Prospective diagnostic performance evaluation of SV-^{1}H-MRS for typing and grading brain tumours

Short title: Prospective diagnostic performance evaluation of SV-^{1}H-MRS for brain tumours

JULIÀ-SAPÉ, MARGARIDA, [1, 2,3], CORONEL, INDIRA [2], MAJÓS, CARLES [4,1], CANDIOTA, ANA PAULA [1, 2], SERRALLONGA, MARTA [4], COS, MÒNICA [4], AGUILERA, CARLES [4,1], ACEBES, JUAN JOSÉ [5,1], GRIFFITHS, JOHN R.[6], ARÚS, CARLES [2, 1, 3]

[1] *Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Cerdanyola del Vallès, Spain*

[2] *Departament de Bioquímica i Biologia Molecular, Unitat de Bioquímica de Biociències, Edifici Cs, Universitat Autònoma de Barcelona, UAB, 08193, Cerdanyola del Vallès, Spain*

[3] *Institut de Biotecnologia i de Biomedicina (IBB), Universitat Autònoma de Barcelona, UAB, 08193, Cerdanyola del Vallès, Spain*

[4] *Institut de Diagnòstic per la Imatge (IDI), CSU de Bellvitge, Autovia de Castelldefels km 2.7, 08907, L'Hospitalet de Llobregat, Barcelona, Spain*

[5] *Department of Neurosurgery, Hospital Prínceps d'Espanya, CSU de Bellvitge, Feixa Llarga s/n, 08907, L'Hospitalet de Llobregat, Barcelona, Spain*

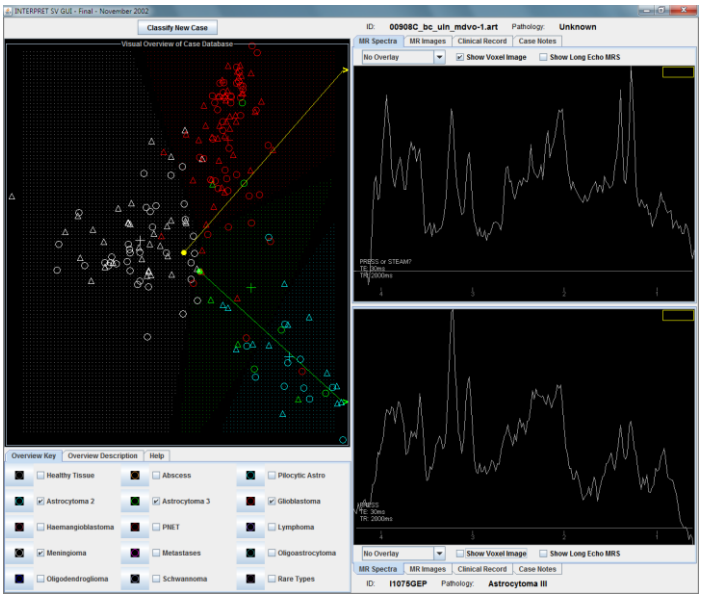
[6] *Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 ORE, UK*

Corresponding author: Carles Arús.

Carles.Arus@uab.es
Telephone: +34 93 581 1257
Fax: +34 93 581 1264

Graphical abstract

Analysis of single-voxel ^{1}H-MRS data from a prospective patient cohort, by spectroscopists blind to other information, gave comparable diagnostic performance to that of expert radiologists using MRI images and clinical data. MRS-based data also helped the radiologists to improve their diagnosis of grade IV glioblastomas, metastases, medulloblastomas and lymphomas, and of glial tumours in general, without negatively influencing other radiological classifications. Of the 4 MRS classifiers tested, the INTERPRET decision-support system performed best.



73 words

Abstract

The purpose of this study was to evaluate whether single-voxel proton magnetic resonance spectroscopy could add useful information to conventional MRI in the preoperative characterisation of the type and grade of brain tumours.

MRI and MRS exams from a prospective cohort of 40 consecutive patients were analysed double-blind by radiologists and spectroscopists before the histological diagnosis was known. The spectroscopists had only the MR spectra, whereas the radiologists had both the MRI images and basic clinical details (age, sex and presenting symptoms). Then, the radiologists and spectroscopists exchanged their predictions and re-evaluated their initial opinions, taking into account the new evidence. Spectroscopists used four different systems of analysis for ^{1}H-MRS data, and the efficacy of each of these methods was also evaluated.

Information extracted from ^{1}H-MRS significantly improved the radiologists' MRI-based characterisation of the grade IV tumours (glioblastomas, metastases, medulloblastomas and lymphomas) in the cohort, (Area Under the Curve (AUC) in the MRI re-evaluation, 0.93, vs. AUC in the MRI evaluation, 0.85), and also of the less malignant glial tumours (AUC in the MRI re-evaluation, 0.93, vs. AUC in the MRI evaluation, 0.81).

One of the MRS analysis systems used, the INTERPRET decision-support system, out-performed the others, as well as being better than the MRI evaluation for the characterisation of grade III astrocytomas.

Thus preoperative MRS data improves radiologists' performance in diagnosing grade IV tumours, and for those of grade II-III, MRS helped them to recognise the glial lineage. Even in cases when it did not improve their diagnoses, provision of MRS data to the radiologists had no negative influence on their predictions.

263 words

Keywords

"Diagnostic Techniques and Procedures", "Diagnosis, Computer-Assisted", "ROC Curve", "Brain Neoplasms", "Magnetic Resonance Spectroscopy", "Magnetic Resonance Imaging".

Abbreviation key

Ala = Alanine

AUC = Area under the curve

CD = Chance Diagonal

Cho = Choline

CI = Confidence Interval

Cre = Creatine

DSS=Decision-support system

DWI =Diffusion Weighted Imaging

EBM = Evidence Based Medicine

Gly = Glycine

Glx = Glutamate and Glutamine

^{1}H-MRS = Proton magnetic resonance spectroscopy.

INTERPRET = International network for Pattern-Recognition of Tumours Using Magnetic Resonance

Lac = Lactate

MI =*myo*-Inositol

MI/Gly= *myo*-Inositol/Glycine index

MRI-E = Magnetic Resonance Imaging Evaluation

MRI-R = Magnetic Resonance Imaging Re-evaluation

MRS-E = Magnetic Resonance Spectroscopy Evaluation

MRS-R = Magnetic Resonance Spectroscopy Re-evaluation

PNET = Primitive Neuroectodermal Tumour

ROC = Receiver-operating characteristic

SNR = Signal-to-Noise Ratio

SV = Single-Voxel

STARD = Standards for the Reporting of Diagnostic accuracy studies

Tau = Taurine

TE = Echo Time

VOI = Volume of Interest

WHO = World Health Organisation

WLW = Water Linewidth

Introduction

Diagnosis of a brain tumour requires the correct assignment of both the cell type from which it arose and its grade of malignancy. An MRI exam only provides the definitive diagnosis of type and grade in a few brain tumours (1) such as low-grade meningiomas, so brain tumour patients routinely undergo stereotactic brain biopsy for histopathological diagnosis. Unfortunately, brain biopsy has been reported to cause serious neurological damage in 2-3% of patients and death in 0.2-1.5% (2-5) so a less harmful alternative that might reduce the need for a biopsy, even in a proportion of cases, would be highly desirable.

Common brain tumour types have characteristic ^{1}H-MRS patterns (6,7) Since MRS can be performed during routine MRI scans, non-invasive MRS data could easily be included when making a radiological evaluation. However, MRS of brain tumours is currently little used in daily clinical practice, partly because several Evidence Based Medicine (EBM) assessments on its diagnostic impact have reported negatively on its clinical utility, resulting in denial of financial reimbursement for MRS in the USA (8-10) and partly because its diagnostic performance has rarely been formally evaluated, and then only for specific tumour types (11-14).

Another problem is that few radiologists are trained to quantitatively evaluate MRS, so ways to minimise the need for such knowledge should be explored. For these reasons, prospective studies on the diagnostic performance and/or impact of MRS over MRI-alone in brain tumours should be performed (15). To ensure that these studies contribute to future EBM meta-analyses, they should adhere to the STARD guidelines for reporting of diagnostic accuracy (8,9,15).

The simplest method for analysing ^{1}H-MRS is assignment of the peaks to known substances (16), followed by quantification of the peak areas relative either to one another (17) or to tissue water, which is assumed to be effectively invariant. However, this is laborious and time-consuming, even for single-voxel (SV) spectra, and there are no "marker" resonances that characterise individual tumour types. Instead, many peak intensities change simultaneously, depending on tumour type or grade. Multivariate analysis techniques are therefore widely employed (18-25), although they have so far had limited clinical use. To further simplify data interpretation, a multivariate analysis program can be included in a decision-support system (DSS) such as the one developed by the INTERPRET consortium (26,27). The INTERPRET DSS works in conjunction with a large database of 304 SV ^{1}H spectra from brain tumours (and also some tumour-like lesions and normal volunteers) that were acquired in the early 2000s by a collaboration of European research hospitals and are now available on-line at http://gabrmn.uab.es/interpretvalidateddb (28). The database contains spectra and clinical data from 43 meningiomas WHO grade I, 84 glioblastomas, 21 astrocytomas WHO grade II, 38 metastases, among a total of 34 different types of tumours and tumour-like lesions. The diagnosis of the lesion giving rise to each spectrum in the database was established by histopathology (except for some tumour-like lesions such as multiple sclerosis) and verified by committees of pathologists and clinicians. Although the spectra were obtained by SV MRS at 1.5T it has subsequently been shown that it would be possible to use 3T spectra with classifiers trained at 1.5T (29).

The present study aimed to test, in a prospective cohort of patients, (a) whether SV MRS could add information to conventional radiology for non-invasively typing and

grading brain tumours; (b) if so, if it was equally helpful for all brain tumour types and (c) which of the four classification systems was best for providing this information.

Experimental

Patients

The study was approved by the institutional review board of the Hospital Prínceps d'Espanya, CSU de Bellvitge, and written informed consent was obtained from all patients. Between January 2004 and March 2005, MRI and MRS data from 54 consecutive patients with suspected brain tumours were prospectively acquired at the MR unit of the Hospital (Table 1).

MRI and MRS acquisition methods

MRI and MRS were acquired on consecutive patients on a 1.5-T MR imaging unit (ACS-NT or Intera-Master, Philips Medical Systems, Best, The Netherlands). The standard MRI examination consisted of: sagittal T1 spin-echo (SE), axial T2-SE, axial fluid-attenuated inversion recovery (FLAIR), axial T1-SE, axial T1 after administration of a gadolinium contrast agent (Gd) and coronal T1 post-Gd, and diffusion-weighted images (DWI). Single voxel (SV) point-resolved ^{1}H MR spectroscopy (MRS) was performed after the conventional MRI, using the standard receiver head-coil in all cases. The volume of interest (VOI) for acquiring MRS data was chosen by one of 4 radiologists (A-D) in order to obtain an average spectroscopic representation of the largest possible part of the tumour while avoiding contamination of the sample by extra-tumoural tissue. Therefore, the following criteria were used for voxel positioning: include the largest possible voxel within the solid tumoural area, as judged by

inspection of the full unenhanced and contrast-enhanced MR images; avoid areas of cysts or necrosis and minimise contamination from the surrounding non-tumoural tissue, as in (26,30). The VOI ranged between 1.5 x 1.5 x 1.5 $(cm)^3$ (3.4 ml) and 2 x 2 x 2 $(cm)^3$ (8 ml), depending on tumour dimensions (26). Spectra at two TE were obtained from the same VOI, for every patient, as in (26,30): 1st, water-suppressed spin-echo short TE (2000/30/ 92–184) (TR(ms)/TE(ms)/averages); 2nd, water-suppressed spin-echo long TE (2000/136/126 –252); 3rd, unsuppressed water spin-echo long TE (2000/136/16); and 4th, unsuppressed water spin-echo short TE (2000/30/16). A total of 512 data points were collected over a spectral width of 1000 Hz.

Spectra were processed with jMRUI (31) and the INTERPRET Data Manipulation software (DMS) (http://gabrmn.uab.es/DMS) using the same parameters as in (26)

Study protocol: MRI and MRS analysis methods

MRI and MRS data were evaluated by a 4-phase procedure (Figure 1): 1st phase, MRI evaluation (MRI-E): an evaluator radiologist analysed MRI data and routine clinical information (age, sex and presenting symptoms) according to the usual institutional protocol; 2nd phase, MRS evaluation (MRS-E): the panel of expert spectroscopists analysed the MRS data blind to any other clinical or image information; 3rd phase, MRS re-evaluation (MRS-R): the panel of spectroscopists saw the MRI-E results and re-evaluated the MRS; 4th phase, MRI re-evaluation (MRI-R): the radiologist who previously performed the MRI-E saw the MRS-E results and re-evaluated the MRI diagnosis. Evaluation results were not used in clinical decision-making.

Evaluation criteria: cases for which the MRI and MRS exams were available to the radiologists and spectroscopists and which could undergo the 4 phases of the study protocol before the histopathological diagnosis was known.

Inclusion criteria: cases for which MRI and MRS were of sufficient technical quality, and in which all the 4 phases had been completed and a diagnosis confidently established (see "Reference standard" section).

Each case was only judged once. No replicate judgments were performed by parallel teams of radiologists and spectroscopists. Ratings were given using a 5-point confidence scale ("0": definitely not; "4": definitely yes). The diagnoses were (32): Meningioma WHO grade I, astrocytoma WHO grade II, astrocytoma WHO grade III, glioblastoma, metastasis, abscess, lymphoma, primitive neuroectodermal tumour (PNET), oligodendroglioma WHO grade II, and oligodendroglioma WHO grade III. It was possible to give ratings for more than one diagnosis, and up to two additional diagnoses could be suggested and rated using the same scale if it was believed that the tumour type was not on the list. In this way, the study allowed differential diagnoses. ~~Radiologists also recorded (by direct question) whether the tumour was of high grade and if it was glial, and whether it was intra- or extra-axial.~~

Four radiologists (A-D) participated in the study. Due to constraints on the time contribution they were allowed to make to research protocols, they requested that the study should span no longer than one year. They either acquired the MRI and MRS (distributor role) or participated in the evaluations (evaluator role), and they took turns assuming the two roles. The distributor radiologist supervised scanning of the patient, reported the MRI study to the clinical board at the hospital and was medically responsible for the patient. The distributor radiologist then sent the MRI study and some

routine clinical information (age, sex) to an evaluator radiologist in the same institution, and sent the MRS study to the team of spectroscopists at the Biochemistry Department.

Four spectroscopists (E-H) participated in the study. They received no other information than the MR spectra, which were performed at both short and long TE.

The Magnetic Resonance Spectroscopy Evaluation (MRS-E) and the Magnetic Resonance Spectroscopy Re-evaluation (MRS-R) were consensus diagnoses established by the spectroscopists from the results of the four different methods (Figure 1), using a standardised decision protocol to ensure uniform rating. The protocol was applied in all cases in the following order:

(1st) -The consensus opinion of the expert spectroscopists (33). The spectra were examined visually and a structured, written summary was created listing the following characteristics on both short and long TE spectra: Presence/absence of necrotic lipids, and at which TE. Macromolecules between 2.0-2.5 ppm at short TE, modulated glutamine + glutamate (Glx) signal at long TE, or reduction or absence of the NAA singlet at 2.0 ppm. Cho/Cre ratio (1, <1, >1); presence/absence of Ala, Lac, Tau, MI/Gly and its variation in relative intensity with TE with respect to the Cre signal, peaks at 5.3 ppm or any other signal of interest. Spectral quality was also evaluated based on SNR and WLW using the same criteria as in (26,34).

(2nd) The result from the INTERPRET-DSS (26,27). Short TE spectra were processed off-line with the INTERPRET data manipulation software (26) (http://gabrmn.uab.es/dms) on a Linux workstation and entered into the INTERPRET-DSS v1.0 (27) (http://gabrmn.uab.es/dss), which had an embedded classifier trained to distinguish among the most common brain tumour types (low-grade glial tumours vs.

low-grade meningiomas vs. glioblastomas and metastases). The rating algorithm is summarised in Figure 2.

(3rd) Area classifiers based on long and short TE spectra. Integrated peak areas were obtained from jMRUI-processed spectra and used to create two 4-class classifiers (one at short TE and the other at long TE) for distinguishing among (glioblastoma and metastasis) vs. meningioma vs. astrocytoma WHO grade III vs. astrocytoma WHO grade II (30). The rating algorithm is summarised in Figure 3.

(4th) A MI/Gly index. This was calculated according to the following formula (35), and the criteria in Table 2, from the jMRUI processed spectra:

MI/Gly index = [(peak height at 3.55 ppm, short TE)/(peak height at 3.03 ppm, short TE)]/[(peak height at 3.55 ppm, long TE)/(peak height at 3.03 ppm, long TE)]

After applying the four methods of spectral analysis, a consensus MRS-E with confidence values for one or more tumour classes was obtained, using the following criteria: For each case and class, the mean confidence value was calculated. If a method was not applicable to a class, it was not used when calculating the mean for that class (for example, the MI/Gly index was not applicable to non-glial tumours). However, if any of the four methods absolutely discarded the possibility of the case being a member of a given class, a value of "0" could be assigned as summary rating for the MRS-E of that class. The expert spectroscopists' opinion could prevail for those pathologies in which no objective classifier system was available (i.e. pyogenic brain abscesses (36)).

As cases had to undergo all of the 4 phases before the histopathological diagnosis was known, the number of spectroscopists evaluating each case varied, although a minimum of two persons was always present. To justify the confidence values given after MRS evaluation a final "spectroscopic report", including a qualitative

biochemical description of the short and long TE spectra, was written by two of the four spectroscopists and reviewed by the senior spectroscopist (H).

The radiologists' experience in neuroradiology was: A, 11 years; B, 8 years; C, 3 years, D, 13 years. The spectroscopists' experience of (*in-vivo*) brain tumour MRS was: E, 4 years; F, 2 years; G, 5 years; H, 14 years. Both readers of MRI and MRS were blind to each other and to the results of the histopathology examination of the biopsy sample.

Reference standard

The definitive diagnosis was achieved by histological evaluation of a biopsy by a pathologist. A diagnosis of metastasis was accepted from MRI in cases where there was no biopsy available, only if a clinical committee of neuroradiologists and neurosurgeons confirmed the presence of three or more brain lesions and the patient had clinical evidence of neoplasia (Table 1). The MRI and MRS study and evaluations were always performed a few days before the operation or the brain biopsy, and the pathologist had no contact either with the participant radiologists or spectroscopists.

Statistics

Diagnostic accuracy values were calculated from ROC curves (SPSS 14.0) under the non-parametric assumption and with a 95% confidence limit for testing the null hypothesis that there were no differences with a random test (i.e. chance diagonal (CD) equal to 0.5). ROC curves (37) were chosen since they provide the best estimate of the accuracy of a diagnostic test, and they are suited to quantitative, non- binary outcome results, like the one used in our protocol (evaluation results were given as 5-point

confidence scales). Differences between AUCs were tested with the Hanley-McNeil test (38). P values lower than 0.05 were considered significant.

As up to 10 or 12 different diagnoses ("classes") and ratings were possible at each phase (i.e., participants were allowed to give several posssible diagnoses), results were analysed by dichotomisation (arranged in two classes) (39), as follows:

$$\text{Dichotomised score class } i = (\text{score class } i / \sum \text{scores}) \times 100$$

In addition, as a limited number of cases per class were available in this study, classes were grouped into higher-order "superclasses" (Figure 4) as in (1), in order to maximise the information obtained from the ratings. For example: glioblastomas, metastases, medulloblastomas and lymphomas were pooled into a "WHO grade IV" superclass before ROC curve analysis by dichotomisation. Then, the pooled dichotomised score was calculated as follows:

$$\text{Pooled dichotomised score superclass } I = [(\text{score class } i_1 + \text{score class } i_2 + \ldots + \text{score class } i_n)/\sum \text{scores}] \times 100$$

when,

$$\text{superclass } I = \text{class } i_1 + \text{class } i_2 + \ldots + \text{class } i_n$$

Results

Fifty-four consecutive patients satisfied the evaluation criteria and forty satisfied the inclusion criteria (Table 1). Seven cases were excluded because of absence of a biopsy (one also had poor MRS data), one because the biopsy was not diagnostic, three because a single histopathological diagnosis could not be established, and three because the four phases of the evaluation could not be completed (Table 1).

Results from the four phases are displayed in Table 3. The most important question was whether a radiologist's classification was improved by knowing the results of the MRS classification. We found that knowledge of the MRS-E scores and report significantly improved the radiological classification of three superclasses (Table 3, MRI-E vs. MRI-R column): WHO grade IV tumours ((glioblastomas, metastases, medulloblastomas and lymphomas); glioblastomas and metastases as a superclass, and glial WHO grade II-III tumours as a superclass. In situations where knowledge of the spectroscopic reports and scores did not improve diagnostic accuracy, the radiologists' ratings were not negatively influenced (no statistically significant differences were found when comparing AUCs using the Hanley-McNeil test).

In general, initial spectroscopic ratings were neither better nor worse than those from the radiologists (Table 3, MRI-E vs. MRS-E column), and no significant differences between MRI-E and MRS-E were found in most comparisons. However, in three categories the radiologists were better than the spectroscopists: Glial WHO grades III-IV as a superclass, aggressive tumours (grades III-IV) as a superclass, and glial tumours.

When MRS-E had no effect on radiologist's re-evaluations, it was for one of three reasons:

(1) Both MRI and MRS already had near-perfect discriminatory power (AUC=1). In meningiomas of WHO grade I and in the low-grade meningioma superclass, the AUC was between 0.98 and 0.99 in the initial ratings, and therefore no improvement was to be expected when adding MRS information.

(2) Neither MRI nor MRS had a good discriminatory power, i.e. the AUC is not significantly different from a chance test (Table 3, MRI-E vs. CD and MRS-vs. CD column). This occurred in metastases, and in some glial types and grades (astrocytomas WHO grade III, and oligoastrocytomas WHO grade III), and meningiomas of WHO grade II.

(3) Radiologists performed better than the spectroscopists: Glial WHO grades III-IV as a superclass, aggressive tumours (grades III-IV) as a superclass, and glial tumours.

~~With respect to the way the ratings were obtained, (by direct question or by pooling results, Table 4b), no significant differences were found.~~

When the individual performances of the four methods used in MRS-E were retrospectively evaluated, the INTERPRET-DSS gave the best results (Table 4), allowing calculation of ROC curves for 5 tumour types and all the 10 superclasses. It performed better than chance in meningioma WHO grade I, glioblastoma and astrocytoma WHO grade III and in all superclasses except "tumour". For astrocytomas of WHO grade III it gave a significantly different - and higher - AUC than MRI-E and MRS-E (0.87 with the INTERPRET-DSS vs. 0.66 of the MRI-E and 0.71 of the MRS-E) (Tables 3 and 5), which means that for this class, the sole use of the INTERPRET DSS for evaluation of a short TE spectrum was better than the combined use of all the

methods for spectroscopic analysis (MRS-E) or the straightforward interpretation of MRI and clinical data by radiologists (MRI-E). An example of the series, an astrocytoma of WHO grade III (case number 4, Table 1), is presented in Figure 5. The INTERPRET-DSS was also better than the consensus of spectroscopic evaluation methods for MRS-E in the glial WHO grade III and glial WHO grades III-IV superclasses. In the rest of classes and superclasses, no differences were found between the DSS and either the radiologists (MRI-E) or the spectroscopists (MRS-E), except for glial WHO grade III and the glial WHO III-IV superclasses, in which, again, the DSS performed significantly better than the consensus of the four MRS methods for MRS-E, and also better than the radiologists (AUC (CI), 0.87 (0.72-1.02) vs. 0.66 (0.44-0.88) MRI-E and vs. 0.71 (0.54-0.88) MRS-E). Surprisingly, the automated DSS performed better than the consensus of the spectroscopists, even though the spectroscopists had knowledge of the DSS predictions. This suggests that either the combined use of spectroscopic evaluation methods or the way in which the consensus was obtained had a negative effect on the theoretical helpfulness of the INTERPRET DSS in the discrimination of astrocytoma WHO grade III tumours.

Expert spectroscopic judgment was the second-best system (Table 4), although not all 40 cases received class ratings. The spectroscopists scored better than a chance test in meningiomas WHO grade I, low-grade meningiomas, glial WHO grades II-III, glial and aggressive tumours. However, they performed significantly different (worse) than the radiologists in aggressive tumours (AUC (CI), 0.70 (0.55-0.86) vs. 0.94 (0.87-1.01) MRI-E) and glial tumours (AUC (CI), 0.70 (0.53-0.87) vs. 0.95 (0.87-1.01) MRI-E), and worse than the consensus of all four MRS methods in glial WHO grades II-III

(AUC (CI), (0.74 (0.55-0.93) vs. 0.90 (0.78-1.02)) and glial tumours (AUC (CI), 0.70 (0.53-0.87) vs. 0.84 (0.73-0.97)).

The area classifiers performed better than a chance test for three of the classes in which they were applicable (meningioma WHO grade I, astrocytoma WHO grade III and glioblastomas) and two superclasses (glial and aggressive). The performance in these classes and superclasses however, was not significantly different from either the MRI-E or the MRS-E.

The MI/Gly ratio did not perform better than a chance test in any of the classes to which it was applicable.

Discussion

For most tumour types, MRS analysis, blind to any other information about the patient, predicted the type of brain mass as well as an expert neuroradiologist who had had access both to the MRI images and the patient's age, sex and presenting symptoms. Information from ^{1}H-MRS helped radiologists to increase their performance in diagnosing tumours (glioblastomas, metastases, medulloblastomas and lymphomas) that displayed the highest grade of malignancy (WHO grade IV), and also in less malignant glial tumours of WHO grades II and III. Even in cases where it did not help, presentation of ^{1}H-MRS-based data to radiologists had no negative influence on their judgements.

The best system for MRS evaluation appeared to be the INTERPRET DSS, which outperformed subjective evaluation by expert spectroscopists, as judged by the AUC values. In astrocytomas of WHO grade III, the INTERPRET-DSS had better performance than either the combined spectroscopic evaluation (MRS-E) or the

radiologists (MRI-E). The performance of the INTERPRET-DSS was remarkably good, as it had only been "trained" to distinguish between three classes: meningiomas, glioblastomas plus metastases, and low grade glial tumours. Despite this limitation, its design, which includes an overview space of the whole database of pre-classified cases, allowed its practical use as a system for comparing similarities between any two spectra, regardless of their underlying pathologies (26).

The most evident limitation of the other MRS evaluation methods used in this study was related to the classes that the different systems were designed to distinguish. For example, the MI/Gly ratios only allowed predictions for glial tumours whereas the area-based classifiers were only suitable for the five most common tumour classes. In fact, when using a mathematical classifier system for analysis, an inclusion criterion is normally introduced to restrict the number of classes to those that the classifier is trained to handle (14), except when the MRS analysis method is based on expert spectroscopic knowledge (12). Our study showed that, in a clinical setting, methods that allow analysing any type of mass are preferrable to those designed to specific diseases. None of the four methods tested distinguished glioblastomas from metastases, despite the clinical interest in that distinction and conflicting evidence from other reports: in one (40) high-grade glial tumours (11 astrocytomas of WHO grade III and 20 glioblastomas) were differentiated from 25 metastases. In another (41), 25 glioblastomas were distinguished from 34 metastases on the basis of a lipid peak. Other approaches, such as multivoxel spectroscopy (42) or acquisition of adjacent voxels (43) may be suitable for distinguishing these two types by MRS.

The semi-automatic DSS gave the best MRS interpretation and was simple to use. The current version v3.0 (27) automatically pre-processes the raw SV ^{1}H-MR

spectrum and positions the case in a latent space (http://gabrmn.uab.es/DSS). The protocol used for manual MRS analysis was similar to that in (14), but we found it much more laborious than the 20 minutes for MRI and MRS analysis described therein. In our hands a trained spectroscopist (F) took about one hour to process each case, followed by another hour for evaluation by the team of spectroscopists (E, F, G, H). A recent study performed using the INTERPRET-DSS v2.0 (13) benefited from recent improvements within the system and did not require the offline DMS spectral processing that was necessary with the earlier version used in our study.

Two earlier prospective studies tested the applicability of MRS to brain tumour diagnosis. One compared the diagnostic performance of MRI with MRI plus MRS (12) (using expert spectroscopic interpretation of the MRS data) and found 15.4% more correct diagnoses when MRS was included. However, that study differed from ours by including only those cases for which the radiologists provided a single diagnosis, and excluding those in which they gave differential diagnoses; also it did not specify the cancer types in which 1H-MRS was helpful. Another prospective study (14) classified 164 patients with three different tumour types (glial, metastasis or lymphomas); a diagnostic performance of 90% was obtained on a test set of 43 patients, using information from MRI and MRS. Although our study was performed on a slightly smaller cohort (40 cases vs. 43) , it adhered to STARD guidelines (44), included all patients presenting in the defined time window, and provided evidence of the performance with different tumour types. In addition, the INTERPRET DSS used a substantially larger training dataset (304 cases vs. 164), while the training dataset used for the MI/Gly classifier and the area-based classifiers (151 cases) was comparable in size to that used in (14). The present study was performed on SV spectra, but future

work may utilise voxels from multi-voxel MRSI datasets, as was done in an earlier INTERPRET study (26).

Two types of evaluations of the value of MRS are possible: diagnostic impact and therapeutic impact. Our study concentrated on the first one, like (11) but differently from (13), in that it evaluated MRS as an intervention-limiting diagnostic tool. In the context of therapeutic impact, it is interesting to note that 11/54 (~20%) patients in this series had no biopsy (too old, in bad physical condition, refused treatment or were sent directly to palliative care) or a non-diagnostic biopsy, so in these patients, treatment decisions were primarily based on radiological criteria. Furthermore, even in this small series there were three non-malignant brain masses (two abscesses and one multiple sclerosis lesion) that had to be diagnosed without biopsy. If in future, as seems likely, non-invasive treatments take more prominent therapeutic roles in brain tumour management, the risks associated with biopsy-based diagnosis will become less acceptable. The combination of MRI and ^{1}H MRS interpreted by a DSS therefore seems to be a practical step forward. Larger prospective studies, adhering to the STARD guidelines (43), are now needed to test the performance of MRS (44).

Finally, we will show an example that illustrates the contribution of the MRS methods to the diagnosis of a difficult case (Fig 5). The initial radiological evaluation (MRI-E) of this patient included eight differential diagnoses: 21% score for oligodendroglioma WHO grade II, 14% for astrocytoma WHO grade III, 14% for astrocytoma WHO grade II, 14% for lymphoma, 14% for oligodendroglioma WHO grade III; 7% for glioblastoma, 7% for abscess and 7% for PNET (Note that the percentages represent dichotomised scores and not diagnostic probabilities).The evaluation of this case with the INTERPRET DSS is shown in Fig 5(c). The unknown

case is shown as a yellow dot in the middle of the "case overview panel" on the left panel of the DSS. Neighbouring cases from several pathologies are shown: glioblastoma (red), meningioma (white) or astrocytoma WHO grade III (green). This positioning in an ambiguous region of the dataspace suggests that the case is quite atypical from the spectroscopic point of view, just as it was by MRI. However, although the unknown case showed several spectroscopic similarities when compared to the nearest astrocytoma WHO grade III, the presence of lactate is a distinctive feature, being also seen in its closest astrocytoma grade III neighbour (spectrum in bottom right panel). Using the DSS, the team of spectroscopists scored 4 pathologies: 28% score for astrocytoma WHO grade III, 28% for glioblastoma, 28% for meningioma; and 14% for metastasis. The combined use of the four MRS evaluation systems and the application of the evaluation protocol yielded an MRS-E with 5 differential diagnoses: 30% score for glioblastoma; 20% for astrocytoma WHO grade III, 20% for metastasis, 20% for "cystic liquid"; and 10% for meningioma. In the final radiological evaluation (MRI-R), the radiologists´ confidence scores shifted towards grade III and from oligodendrogliomas as follows: 25% score for astrocytoma WHO grade III, 25% for glioblastoma, 17% for oligodendroglioma WHO grade III, 8% score for astrocytoma WHO grade II, 8% for oligodendroglioma WHO grade II, 8% for lymphoma and 8% for PNET. The histopathological diagnosis was Astrocytoma WHO grade III.

This was a particularly difficult case in which neither the radiologists nor the spectroscopists were able to produce a definitive diagnosis. Nevertheless, even here the spectroscopic data, as interpreted by the spectroscopists, enabled the radiologists to develop a differential diagnosis that came closer to the histopathological diagnosis

Conclusion

^{1}H-MRS interpreted by a team of spectroscopists improved the performance of radiologists in the typing and grading of some brain tumours while not having a negative effect on radiologists' predictions on other types. The combination of MRI and ^{1}H MRS interpreted with the help of a DSS may be a practical solution for a radiological department

Acknowledgements

Guillem Mercadal and Miquel Cabañas for recompiling the DMS processing software.

This work was also partially funded by CIBER-BBN, which is an initiative funded by the VI National R&D&i Plan *2008-2011, CIBER Actions* and financed by the *Instituto de Salud Carlos III* with assistance from the European Regional Development Fund. Funding also from projects (MEDIVO2: SAF2005-03650) from Ministerio de Educación y Ciencia; and (PHENOIMA: SAF2008-03323) from Ministerio de Ciencia e Innovación in Spain. Partial funding also contributed by the European Commission: eTUMOUR (contract no. FP6-2002-LIFESCIHEALTH 503094), and HealthAgents (contract no.: IST2004-27214).

Authors' Contributions statement

MJS designed the study, processed MRS data, performed MRS-E and MRS-R, analysed data and wrote the paper; IC collected results of all evaluations, coordinated the MRS-E and MRS-R, processed MRS data and performed MRS-E and MRS-R; CM designed the study, collected MRS and MRI data, coordinated the neuroradiologists' team, performed MRI-E and MRI-R and critically revised the results and the paper; AC processed MRS data and performed MRS-E and MRS-R, MS collected MRS and MRI data and performed MRI-E and MRI-R, MC, collected MRS and MRI data and performed MRI-E and MRI-R; CAG, collected MRS and MRI data and performed MRI-E and MRI-R, JJA, collected biopsy results , coordinated the clinical committee and critically revised the results and the paper, JRG designed the study and critically revised the results and the paper; CA coordinated the spectroscopists' team, designed

the study, performed MRS-E and MRS-R and critically revised the results and the paper. All authors discussed the results and commented on the manuscript.

Table 1. Clinical demographic characteristics of the study population~~, as well as whether they were included or not into the study and the reason why~~. M: Male, F: Female, ~~ADD : Absence of a definitive diagnosis~~, OS: Open surgery, OB: Open biopsy, SB: Stereotactic biopsy, C: Clinical.

Case	Sex	Age	Presenting symptoms	Tumour location	Diagnosis reached	Diagnosis
1	M	60	Right hemiparesis	Left frontal	OS	Glioblastoma Multiforme
2	F	31	Seizures	Left frontal	OS	Astrocytoma WHO grade III
3	F	73	Disturbed mental status, left hemiparesis	Right temporal	SB	Oligoastrocytoma WHO grade III
4	M	23	Seizures	Left parietal	SB	Astrocytoma WHO grade III
5	F	59	Seizures	Right frontal	SB	Oligodendroglioma WHO grade III
6	M	62	Headache, syncope	Left temporal	OB	Glioblastoma Multiforme
7	M	28	Gait disturbance	Right cerebellum	OB	Medulloblastoma
8	M	58	Left hemiparesis	Right temporoparietal	OB	Glioblastoma Multiforme
9	M	62	Headache , visual disturbances	Right temporal	OS	Glioblastoma Multiforme
10	M	58	Asthenia	Both frontal lobes	SB	Astrocytoma WHO grade II
11	M	48	Headache , gait disturbance	Right cerebellum	OB	Metastasis
12	F	60	Right homonymous hemianopsia	Left occipital	OB and necropsy	Glioblastoma Multiforme
13	M	68	Headache, speech disturbance and right hemiparesis	Left temporoparietal	SB	Glioblastoma Multiforme
14	F	59	Seizures	Corpus callosum	SB	Astrocytoma WHO grade III
15	M	75	Headache and vertigo	Cerebellar vermis	OB	Metastasis
16	M	75	Left hemiparesis	Right frontoparietal	OB	Meningioma WHO grade I
17	M	35	Headache , visual disturbance	Both frontal lobes	SB	Metastasis
18	M	47	Headache , seizures	Left parietooccipital	OB	Metastasis
19	M	65	Gait disturbance	Right frontal	OB	Meningioma WHO grade I
20	M	50	Fever , seizures	Right frontal	SB	Abscess
21	M	58	Seizures	Left frontal	OS	Meningioma WHO grade II
22	F	78	Left hemiparesis	Right frontoparietal	OB	Astrocytoma WHO grade III
23	F	50	Seizures	Left frontotemporal	OB	Astrocytoma WHO grade III
24	F	47	Seizures	Left frontal	OB	Meningioma WHO grade I
25	F	58	Headache , disturbed mental status	Left frontotemporal	OB	Astrocytoma WHO grade III
26	M	67	Right hemiparesis , disturbed mental status	Left temporal	OB	Astrocytoma WHO grade III
27	F	31	Increased intracranial pressure	Right thalamus	SB	Astrocytoma WHO grade III
28	M	54	Right hemiparesis	Left frontal	OB	Meningioma WHO grade I
29	M	50	Headache, toxic syndrome	Both frontal lobes	SB	Astrocytoma WHO grade III
30	M	60	Right hemiparesis	Both frontal lobes	C	Abscess
31	F	77	Speech disturbance	Left temporal	SB	Glioblastoma Multiforme
32	M	24	Increased intracranial pressure	Right cerebellum	OB	Melanocytoma
33	M	27	Headache	Left temporoparietal	SB	Meningioma WHO grade II
34	M	52	Vertigo	Cerebellar vermis	C	Metastasis
35	F	34	Headache	Left parietal	C	Multiple sclerosis
36	F	71	Gait disturbance	Right temporal	C	Meningioma WHO grade I
37	M	63	Syncope	Pituitary	OS	Pituitary adenoma
38	F	75	Left hemiparesis	Right frontoparietal	C	Meningioma WHO grade I
39	M	45	Seizures	Left frontal	OB	Oligoastrocytoma WHO grade III
40	M	69	Headache and gait disturbance	Right temporal	C	Lymphoma

Table 2. Decision protocol for the MI/Gly index.

	MI/Gly index interval value			
	[1.11-1.92]	[1.92-2.85]	[2.85-4.87]	[4.87-6.07]
GLIOBLASTOMA	3	2	0	0
ASTROCYTOMA WHO GRADE III	0	2	1	0
ASTROCYTOMA WHO GRADE II	0	0	2	2
OLIGODENDROGLIOMA WHO GRADE II	0	0	2	3

Table 3. Results of the 4 evaluation phases. Comparison between phases, taking into account the classes and the superclasses. AUC: Area Under the Curve, CI: 95% Confidence Interval, vs.: versus, CD: Chance Diagonal, S: Significant differences found ($P < 0.05$). -: No significant differences found ($P \geq 0.05$).

		AUC (CI) at PHASE											
	n	**MRI-E**	**MRS-E**	**MRS-R**	**MRI-R**	**MRI-E vs. CD**	**MRS-E vs. CD**	**MRS-R vs. CD**	**MRI-R vs. CD**	**MRI-E vs. MRS-E**	**MRI-E vs. MRS-R**	**MRI-E vs. MRI-R**	**MRS-E vs. MRS-R**
Classes													
MENINGIOMA WHO GRADE I	6	0.98 (0.95-1.01)	0.99 (0.96-1.02)	0.99 (0.97-1.01)	0.99 (0.96-1.02)	***S***	***S***	***S***	***S***	-	-	-	-
MENINGIOMA WHO GRADE II	2	0.50 (0.08-0.92)	0.74 (0.28-1.20)	1.00 (1.00-1.00)	0.73 (0.29-1.17)	-	-	***S***	-	-	***S***	-	-
METASTASIS	5	0.72 (0.48-0.96)	0.74 (0.49-0.99)	0.75 (0.51-1.00)	0.89 (0.78-1.00)	-	-	-	***S***	-	-	-	-
GLIOBLASTOMA	7	0.91 (0.82-1.00)	0.81 (0.66-0.96)	0.84 (0.72-0.86)	0.88 (0.76-1.00)	***S***	***S***	***S***	***S***	-	-	-	-
ASTROCYTOMA WHO GRADE III	9	0.66 (0.44-0.88)	0.71 (0.54-0.88)	0.69 (0.51-0.87)	0.78 (0.60-0.96)	-	-	-	***S***	-	-	-	-
OLIGOASTROCYTOMA WHO GRADE III	2	0.50 (0.08-0.92)	0.50 (0.08-0.92)	0.49 (0.08-0.90)	0.50 (0.08-0.92)	-	-	-	-	-	-	-	-
ABSCESS	2	1.00 (1.00-1.00)	0.74 (0.28-1.20)	1.00 (1.00-1.00)	1.00 (1.00-1.00)	***S***	-	***S***	***S***	-	-	-	-
Superclasses													
LOW-GRADE MENINGIOMAS	8	0.98 (0.94-1.02)	0.98 (0.95-1.01)	1.00 (1.00-1.00)	1.00 (1.00-1.00)	***S***	***S***	***S***	***S***	-	-	-	-
GLIOBLASTOMAS AND METASTASES	12	0.83 (0.70-0.96)	0.88 (0.77-0.99)	0.89 (0.79-0.99)	0.93 (0.85-1.01)	***S***	***S***	***S***	***S***	-	-	***S***	-
WHO GRADE IV TUMORS	14	0.85 (0.73-0.97)	0.85 (0.73-0.97)	0.87 (0.76-0.98)	0.93 (0.85-1.01)	***S***	***S***	***S***	***S***	-	-	***S***	-
GLIALWHO GRADE III	12	0.70 (0.52-0.89)	0.73 (0.57-0.89)	0.77 (0.59-0.91)	0.84 (0.70-0.98)	***S***	***S***	***S***	***S***	-	-	-	-
GLIAL WHO GRADES II-III	13	0.81 (0.65-0.97)	0.90 (0.78-1.02)	0.92 (0.80-1.04)	0.93 (0.81-1.05)	***S***	***S***	***S***	***S***	-	-	***S***	-
GLIAL WHO GRADES III-IV	19	0.89 (0.79-0.99)	0.63 (0.45-0.81)	0.70 (0.70-1.00)	0.91 (0.82-1.00)	***S***	-	***S***	***S***	***S***	***S***	-	-
OLIGODENDROGLIOMAS AND OLIGOASTROCYTOMAS	3	0.84 (0.68-1.00)	0.87 (0.71-1.03)	0.85 (0.79-0.99)	0.91 (0.82-1.00)	-	***S***	-	***S***	-	-	-	-
GLIAL	20	0.95 (0.87-1.01)	0.84 (0.73-0.97)	0.89 (0.79-0.99)	0.97 (0.91-1.01)	***S***	***S***	***S***	***S***	***S***	-	-	-
AGGRESSIVE (WHO GRADES III-IV)	26	0.94 (0.87-1.01)	0.77 (0.59-0.94)	0.81 (0.66-0.96)	0.95 (0.88-1.02)	***S***	***S***	***S***	***S***	***S***	-	-	-
TUMOUR	37	1.00 (1.00-1.00)	0.92 (0.78-1.06)	0.90 (0.73-1.07)	1.00 (1.00-1.00)	***S***	***S***	***S***	***S***	-	-	-	-

Table 4. Results of the 4 evaluation phases for the glial and aggressive superclasses. a) Results for direct questions, b) Comparison between pooling ratings or by direct question. AUC: Area Under the Curve, CI: 95% Confidence Interval, vs.: versus, CD: Chance Diagonal, S: Significant differences found ($P < 0.05$). -: No significant differences found ($P \geq 0.05$).

a)

direct questions	n	*AUC (CI) at PHASE* MRI-E	MRS-E	MRS-R	MRI-R	MRI-E vs. CD	MRS-E vs. CD	MRS-R vs. CD	MRI-R vs. CD	MRI-E vs. MRS-E	MRI-E vs. MRS-R	MRI-E vs. MRI-R	MRS-E vs. MRS-R
GLIAL	20	0.90 (0.81-1.00)	0.81 (0.67-0.94)	0.80 (0.66-0.93)	0.87 (0.76-0.98)	*S*	*S*	*S*	*S*	-	-	-	-
AGGRESSIVE(WHO GRADES III-IV)	26	0.88 (0.76-1.01)	0.82 (0.66-0.98)	0.80 (0.65-0.96)	0.94 (0.86-1.01)	*S*	*S*	*S*	*S*	-	-	-	-

b)

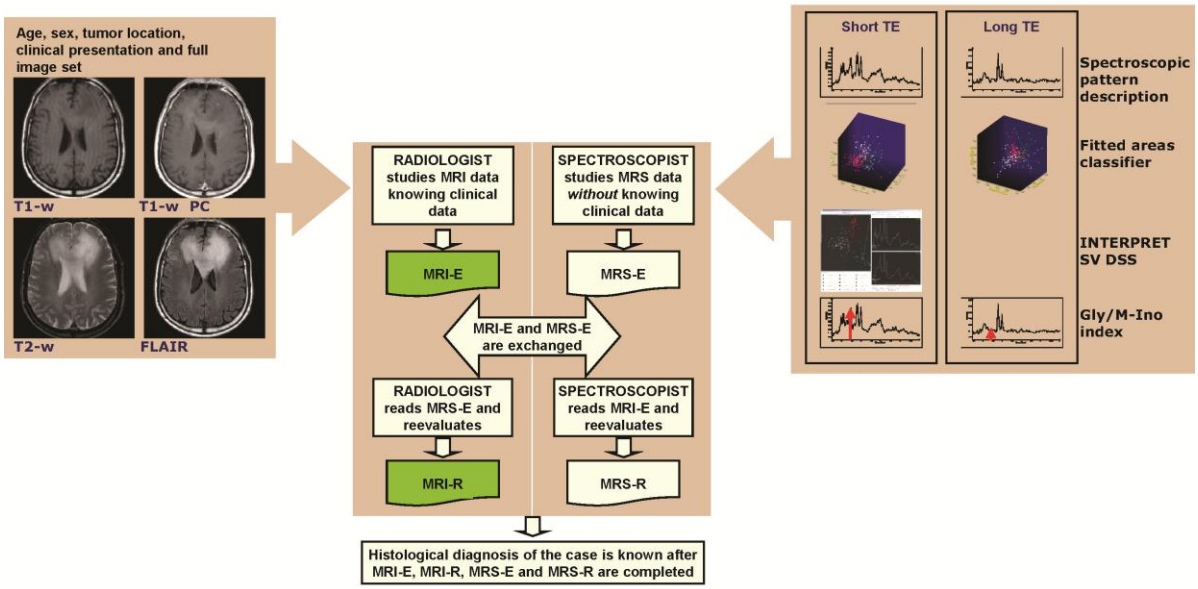
	n	*P for pooled ratings vs. direct questions* MRI-E	MRS-E	MRS-R	MRI-R
GLIAL	20	-	-	-	-
AGGRESSIVE(WHO GRADES III-IV)	26	-	-	-	-

Table 4. Results of the four methods for performing the MRS-E, in comparison with MRI-E and global MRS-E alone. AUC: Area Under the Curve, CI: 95% Confidence Interval, vs.: versus, CD: Chance Diagonal, P: Probability, S: Significant differences found ($P < 0.05$). -: No significant differences found ($P \geq 0.05$). #: Not calculated.

Consensus opinion of expert spectroscopists	***n***	***AUC (CI)***	***vs. CD***	***vs. MRI-E***	***vs. MRS-E***
Classes					
MENINGIOMA WHO GRADE I	6	0.95 (0.89-1.01)	*S*	-	-
GLIOBLASTOMA	7	0.53 (0.28-0.78)	-	#	#
METASTASIS	5	0.49 (0.22-0.75)	-	#	#
ASTROCYTOMA WHO GRADE III	9	0.68 (0.45-0.90)	-	#	#
ABSCESS	2	0.75 (0.30-1.20)	-	#	#
Superclasses					
LOW-GRADE MENINGIOMAS	8	0.89 (0.74-1.04)	*S*	-	-
GLIOBLASTOMAS AND METASTASES	12	0.55 (0.35-0.75)	-	#	#
WHO GRADE IV TUMORS	14	0.57 (0.38-0.77)	-	#	#
GLIAL WHO GRADE III	12	0.61 (0.41-0.82)	-	#	#
GLIAL WHO GRADES II-III	13	0.74 (0.55-0.93)	*S*	-	*S*
GLIAL WHO GRADES III-IV	24	0.60 (0.43-0.78)	-	#	#
OLIGODENDROGLIOMAS AND OLIGOASTROCYTOMAS	3	0.64 (0.25-1.02)	-	#	#
GLIAL	20	0.70 (0.53-0.87)	*S*	*S*	*S*
AGGRESSIVE (WHO GRADES III-IV)	26	0.57 (0.39-0.75)	-	#	#
TUMOUR	25	0.74 (0.28-1.20)	-	#	#
~~***Direct questions***~~					
~~*GLIAL*~~	~~20~~	~~0.65 (0.48-0.82)~~	~~-~~	~~#~~	~~#~~
~~*AGGRESSIVE (WHO GRADES III-IV)*~~	~~26~~	~~0.70 (0.55-0.86)~~	~~*S*~~	~~*S*~~	~~-~~
INTERPRET-DSS	***n***	***AUC (CI)***	***vs. CD***	***vs. MRI-E***	***vs. MRS-E***
Classes					
MENINGIOMA WHO GRADE I	6	0.99 (0.97-1.02)	*S*	-	-
MENINGIOMA WHO GRADE II	2	0.72 (0.28-1.17)	-	#	#
METASTASIS	5	0.75 (0.50-099)	-	#	#
GLIOBLASTOMA	7	0.77 (0.60-0.95)	*S*	-	-
ASTROCYTOMA WHO GRADE III	9	0.87 (0.72-1.02)	*S*	*S*	*S*
Superclasses					
LOW-GRADE MENINGIOMAS	8	0.98 (0.93-1.02)	*S*	-	-
GLIOBLASTOMAS AND METASTASES	12	0.89 (0.77-1.00)	*S*	-	-
WHO GRADE IV TUMORS	14	0.85 (0.73-0.97)	*S*	-	-
GLIAL WHO GRADE III	12	0.87 (0.74-1.00)	*S*	-	*S*
GLIAL WHO GRADES II-III	13	0.91 (0.79-1.04)	*S*	-	-
GLIAL WHO GRADES III-IV	24	0.87 (0.73-1.00)	*S*	-	*S*
OLIGODENDROGLIOMAS AND OLIGOASTROCYTOMAS	3	0.88 (0.76-1.00)	*S*	-	-
GLIAL	20	0.88 (0.77-1.00)	*S*	-	-
AGGRESSIVE (WHO GRADES III-IV)	26	0.78 (0.59-0.96)	*S*	-	-
TUMOUR	36	0.74 (0.29-1.20)	-	#	#
~~***Direct questions***~~					
~~*GLIAL*~~	~~20~~	~~0.84 (0.71-0.96)~~	~~*S*~~	~~-~~	~~-~~

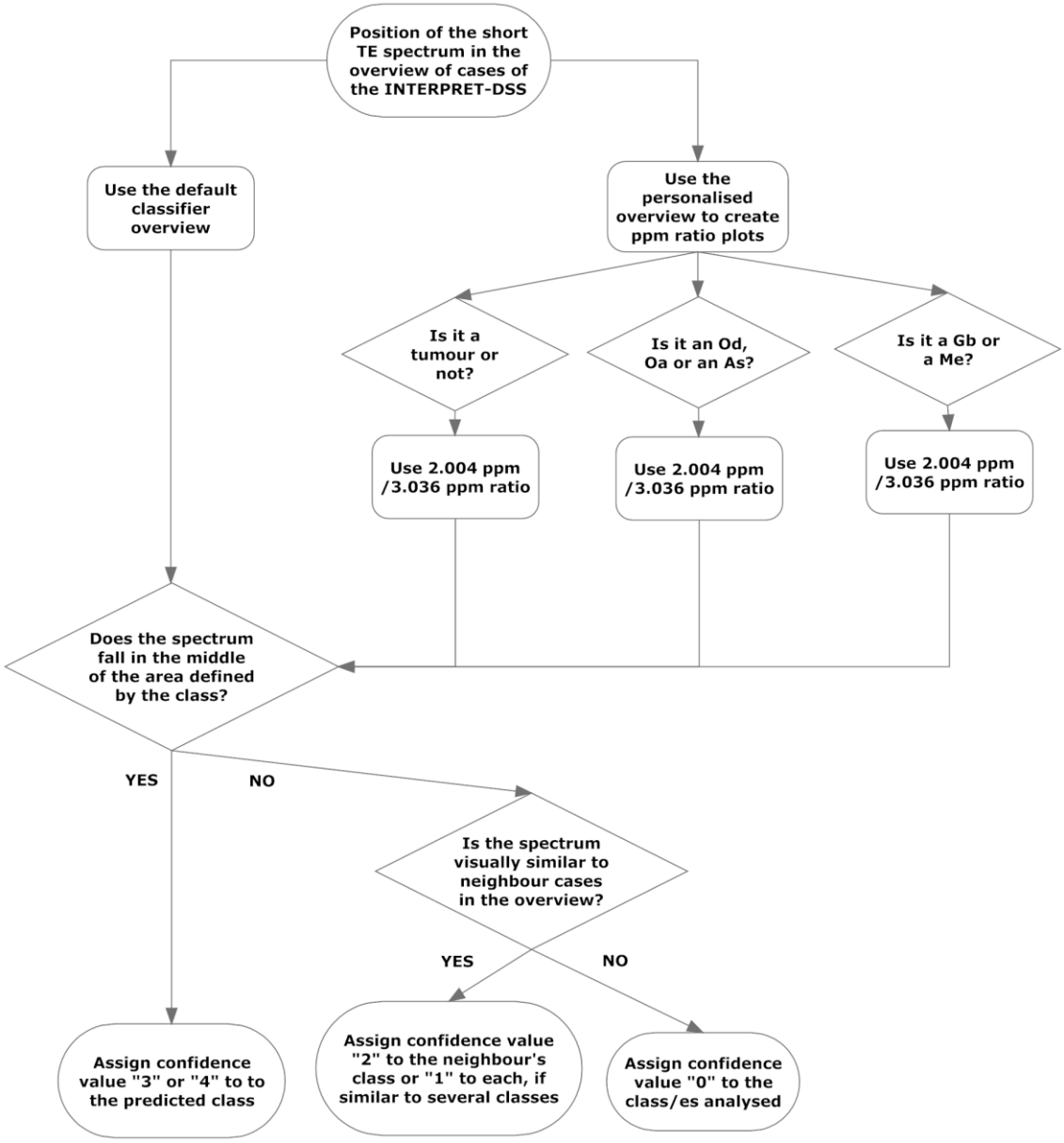
~~*AGGRESSIVE (WHO GRADES III-IV)*~~	~~25~~	~~0.77 (0.59-0.96)~~	~~*S*~~	~~-~~	~~-~~
Area classifiers	***n***	***AUC (CI)***	***vs. CD***	***vs. MRI-E***	***vs. MRS-E***
Classes					
MENINGIOMA WHO GRADE I	6	0.97 (0.92-1.02)	*S*	-	-
ASTROCYTOMA WHO GRADE III	9	0.85 (0.70-1.00)	*S*	-	-
GLIOBLASTOMAS	7	0.78 (0.59-0.95)	*S*	-	-
METASTASES	5	0.75 (0.51-0.99)	-	#	#
Superclasses					
GLIOBLASTOMAS and METASTASES	16	0.61 (0.43-0.79)	-	#	#
~~***Direct questions***~~					
~~*GLIAL*~~	~~20~~	~~0.85 (0.73-0.97)~~	~~*S*~~	~~-~~	~~-~~
~~*AGGRESSIVE (WHO GRADES III-IV)*~~	~~26~~	~~0.79 (0.61-0.96)~~	~~*S*~~	~~-~~	~~-~~
MI/Gly index	***n***	***AUC (CI)***	***vs. CD***	***vs. MRI-E***	***vs. MRS-E***
Classes					
GLIOBLASTOMA	7	0.63 (0.38-0.88)	-	#	#
ASTROCYTOMA III	9	0.68 (0.46-0.90)	-	#	#
~~***Direct questions***~~					
~~*AGGRESSIVE (WHO GRADES III-IV)*~~	~~26~~	~~0.67 (0.49-0.84)~~	~~-~~	~~#~~	~~#~~

Figure 1. Summary of protocol and main results.



PC, post-contrast administration.

Figure 2. Decision tree for using the INTERPRET-DSS.



Od, oligodendroglioma WHO grade II; Oa, oligoastrocytoma WHO grade II; As, Astrocytoma WHO grade II; Gb, Glioblastoma; Me, Metastasis.

Figure 3. Decision tree for assigning confidence values to the area-based classifier results.

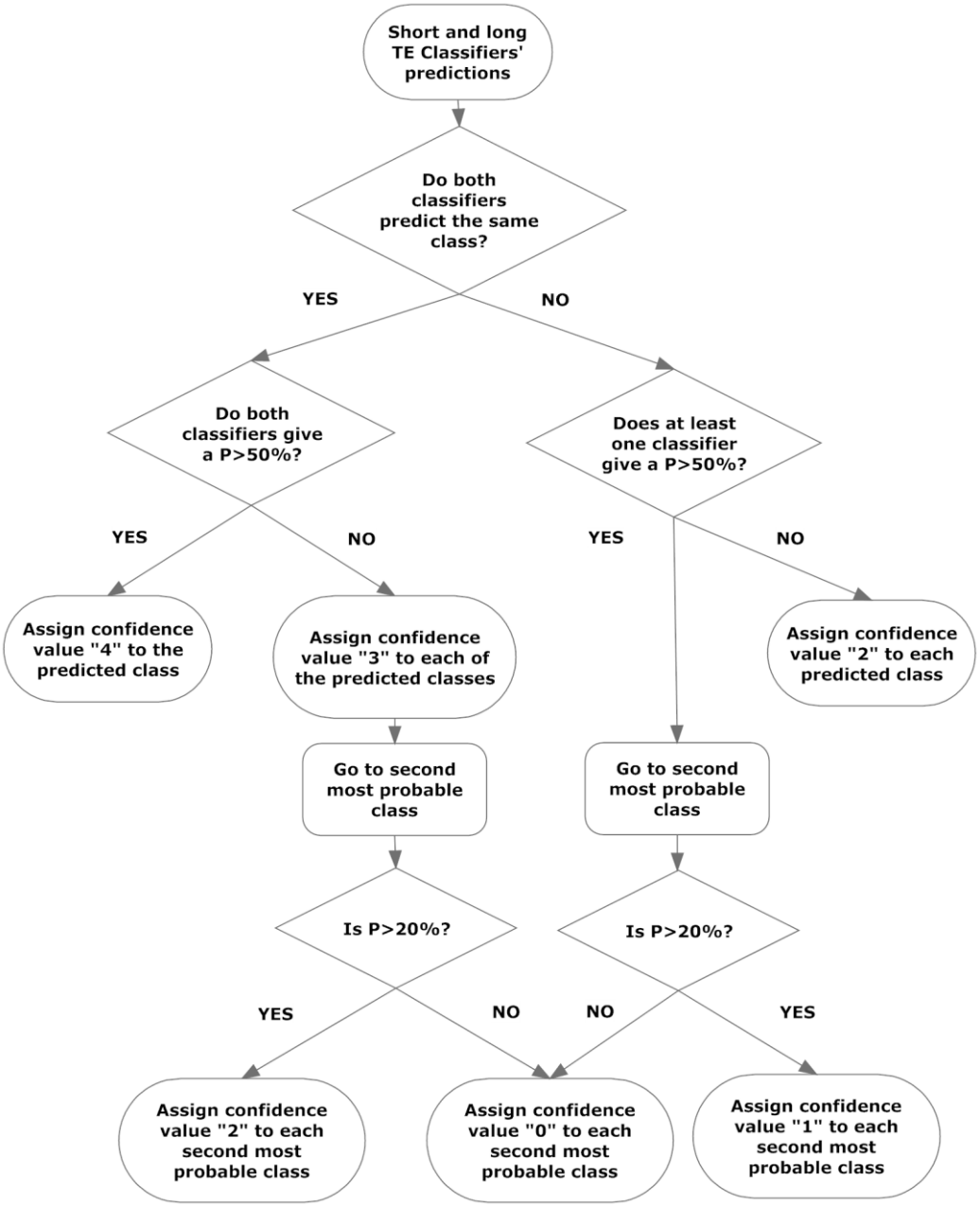


Figure 4. Classes and superclasses defined for analysis.

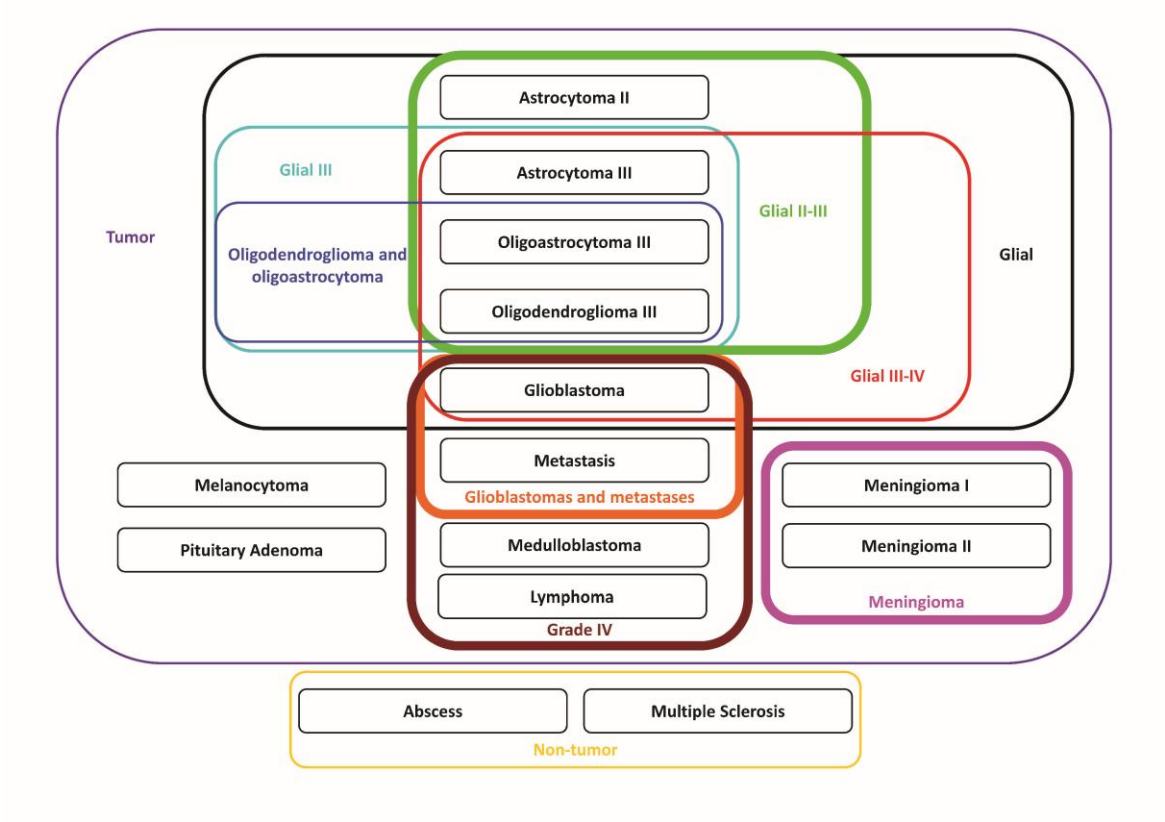
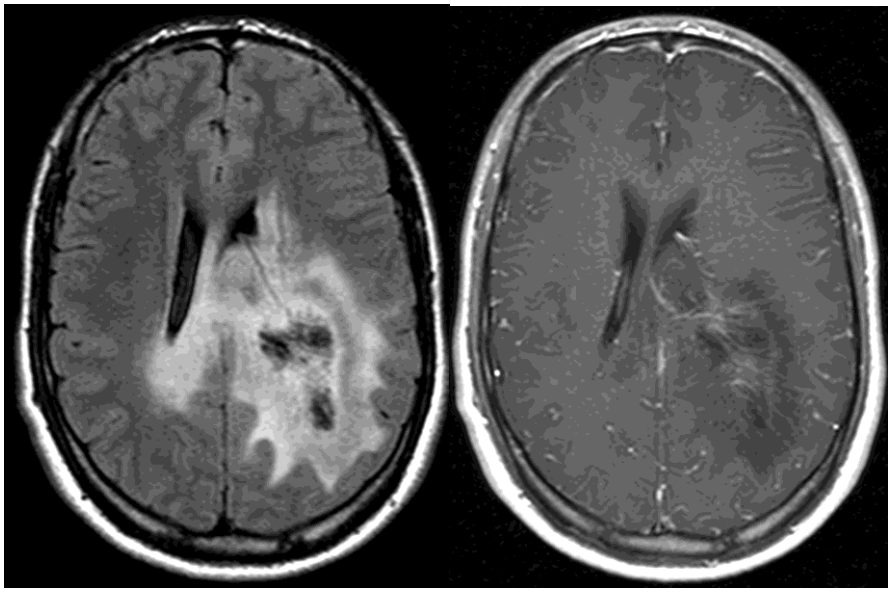
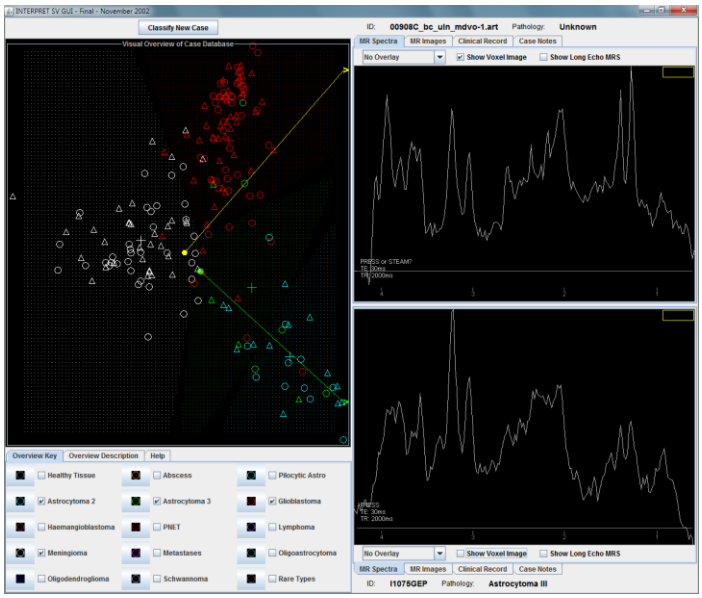


Figure 5. MRI-E, the INTERPRET DSS and MRS-E results for Case 4, an Astrocytoma WHO grade III. MRI-E, the INTERPRET DSS and MRS-E results for Case 4, an Astrocytoma WHO grade III. a: Axial FLAIR-weighted MR images in a 23 year-old male showed a heterogeneous mass in the left hemisphere that extends into the corpus callosum and the contralateral white matter. b: Axial T1-weighted images after contrast administration showing irregular linear-shaped enhancing areas within the mass. c: Use of the INTERPRET DSS. Yellow dot, unknown case; red dot, glioblastoma, white dot, meningioma; green dot, astrocytoma WHO grade III.

a b



c


Classify New Case
Visual Overview of Case Database
ID: 00908C_bc_uin_mdvo-1.art
Pathology: Unknown
MR Spectra
MR Images
Clinical Record
Case Notes
No Overlay
Show Voxel Image
Show Long Echo MRS
Overview Key
Overview Description
Help
Healthy Tissue
Abscess
Pilocytic Astro
Astrocytoma 2
Astrocytoma 3
Glioblastoma
Haemangioblastoma
PNET
Lymphoma
Meningioma
Metastases
Oligoastrocytoma
Oligodendroglioma
Schwannoma
Rare Types
ID: I1075QEP
Pathology: Astrocytoma III

References

1. Julia-Sape M, Acosta D, Majos C, Moreno-Torres A, Wesseling P, Acebes JJ, Griffiths JR, Arus C. Comparison between neuroimaging classifications and histopathological diagnoses using an international multicenter brain tumor magnetic resonance imaging database. J Neurosurg 2006;105(1):6-14.
2. Favre J, Taha JM, Burchiel KJ. An analysis of the respective risks of hematoma formation in 361 consecutive morphological and functional stereotactic procedures. Neurosurgery 2002;50(1):48-56; discussion 56-47.
3. Hall WA. The safety and efficacy of stereotactic biopsy for intracranial lesions. Cancer 1998;82(9):1749-1755.
4. Field M, Witham TF, Flickinger JC, Kondziolka D, Lunsford LD. Comprehensive assessment of hemorrhage risks and outcomes after stereotactic brain biopsy. J Neurosurg 2001;94(4):545-551.
5. Dammers R, Haitsma IK, Schouten JW, Kros JM, Avezaat CJ, Vincent AJ. Safety and efficacy of frameless and frame-based intracranial biopsy techniques. Acta Neurochir (Wien) 2008;150(1):23-29.
6. Langkowski JH, Wieland J, Bomsdorf H, Leibfritz D, Westphal M, Offermann W, Maas R. Pre-operative localized in vivo proton spectroscopy in cerebral tumors at 4.0 Tesla--first results. Magn Reson Imaging 1989;7(5):547-555.
7. Bruhn H, Frahm J, Gyngell ML, Merboldt KD, Hanicke W, Sauter R, Hamburger C. Noninvasive differentiation of tumors with use of localized H-1 MR spectroscopy in vivo: initial experience in patients with cerebral tumors. Radiology 1989;172(2):541-548.
8. (BCBS) BCBSA. Magnetic resonance spectroscopy for evaluation of suspected brain tumor. TEC Bull (Online). 2003/07/04 ed. Volume 20; 2003. p 23-26.
9. Lin AP, Tran TT, Ross BD. Impact of evidence-based medicine on magnetic resonance spectroscopy. NMR Biomed 2006;19(4):476-483.
10. Hollingworth W, Medina LS, Lenkinski RE, Shibata DK, Bernal B, Zurakowski D, Comstock B, Jarvik JG. A systematic literature review of magnetic resonance spectroscopy for the characterization of brain tumors. AJNR Am J Neuroradiol 2006;27(7):1404-1411.
11. Murphy M, Loosemore A, Clifton AG, Howe FA, Tate AR, Cudlip SA, Wilkins PR, Griffiths JR, Bell BA. The contribution of proton magnetic resonance spectroscopy (1HMRS) to clinical brain tumour diagnosis. Br J Neurosurg 2002;16(4):329-334.
12. Moller-Hartmann W, Herminghaus S, Krings T, Marquardt G, Lanfermann H, Pilatus U, Zanella FE. Clinical application of proton magnetic resonance spectroscopy in the diagnosis of intracranial mass lesions. Neuroradiology 2002;44(5):371-381.
13. Fellows GA, Wright AJ, Sibtain NA, Rich P, Opstad KS, McIntyre DJ, Bell BA, Griffiths JR, Howe FA. Combined use of neuroradiology and 1H-MR spectroscopy may provide an intervention limiting diagnosis of glioblastoma multiforme. J Magn Reson Imaging 2010;32(5):1038-1044.
14. Galanaud D, Nicoli F, Chinot O, Confort-Gouny S, Figarella-Branger D, Roche P, Fuentes S, Le Fur Y, Ranjeva JP, Cozzone PJ. Noninvasive diagnostic

assessment of brain tumors using combined in vivo MR imaging and spectroscopy. Magn Reson Med 2006;55(6):1236-1245.
15. Sardanelli F, Hunink MG, Gilbert FJ, Di Leo G, Krestin GP. Evidence-based radiology: why and how? Eur Radiol 2010;20(1):1-15.
16. Danielsen ER, Ross B. Magnetic resonance spectroscopy diagnosis of neurological diseases. New York: M. Dekker; 1999. xi, 327 p. p.
17. Herminghaus S, Pilatus U, Moller-Hartmann W, Raab P, Lanfermann H, Schlote W, Zanella FE. Increased choline levels coincide with enhanced proliferative activity of human neuroepithelial brain tumors. NMR Biomed 2002;15(6):385-392.
18. Negendank W. Studies of human tumors by MRS: a review. NMR Biomed 1992;5(5):303-324.
19. Howells SL, Maxwell RJ, Griffiths JR. Classification of tumour 1H NMR spectra by pattern recognition. NMR Biomed 1992;5(2):59-64.
20. Opstad KS, Ladroue C, Bell BA, Griffiths JR, Howe FA. Linear discriminant analysis of brain tumour (1)H MR spectra: a comparison of classification using whole spectra versus metabolite quantification. NMR Biomed 2007;20(8):763-770.
21. Devos A, Lukas L, Suykens JA, Vanhamme L, Tate AR, Howe FA, Majos C, Moreno-Torres A, van der Graaf M, Arus C, Van Huffel S. Classification of brain tumours using short echo time 1H MR spectra. J Magn Reson 2004;170(1):164-175.
22. Lukas L, Devos A, Suykens JA, Vanhamme L, Howe FA, Majos C, Moreno-Torres A, Van der Graaf M, Tate AR, Arus C, Van Huffel S. Brain tumor classification based on long echo proton MRS signals. Artif Intell Med 2004;31(1):73-89.
23. Garcia-Gomez JM, Luts J, Julia-Sape M, Krooshof P, Tortajada S, Robledo JV, Melssen W, Fuster-Garcia E, Olier I, Postma G, Monleon D, Moreno-Torres A, Pujol J, Candiota AP, Martinez-Bisbal MC, Suykens J, Buydens L, Celda B, Van Huffel S, Arus C, Robles M. Multiproject-multicenter evaluation of automatic brain tumor classification by magnetic resonance spectroscopy. Magn Reson Mater Phy 2009;22(1):5-18.
24. Garcia-Gomez JM, Tortajada S, Vidal C, Julia-Sape M, Luts J, Moreno-Torres A, Van Huffel S, Arus C, Robles M. The effect of combining two echo times in automatic brain tumor classification by MRS. NMR Biomed 2008;21(10):1112-1125.
25. Alusta P, Im I, Pearce BA, Beger RD, Kretzer RM, Buzatu DA, Wilkes JG. Improving proton MR spectroscopy of brain tissue for noninvasive diagnostics. J Magn Reson Imaging 2010;32(4):818-829.
26. Tate AR, Underwood J, Acosta DM, Julia-Sape M, Majos C, Moreno-Torres A, Howe FA, van der Graaf M, Lefournier V, Murphy MM, Loosemore A, Ladroue C, Wesseling P, Luc Bosson J, Cabanas ME, Simonetti AW, Gajewicz W, Calvar J, Capdevila A, Wilkins PR, Bell BA, Remy C, Heerschap A, Watson D, Griffiths JR, Arus C. Development of a decision support system for diagnosis and grading of brain tumours using in vivo magnetic resonance single voxel spectra. NMR Biomed 2006;19(4):411-434.
27. Perez-Ruiz A, Julia-Sape M, Mercadal G, Olier I, Majos C, Arus C. The INTERPRET Decision-Support System version 3.0 for evaluation of Magnetic

Resonance Spectroscopy data from human brain tumours and other abnormal brain masses. BMC Bioinformatics 2010;11(1):581.
28. Julia-Sape M, Acosta D, Mier M, Arus C, Watson D. A multi-centre, web-accessible and quality control-checked database of in vivo MR spectra of brain tumour patients. Magn Reson Mater Phy 2006;19(1):22-33.
29. Fuster-Garcia E, Navarro C, Vicente J, Tortajada S, García-Gómez J, Sáez C, Calvar J, Griffiths J, Julià-Sapé M, Howe F, Pujol J, Peet A, Heerschap A, Moreno-Torres À, Martínez-Bisbal M, Martínez-Granados B, Wesseling P, Semmler W, Capellades J, Majós C, Alberich-Bayarri À, Capdevila A, Monleón D, Martí-Bonmatí L, Arús C, Celda B, Robles M. Compatibility between 3T 1H SV-MRS data and automatic brain tumour diagnosis support systems based on databases of 1.5T 1H SV-MRS spectra. Magnetic Resonance Materials in Physics, Biology and Medicine 2011:1-8.
30. Majos C, Julia-Sape M, Alonso J, Serrallonga M, Aguilera C, Acebes JJ, Arus C, Gili J. Brain tumor classification by proton MR spectroscopy: comparison of diagnostic accuracy at short and long TE. AJNR Am J Neuroradiol 2004;25(10):1696-1704.
31. Naressi A, Couturier C, Castang I, de Beer R, Graveron-Demilly D. Java-based graphical user interface for MRUI, a software package for quantitation of in vivo/medical magnetic resonance spectroscopy signals. Comput Biol Med 2001;31(4):269-286.
32. Kleihues P, Cavenee WK. Pathology and genetics of tumours of the nervous system. Lyon: IARC Press; 2000. 314 p. p.
33. Julià-Sapé M, Majós C, Arús C. Diagnosis and Staging of Brain Tumours: Magnetic Resonance Single Voxel Spectra. In: Hayat MA, editor. Methods of Cancer Diagnosis, Therapy, and Prognosis. Volume 8, Methods of Cancer Diagnosis, Therapy and Prognosis: Springer Netherlands; 2010. p 227-243.
34. van der Graaf M, Julia-Sape M, Howe FA, Ziegler A, Majos C, Moreno-Torres A, Rijpkema M, Acosta D, Opstad KS, van der Meulen YM, Arus C, Heerschap A. MRS quality assessment in a multicentre study on MRS-based classification of brain tumours. NMR Biomed 2008;21(2):148-158.
35. Candiota A, Majós M, Julia-Sapé M, Cabañas M, Mercadal G, Acebes J, Moreno-Torres A, Griffiths R, Arús C. M-Inositol and glycine content measurement for grading astrocytic tumours with in vivo MRS. Magn Reson Mater Phy 2005;18 (1):193.
36. Grand S, Passaro G, Ziegler A, Esteve F, Boujet C, Hoffmann D, Rubin C, Segebarth C, Decorps M, Le Bas JF, Remy C. Necrotic tumor versus brain abscess: importance of amino acids detected at 1H MR spectroscopy--initial results. Radiology 1999;213(3):785-793.
37. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143(1):29-36.
38. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148(3):839-843.
39. Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters 2006;27(8):861-874.

40. Ishimaru H, Morikawa M, Iwanaga S, Kaminogo M, Ochi M, Hayashi K. Differentiation between high-grade glioma and metastatic brain tumor using single-voxel proton MR spectroscopy. Eur Radiol 2001;11(9):1784-1791.
41. Opstad KS, Murphy MM, Wilkins PR, Bell BA, Griffiths JR, Howe FA. Differentiation of metastases from high-grade gliomas using short echo time 1H spectroscopy. J Magn Reson Imaging 2004;20(2):187-192.
42. Law M, Cha S, Knopp EA, Johnson G, Arnett J, Litt AW. High-grade gliomas and solitary metastases: differentiation by using perfusion and proton spectroscopic MR imaging. Radiology 2002;222(3):715-721.
43. Burtscher IM, Skagerberg G, Geijer B, Englund E, Stahlberg F, Holtas S. Proton MR spectroscopy and preoperative diagnostic accuracy: an evaluation of intracranial mass lesions characterized by stereotactic biopsy findings. AJNR Am J Neuroradiol 2000;21(1):84-93.
44. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Radiology 2003;226(1):24-28.

List of figure legends

Figure 1. Summary of protocol and main results.

Figure 2. Decision tree for using the INTERPRET-DSS.

Figure 3. Decision tree for assigning confidence values to the area-based classifier results.

Figure 4. Classes and superclasses defined for analysis.

Figure 5. MRI-E, the INTERPRET DSS and MRS-E results for Case 4, an Astrocytoma WHO grade III.