assessment of the hemispheric predominance for language in epileptic patients using a simple rhyme detection task. *Epileptic Disord* 2001;3(3):117–124.
19. Liégeois F, Connelly A, Salmond CH, Gadian DG, Vargha-Khadem F, Baldeweg T. A direct test for lateralization of language activation using fMRI: comparison with invasive assessments in children with epilepsy. *Neuroimage* 2002;17(4):1861–1867.
20. Gaillard WD, Balsamo L, Xu B, et al. Language dominance in partial epilepsy patients identified with an fMRI reading task. *Neurology* 2002;59(2):256–265.
21. Rutten GJ, Ramsey NF, van Rijen PC, Alpherts WC, van Veelen CW. fMRI-determined language lateralization in patients with unilateral or mixed language dominance according to the Wada test. *Neuroimage* 2002;17(1):447–460.
22. Spreer J, Arnold S, Quiske A, et al. Determination of hemisphere dominance for language: comparison of frontal and temporal fMRI activation with intracarotid amytal testing. *Neuroradiology* 2002;44(6):467–474.
23. Woermann FG, Jokeit H, Luerding R, et al. Language lateralization by Wada test and fMRI in 100 patients with epilepsy. *Neurology* 2003;61(5):699–701.
24. Adcock JE, Wise RG, Oxbury JM, Oxbury SM, Matthews PM. Quantitative fMRI assessment of the differences in lateralization of language-related brain activation in patients with temporal lobe epilepsy. *Neuroimage* 2003;18(2):423–438.
25. Sabbah P, Chassoux F, Leveque C, et al. Functional MR imaging in assessment of language dominance in epileptic patients. *Neuroimage* 2003;18(2):460–467.
26. Gaillard WD, Balsamo L, Xu B, et al. fMRI language task panel improves determination of language dominance. *Neurology* 2004;63(8):1403–1408.
27. Deblaere K, Boon PA, Vandemaele P, et al. MRI language dominance assessment in epilepsy patients at 1.0 T: region of interest analysis and comparison with intracarotid amytal testing. *Neuroradiology* 2004;46(6):413–420.
28. Baciú MV, Watson JM, Maccotta L, et al. Evaluating functional MRI procedures for assessing hemispheric language dominance in neurosurgical patients. *Neuroradiology* 2005;47(11):835–844.
29. Benke T, Köylü B, Visani P, et al. Language lateralization in temporal lobe epilepsy: a comparison between fMRI and the Wada Test. *Epilepsia* 2006;47(8):1308–1319.
30. Chlebus P, Mikl M, Brázdil M, Pazourková M, Krupa P, Rektor I. fMRI evaluation of hemispheric language dominance using various methods of laterality index calculation. *Exp Brain Res* 2007;179(3):365–374.
31. Kesavadas C, Thomas B, Sujesh S, et al. Real-time functional MR imaging (fMRI) for presurgical evaluation of paediatric epilepsy. *Pediatr Radiol* 2007;37(10):964–974.
32. Szafarski JP, Holland SK, Jacola LM, Lindsell C, Privitera MD, Szafarski M. Comprehensive presurgical functional MRI language evaluation in adult patients with epilepsy. *Epilepsy Behav* 2008;12(1):74–83.
33. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151(4):264–269, W64.
34. Hirsch J, Ruge MI, Kim KH, et al. An integrated functional magnetic resonance imaging procedure for preoperative mapping of cortical areas associated with tactile, motor, language, and visual functions. *Neurosurgery* 2000;47(3):711–721; discussion 721–722.
35. Worthington C, Vincent DJ, Bryant AE, et al. Comparison of functional magnetic resonance imaging for language localization and intracarotid speech amytal testing in pre-surgical evaluation for intractable epilepsy: preliminary results. *Stereotact Funct Neurosurg* 1997;69(1-4 Pt 2):197–201.
36. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
37. Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Chapter 9: Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C (eds). *Cochrane handbook for systematic reviews of diagnostic test accuracy*. Version 1.0.0. The Cochrane Collaboration, 2009. <http://srdta.cochrane.org/>. Edited October 27, 2009. Accessed September 21, 2010.
38. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58(10):982–990.
39. Agresti A, Coull BA. Approximate is better than ‘exact’ for interval estimation of binomial proportions. *Am Stat* 1998;52(2):119–126.
40. Medina LS, Bernal B, Ruiz J. Role of functional MR in determining language dominance in epilepsy and nonepilepsy populations: a Bayesian analysis. *Radiology* 2007;242(1):94–100.
41. Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T. Human brain language areas identified by functional magnetic resonance imaging. *J Neurosci* 1997;17(1):353–362.
42. Gernsbacher MA, Kaschak MP. Neuroimaging studies of language production and comprehension. *Annu Rev Psychol* 2003;54:91–114.
43. Kurthen M, Helmstaedter C, Linke DB, Solymosi L, Elger CE, Schramm J. Inter-hemispheric dissociation of expressive and receptive language functions in patients with complex-partial seizures: an amobarbital study. *Brain Lang* 1992;43(4):694–712.