Are your MRI contrast agents cost-effective?
Learn more about generic Gadolinium-Based Contrast Agents.
FRESENIUS KABI
caring for life

AJNR

A Systematic Literature Review of Magnetic Resonance Spectroscopy for the Characterization of Brain Tumors

W. Hollingworth, L.S. Medina, R.E. Lenkinski, D.K. Shibata, B. Bernal, D. Zurakowski, B. Comstock and J.G. Jarvik

This information is current as of May 12, 2024.

AJNR Am J Neuroradiol 2006, 27 (7) 1404-1411
http://www.ajnr.org/content/27/7/1404

ORIGINAL RESEARCH

A Systematic Literature Review of Magnetic Resonance Spectroscopy for the Characterization of Brain Tumors

W. Hollingworth
L.S. Medina
R.E. Lenkinski
D.K. Shibata
B. Bernal
D. Zurakowski
B. Comstock
J.G. Jarvik

BACKGROUND AND PURPOSE: Proton MR spectroscopy (^{1}H-MR spectroscopy) is a potentially useful adjunct to anatomic MR imaging in the characterization of brain tumors. We performed an updated systematic review of the evidence.

METHODS: We employed a standardized search strategy to find studies published during 2002–2004. We reviewed studies measuring diagnostic accuracy and diagnostic, therapeutic, or health impact of ^{1}H-MR spectroscopy. We abstracted information on study design, ^{1}H-MR spectroscopy technique, and methodologic quality. We categorized studies into 5 subgroups: (1) metastasis versus high-grade tumor; (2) high-versus low-grade tumor; (3) recurrent tumor versus radiation necrosis; (4) tumor extent; and (5) tumor versus non-neoplastic lesion.

RESULTS: We identified 26 studies evaluating diagnostic performance, diagnostic impact, or therapeutic impact. No articles evaluated patient health or cost-effectiveness. Methodologic quality was mixed; most used histopathology as the reference standard but did not specify blinded interpretation of histopathology. One large study demonstrated a statistically significant increase in diagnostic accuracy for indeterminate brain lesions from 55%, based on MR imaging, to 71% after analysis of ^{1}H-MR spectroscopy. Several studies have found that ^{1}H-MR spectroscopy is highly accurate for distinguishing high- and low-grade gliomas, though the incremental benefit of ^{1}H-MR spectroscopy in this setting is less clear. Interpretation for the other clinical subgroups is limited by the small number of studies.

CONCLUSION: The current evidence on the accuracy of ^{1}H-MR spectroscopy in the characterization of brain tumors is promising. However, additional high-quality studies are needed to convince policy makers. We present guidelines to help focus future research in this area.

Conventional MR imaging provides highly detailed anatomic information and has become a mainstay in the diagnosis of suspicious brain lesions.[1] Several advances, most notably the development of contrast-enhanced MR imaging, have greatly improved the diagnostic accuracy of MR imaging. Despite this progress, the accurate characterization of brain lesions with MR imaging remains problematic in many cases.[2]

Proton MR spectroscopy (^{1}H-MR spectroscopy) provides additional information on the metabolic composition within an area of tissue. By comparing the relative concentration of these metabolites, clinicians can judge factors such as neuronal viability, neurotoxins, and membrane turnover within the volume of interest and, thereby, the likely underlying pathology.[3] The collection of ^{1}H-MR spectroscopy data requires that the MR imaging time is extended for 15 to 30 minutes while additional acquisition sequences are performed. ^{1}H-MR spectroscopy is an appealing, noninvasive adjunct to MR imaging.

In August 2002, the American College of Radiology requested that the Center for Medicare and Medicaid Services (CMS) reconsider the 1994 noncoverage decision for ^{1}H-MR spectroscopy. In September 2004, based in large part on 2 technology assessments,[4,5] CMS reaffirmed the existing noncoverage policy, concluding that "... the evidence is not adequate to conclude that ^{1}H-MR spectroscopy is reasonable and necessary... for use in the diagnosis of brain tumors." Several subsidiaries of large managed care organizations have reached similar noncoverage decisions, though this is far from universal. In the long run, noncoverage decisions are likely to discourage the uptake and use of ^{1}H-MR spectroscopy.

The first aim of this study is to provide an updated systematic review of the value of ^{1}H-MR spectroscopy for characterizing brain tumors. The second aim is to develop methodologic guidelines for measuring the efficacy of ^{1}H-MR spectroscopy to help focus future research in this area.

Methods

Defining Study Type and Clinical Subgroups

In this systematic review, we elected to include all studies that assessed the diagnostic performance (eg, sensitivity, specificity) or the impact of ^{1}H-MR spectroscopy on subsequent diagnostic testing, treatment choices, patient health, or cost effectiveness of care. It would be inappropriate to combine diagnostic accuracy results from diverse clinical applications of ^{1}H-MR spectroscopy. Therefore, we categorized publications according to the following 5 main clinical subgroups: 1) metastasis versus high-grade astrocytoma; 2) high- versus low-grade astrocytoma; 3) tumor extent before treatment; 4) neoplastic versus non-neoplastic lesions; 5) recurrent or residual tumor versus treatment-related change.

Received September 22, 2005; accepted after revision December 2.

From the Departments of Radiology (W.H., D.K.S., J.G.J.) and General Internal Medicine (B.C.), University of Washington, Seattle, Wash; Department of Radiology (L.S.M., B.B.), Miami Children's Hospital, Miami, Fla; Department of Radiology (R.E.L.), Beth Israel Deaconess Medical Center, Boston, Mass; and Departments of Orthopaedic Surgery and Biostatistics (D.Z.), Harvard Medical School, Boston, Mass.

This research was funded by a grant from the Neuroradiology Education and Research Foundation.

A preliminary summary of these data were presented at the American Society of Neuroradiology conference, Toronto, Canada, May 2005. Some of the data contained within this paper were also discussed at the American College of Radiology Imaging Network meeting, Washington, DC, September 2005.

Address correspondence to William Hollingworth, PhD, Department of Radiology, Box 359960, 325 Ninth Ave, Seattle, WA 98104-2499; e-mail: willh@u.washington.edu

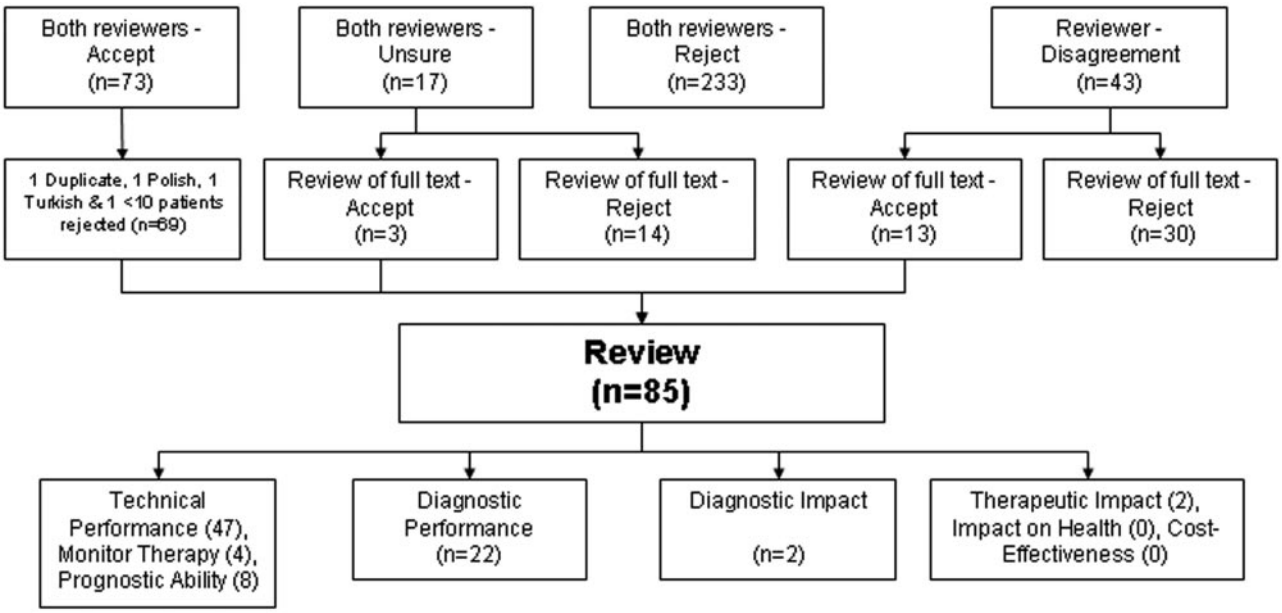

Fig 1. Studies identified by the systematic review.

Search Strategy

We searched Medline via the Pubmed interface, Embase via the Dialog interface, and the Cochrane Library data bases for relevant articles. Because the primary focus of this project was to update previous technology assessments,[4,5] we limited our search strategy to articles published between January 1, 2002, and December 31, 2004. The search strategy was tailored for each data base. The Medline search strategy is presented in Appendix 1. We excluded all Embase titles already identified by the Medline search. Two authors (J.G.J. and W.H.) then reviewed each Embase title to reach consensus on whether to purchase the abstract. The Cochrane library data base was searched by using the "Brain neoplasms" and "Magnetic Resonance Spectroscopy" medical subject headings.

All selected abstracts were independently screened by 2 authors based on the following 6 exclusion criteria: (1) does not use ^{1}H-MR spectroscopy; (2) not focused on brain tumors; (3) less than 10 patients with suspected tumors get ^{1}H-MR spectroscopy; (4) uses ^{1}H-MR spectroscopy to study the effect of therapy on normal brain tissue; (5) includes only patients with HIV/AIDS; and (6) a review paper reporting no new data. We obtained the full text of each article when one or both reviewers were "unsure" or recommended "full text review." Two additional exclusion criteria were applied on reviewing the full text: (1) duplicate publications and (2) articles not published in English, French, Spanish, German, or Japanese. We hand-searched citations of all eligible articles and sent e-mails to corresponding authors to identify additional articles initially overlooked.

Diagnostic performance studies were distributed to 2 reviewers (B.B., D.K.S., J.G.J., R.E.L., L.S.M., or W.H.) for independent review. Non-English language articles were reviewed by one reviewer fluent in that language—French (J.G.J.), Spanish (L.S.M.), German (K.F.L.), and Japanese (Y.A.). Each reviewer abstracted study information on a standardized Microsoft Excel spreadsheet. Reviewers recorded details about the dates of patient recruitment, sample size, other imaging tests used, reference standard, and the ^{1}H-MR spectroscopy technique. In particular, we recorded the metabolites evaluated, spectral analysis methods, single or multivoxel spectroscopy, imaging field strength, repetition time, echo time, and pulse sequence. Any differences between the 2 reviewers were resolved by a third reviewer through recourse to the original text.

Study Quality

We used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS)[6] tool to measure methodologic quality. QUADAS contains 14 items, including questions about the spectrum of patients, the validity of the reference standard, and the potential existence of disease progression, verification, review, and incorporation biases.[7] We added 1 item to the standard QUADAS tool: "Was the reproducibility (inter-radiologist or intertechnologist) of MR spectroscopy described?" The reviewers coded each item as "yes," "no," or "unclear." In our analysis, we interpreted both "no" and "unclear" responses as indicating that the quality criterion was not met.

Data Analysis

For each clinical subgroup, we tabulated estimates of sensitivity, specificity, percentage of correct diagnoses, and area under the receiver operating characteristic (ROC) curve. If data on statistical uncertainty were missing or incorrect, we calculated confidence intervals from the raw data. We plotted sensitivity, specificity, and ROC curve results to aid interstudy comparisons. ROC curves were calculated from the published area under the curve estimates by using PlotROC software.[8] This method assumes a bi-normal model for sensitivity and specificity and produces an ROC curve that is an approximation, though not identical, to the original data.

Results

The Medline search strategy identified 323 abstracts. After exclusion of duplicate and irrelevant Embase titles, 37 Embase abstracts were obtained for review. The search of the Cochrane library data base revealed no additional abstracts. The hand search of the citations and request to corresponding authors revealed 6 additional abstracts. Therefore, a total of 366 abstracts were reviewed. Reviewers agreed on the eligibility of the abstract in 323 of 366 cases (88%) (Figure 1).

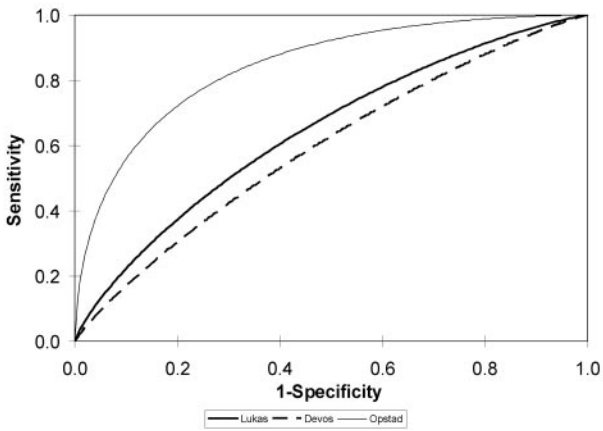

Fig 2. Receiver operating characteristic (ROC) curves measuring the sensitivity and specificity of ^{1}H-MR spectroscopy for distinguishing metastases from high-grade astrocytomas. The ROC curves are back-calculated from the area-under-the-curve figures provided by the authors. They approximate, but are not perfect matches, for the ROC curves based on the individual patient data.

Of the 85 eligible articles, 47 were considered to be technical feasibility studies that provided no estimate of diagnostic accuracy. Eight studies correlated ^{1}H-MR spectroscopy findings with survival to quantify the prognostic value. Four studies used ^{1}H-MR spectroscopy to monitor the success of therapy in changing the metabolic profile of brain tumors. These studies were not reviewed further. Of the remainder, 22[9-30] examined the diagnostic performance of ^{1}H-MR spectroscopy, 2 examined diagnostic impact,[31,32] and 2 measured the impact on radiation therapy.[33,34] No articles were found that evaluated the changes in patient health or the cost-effectiveness of health care due to ^{1}H-MR spectroscopy.

Technical Details

^{1}H-MR spectroscopy was most frequently evaluated for differentiation of high- and low-grade astrocytomas (Table 1). Most studies did not report the enrollment dates, making it difficult to judge whether multiple publications from the same researchers report on mutually exclusive patient cohorts. Biopsy or surgical resection was the sole reference standard in most studies. A substantial minority also used clinical and radiologic follow up to determine the final diagnosis. Several studies used automated analysis of the complete spectrum of metabolites to diagnostically categorize the MR spectra. The remaining studies focus on a handful of metabolites, most commonly choline (Cho), creatine (Cr), *N*-acetylaspartate (NAA), lactate, lipids, and myo-inositol. Most studies used single voxel spectroscopy; only one study[26] used 3T field strength MR.

Study Quality

On average, reviewers considered that 90% of studies used an accurate reference standard (Table 2). The same proportion also used a consistent reference standard in all patients, thereby minimizing verification bias. Very few studies (12%) were judged to have adequately addressed the issue of interradiologist variation. Likewise, authors were generally poor at reporting the median time delay between the index test, ^{1}H-MR spectroscopy, and the reference standard. Other areas of weakness included failure to explain the reason for patient withdrawals; lack of clarity about the pre-^{1}H-MR spectroscopy diagnostic tests; and failure to state that the reference standard results were interpreted independently from ^{1}H-MR spectroscopy.

Diagnostic Performance: MR Imaging & ^{1}H-MR Spectroscopy versus MR Imaging Alone

Moller-Hartmann et al[19] reported on 176 consecutive patients. The final diagnosis in most patients was established by histology within 10 days of single-voxel ^{1}H-MR spectroscopy. One pair of radiologists interpreted only the MR images; a second pair examined the MR imaging and MR spectra based on a qualitative interpretation of the metabolite peaks. All radiologists were unaware of the final diagnosis. The type and grade of lesion were correctly identified in 97 of 176 (55%) cases based on MR imaging alone. The remaining diagnoses were incorrect (15%) or indeterminate (30%). The addition of ^{1}H-MR spectroscopy information statistically significantly increased the proportion of correctly diagnosed cases to 71% (124/176) ($P < .01$). There were no cases where a correct diagnosis on MR imaging was mistakenly discarded due to the ^{1}H-MR spectroscopy findings.

A second, smaller, study by Ando et al[22] compared contrast-enhanced MR imaging (CE-MR imaging) to CE-MR imaging and ^{1}H-MR spectroscopy in 20 patients with suspected residual or recurrent tumor after therapy. The method of final diagnosis was inconsistent between patients, relying on either pathologic or clinical findings. Fourteen patients had a final diagnosis of residual or recurrent tumor, and 6 had treatment-related changes. The authors retrospectively selected a Cho/Cr ratio of greater than 1.5 to be indicative of tumor. Based on this threshold, the addition of ^{1}H-MR spectroscopy information to CE-MR imaging findings marginally increased sensitivity from 12 of 14 (86%) to 14 of 14 (100%) ($P = .79$) without altering specificity (4 of 6; 67%).

Metastasis versus High-Grade Astrocytoma

Two studies from the same research group evaluated the differentiation of metastases from high-grade astrocytoma by using long[24] and short[9] echo time ^{1}H-MR spectroscopy. The extent of any overlap in the 2 patient cohorts is unclear (Fig. 2). Both studies retrospectively assembled a cohort of patients from multiple hospitals. The MR hardware varied between hospitals, but spectroscopy was performed using standardized protocols. Both studies used automated spectral analysis for diagnostic classification. At long and short echo times, the area under the ROC curve (AUC) for differentiating glioblastomas from metastases was relatively poor (AUC = 64% [0.10 SE] and 59% [0.10 SE], respectively). This is statistically significantly better than chance alone; however, it is not high enough to suggest that ^{1}H-MR spectroscopy can be relied upon to differentiate metastases from glioblastomas.

Opstad et al,[25] prospectively recruited 47 patients with pathologically proved glioblastomas[23] or metastases[24]; 7 patients were later excluded due to poor quality spectra. The authors focused on the lipid peak-area ratio derived from short echo time, single-voxel ^{1}H-MR spectroscopy. They defined this as the ratio of L1 (the combined alanine, lactate, δ1.4 macromolecule, and δ1.3 lipid peak) to L2 (the combined δ0.9 lipid and δ0.87 macromolecule peaks). Using this ratio, they reported an AUC of 84% with both sensitivity and specificity equal to 80% at a threshold value of 2.9. The authors specu-

Table 1: Study description

Study	Clinical Subgroup	Start/End Year	No. of Patients	Reference Standard(s)	Metabolites	Diagnostic Categorization*	Voxels	TR (ms)	TE	Pulse Sequence
Ando et al, 2004[22]	Residual-recurrent/ necrosis	Unclear	20	Biopsy/resection, clinical & radiologic follow-up	Cho, Cr, NAA, Lac, Lip	Quantitative	Single	1500	270	Unclear
Astrakas et al, 2004[28]	Tumor grading	Unclear	66	Biopsy/resection	Cho, Cr, NAA, Lip, Lac	Quantitative	Multiple	1000	65	PRESS
Devos et al, 2004[9]	Tumor grading, primary/met	Unclear	205	Biopsy/resection	Complete spectrum	Automated	Single	1600, 2000, 2018, 2020	20, 30, 31, 32	STEAM and PRESS
Fountas et al, 2004[10]	Tumor grading	2000/2001	71	Biopsy/resection	Cho, Cr, NAA, Lac, Lip, mIns/Gly	Quantitative	Single	1600	135	PRESS
Gajewicz et al, 2003[23]	Tumor/tumorlike; tumor grading, primary/met	Unclear	29	Biopsy/resection	Cho, Cr, NAA, Lac, Lip, mIns/Gly, Glx, Ala	Automated	Single	2000	20, 136	STEAM and PRESS
Herminghaus et al, 2003[11]	Tumor grading	Unclear	94	Biopsy/resection	Cho, Cr, NAA, Lac, Lip	Automated	Single	1500	135	PRESS
Huang et al, 2003[12]	Tumor grading	Unclear	41 + 11 (test set)	Biopsy/resection	Complete spectrum	Automated	Single	1600	135	PRESS
Law et al, 2003[13]	Tumor grading	1999/2002	160	Biopsy/resection	Cho, Cr, NAA, Lac, Lip	Quantitative	Multiple	1500	144	PRESS
Lichy et al, 2004[29]	Residual-recurrent/ necrosis	Unclear	24	Clinical and radiologic follow-up	Cho, Cr, NAA, Lac, Lip	Quantitative	Multiple	1500	135	PRESS
Lukas et al, 2004[24]	Tumor grading	Unclear	183	Biopsy/resection	Complete spectrum	Automated	Single	1500–2020	135 or 136	PRESS
Majos et al, 2002[14]	Tumor grading, primary/met	Unclear	95 + 24 (test set)	Biopsy/resection	Cho, Cr, NAA, Lac, Lip, mIns/Gly, Glx, Ala	Quantitative	Single	2000	136	PRESS
Majos et al, 2003[16]	Tumor grading, primary/met	Unclear	108 + 25 (test set)	Biopsy/resection, clinical and radiologic follow-up	Cho, Cr, NAA, Lac, Lip, mIns/Gly, Glx, Ala	Quantitative	Single	2000	136	PRESS
Majos et al, 2003[15]	Tumor grading, primary/met	Unclear	130	Biopsy/resection, clinical and radiologic follow-up	Cho, Cr, NAA, Lac, Lip, Glx, Ala	Quantitative	Single	2000	136	PRESS
Majos et al, 2004[30]	Tumor grading	1998/2003	151	Biopsy/resection, clinical and radiologic follow-up	Cho, Cr, NAA, Lac, Lip, Glx, Ala, mIns/Gly	Automated	Single	2000	136 and 30	PRESS
McKnight et al, 2002[17]	Tumor extent	Unclear	44 (100 biopsies)	Biopsy/resection	Cho, NAA	Quantitative	Multiple	1000	144	PRESS
Mishra et al, 2004[18]	Tumor/tumorlike	Unclear	52	Biopsy/resection	Cho, Cr, NAA, Lac, Lip, mIns/Gly, Ala, Suc, Ace	Qualitative	Single	3000	144	Unclear
Moller-Hartmann et al, 2002[19]	Tumor/tumorlike; tumor grading, primary/met	Unclear	176	Biopsy/resection, clinical and radiologic follow-up, CSF and laboratory tests	Cho, Cr, NAA, Lac, Lip	Qualitative	Single	1500	135	PRESS
Nafe et al, 2003[20]	Tumor grading	Unclear	46	Biopsy/resection	Cho, Cr, NAA, Lac, Lip	Automated	Single	1500	135	PRESS
Opstad et al, 2004[25]	Primary/met	Unclear	47	Biopsy/resection	Cho, Cr, NAA, Lac, Lip, mIns/Gly, Glx, Others	Quantitative	Single	2000	30	STEAM or PRESS
Plotkin et al, 2004[26]	Residual-recurrent/ necrosis	Unclear	25	Clinical and radiologic follow-up	Cho, Cr, NAA	Quantitative	Single	6000	30	PRESS
Tate et al, 2003[21]	Tumor grading, primary/met	Unclear	144	Biopsy/resection	Complete spectrum	Automated	Single	2000, 1600	30 or 20	STEAM or PRESS
Traber et al, 2002[27]	Residual-recurrent/ necrosis	Unclear	54	Biopsy/resection, clinical and radiologic follow-up	Cho, Cr, NAA, Lac	Quantitative	Multiple	2000	272	Unclear

Note:—TR indicates repetition time; TE, echo time; Ala, alanine; Cho, choline; Cr, creatine; Gly, glycine; Glx, glutamate and glutamine; Lac, lactate; Lip, lipids; mIns, myo-inositol; NAA, *N*-acetylaspartate; Suc, succinate; PRESS, point-resolved spectroscopy sequence; STEAM, stimulated echo acquisition mode; primary/met, metastasis versus high grade tumor; Ace, acetate.

* Authors who made diagnostic classifications based on visualizing the spectra are categorized as "qualitative"; authors who present specific ratios or threshold values for distinguishing lesions are categorized as "quantitative"; authors who used statistical modeling, such as linear discriminant analysis, are categorized as "automated".

Table 2: Methodologic quality

Quality item	%*
Is the reference standard likely to correctly classify the target condition?	90
Did the whole sample or a random selection of the sample receive verification using a reference standard?	90
Did patients receive the same reference standard regardless of the index test result?	80
Were selection criteria clearly described?	76
Was the spectrum of patients representative of the patients who will receive the test in practice?	73
Were the MRS results interpreted without knowledge of the results of the reference standard?	71
Was the execution of MRS described in sufficient detail to permit replication of the test?	68
Was the reference standard independent of the MRS (ie, MRS did not contribute to the reference standard)?	66
Was the execution of the reference standard described in sufficient detail to permit its replication?	63
Were uninterpretable/intermediate test results reported?	59
Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	49
Were withdrawals from the study explained?	49
Were the reference test results interpreted without knowledge of the results of MRS?	41
Is the time period between MRS and the reference standard short enough to be reasonably sure that the target condition did not change between the 2 tests?	34
Was the reproducibility of (inter-radiologist or inter-technologist) MRS described?	12

Note:—MRS indicates magnetic resonance spectroscopy.
* Each of the quality items were assessed by 2 reviewers for English language articles and by one reviewer for the foreign language articles. Percentages represent the proportion of these assessments which judged the article to have met the quality criterion.

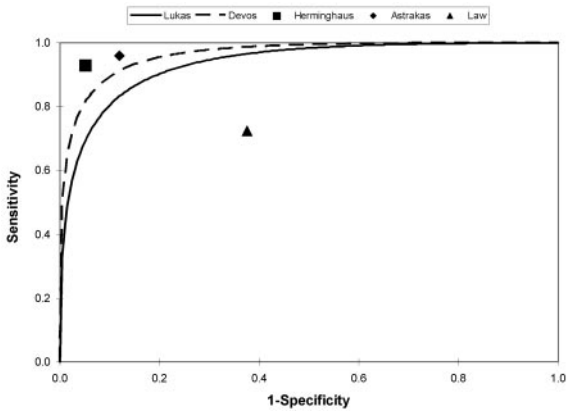

Fig 3. Receiver operating characteristic (ROC) curves and point estimates of sensitivity and specificity of ^{1}H-MR spectroscopy for distinguishing high- and low-grade astrocytomas. The ROC curves are back-calculated from the area-under-the-curve figures provided by the authors. They approximate, but are not perfect matches, for the ROC curves based on the individual patient data.

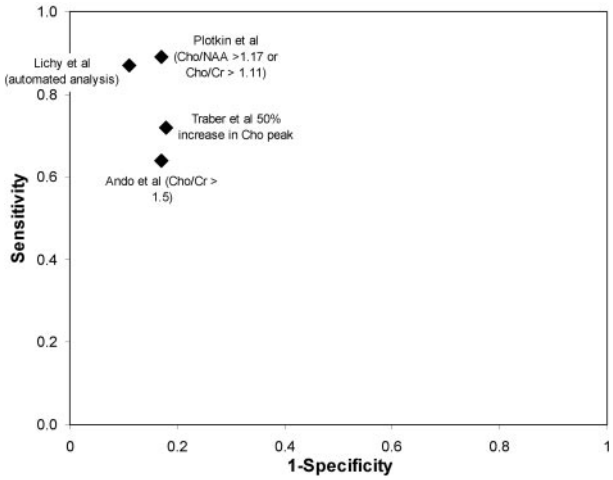

Fig 4. Sensitivity and specificity of ^{1}H-MR spectroscopy for differentiating recurrent or residual tumor from treatment-related changes.

lated that the difference in lipid profiles may be related to differences of membrane structure of infiltrative versus migratory tumor cells or to lipid metabolism.

High- versus Low-Grade Astrocytoma

Five studies examined the sensitivity and specificity of ^{1}H-MR spectroscopy for differentiating high- from low-grade tumors (Fig 3). Two studies described in the previous section[9,24] also provide information on tumor grading. In both studies, ^{1}H-MR spectroscopy was very accurate in differentiating high- and low-grade tumors, achieving an AUC of 94% (0.05 SE) and 96% (0.03 SE) in the studies that used long and short echo times, respectively.

Herminghaus et al[11] also used automated spectral analysis derived from a training set of 126 patients. This algorithm was validated in an independent cohort of 90 patients with histopathologically graded tumors (30 grade I/II, 29 grade III, 31 grade IV). The sensitivity and specificity of ^{1}H-MR spectroscopy for differentiating high- and low-grade tumors in this independent cohort was 95% (86%–98%; 95% confidence interval [95% CI]) and 93% (95% CI, 79%–98%) respectively (Fig 3). This diagnostic accuracy diminished in the differentiation grade III and grade IV tumors, with 6 of 31 grade IV tumors mistakenly assigned to grade III status.

Astrakas et al[28] prospectively recruited 66 patients with histologically confirmed brain tumors (grade I, 13; grade II, 30; grade III, 7; grade IV, 16). Multivoxel ^{1}H-MR spectroscopy analysis focused on the voxel with the highest Cho. The best diagnostic accuracy was achieved by an amalgam of Cho, Cr, and lipids and/or lactate (L) (Cho/Cr +0.49 L/Cr). This linear combination resulted in an AUC of 96% (0.02 SE). At a threshold value of 1.8, the sensitivity and specificity of ^{1}H-MR spectroscopy for diagnosing high-grade tumors were 96% (95% CI, 78%–100%) and 88% (95% CI, 75%–96%), respectively.

In contrast to the preceding work, Law et al[13] observed much lower diagnostic performance in a retrospective cohort of 160 patients with histopathologically confirmed lesions (120 grade III/IV, 40 low-grade) evaluated with multivoxel ^{1}H-MR spectroscopy, CE and perfusion MR imaging. A blinded interpretation of the Cho/NAA ratio had a sensitivity of 73% (95% CI, 64%–80%) and specificity of 63% (95% CI,

47%–76%) at a threshold value of 1.66; this was less accurate than MR perfusion and no better than CE-MR imaging.

Eight other studies have examined tumor grading but did not report sensitivity, specificity, or AUC estimates.[10,12,14-16,20,21,30]

These studies indicated that ^{1}H-MR spectroscopy resulted in accurate diagnoses in 78% to 96% of cases, though these accuracy figures will be dependent upon case mix.

Recurrent/Residual Tumor versus Treatment-Related Change

We identified 4 small studies examining the diagnostic performance of ^{1}H-MR spectroscopy in distinguishing recurrent tumor from treatment-related changes (Fig 4). Traber et al[27] presented data on 43 patients, with high-grade astrocytomas sequentially tracked with multiple-voxel ^{1}H-MR spectroscopy until completion of radiation therapy. An increased Cho peak (50% higher than contralateral tissue) was 72% (95% CI, 53%–86%) sensitive and 82% (95% CI, 48%–98%) specific in distinguishing tumor from radiation-induced necrosis. Ando et al,[22] in a study described in more detail previously, examined 20 patients with CE-MR imaging and ^{1}H-MR spectroscopy. Based on a choline-to-creatine ratio diagnostic threshold of 1.5, ^{1}H-MR spectroscopy had a sensitivity of 64% (95% CI, 35%–87%) and a specificity of 83% (95% CI, 36%–100%).

Lichy et al[29] used multivoxel spectroscopy in 24 patients with irradiated gliomas and a suspicious lesion on gadolinium-enhanced MR imaging. The final diagnosis was determined by clinical and imaging follow-up. Using the Cho/Cr ratio with a diagnostic threshold of 2, the authors identified 13 of 15 (87% [95% CI, 60%–98%] sensitivity) recurrent or residual tumors and 8 of 9 (89% [95% CI, 52%–100%] specificity) radiation-related changes. Plotkin et al[26] investigated the value of single-voxel ^{1}H-MR spectroscopy at 3T in a prospective study of 25 patients with suspected recurrent glioma based on MR imaging after treatment with surgery, interstitial radiation therapy, external radiation therapy, or chemotherapy. The final diagnoses were based on a minimum of 6 months' clinical follow-up and repeat MR imaging examinations. A combined diagnostic threshold of Cho/NAA (>1.17) and Cho/Cr (>1.11), resulted in 89% sensitivity and 83% specificity for identifying tumor. However, the authors also observed that sensitivity (95%) and specificity (100%) were higher still with single-photon emission CT.

Tumor Extent before Treatment

McKnight et al[17] prospectively recruited 44 patients with suspected glioma before image-guided resection or stereotactic biopsy of the tumor. Data from the preoperative multivoxel ^{1}H-MR spectroscopy study was used to select 4 potential targets for biopsy in each patient. In practice, the authors were unable to obtain biopsy samples at each target, and their analysis was based on 100 samples, of which only 7 were classified as nontumor. The authors based diagnosis on the Cho-NAA index (CNI), where CNI is the number of standard deviations between the Cho to NAA ratio within a given voxel and that of the control voxels. At a threshold CNI of greater than 2.5, the authors reported 90% (95% CI, 84%–96%) sensitivity and 86% (95% CI, 56%–100%) specificity for predicting the presence of tumor in the biopsy sample. The overall AUC for CNI was 94% (95% CI, 87%–99%). Up to half of the T2-hyperintense lesion outside of the gadolinium-enhanced lesion contained CNI greater than 2.5. This suggests that ^{1}H-MR spectroscopy might have a considerable therapeutic impact on surgical and radiation target volumes.

Tumor versus Non-Neoplastic Lesions

Two studies measured the diagnostic performance of ^{1}H-MR spectroscopy for distinguishing tumor from non-neoplastic lesions.[18,23] However, the study Gajewicz et al[23] included only 2 non-neoplastic lesions and was of limited value. Mishra et al[18] differentiated 52 histopathologically proved tumor cysts, abscesses, or benign cysts by using single voxel ^{1}H-MR spectroscopy and diffusion-weighted MR imaging. The authors reported the sensitivity and specificity of ^{1}H-MR spectroscopy to be 96% (95% CI, 83%–99%) and 100% (95% CI, 86%–100%) respectively. This compares favorably with diffusion-weighted imaging where specificity remained high (100%), but sensitivity was diminished (72%).

Diagnostic Impact

Murphy et al performed ^{1}H-MR spectroscopy in 100 consecutive patients with suspected brain tumors.[31] The authors highlighted 2 cases incorrectly classified as glioblastoma based on conventional imaging that were correctly down-graded to grade 2/3 astrocytomas based on ^{1}H-MR spectroscopy. In a further 4 cases, the differential diagnoses of arachnoid cyst, infection, stroke, or meningiomas were correctly excluded on the basis of ^{1}H-MR spectroscopy. The authors concluded that in 6 of 100 (6%) cases, ^{1}H-MR spectroscopy could have made a significant contribution to the preoperative diagnosis. It was unclear whether ^{1}H-MR spectroscopy provided confirmatory or contradictory information in the remaining patients.

The study by Hall et al[32] looked at the impact of ^{1}H-MR spectroscopy on the diagnostic yield of biopsy. ^{1}H-MR spectroscopy was used in 42 patients and all 42 biopsies yielded diagnostic tissue; however, the lack of a control group who did not get ^{1}H-MR spectroscopy-guided biopsy limits the value of these results. No firm conclusions can be drawn about the incremental benefit of ^{1}H-MR spectroscopy in guiding biopsy.

Therapeutic Impact

Pirzkall et al[34] recruited 20 patients with grade II gliomas who underwent both MR imaging and 3D multivoxel ^{1}H-MR spectroscopy before surgery. The target volume based on MR imaging was compared with a ^{1}H-MR spectroscopy target volume based on areas of the tumor with a Cho/NAA index greater than 2 (CNI2). Due to technical limitations, spectra were only available for an average of 68% of the tumor volume. In the tumor regions that were assessed by both MR imaging and ^{1}H-MR spectroscopy, 96% of the ^{1}H-MR spectroscopy-defined tumor volume was contained within the MR imaging-defined volume. Despite this, in 45% of patients, some portion of the ^{1}H-MR spectroscopy-defined tumor margin extended beyond the MR imaging-defined volume. The authors suggested that MR spectra can be used in conjunction with MR images to fine tune the clinical target volume.

The same researchers published a study of multivoxel ^{1}H-MR spectroscopy on 30 high-grade gliomas after surgery but before adjuvant radiation therapy.[33] MR imaging target volumes were compared with the ^{1}H-MR spectroscopy-derived CNI2 volume. By adding the area of metabolic tumor cell infiltration, defined by the CNI2 to the MR imaging target

volume, the authors found a 14% increase in the clinical target volume for radiation therapy. Of 10 patients who had had ostensibly total tumor resection, all had areas of residual elevated CNI and 8 had a new onset of contrast enhancement during follow-up. In 4 of these 8 cases, the area of contrast enhancement was located within the area of elevated CNI2. There was also a direct relationship between a large volume of residual CNI2 and a shorter time to occurrence of new contrast enhancement, though this was of borderline statistical significance.

Discussion

We conducted a systematic review of studies of ^{1}H-MR spectroscopy for the characterization of brain tumors published between 2002 and 2004. Many relevant studies on ^{1}H-MR spectroscopy were published before 2002,[17,19,35-45] but we elected to focus our review on research published since the technology assessments that formed the basis of the CMS non-coverage decision. Our search was restricted to publications in peer-reviewed medical journals because we believe that this evidence carries the most weight. We may have missed some relevant data that appeared in books, reports, or conference proceedings. As with all literature reviews, our results might be subject to publication bias, whereby positive findings on ^{1}H-MR spectroscopy get submitted and published while negative findings do not. Tests for publication bias are available,[46] but with the small sample of articles in each clinical subgroup, such tests would have limited statistical power.

Despite the steady accumulation of evidence, many policy makers remain unconvinced about the value of ^{1}H-MR spectroscopy. The following list provides guidance that we believe would make future research more valuable for policy makers.

- Diagnostic performance studies should include a blinded assessment of the sensitivity and specificity of MR imaging and, if relevant, contrast enhanced MR imaging as a benchmark against which to compare ^{1}H-MR spectroscopy. Policy makers are being asked to pay for ^{1}H-MR spectroscopy on top of standard MR imaging sequences; therefore, it is reasonable for them to expect evidence to demonstrate incremental benefit of ^{1}H-MR spectroscopy. Likewise, radiologists should be interested in evaluating whether the extra scanner and interpretation time is justified by improved diagnoses and patient care.
- Diagnostic performance studies should evaluate the accuracy of ^{1}H-MR spectroscopy *in combination with* MR imaging. In the near future, it is unlikely that radiologists will make a diagnosis based solely on an automated decision rule. Where MR imaging findings are highly suggestive and ^{1}H-MR spectroscopy is equivocal, radiologists will naturally place more weight on the former. Diagnostic performance studies looking at ^{1}H-MR spectroscopy in isolation are of scientific value, but of less clinical significance. In our review, the study by Moller-Hartmann et al[19] provides the best role model for diagnostic performance studies.
- Standardized diagnostic thresholds would aid the interpretation of the literature. Many authors selected post hoc thresholds that maximized accuracy. It is likely that these thresholds will not perform as well in independent patient cohorts.
- For statistically precise conclusions, studies of ^{1}H-MR spectroscopy should have large sample sizes.[47] For example, if ^{1}H-MR spectroscopy changes the diagnosis in approximately 10% of patients, a sample size of about 160 would be needed to estimate this proportion to within $\pm$5%. In our review, few studies recruited sufficiently large patient cohorts. In some clinical subgroups, there were no large studies.
- Researchers, reviewers, and editors should ensure that published studies adhere to the STARD guidelines.[48] These guidelines are recent; therefore, it is understandable that many articles fell short of their high standards. It is unclear whether the apparent poor quality was due to poor research methods or simply ambiguous descriptions of methods.
- Diagnostic impact studies are a very important element of the case for ^{1}H-MR spectroscopy. Even if ^{1}H-MR spectroscopy does not change the leading diagnosis, it may rule out differential diagnoses and thereby reduce the need for biopsy. Diagnostic impact studies measuring radiologist confidence in the leading differential diagnoses and the perceived need for biopsy before and after ^{1}H-MR spectroscopy are warranted.
- Important therapeutic impact studies have been done and have suggested that 3D ^{1}H-MR spectroscopy imaging can significantly alter radiation therapy target volumes. The next step should be a randomized controlled study to assess whether ^{1}H-MR spectroscopy actually does influence management decisions and patient outcomes.
- There has been little discussion of the cost effectiveness of ^{1}H-MR spectroscopy. This is a relevant factor for policy makers. This metric can be calculated in many ways, including cost per additional case correctly diagnosed, cost per biopsy avoided, cost per year of survival, and cost per quality adjusted life year.

Conclusions

Of the 22 studies that measured diagnostic performance, the largest head-to-head comparison of MR imaging alone versus MR imaging and ^{1}H-MR spectroscopy provided encouraging findings that ^{1}H-MR spectroscopy can make a significant contribution to diagnosis for patients with indeterminate brain lesions.[19] The conduct of additional, well-designed, prospective studies aiming to replicate this head-to-head comparison will provide the more definitive evidence that policy makers seek before making coverage decisions.

A number of large diagnostic performance studies have demonstrated that ^{1}H-MR spectroscopy can accurately distinguish between high- and low-grade astrocytomas. This work now needs to be extended to demonstrate: (1) diagnostic thresholds selected a priori, rather than post hoc, can achieve similar diagnostic accuracy, (2) the incremental diagnostic yield of ^{1}H-MR spectroscopy compared with anatomic MR imaging, and (3) that any improvement in tumor grading by ^{1}H-MR spectroscopy leads to a reduction in biopsy rates or changes in therapy. Evidence in other clinical subgroups, such as the use of ^{1}H-MR spectroscopy to distinguish neoplastic and non-neoplastic lesions or to differentiate recurrent tumors from radiation necrosis, is limited by the small number of studies.

Appendix 1

Medline Search Strategy

MEDLINE < Searched on 22nd February 2005>

Search Terms Results

1. Magnetic Resonance Spectroscopy [MH] OR "Magnetic Resonance Spectroscopy" [TW] OR "MR spectroscopy"[TW]: 112,283

2. Human [MH] AND (2002 [DP] OR 2003 [DP] OR 2004 [DP]) NOT Case reports[PT]: 1,009,999

3. Neoplasms [MH] OR tumor [TW] OR cancer* [TW] OR neoplasm* [TW] OR neoplas*[TW] OR lesion* [TW] OR mass [TW]: 2,480,174

4. Brain [TW] OR cranial [TW] OR cerebr* [TW]: 913,827

5. Brain neoplasms [MH]: 77,290

6. #1 AND #2 AND (#5 OR (#3 AND #4)): 323

Acknowledgments

We thank Ken F Linnau, MD, and Yoshimi Anzai, MD, for their assistance in translating the German and Japanese language literature included in this systematic review.

References

1. Sartor K. **MR imaging of the brain: tumors.** *Eur Radiol* 1999;9:1047–54
2. Nelson SJ. **Multivoxel magnetic resonance spectroscopy of brain tumors.** *Mol Cancer Ther* 2003;2:497–507
3. Barker PB. **Fundamentals of MR spectroscopy.** In: Gillard JH, Waldman AD, Barker PB, eds. *Clinical MR Neuroimaging: Diffusion, Perfusion and Spectroscopy.* 1st ed. Cambridge, UK: Cambridge University Press; 2005
4. New England Medical Center EPC. Magnetic resonance spectroscopy for brain tumors. In: EPC Technical Support of the CPTA Technology Assessment Program. Prepared for the Agency for Healthcare Research and Quality (AHRQ). Contract No. 290-02-0022 TO. Boston, MA; 2003
5. Technology Evaluation Center. Magnetic resonance spectroscopy for evaluation of suspected brain tumor. *TEC Bull (Online)* 2003;20:23–26
6. Whiting P, Rutjes AW, Reitsma JB, et al. **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
7. Kelly S, Berry E, Roderick P, et al. **The identification of bias in studies of the diagnostic performance of imaging modalities.** *Br J Radiol* 1997;70:1028–35
8. Kurt Rossman Laboratories, University of Chicago. *PlotROC Excel macro* [computer program]. Version 1.0.0. Chicago: University of Chicago.
9. Devos A, Lukas L, Suykens JA, et al. **Classification of brain tumours using short echo time 1H MR spectra.** *J Magn Reson* 2004;170:164–75
10. Fountas KN, Kapsalaki EZ, Vogel RL, et al. **Noninvasive histologic grading of solid astrocytomas using proton magnetic resonance spectroscopy.** *Stereotact Funct Neurosurg* 2004;82:90–97
11. Herminghaus S, Dierks T, Pilatus U, et al. **Determination of histopathological tumor grade in neuroepithelial brain tumors by using spectral pattern analysis of in vivo spectroscopic data.** *J Neurosurg* 2003;98:74–81
12. Huang Y, Lisboa PJ, El-Deredy W. **Tumour grading from magnetic resonance spectroscopy: a comparison of feature extraction with variable selection.** *Stat Med* 2003;22:147–64
13. Law M, Yang S, Wang H, et al. **Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging.** *AJNR Am J Neuroradiol* 2003;24:1989–98
14. Majos C, Alonso J, Aguilera C, et al. **Adult primitive neuroectodermal tumor: proton MR spectroscopic findings with possible application for differential diagnosis.** *Radiology* 2002;225:556–66
15. Majos C, Alonso J, Aguilera C, et al. **Utility of proton MR spectroscopy in the diagnosis of radiologically atypical intracranial meningiomas.** *Neuroradiology* 2003;45:129–36
16. Majos C, Alonso J, Aguilera C, et al. **Proton magnetic resonance spectroscopy ((1)H MRS) of human brain tumours: assessment of differences between tumour types and its applicability in brain tumour categorization.** *Eur Radiol* 2003;13:582–91
17. McKnight TR, von dem Bussche MH, Vigneron DB, et al. **Histopathological validation of a three-dimensional magnetic resonance spectroscopy index as a predictor of tumor presence.** *J Neurosurg* 2002;97:794–802
18. Mishra AM, Gupta RK, Jaggi RS, et al. **Role of diffusion-weighted imaging and in vivo proton magnetic resonance spectroscopy in the differential diagnosis of ring-enhancing intracranial cystic mass lesions.** *J Comput Assist Tomogr* 2004;28:540–47
19. Moller-Hartmann W, Herminghaus S, Krings T, et al. **Clinical application of proton magnetic resonance spectroscopy in the diagnosis of intracranial mass lesions.** *Neuroradiology* 2002;44:371–81
20. Nafe R, Herminghaus S, Raab P, et al. **Preoperative proton-MR spectroscopy of gliomas–correlation with quantitative nuclear morphology in surgical specimen.** *J Neurooncol* 2003;63:233–45
21. Tate AR, Majos C, Moreno A, et al. **Automated classification of short echo time in in vivo 1H brain tumor spectra: a multicenter study.** *Magn Reson Med* 2003;49:29–36
22. Ando K, Ishikura R, Nagami Y, et al. [**Usefulness of Cho/Cr ratio in proton MR spectroscopy for differentiating residual/recurrent glioma from non-neoplastic lesions**]. *Nippon Igaku Hoshasen Gakkai Zasshi* 2004;64:121–26
23. Gajewicz W, Papierz W, Szymczak W, et al. **The use of proton MRS in the differential diagnosis of brain tumors and tumor-like processes.** *Med Sci Monit* 2003;9:MT97–105
24. Lukas L, Devos A, Suykens JA, et al. **Brain tumor classification based on long echo proton MRS signals.** *Artif Intell Med* 2004;31:73–89
25. Opstad KS, Murphy MM, Wilkins PR, et al. **Differentiation of metastases from high-grade gliomas using short echo time 1H spectroscopy.** *J Magn Reson Imaging* 2004;20:187–92
26. Plotkin M, Eisenacher J, Bruhn H, et al. **123I-IMT SPECT and 1H MR-spectroscopy at 3.0 T in the differential diagnosis of recurrent or residual gliomas: a comparative study.** *J Neurooncol* 2004;70:49–58
27. Traber F, Block W, Flacke S, et al. [**1H-MR Spectroscopy of brain tumors in the course of radiation therapy: use of fast spectroscopic imaging and single-voxel spectroscopy for diagnosing recurrence**]. *Rofo* 2002;174:33–42
28. Astrakas LG, Zurakowski D, Tzika AA, et al. **Noninvasive magnetic resonance spectroscopic imaging biomarkers to predict the clinical grade of pediatric brain tumors.** *Clin Cancer Res* 2004;10:8220–28
29. Lichy MP, Henze M, Plathow C, et al. [**Metabolic imaging to follow stereotactic radiation of gliomas—the role of 1H MR spectroscopy in comparison to FDG-PET and IMT-SPECT**]. *Rofo* 2004;176:1114–21
30. Majos C, Julia-Sape M, Alonso J, et al. **Brain tumor classification by proton MR spectroscopy: comparison of diagnostic accuracy at short and long TE.** *AJNR Am J Neuroradiol* 2004;25:1696–704
31. Murphy M, Loosemore A, Clifton AG, et al. **The contribution of proton magnetic resonance spectroscopy (1HMRS) to clinical brain tumour diagnosis.** *Br J Neurosurg* 2002;16:329–34
32. Hall WA, Liu H, Maxwell RE, et al. **Influence of 1.5-Tesla intraoperative MR imaging on surgical decision making.** *Acta Neurochir Suppl* 2003;85:29–37
33. Pirzkall A, Li X, Oh J, et al. **3D MRSI for resected high-grade gliomas before RT: tumor extent according to metabolic activity in relation to MRI.** *Int J Radiat Oncol Biol Phys* 2004;59:126–37
34. Pirzkall A, Nelson SJ, McKnight TR, et al. **Metabolic imaging of low-grade gliomas with three-dimensional magnetic resonance spectroscopy.** *Int J Radiat Oncol Biol Phys* 2002;53:1254–64
35. Adamson AJ, Rand SD, Prost RW, et al. **Focal brain lesions: effect of single-voxel proton MR spectroscopic findings on treatment decisions.** *Radiology* 1998;209:73–78
36. Butzen J, Prost R, Chetty V, et al. **Discrimination between neoplastic and nonneoplastic brain lesions by use of proton MR spectroscopy: the limits of accuracy with a logistic regression model.** *AJNR Am J Neuroradiol* 2000;21:1213–19
37. Hall WA, Martin A, Liu H, et al. **Improving diagnostic yield in brain biopsy: coupling spectroscopic targeting with real-time needle placement.** *J Magn Reson Imaging* 2001;13:12–15
38. Kimura T, Sako K, Gotoh T, et al. **In vivo single-voxel proton MR spectroscopy in brain lesions with ring-like enhancement.** *NMR Biomed* 2001;14:339–49
39. Lin A, Bluml S, Mamelak AN. **Efficacy of proton magnetic resonance spectroscopy in clinical decision making for patients with suspected malignant brain tumors.** *J Neurooncol* 1999;45:69–81
40. Rand SD, Prost R, Haughton V, et al. **Accuracy of single-voxel proton MR spectroscopy in distinguishing neoplastic from nonneoplastic brain lesions.** *AJNR Am J Neuroradiol* 1997;18:1695–704
41. Roser W, Hagberg G, Mader I, et al. **Assignment of glial brain tumors in humans by in vivo 1H-magnetic resonance spectroscopy and multidimensional metabolic classification.** *Magma* 1997;5:179–83
42. Shukla-Dave A, Gupta RK, Roy R, et al. **Prospective evaluation of in vivo proton MR spectroscopy in differentiation of similar appearing intracranial cystic lesions.** *Magn Reson Imaging* 2001;19:103–10
43. Tedeschi G, Lundbom N, Raman R, et al. **Increased choline signal coinciding with malignant degeneration of cerebral gliomas: a serial proton magnetic resonance spectroscopy imaging study.** *J Neurosurg* 1997;87:516–24
44. Taylor JS, Langston JW, Reddick WE, et al. **Clinical value of proton magnetic resonance spectroscopy for differentiating recurrent or residual brain tumor from delayed cerebral necrosis.** *Int J Radiat Oncol Biol Phys* 1996;36:1251–61
45. Wilken B, Dechent P, Herms J, et al. **Quantitative proton magnetic resonance spectroscopy of focal brain lesions.** *Pediatr Neurol* 2000;23:22–31
46. Deeks JJ, Macaskill P, Irwig L. **The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed.** *J Clin Epidemiol* 2005;58:882–93
47. Medina LS, Zurakowski D. **Measurement variability and confidence intervals in medicine: why should radiologists care?** *Radiology* 2003;226:297–301
48. Bossuyt PM, Reitsma JB, Bruns DE, et al. **Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative.** *Radiology* 2003;226:24–28