

Document downloaded from:

<http://hdl.handle.net/10251/39576>

This paper must be cited as:

Fuster García, E.; Navarro., C.; Vicente Robledo, J.; Tortajada Velert, S.; García Gómez, JM.; Sáez Silvestre, C.; Calvar ., J.... (2011). 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*. 24(1):35-42. doi:10.1007/s10334-010-0241-8.



The final publication is available at

<http://link.springer.com/article/10.1007/s10334-010-0241-8>

Copyright Springer Verlag (Germany)

Compatibility between 3T ¹H SV-MRS data and automatic brain tumour diagnosis support systems based on databases of 1.5T ¹H SV-MRS spectra

Elies Fuster-Garcia^{1,2}, Clara Navarro³, Javier Vicente², Salvador Tortajada², Juan M. García-Gómez², Carlos Sáez², Jorge Calvar⁴, John Griffiths⁵, Margarida Julià-Sapé^{6,7,8}, Franklyn A. Howe⁹, Jesús Pujol^{6,10}, Andrew C. Peet¹¹, Arend Heerschap¹², Àngel Moreno-Torres^{13,6}, M. C. Martínez-Bisbal^{6,14}, Beatriz Martínez-Granados¹⁴, Pieter Wesseling¹², Wolfhard Semmler¹⁵, Jaume Capellades¹⁶, Carles Majós^{6,17}, Àngel Alberich-Bayarri³, Antoni Capdevila¹³, Daniel Monleon¹⁸, Luis Martí-Bonmatí³, Carles Arús^{6,7,8}, Bernardo Celda¹⁴, Montserrat Robles²

Abstract

Object This study demonstrates that 3T SV-MRS data can be used with the currently available automatic brain tumour diagnostic classifiers, which were trained on databases of 1.5T spectra. This will allow the existing large databases of 1.5T MRS data to be used for diagnostic classification of 3T spectra, and perhaps also the combination of 1.5T and 3T databases.

Materials and Methods Brain tumour classifiers trained with 154 1.5T spectra to discriminate among high grade malignant tumours and common grade II glial tumours were evaluated with a subsequently-acquired set of 155 1.5T and 28 3T spectra. A similarity study between spectra and main brain tumour metabolite ratios for both field strengths (1.5T and 3T) was also performed.

Results Our results showed that classifiers trained with 1.5T samples had similar accuracy for both test datasets (0.87 ± 0.03 for 1.5T and 0.88 ± 0.03 for 3.0T). Moreover non-significant differences were observed with most metabolite ratios and spectral patterns.

Conclusion These results encourage the use of existing classifiers based on 1.5T datasets for diagnosis with 3T ¹H SV-MRS. The large 1.5T databases compiled throughout many years and the prediction models based on 1.5T acquisitions can therefore continue to be used with data from the new 3T instruments

Keywords

Brain Tumours — Magnetic Resonance Spectroscopy — Clinical Decision Support Systems

¹ Universitat Internacional Valenciana, Spain

² IBIME, Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas (ITACA), Universidad Politécnica de Valencia, Spain

³ Quirón Valencia Hospital, Radiology Department, Spain

⁴ Institute for Neurological Research (FLENI), Buenos Aires, Argentina

⁵ CR UK Cambridge Research Institute Cambridge, United Kingdom

⁶ Networking Research Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Spain

⁷ Universitat Autònoma de Barcelona, Department of Biochemistry and Molecular Biology, Cerdanyola del Vallès, Spain

⁸ Institut de Biotecnologia i Biomedicina (IBB), Universitat Autònoma de Barcelona (UAB), Cerdanyola del Vallès, Spain

⁹ St. George's, University of London, Cardiac and Vascular Sciences, London, United Kingdom

¹⁰ CRC Corporació Sanitària, Institut d'Alta Tecnologia-PRBB, Barcelona, Spain

¹¹ University of Birmingham, Academic Department of Paediatrics and Child Health, Birmingham, United Kingdom

¹² Radboud University, Nijmegen Medical Centre, Radiology and Pathology, Nijmegen, Netherlands

¹³ Research Department, Centre Diagnòstic Pedralbes, Esplugues de Llobregat, Spain

¹⁴ University of Valencia, Physical-Chemistry, Burjassot, Valencia, Spain

¹⁵ German Cancer Research Center (DKFZ), Department of Medical Physics in Radiology, Germany

¹⁶ Hospital Universitari Germans Trias i Pujol, Badalona, Spain

¹⁷ Hospital Universitari de Bellvitge, Institut de Diagnòstic per la Imatge, L'Hospitalet de Llobregat, Spain

¹⁸ Fundación Investigación Hospital Clínico Valencia / INCLIVA, Valencia, Spain

Contents

1	Introduction	2
2	Materials and methods	3
2.1	<i>In vivo</i> ¹ H SV-MRS datasets	3
2.2	MRS processing	3
2.3	Feature extraction method	3
2.4	Classifiers	4
2.5	Performance Measures	4
2.6	Statistical Analysis	5
3	Results	5
4	Discussion	5
5	Conclusions	7
6	Acknowledgements	7
	References	7

Abbreviations

¹H SV-MRS	single voxel proton magnetic resonance spectroscopy
ACC	accuracy
ANN	artificial neural networks
CDSS	clinical decision support systems
CG2G	common grade II glial
G	geometric mean of recalls
HGM	high grade malignant
KNN	k-nearest neighbors
LDA	linear discriminant analysis
PI	peak integration
SNR	signal-to-noise ratio
SW	stepwise
TE	echo time

1. Introduction

The current development of automatic brain tumour classifiers based on single voxel proton magnetic resonance spectroscopy (¹H SV-MRS) data has provided promising results for brain tumour diagnostic support [1, 2, 3, 4]. A growing number of studies and applications have been presented in the last few years showing the ability of MRS to distinguish among different brain tumour tissue types [5, 6, 7, 8, 9, 10, 11, 12]. These systems are mostly based on the pattern recognition approach, where classification models have been inferred from experimental data, after the extraction of relevant features [13, 14, 15].

The learning procedures commonly used in pattern recognition assume that samples are independent and identically distributed; therefore, these classifiers are expected to be useful when classifying spectra acquired in similar configurations to those in the training data. This assumption represents a challenge when new spectra are acquired with an evolving technology, such as changing from 1.5T to 3T MR scanners. 3T scanners are becoming widely available in the clinical environment, complementing the more common 1.5T scanners. Their increased magnetic field improves signal-to-noise ratio (SNR), and spectral resolution: the latter is particularly important for short echo time (TE) spectra [16, 17] as fine structure in the Gln/Glu region of the spectrum downfield from NAA is better resolved, and resonances downfield from water are better visualized [17]. This better resolution of overlapping signals from coupled spin systems also improves metabolic characterization, thus enhancing the diagnostic abilities of MRS

Despite these advantages, it would take many years to develop databases of 3T brain tumour spectra comparable to those currently available at 1.5T, so there is a strong incentive to use 1.5T-based classifiers to characterize 3T spectra. However, the currently available 1.5T based classifiers have not been validated on 3T data and it is not yet known whether we can expect a decrease in their level of performance due to differences in the overall spectral patterns. Such differences may arise from a variety of factors, both biophysical and instrumental: differences due to coupling or T2 relaxation times; and artefacts arising from water residuals or chemical shift displacement across the localization voxel. The study of Baker et al. [17] showed subtle differences between the spectra obtained at the two field strengths, including better-resolved peaks of the NAA amide and glutamate/glutamine region at 3T compared to 1.5T and a peak at 3.3 ppm clearly observable in some subjects at 1.5T, which was less prominent at 3T.

Some authors have suggested the application of established 1.5T metabolite ratios for the evaluation of brain tumours at 3T [16]. Additionally, Roser et al. [18] concluded that the change from 2T to 1.5T had no measurable deleterious effect on multidimensional metabolic classification for assignment of glial brain tumours.

To analyze this behavior, we have performed a study of the performance of 1.5T based classifiers when tested with 3T spectra. The main goal of our study was to test the compatibility of 1.5 and 3T data when they are used on automatic brain tumour clinical decision support systems (CDSS) [19, 20]. To achieve that goal, a standard spectrum-processing protocol has been applied, and the performance of different feature extraction methods and classification algorithms has been analyzed. Classifiers already included in a CDSS [21, 22] have been tested on new 1.5T and 3T acquisitions, and their performance has been evaluated on the well-known test problem [23] to discriminate tumours high grade malignant (HGM) - glioblastoma and metastases -, from common grade II glial (CG2G)

- astrocytomas grade II, oligodendrogliomas and oligoastrocytomas –. A hypothesis testing of the accuracy of each combination of feature extraction and classification methods was performed to evaluate the compatibility between 3T data and the existing classifiers based on 1.5T data. Finally, a study on the differences between spectra obtained at both field strengths was performed, analyzing the differences between the main brain metabolite ratios and spectra shapes.

2. Materials and methods

2.1 *In vivo* ¹H SV-MRS datasets

Three datasets were used in our study. The first one is a training 1.5T ¹H SV-MRS dataset accrued during the INTERPRET EU project [5]. The second dataset is part of the 1.5T dataset compiled in the eTUMOUR EU project [24], and was used as an independent test set to evaluate the performance of the classifiers. The third dataset is a new 3T dataset used to evaluate the performance of the 1.5T classifiers on 3T cases. All the datasets were obtained using magnetic resonance (MR) scanners of three major manufacturers (Philips, General Electric and Siemens) in ten international centers. The number of cases used in each dataset is shown in Table 1.

The 1.5T training dataset included 154 1.5T short TE SV-MRS spectra. The acquisition protocols included PRESS or STEAM sequences with the following spectral parameters: Repetition Time (TR) of 1600-2020ms, echo time (TE) of 20 or 30-32ms, spectral width (SW) of 1000 - 2500Hz, 512, 1024, or 2048 data-points, as described in previous studies [4]. All these cases were validated using a standard quality control protocol carried out by the INTERPRET Clinical Data Validation Committee and expert spectroscopists [5, 25], and all had a histopathological diagnosis.

The 1.5T test dataset included 155 1.5T short TE SV-MRS spectra. These spectra were validated by an expert spectroscopist panel, and the histopathological diagnosis of these cases was also available. The acquisition protocols included PRESS and STEAM with spectral parameters: TR of 1500-2000ms, TE of 30 – 31ms, SW of 500-2500Hz, 512, 1024, or 2048 data-points.

The 3T test dataset included 37 spectra and came from two different sources. The first 21 spectra were obtained in the eTUMOUR project, including 4 CG2G tumours and 17 HGM tumours. Their histopathological diagnoses were also available. The scanner used was GE Signa 3T. The acquisition parameters included PRESS sequences with spectral parameters: TR of 2000 - 5000ms, TE of 30ms, SW of 1000Hz, 2048 data-points. The remaining 16 spectra were acquired at the *Hospital Quirón of Valencia* on a Philips scanner. There were 11 histopathology proven HGM tumours and 5 CG2G tumours in which the diagnosis was made on clinical grounds, radiological appearance, and follow up. The acquisition protocols included a PRESS sequence with spectral parameters: TR of 1800-2000ms, TE of 32 ms, SW of 2000Hz, 1024 data-points. The TE's were optimized for a satisfactory SNR without losing any metabolite resonances or showing coupling variations.

Table 1. Number of 1H SV-MRS spectra in each dataset per tumour type. High Grade Malignant (HGM), consisting of glioblastomas and metastases, and common grade II glial (CG2G) consisting of astrocytomas grade II, oligodendrogliomas and oligoastrocytomas.

Dataset	CG2G	HGM	Total
1.5T Training	34	120	154
1.5T Test	53	102	155
3T Test	9	28	37

In Figure 2, spectra sample of the two tumour types included in the study are presented from 1.5T and 3T MR scanners.

2.2 MRS processing

A common MRS processing pipeline, previously used in the INTERPRET and eTUMOUR EC projects, and included in the CDSS software they developed [22, 26, 24], was applied in the present study. Each spectrum was semi-automatically pre-processed using a pipeline consisting of 1) eddy current correction applied to the water-suppressed free induction decay of each case using the Klose algorithm [27]; 2) zero and first order manual phase correction; 3) residual water resonance suppression by the Hankel-Lanczos singular value decomposition time-domain selective filtering using 10 singular values and a water region of [4.33, 5.07] ppm; 4) an apodization with a Lorentzian function of 1Hz of damping; 5) zero filling, to increase the number of points of the low resolution spectra to the maximum number used in the acquisition protocols (2048); 6) baseline offset subtraction, estimated as the mean value of the regions [11, 9] and [-2,-1] ppm; 7) normalization of spectra to the Euclidean norm using the regions [-2.7, 4.33] [5.07, 7.1] ppm; 8) additional frequency alignment check of the spectrum by referencing the ppm-axis to the total creatine at 3.03ppm or to the choline-containing compounds at 3.21 ppm or the mobile lipids at 1.29ppm, depending on the SNR and the tumor pattern; and finally 9) reduction of the number of points of the spectra, using 512 points for the defined region of [-2.7, 7.1] ppm. No corrections for T1 or T2 relaxation effects were made to the spectra prior to the pattern recognition analysis.

The software used to pre-process the spectra was jMRUI 3.0 [28] in batch mode (steps 1-5) and jDMS [26] (steps 6-9).

2.3 Feature extraction method

One of the major problems in spectral classification arises from the number of variables that represent the full region of interest (190 data points in this case). The use of a large number of variables in classification problems generally overfits the training sample and generalizes poorly to new samples. To overcome this problem, the variables are usually transformed

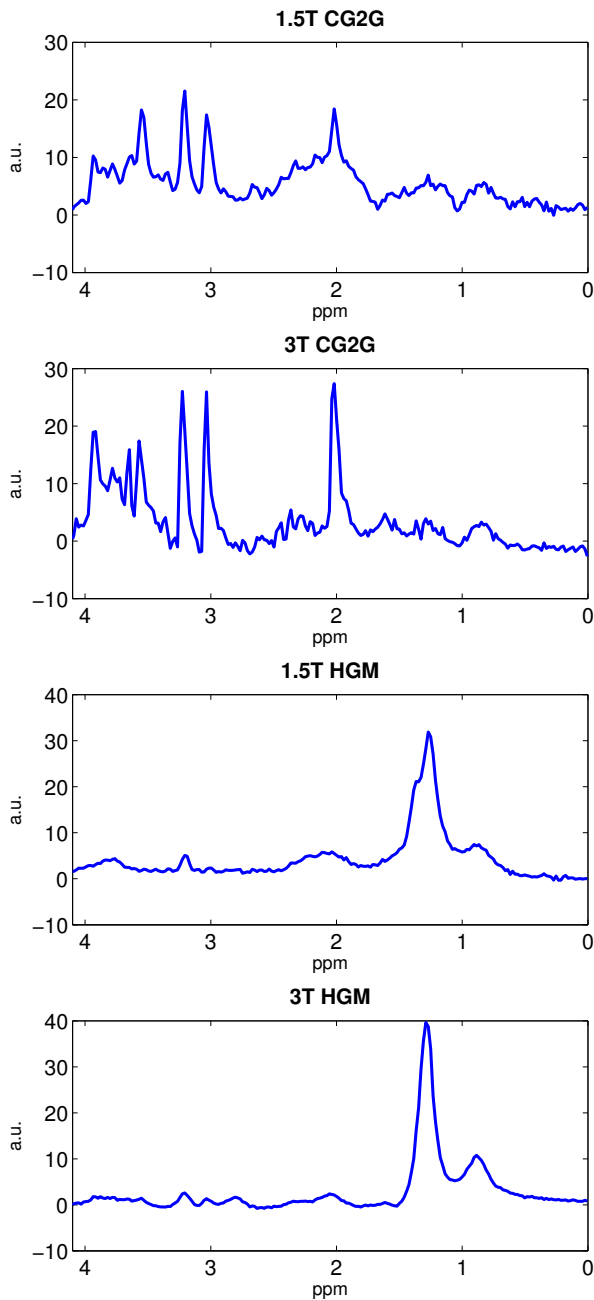


Figure 1. Example short TE spectra from 1.5 and 3T MR scanners of the two tumour types used in this study. Spectra have been processed as described in section 2.2.

into a reduced representation set of features that maintains the most representative information [29].

Two different feature extraction methods were used in this study: stepwise (SW) and peak integration (PI). SW is a sub-

optimal greedy hill climbing approach [30]. This algorithm was applied with the Mahalanobis squared distance, for selecting relevant subsets of features based on the performance measure of the training classification. PI computes the value of the area under the peaks of the most relevant metabolites as a representation of the significant information contained in the spectra. To obtain the areas under the peaks we considered an interval of 0.15 ppm from the assumed peak centre (Figure 2). The metabolites used were mobile lipids, lactate, alanine, N-acetyl aspartate, creatine, total choline, glutamine, glutamate, myo-inositol+glycine, and taurine.

SW is an automatic feature selection method that does not assume any a priori knowledge; in contrast PI uses the knowledge of the experts to select the potentially most relevant parts of the spectra for discrimination purposes.

2.4 Classifiers

Fisher’s linear discriminant analysis (LDA), k-nearest neighbors (KNN) and artificial neural networks (ANN) were used for the classification. All of them have been successfully applied in a CDSS for brain tumor diagnosis based on MRS [1, 5, 31].

Fisher’s LDA [32] is a classification technique that finds the linear combination of features that best separates the classes of objects. It consists of a ratio between the difference of the projected means and a measure of dispersion of each class. This function is optimal when the distance between means is maximal and the inside class dispersion is minimal.

ANN are data models composed by an interconnected group of simple processors that work in parallel to process the information from the input to the output [14]. A multi-layer perceptron trained with the back-propagation algorithm with Bayesian regularization was used. The architecture of the network considered here had two hidden layers with 10 perceptrons in each layer. The activation function for each neuron or processor was the hyperbolic tangent function.

The KNN algorithm is an instance-based method for classifying objects based on the closest training examples in the feature space given a metric. A number of $k = 8$ has been chosen for this study after carrying out an empirical tuning using the training dataset.

2.5 Performance Measures

To determine the performance of a classifier, the following evaluation metrics were selected:

- accuracy (ACC) : Defined as $\frac{N^+}{N}$, where N^+ is the number of samples correctly predicted by the classifier and N is the total number of samples used for testing.
- geometric mean of recalls (G): Defined as the $|C| - th$ root of the product of all the successful predictions for each type of class, where $|C|$ is the total number of classes, N_c^+ is the number of samples of class C correctly predicted by the classifier and N_c is the total number of samples of class C used for testing (Equation 1). This nonlinear metric is especially useful for

determining the average success of every discriminated class. This estimator is more pessimistic than the commonly used balanced accuracy rate (BAR), being high if and only if the accuracy of each class is high and they are in equilibrium [33].

$$G = \sqrt[|c|]{\prod_c \frac{N_c^+}{N_c}} \quad (1)$$

2.6 Statistical Analysis

To evaluate the compatibility between 3T data and the classifiers based on 1.5T data, the Pearson's χ^2 test ($\alpha = 0.05$) for a contrast hypothesis was performed on the accuracy for each combination of feature extraction and classification methods.

To complete the compatibility study between 1.5 and 3T SV-MRS, the differences between the main brain metabolite ratios were compared for both field strengths. Hence, non-parametric Mann-Witney U test and box-and-whisker diagrams were calculated. The peak area ratios included in this study were: Myo-Inositol / Creatine (MI/Cr, where MI at short TE may also include signals from glycine which overlaps the myo-inositol peak), Choline / Creatine (Cho/Cr), Choline / N-acetyl aspartate (Cho/NAA), and (Lipids+Lactate) / Creatine ((Lip+Lac)/Cr).

The software used to perform the statistical analysis was MATLAB 2008.

3. Results

Every MR spectrum was processed by the above-mentioned pipeline, and the feature extraction methods were applied to the 1.5T training dataset. The significant points selected by the SW algorithm from the spectra region of interest ([4.1-0.5] ppm) were 3.97, 3.76, 3.57, 3.30, 3.11, 3.03, 2.34, 1.25, 0.98, 0.85 ppm (Figure 2).

Before dealing with performances of the classifiers, an analysis of MRS patterns at 1.5T and 3T was performed. A qualitative comparison between the mean spectra of each class (HGM, CG2G) for the three datasets (1.5T train, 1.5T test, 3T test) showed that the mean spectra tend to fall inside the region of coincidence among the three patterns (Figure 2). Also an analysis of the differences in metabolite ratios has been done performing the Mann-Witney U nonparametric test and using box-and-whisker diagrams (see Figure 3). In all cases, the p-values obtained were greater than 0.05, which indicates no significant difference among the datasets.

The performance of the classifiers on the 1.5T training dataset was estimated by a 10-fold cross-validation. The results are presented in the first row of Table 2.

In order to evaluate the 1.5T-based classifiers, the 1.5T and 3T test datasets were used as independent tests. Their performance values are presented in the second and third rows of Table 2. In both cases, the classifiers based on Knn+PI and ANN+PI gave better performance in terms of G and

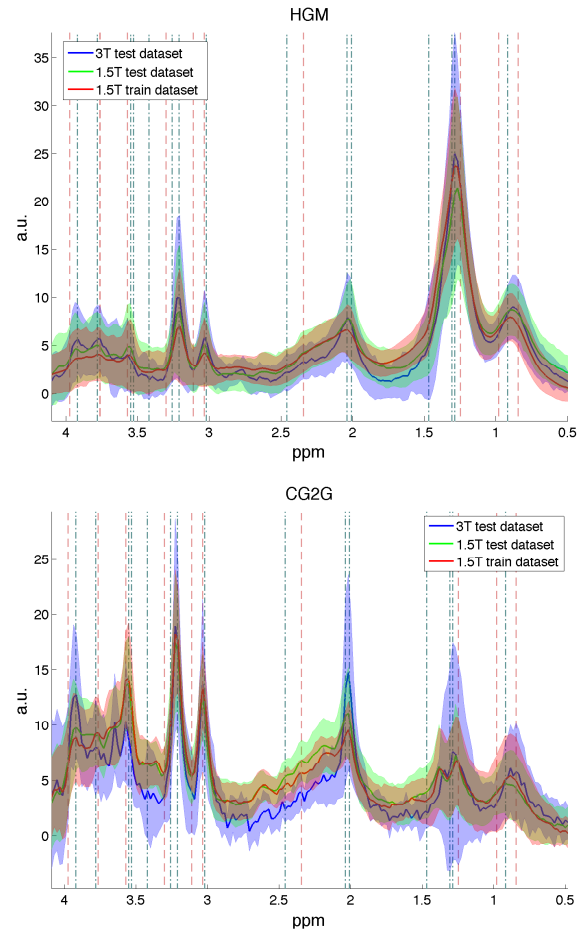


Figure 2. Spectral patterns of each of the classes (HGM, CG2G) for the 1.5T training, 1.5T test and 3T test data sets. The mean spectrum of each dataset is shown as a solid line and its standard deviation in shadowed colour. The darker brown region corresponds to the region of coincidence among the three patterns. The points selected by the stepwise algorithm are shown as vertical red dashed lines. The points selected as centroids for the peak integration method are shown as vertical green dash-dot lines.

ACC; however these differences were non-significant. The results obtained showed a p-value greater than 0.1 for every hypothesis contrast.

4. Discussion

We have tested the compatibility between the two currently coexisting clinical MR scanners of 1.5T and 3T, both for the development of new classifiers for tumour diagnosis support and also for the use of the existing ones based on 1.5T spectra. Although our present study was focused on two tumour classes (HGM, CG2G), the results of Kim et al. [16] also suggest that we may apply the established methods concerning the metabolite ratios obtained from 1.5T spectra for the evaluation of brain tumours at 3T. Thus it may be possible to extend the results to other focal brain lesions as long as their classification

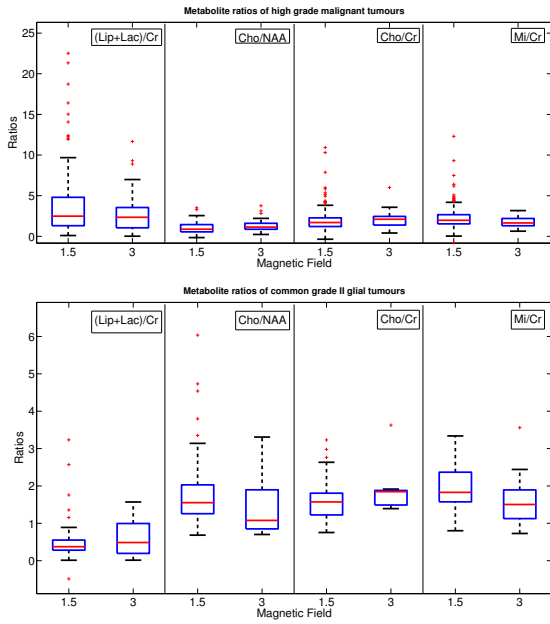


Figure 3. Box-and-whisker diagrams for each tumour class (HGM,CG2G), comparing the main brain metabolite ratios obtained for both field strengths. Outliers are shown as red crosses.

Table 2. Classification results obtained for the three datasets. In the columns, the results for each combination of feature extraction method (SW and PI) and classifier (Knn, LDA and ANN) expressed in terms of accuracy (ACC) and geometric mean of success (G). CV has been applied in the case of the 1.5T training dataset in order to obtain the G and the ACC estimators.

Dataset	Knn+SW		Knn+PI		LDA+SW		LDA+PI		ANN+SW		ANN+PI	
	ACC	G	ACC	G	ACC	G	ACC	G	ACC	G	ACC	G
1.5T train (CV)	0.93	0.90	0.91	0.88	0.93	0.91	0.91	0.90	0.95	0.94	0.93	0.90
1.5T test	0.87	0.87	0.90	0.90	0.84	0.83	0.82	0.79	0.88	0.88	0.91	0.92
3T test	0.89	0.89	0.91	0.90	0.84	0.82	0.84	0.82	0.91	0.90	0.89	0.89

is not heavily dependent on features relating to metabolites that are well resolved in 3T MRS but not in 1.5T (e.g. Glu and Gln). [17].

This initial study was designed to provide evidence of whether we will be able to further benefit from the effort made throughout many years in the compilation of large databases of 1.5T MRS data by using them to analyze results obtained at 3T [24, 5].

The classification algorithms (ANN, LDA, Knn) as well as the feature extraction methods (PI, SW) used in this study have been selected from those used in published works to facilitate validation of the results [1]. Also, the algorithms already implemented in diagnosis support systems [22, 26] have been taken into account to increase the relevance of the conclusions obtained.

The training performances obtained show consistent results. Firstly, most of the points automatically selected by

the SW algorithm are related to or are close to the metabolite peaks chosen by the spectroscopists: creatine (3.03 and 3.93 ppm), alanine (3.77 ppm), myo-Inositol (3.53 and 3.26), choline (3.21 ppm), creatine (3.02 ppm), lactate (1.30 ppm), taurine (3.30ppm), glycine (3.56ppm), N-acetyl-L-aspartic acid(2.02ppm), glutamine (2.14 ppm), glutamate (2.35ppm) and lipids (0.92 and 1.29ppm). Secondly, the performances obtained on the training data using a cross validation method (first row of Table 2) are comparable with the ones reported by García-Gómez et al, in [1, 23].

Since the cross validation method is optimistic compared to an independent test, a performance reduction occurs when using the trained classifiers on the test samples. This expected performance reduction can be observed between the first and second row of Table 2.

Because no significant difference on ACC was achieved between the two test datasets, we consider that the performances of the 1.5T based classifiers when tested on either 1.5T test set or 3T test set are comparable for the tumour classes analyzed. These results establish the possibility of using existing 1.5T based classifiers on 3T SV-MRS spectra. This agrees with the results obtained for 1.5T and 2T by Roser et al. in [18] when using multidimensional metabolic classification for assignment of glial brain tumors.

This conclusion was reinforced by the similarities observed between both spectra types (1.5T and 3T). For each tumour type obtained for 1.5T and 3T the spectral patterns fell inside the coincidence region defined by the spectra standard deviations, and the height of the metabolite peaks were similar (see Figure 2). Nevertheless, there were clear differences in the mean 3T spectra compared to the 1.5T spectra in the region of 3.5 to 4 ppm where strong signals from coupled spins can be found.

Furthermore, a comparison between the main metabolite ratios at the two magnetic field strengths was performed. The Mann-Witney U test showed non-significant differences between metabolite ratios at the two magnetic fields except for the case of the Myo-Inositol / Creatine ratio in HGM tumours. The box-and-whisker diagrams (see Figure 3) showed that in all cases the intervals defined by the first and third quartile contained common values for both magnetic fields. This agrees with the conclusions obtained by Kim et al. in [16].

In future work, incremental learning algorithms [34] will be introduced to generate new classifiers based on 1.5T data that could learn from new 3T cases. These techniques will increase the performance of the classifiers over the course of time, and will provide more reliable results. Also the generalization of this study to the case of multi-voxel MRSI data is an important goal for future work. This is not a trivial problem, because of the substantial differences between the two data types, both in the acquisition and processing of the spectra. Moreover the differences between 1.5T and 3T datasets maybe larger in the case of MRSI data [35] even with long TE if the advantages of 3T are used for rapid MRSI data acquisitions [36].

5. Conclusions

The present study has tested the compatibility of existing classifiers based on 1.5T datasets when used to classify 3T 1H SV-MRS brain tumour spectra. The results obtained suggest that existing classifiers based on 1.5T datasets are applicable to classification of 3T 1H SV-MRS data. Since the methods used in this study are available on existing software [22, 26], the conclusions obtained have immediate implications for the use of the currently-available multi-centre brain tumour datasets and prediction models that are based on 1.5T MR spectra.

6. Acknowledgements

We would like to thank Miriam Camison-Sánchez for their help in the quality control of the MRS data and diagnosis validation. This work was partially funded by the European Commission (FP6-2002-LIFESCIHEALTH 503094) and (IST-2004-27214), the I+D support program of the Universitat Politècnica de València and by the Health Institute Carlos III through the RETICS Combiomed, RD07/0067/2001.CIBER-BBN 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. The authors acknowledge to Programa Torres Quevedo from Ministerio de Educación y Ciencia, co-founded by the European Social Fund (PTQ05-02-03386, PTQ-08-01-06802, PTQ-08-01-06817).

References

- [1] García-Gómez JM, Luts J, Julià-Sapé M, Krooshof P, Tortajada S, Vicente Robledo J, Melssen W, Fuster-García E, Olier I, Postma G, Monleón D, Moreno-Torres A, Pujol J, Candiota AP, Martínez-Bisbal M, Suykens J, Buydens L, Celda B, Van Huffel S, Arús C, Robles M (2009) Multiproject-Multicenter evaluation of automatic brain tumor classification by Magnetic Resonance Spectroscopy. *Magn Reson Mater Phy* 22: 5–18
- [2] Lukas L, Devos A, Suykens JAK, Vanhamme L, Howe FA, Majós C, Moreno-Torres A, Van der Graaf M, Tate AR, Arús C, Van Huffel S (2004) Brain tumor classification based on long echo proton MRS signals. *Artif Intell Med* 31: 73–89
- [3] Devos A, Lukas L, Suykens JAK, Vanhamme L, Tate AR, Howe FA, Majos C, Moreno-Torres A, Van der Graaf M, Arús C, Van Huffel S (2004) Classification of brain tumours using short echo time 1H MR spectra. *J Magn Reson* 170: 164–175
- [4] Tate AR, Majós C, Moreno A, Howe FA, Griffiths JR, Arús C (2003) Automated classification of short echo time in in vivo 1H brain tumor spectra: a multicenter study. *Magn Reson Med* 49: 29–36
- [5] Tate AR, Underwood J, Acosta DM, Julià-Sapé M, Majós 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, Arús C (2006) Development of a decision support system for diagnosis and grading of brain tumours using in vivo magnetic resonance single voxel spectra. *NMR Biomed* 19: 411–434
- [6] Celda B, Monleón D, Martínez-Bisbal MC, Esteve V, Martínez-Granados B, Piñero E, Ferrer R, Piquer J, Martí-Bonmatí L, Cervera J (2006) MRS as endogenous molecular imaging for brain and prostate tumors: FP6 project "eTUMOR". *Adv Exp Med Biol* 587: 285–302
- [7] Laudadio T, Martínez-Bisbal MC, Celda B, Van Huffel S (2008) Fast nosological imaging using canonical correlation analysis of brain data obtained by two-dimensional turbo spectroscopic imaging. *NMR Biomed* 21: 311–321
- [8] Martínez-Bisbal M, Celda B, Martí-Bonmatí L, Ferrer P, Revert A, Piquer J, Mollá E, Arana E, Dosdá R (2002) Contribution of magnetic resonance spectroscopy to the classification of high glial tumours. Predictive value of macromolecules. *Rev Neurol* 34: 309–313
- [9] Opstad K, Provencher S, Bell B, Griffiths J, Howe F (2009) Detection of elevated glutathione in meningiomas by quantitative in vivo 1H MRS. *Magn Reson Med* 49: 632–7
- [10] Sundgren P, Nagesh V, Elias A, Tsien C, Junck L, Gomez Hassan D, Lawrence T, Chenevert T, Rogers L, McKeever P, Cao Y (2003) Metabolic alterations: a biomarker for radiation-induced normal brain injury—an MR spectroscopy study. *J Magn Reson Imaging* 29: 291–7
- [11] Hattingen E, Raab P, Franz K, Lanfermann H, Setzer M, Gerlach R, Zanella F, Pilatus U (2008) Prognostic value of choline and creatine in WHO grade II gliomas. *Neuroradiology* 50: 759–67
- [12] Server A, Josefsen R, Kulle B, Maehlen J, Schellhorn T, Gadmar, Kumar T, Haakonsen M, Langberg C, Nakstad P (2010) Proton magnetic resonance spectroscopy in the distinction of high-grade cerebral gliomas from single metastatic brain tumors. *Acta Radiol* 51: 316–25
- [13] Theodoridis S, Koutroumbas K (2006) Pattern Recognition, Third Edition. Academic Press, San Diego
- [14] Bishop CM (1995) Neural Networks for Pattern Recognition. Oxford University Press, New York
- [15] Duda RO, Hart PE, Stork DG (2000) Pattern Classification (2nd Edition). Wiley-Interscience, New York
- [16] Kim J, Chang K, Na DG, Song IC, Kim SJ, Kwon BJ, Han MH (2006) Comparison of 1.5T and 3T 1H-MR

- Spectroscopy for Human Brain Tumors. *Korean J Radiol* 7: 156–161
- [17] Barker PB, Hearshen DO, Boska MD (2001) Single-Voxel Proton MRS of the Human Brain at 1.5T and 3.0T. *Magn Reson Med* 45: 765–769
- [18] Roser W, Hagberg G, Mader I, Dellas S, Seelig J, Radue E, Steinbrich W (1997) Assignment of glial brain tumors in humans by in vivo 1H-magnetic resonance spectroscopy and multidimensional metabolic classification. *Magn Reson Mater Phy* 5: 179–183
- [19] Berner ES, (Ed) (2007) *Clinical Decision Support Systems: Theory and Practice*. Springer New York
- [20] Sim I, Gorman P, Greenes RA, Haynes RB, Kaplan B, Lehmann H, Tang PC (2001) Clinical decision support systems for the practice of evidence-based medicine. *J Am Med Inform Assn* 8: 527–534
- [21] González-Vélez H, Mier M, Julià-Sapé M, Arvanitis T, García-Gómez JM, Robles M, Lewis P, Dasmahapatra S, Dupplaw D, Peet A, Arús C, Celda B, Van Huffel S, Lluch-Ariet M (2009) HealthAgents: distributed multi-agent brain tumor diagnosis and prognosis. *Appl Intell* 30: 191–202
- [22] Sáez C, García-Gómez JM, Vicente J, Tortajada S, Esparza M, Navarro A, Fuster-Garcia E, Robles M, Martí-Bonmatí L, Arús C (2008) A generic Decision Support System featuring an assembled view of predictive models for Magnetic Resonance and clinical data. In *ESMRMB 25th Annual Meeting, Valencia, 2 – 4*. Springer
- [23] García-Gómez JM, Tortajada S, Vidal C, Julia-Sape M, Luts J, Van Huffel S, Arús C, Robles M (2008) The influence of combining two echo times in automatic brain tumor classification by Magnetic Resonance Spectroscopy. *NMR Biomed* 21: 1112–1125
- [24] The eTUMOUR Consortium (2007) eTUMOUR: Web accessible MR Decision support system for brain tumour diagnosis and prognosis, incorporating in vivo and ex vivo genomic and metabolomic data, VI framework programme, EC, <http://www.etumour.net>. Technical Report FP6-2002-LIFESCIHEALTH 503094
- [25] Van der Graaf M, Julià-Sapé M, Howe FA, Ziegler A, Majós C, Moreno-Torres A, Rijpkema M, Acosta D, Opstad KS, Van der Meulen Y, Arús C, Heerschap A (2008) MRS quality assessment in a multicentre study on MRS-based classification of brain tumours. *NMR Biomed* 21: 148–158
- [26] Pérez-Ruiz A, Olier-Caparroso IA, Julià-Sapé M, Candiota AP, Arús C (2008) Brain Tumor Diagnosis with MRS: The Single Voxel INTERPRET Decision-Support System version 2.0. In *ESMRMB 25th Annual Meeting, Valencia, 2 – 4*. *Magn Reson Mater Phy*
- [27] Klose U (1990) In vivo proton spectroscopy in presence of eddy currents. *Magn Reson Med* 14: 26–30
- [28] Van den Boogaart A, Van Hecke P, Van Huffel S, Graveron-Demilly S, Van Ormondt D, de Beer R (1996) MRUI: a graphical user interface for accurate routine MRS data analysis. In *ESMRMB 13th Annual Meeting, Prague, 12 – 15*. Springer
- [29] Luts J, Pouillet J, Garcia-Gomez J, Heerschap A, Robles M, Suykens J, Van Huffel S (2008) Effect of feature extraction for brain tumor classification based on short echo time 1H MR spectra. *Magn Reson Mater Phy* 60: 88–98
- [30] Russell SJ, Norvig P (1995) *Artificial intelligence: a modern approach*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA
- [31] Opstad KS, Ladroue C, Bell BA, Griffiths JR, Howe FA (2007) Linear discriminant analysis of brain tumour 1H MR spectra: a comparison of classification using whole spectra versus metabolite quantification. *NMR Biomed* 20: 763–770
- [32] Fisher RA (1936) The Use of Multiple Measurements in Taxonomic Problems. *Ann Eugenics* 7: 179–188
- [33] Kubat M, Matwin S (1997) Addressing the curse of imbalanced training sets: one-sided selection. In *Proc. 14th International Conference on Machine Learning*, pp. 179–186. Morgan Kaufmann
- [34] Giraud-Carrier C (2000) A note on the utility of incremental learning. *AI Communications* 13: 215–223
- [35] Gonen O, Gruber S, Li B, Mlynrik V, Moser E (2001) Multivoxel 3D proton spectroscopy in the brain at 1.5 versus 3.0 T: signal-to-noise ratio and resolution comparison. *Am J Neuroradiol* 22: 1727–31
- [36] Dydak U, Meier D, Lamerichs R, Boesiger P (2006) Effect of feature extraction for brain tumor classification based on short echo time 1H MR spectra. *Am J Neuroradiol* 27: 1441–6