Application of pattern recognition techniques for classification of paediatric brain tumours by in vivo 3T 1H MRS – A multi centre study

Zarinabad, Niloufar; Abernethy, Laurence J.; Avula, Shivaram; Davies, Nigel; Gutierrez, Daniel Rodriguez; Jaspan, Tim; Macpherson, Lesley; Mitra, Dipayan; Rose, Heather; Wilson, Martin; Morgan, Paul S.; Bailey, Simon; Pizer, Barry; Arvanitis, Theodoros; Grundy, Richard G.; Auer, Dorothee P.; Peet, Andrew

DOI:
10.1002/mrm.26837

License:
Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- Users may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?).
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBRRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.
Application of Pattern Recognition Techniques for Classification of Pediatric Brain Tumors by In Vivo 3T 1H-MR Spectroscopy—A Multi-Center Study

Niloufar Zarinabad,1,2 Laurence J. Abernethy,3 Shivaram Avula,3 Nigel P. Davies,1,2,4 Daniel Rodriguez Gutierrez,5,6 Tim Jaspan,5,7 Lesley MacPherson,2 Dipayan Mitra,8 Heather E.L. Rose,1,2 Martin Wilson,9 Paul S. Morgan,5,6,10 Simon Bailey,11 Barry Pizer,12 Theodoros N. Arvanitis,2,13 Richard G. Grundy,5 Dorothee P. Auer,5,7,10 and Andrew Peet1,2*

Purpose: 3T magnetic resonance scanners have boosted clinical application of 1H-MR spectroscopy (MRS) by offering an improved signal-to-noise ratio and increased spectral resolution, thereby identifying more metabolites and extending the range of metabolic information. Spectroscopic data from clinical 1.5T MR scanners has been shown to discriminate between pediatric brain tumors by applying machine learning techniques to further aid diagnosis. The purpose of this multi-center study was to investigate the discriminative potential of metabolite profiles obtained from 3T scanners in classifying pediatric brain tumors.

Methods: A total of 41 pediatric patients with brain tumors (17 medulloblastomas, 20 pilocytic astrocytomas, and 4 ependymomas) were scanned across four different hospitals. Raw spectroscopy data were processed using TARQUIN. Borderline synthetic minority oversampling technique was used to correct for the data skewness. Different classifiers were trained using linear discriminative analysis, support vector machine, and random forest techniques.

Results: Support vector machine had the highest balanced accuracy for discriminating the three tumor types. The balanced accuracy achieved was higher than the balanced accuracy previously reported for similar multi-center dataset from 1.5T magnets with echo time 20 to 32 ms alone.

Conclusion: This study showed that 3T MRS can detect key differences in metabolite profiles for the main types of childhood tumors. Magn Reson Med 000:000–000, 2017. © 2017 International Society for Magnetic Resonance in Medicine.

Key words: MR spectroscopy; 3T; pediatric brain tumors; diagnosis; classification

INTRODUCTION

Brain tumors are a significant cause of death and long-term disability in children, with a range of treatment options and outcomes depending on the tumor type, location, and age of the patient. Histopathology following biopsy or surgical resection is the current gold standard for diagnosis of brain tumor type and grade (1). Although surgical resection is often appropriate for many tumors and also provides a histological diagnosis, there are distinct advantages to having a diagnosis before surgery. A pre-operative diagnosis can influence the extent of surgical resection attempted, allow the timely planning of adjuvant treatment and aid discussions with the family. Conventional magnetic resonance imaging (MRI) is commonly used to propose a diagnosis before surgery but is of limited accuracy. A previous study reviewed the radiological reports from a cohort of children with medulloblastoma, pilocytic astrocytoma, or ependymoma and showed an accuracy of diagnosis of 66% (2).

MRI techniques have significantly advanced in recent years with new imaging techniques being able to provide information on tissue properties, structure, and basic metabolic processes. Amongst these new techniques, 1H-MR spectroscopy (MRS) has the ability to provide non-invasive measurements of metabolite profiles with the potential to
Spectroscopic studies on pediatric brain tumors have attempted to characterize different histologic types and predict the degree of malignancy (7,9). Non-invasive grading is especially important for tumors in the eloquent areas or deeply located tumors. Single and multi-center studies have worked on development and optimization of diagnostic classifiers of childhood brain tumors using MRS. This has yielded promising results in discriminating between certain common tumor types (6,7).

With the integration of 3T MRI into clinical practice, there is growing interest in the practical improvement of MRS at a 3T field strength over the more established magnetic field strength of 1.5T, because both the spectral resolution and the spatial resolution depend, in a linear fashion, on the magnetic field (10–12).

Although single-voxel MRS of the human brain has been carried out at many field strengths (from 0.5T to 7T), to date, no study has been reported on application of pattern recognition techniques for classification of pediatric brain tumors using 3T MRS.

The aim of this study was to evaluate the discriminative potential of metabolites obtained from 3T MRI scanners in classifying pediatric brain tumors by comparing the performance of three different pattern recognition techniques.

**METHODS**

**Patients**

This study included 52 patients less than 16 years of age (8.2 ± 5.32 years, 21 female and 31 male) with histologically proven brain tumor collected from four centers in the United Kingdom. Patient data were collected retrospectively from children who underwent single-voxel MRS between November 2009 and April 2016 during a routine MRI for a suspected brain tumor before treatment.

The enrolled cohort consisted of patients with three different tumor types from all regions of the brain, including medulloblastoma (MB) (n = 18), pilocytic astrocytoma (PA) (n = 26), and ependymoma (EP) (n = 8). Histopathological, clinical, and radiological features were used to form a diagnosis agreed by a multidisciplinary team. Approval was obtained from the research ethics committee and informed consent given by parents/guardians.

**Data Acquisition**

All studies were performed using 3T scanners from different manufacturers (Philips Achieva, Siemens MAGNETOM Verio). MRS was performed after conventional MRI that included T1- and T2-weighted and T1-weighted post-contrast sequences. Spectroscopy images were acquired using a point resolved single voxel spectroscopy (PRESS) sequence (time echo = 30–46 ms, pulse repetition time = 2000 ms). Cubic voxels size varied from 3.38 cm³ to 8 cm³ and 128 repetitions were used. A water unsuppressed acquisition was also acquired as a concentration reference. Conventional MRI was used for guiding voxel placement to ensure it is entirely located within the tumor.

**MRS Processing**

Raw spectroscopy data were processed using TARQUIN (version 4.3.6) (13) fitting to a linear combination of 19 metabolite basis functions generated at the correct field strength and echo time with an additional 9 lipid and macromolecular components (13–15). Frequency alignment, zero order phase correction, baseline correction, and water removal using Hankel singular value decomposition (HSVD) methods were applied by TARQUIN. TARQUIN determines the chemical shift offset, phase, and baseline during the fitting process. It then zero fills the time domain data by a factor of 2 (× 2) and converts the time domain signal to spectral domain using a Fourier transform. The obtained spectra are then resampled to 0.49 Hz/point. This is to ensure all cases have a consistent Hz/point. The resampled spectra are used for analysis in this study. The full spectral range (−3.00 to 12.5 ppm) was used for fitting the data and metabolite quantitation. Quantitation was carried out relative to a water reference spectrum. Corrections for T₂ relaxation times of metabolites and water were applied as the default values for TARQUIN (16). To include the metabolites of interest in the study, main metabolites in 0.5 to 4 ppm region were used for classification.

Out of 52 enrolled cases, 41 cases (medulloblastoma n = 17, pilocytic astrocytoma n = 20, and ependymoma n = 4, of which 1 is analplastic ependymoma) passed the following quality control criteria: signal-to-noise ratio obtained from TARQUIN (SNR) ≥ 4 (here SNR is defined as ratio between the maximum in the spectrum minus baseline divided by 2 × root-mean-square of the spectral noise level), full-width half-maximum obtained from TARQUIN (FWHM) ≤ 0.15 ppm, stable baseline, good phasing, adequate water suppression, and absence of artefacts. The voxel position was also reviewed to ensure it was positioned over tumor, did not include significant amounts of normal-appearing brain or cyst, and was at least 3 mm away from lipid-containing bone and scalp. The 11 failed cases include medulloblastoma n = 1, pilocytic astrocytoma n = 6, and ependymoma n = 4. A poor voxel placement that included normal brain, small voxel size, and low quality spectra were the main quality control failure reasons for the latter cases.

For the quantified metabolite profiles, Cramer-Rao lower bounds were calculated to evaluate metabolite accuracy. All metabolites where at least two patients had a Cramer-Rao lower bound < 50 were included (alanine, aspartate and GABA were excluded) (7). The following metabolites were used for classification: citrate (Cit), glucose (Glc), glutamine (Gln), glutathione (Glt), glutamate (Glu), glycerine (Gly), myo-inositol (mIns), lactate (Lac), tCho = glycerophosphocholine (GPC) + phosphocholine (PCh), scylo-inositol (scyllo), taurine (Tau), total creatine (tCr = Cr + PCr), and total NAA (tNAA = NAA + NAAG). The nine macromolecular and lipid components were grouped together to account for three broad resonances at 0.9 ppm (tLM09), 1.3 ppm (tLM13), and 2.0 ppm (tLM20), giving a maximum of 16 variables.

The Kruskal-Wallis test for the analysis of variance (α = 0.05) was applied to determine the significant differences in metabolite concentrations between the three groups. Mann-Whitney U tests were carried out to
compare mean metabolite values between individual pair of groups. Statistical analysis was carried out using SPSS statistics software (version 21.0).

Classification

Borderline synthetic minority oversampling technique (bSMOTE) (17) was used to overpopulate the original ependymoma group by 100% and correct for the skewness and class imbalance in the original data. Over-sampled ependymoma was added to the original data sets to create an overpopulated metabolite feature set. Principal component analysis was used to reduce dimension of this overpopulated set and extract features that best discriminate between the three tumor groups.

Classifiers were trained using linear discriminative analysis (LDA), support vector machine (SVM), and random forest (RF) approaches. These techniques have low, medium, and high model flexibility, respectively. A radial basis (Gaussian) function kernel was chosen for SVM. LDA was trained using diagonal linear learner where all classes have the same diagonal covariance matrix.

Developed classifiers using the bSMOTE-overpopulated sets were tested using the original data and the validity of the classifier was checked by comparison with the known histological assignment. Ten-fold cross-validation was used to evaluate the learning algorithm performance. To evaluate classifiers trained using bSMOTE data, the original data set is partitioned into \( z = 10 \) pools in which one pool is used for evaluation while the remaining \( z - 1 \) pools are added to the bSMOTE ependymoma cases, obtained from oversampling ependymoma cases in the remaining \( z - 1 \) pools, and used for training to produce mean error rates, where \( z \in \{1, 2, \ldots, Z\} \). The pools were then randomly rotated and the subsampling and permutation tests were repeated until all pools had been evaluated exactly.

### Table 1

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Ependymoma ((n = 4))</th>
<th>Medulloblastoma ((n = 17))</th>
<th>Pilocytic Astrocytoma ((n = 20))</th>
<th>(P)-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cit*</td>
<td>0.96 ± 0.21</td>
<td>0.56 ± 0.31</td>
<td>0.47 ± 0.48</td>
<td>0.031</td>
</tr>
<tr>
<td>Gln</td>
<td>1.2 ± 0.27</td>
<td>2.3 ± 1.7</td>
<td>1.07 ± 1.06</td>
<td>0.073</td>
</tr>
<tr>
<td>Gln**</td>
<td>1.69 ± 0.11</td>
<td>2.6 ± 1.53</td>
<td>2.2 ± 2.6</td>
<td>0.429</td>
</tr>
<tr>
<td>Glth*</td>
<td>1.12 ± 0.32</td>
<td>1.8 ± 0.8</td>
<td>0.66 ± 0.6</td>
<td>&lt;0.000</td>
</tr>
<tr>
<td>Glth**</td>
<td>1.14 ± 0.94</td>
<td>2.4 ± 2.2</td>
<td>3.2 ± 2.3</td>
<td>0.055</td>
</tr>
<tr>
<td>Gly</td>
<td>1.92 ± 0.5</td>
<td>3.7 ± 2.08</td>
<td>0.73 ± 0.76</td>
<td>&lt;0.000</td>
</tr>
<tr>
<td>mIns</td>
<td>4.53 ± 2.3</td>
<td>2.29 ± 1.66</td>
<td>1.94 ± 1.73</td>
<td>0.097</td>
</tr>
<tr>
<td>Lac</td>
<td>1.39 ± 0.68</td>
<td>1.65 ± 1.7</td>
<td>0.65 ± 0.74</td>
<td>0.041</td>
</tr>
<tr>
<td>Scyllo</td>
<td>0.008 ± 0.01</td>
<td>0.37 ± 0.47</td>
<td>0.05 ± 0.08</td>
<td>0.009</td>
</tr>
<tr>
<td>Tau</td>
<td>0.6 ± 0.75</td>
<td>6.3 ± 4.1</td>
<td>1.07 ± 1.03</td>
<td>&lt;0.000</td>
</tr>
<tr>
<td>tNAA* (tNAA = NAA + NAAG)</td>
<td>0.53 ± 0.25</td>
<td>1.03 ± 0.44</td>
<td>1.37 ± 0.95</td>
<td>0.007</td>
</tr>
<tr>
<td>tCho*</td>
<td>1.69 ± 0.29</td>
<td>3.94 ± 1.74</td>
<td>1.4 ± 0.709</td>
<td>&lt;0.000</td>
</tr>
<tr>
<td>tCr*</td>
<td>1.69 ± 0.79</td>
<td>3.91 ± 1.5</td>
<td>2.29 ± 2.22</td>
<td>0.003</td>
</tr>
<tr>
<td>tLM09</td>
<td>3.7 ± 1.4</td>
<td>4.94 ± 2.4</td>
<td>3.02 ± 1.87</td>
<td>0.084</td>
</tr>
<tr>
<td>tLM13*</td>
<td>21.4 ± 10.05</td>
<td>19.32 ± 13.01</td>
<td>7.8 ± 3.7</td>
<td>&lt;0.000</td>
</tr>
<tr>
<td>tLM20</td>
<td>7.2 ± 1.8</td>
<td>7.53 ± 4.1</td>
<td>6.3 ± 3.4</td>
<td>0.444</td>
</tr>
</tbody>
</table>

Cit, citrate; Gln, glucose; Glth, glutathione; Glth, glutathione; Gly, glycine; Lac, lactate; mIns, myo-inositol; NAA, N-acetyl aspartate; NAAG, N-acetylaspartylglutamic acid; scyllo, scyllo-inositol; SD, standard deviation; Tau, taurine; tNAA, total N-acetylaspartate; tCho, total choline; tCr, total creatine; tLM09, lipids and macromolecules 0.9; tLM13, lipids and macromolecules 1.3; tLM20, lipids and macromolecules 2.0.

*The \(P\)-value of analysis of variance (calculated using Kruskal-Wallis test with \( \alpha = 0.05 \)) is comparing ependymoma versus medulloblastoma versus pilocytic astrocytoma.

*Metabolites with \(P\)-values less than 0.05.

### Table 2

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Ependymoma ((n = 4))</th>
<th>Pilocytic Astrocytoma ((n = 20))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cit*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tCho**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tCr**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tau**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mIns*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tNAA**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tML13**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tLM90*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Cit, citrate; Gln, glucose; Glth, glutathione; Glth, glutathione; Gly, glycine; Lac, lactate; mIns, myo-inositol; scyllo, scyllo-inositol; Tau, taurine; tNAA, total N-acetylaspartate; tCho, total choline; tCr, total creatine; tLM09, lipids and macromolecules 0.9; tLM13, lipids and macromolecules 1.3.

\(\ast P < 0.05; \ast\ast P < 0.01; \ast\ast\ast P < 0.001.\)
once. This ensured that there were no bSMOTE ependymoma cases in any of the test sets. The analysis have been repeated 100 times and results been averaged.

G-mean and F-measure metrics were used for more precise performance evaluation in class imbalance learning. The balanced accuracy rate (BAR) of the learning algorithm, calculated as the mean of the accuracies for the three tumor types, is also reported here as a performance measure metric. In this study, all learning algorithms were developed in python 2.7 (Python Software Foundation, Wilmington, DE) using Scikit-learn (Version 0.16.1) and Orange (Version 2.6a2) libraries.

RESULTS

Mean metabolite concentrations (mM) ± standard deviations (SD) for each diagnostic group are represented in Table 1. Analysis of variance revealed significant differences between ependymoma, pilocytic astrocytoma, and medulloblastoma in ten of the individual metabolites and combined macromolecules and lipids at 1.3 ppm. Glc, Gln, Glu, mIns, tLM09, and tLM20 had P-values > 0.05. Results comparing metabolites between specific pairs of tumor types are reported in Table 2. MB demonstrated increased Tau and decreased mIns compared to EP. Compared with PA, MB had a number of significantly higher metabolites including tCho, Tau, scyllo, tCr, Gly, Glth, and tLM1.3. Compared with EP, PA demonstrated elevated tNAA and decreased tLM13 and mIns. Mean MRS spectra of the three tumor types is presented in Figure 1.

Principal component analysis was carried out and principal components accounting for 95% of variance were extracted, giving four principal components. Figure 2 shows the three-dimensional scatter plot of the three tumor groups using the first three principal components. Data points of each tumor types demonstrated a good degree of data clustering and separation. The four principal components were then submitted to SVM, RF, and LDA learning algorithms and classifiers were compared in view of their performance.

The bar plot in Figure 3 characterizes PCA loadings of the metabolite set for the four main components. tLM13 has dominated the first principal component. Except Glu and tNAA, all loadings on the first principal component are positive. tLM20, Tau, Gly, tCho, and tCr were highly loaded on the second principal component. For the third principal component, Glc, Glu, mIns, tLM20, and Tau with a negative load were the most dominant. In the fourth principal component, the majority of the

FIG. 1. Mean 3T short time echo spectra for (a) ependymoma (n = 4), (b) medulloblastoma (n = 17), and (c) pilocytic astrocytoma (n = 20) with solid black line indicating the mean spectra and SD indicated by the shaded region.

FIG. 2. 3D scatter plot presents principal component analysis (PCA) space for the discrimination of medulloblastoma (MB), pilocytic astrocytoma (PA), and ependymoma (EP).
metabolites were negatively loaded, however, except Gln, they all were <0.4.

BAR of the learning algorithms and their corresponding individual tumor types, F-measure and G-mean, are presented in Table 3. SVM (BAR = 0.86) performed favorably in comparison to RF (BAR = 0.84) and LDA (BAR = 0.81) in discriminating between the three tumor types. Figure 4 represents the BAR for the three pattern-recognition techniques with their accuracy in discriminating individual tumor groups. To further evaluate and describe the differences between the three learning algorithms performance, analysis of variance was carried out on the BAR data obtained from 100 runs of samplings. The three methods showed to have BAR mean of significantly different from each other ($P<0.0001$) (Fig. 5). RF had higher variance ($\sigma^2 = 0.002$) in comparison to the LDA ($\sigma^2 = 0.0016$) and SVM ($\sigma^2 = 0.0011$).

Comparing methods in view of their performance in classifying EP, all methods had similar accuracy (EP accuracy = 0.75) and misclassified one case as PA.

Table 3

<table>
<thead>
<tr>
<th></th>
<th>F-Measure</th>
<th>G-Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EP</td>
<td>MB</td>
</tr>
<tr>
<td>LDA</td>
<td>0.54</td>
<td>0.82</td>
</tr>
<tr>
<td>SVM</td>
<td>0.85</td>
<td>0.9</td>
</tr>
<tr>
<td>RF</td>
<td>0.75</td>
<td>0.9</td>
</tr>
</tbody>
</table>

LDA, linear discriminate analysis; RF, Random forest; SVM, support vector machine; MB, medulloblastoma; PA, pilocytic astrocytoma; EP, ependymoma.
However, EP F-measure and G-mean varied amongst the three methods with SVM having the highest values. This finding demonstrates SVM’s better ability in balancing the classification performance between the three groups in comparison to the other two methods.

For PA, SVM, and LDA achieved the highest accuracy (F-measure = 0.9, PA accuracy = 0.95). For the MB group, both SVM and RF had highest number of correctly classified cases (F-measure = 0.9, MB accuracy = 0.88). The misclassification spread across different tumor types, data sites, and scanner types is summarized in Table 4.

The 3D scatter plot in Figure 6 represents LDA and SVM estimated regions. SVM provided improved estimated class boundaries for the three tumor groups that resulted in less overlap between clustered data points and estimated regions from dissimilar tumor types. This difference is more apparent in MB where three cases assigned incorrectly by LDA were correctly classified by SVM.

**DISCUSSION**

This is the first multi-center study investigating 3T MRS and pattern classification for pediatric tumors. Data represented here are from four different centers collected from 3T scanners at short echo time from three common types of childhood tumors.

Mean metabolite concentrations were shown to differ between tumor types with some individual metabolites differing significantly (Mann-Whitney U test $P < 0.05$) between specific pairs of tumors. The metabolites identified to be significantly different between the three tumor types at 3T in this study generally agrees with previously reported metabolites for the same tumors at 1.5T (6).

**Table 4**

Summary of the Misclassification Spread across Different Tumor Types, Data Sites, and Scanner Types.

<table>
<thead>
<tr>
<th>Tumor type</th>
<th>Ependymoma</th>
<th>Medulloblastoma</th>
<th>Pilocytic astrocytoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misclassified cases (n)</td>
<td>1</td>
<td>3–5</td>
<td>1–2</td>
</tr>
<tr>
<td>Site of misclassified cases</td>
<td>Alder Hey Children’s Hospital (Liverpool)</td>
<td>Birmingham Children’s Hospital</td>
<td>Birmingham Children’s Hospital Alder Hey Children’s Hospital (Liverpool)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The Great North Children’s Hospital</td>
<td>Nottingham Children’s Hospital</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scanner manufacturer</td>
<td>Philips</td>
<td>Philips, Siemens</td>
<td>Philips</td>
</tr>
</tbody>
</table>

**FIG. 5.** Comparison of the three learning algorithms balanced accuracy rate. Box plot represents the distribution of BAR obtained from 100 runs of oversampling.

**FIG. 6.** Three-dimensional scatter plot shows the data clustering for the different tumor types with (a) LDA and (b) SVM decision boundaries. The graph is oriented so that each clustering can be seen from the best angle with a clear boundary.
Exceptions are metabolites such as Glc, Glu, tLM09, and tLM20 that were reported to be significantly different at 1.5T, and Lac that was found to be significantly different at 3T.

For the 3T data set under study, tCho was higher ($P = 0.01$) in MB compared to PA in agreement with tCho as an indicator of cell proliferation and tumor malignancy (18). EP and MB had higher concentrations of lipids and macromolecules compared with PA that is associated with hypoxia, apoptosis, and necrosis and linked to high malignancy and poor survival (3,19). Tau concentration was significantly higher in MB than PA ($P = 0.0001$) and EP ($P = 0.001$) (20,21). Similarly mlns concentration was significantly higher in EP than MB and PA ($P = 0.05$).

Glc and Glu were not significantly different between the three tumor types as reported in previous 1.5T studies, but Lac was significantly different between PA and MB ($P = 0.05$). This finding can be because of difference in sample size or increased resolution at 3T allowing more accurate quantitation.

Applying pattern recognition techniques to the metabolite profile data, a maximum balanced accuracy of 86% was achieved for discriminating between astrocytoma, ependymoma, and medulloblastoma. High classification rates are seen despite the relatively small number of cases used to train the classifiers and the inclusion of data from more than one scanner type. The achieved accuracy was better than the reported BAR obtained from short time echo 1.5T data from ten different international centers where similar tumor types and similar data skewness were studied (6). Using LDA for classification, Vicente et al. (6) reported a BAR of 0.79 for tumors from all regions of the brain. However, the reported 1.5T BAR for LDA is in a very close agreement with LDA accuracy when 3T data is used (3T LDA BAR = 0.81).

The discrimination power of the pattern recognition techniques in classifying the ependymoma group alone was on average 75%. This is mainly because of the imbalanced nature of the data. Although SMOTE was used to overpopulate the ependymoma group by 100% to correct for this skewness, the data distribution still remained imbalanced because of the small number of tumors in this group.

In comparing the performance of different pattern recognition techniques, SVM was a fast, easily trained, and reliable discriminator with the highest balanced accuracy rate. This method, with medium flexibility, avoids under- or overfitting of the data. SVMs strongly draw on variation methods in their construction and are designed to yield the best estimate of the optimal separating hyperplane (Fig. 6b). SVM outperforms conventional pattern recognition methods especially when the number of training data is small and number of input variables is large. This is because the conventional pattern recognition methods do not have the mechanism to maximize the margins of class boundaries. Therefore, if some mechanism is introduced to maximize the margins, the generalization ability is improved (22).

Future work should focus on optimizing pattern recognition techniques to classify a wider range of tumor types using 3T MRS. Moreover, a separate prospective study based on an independent test set to ascertain the accuracy of non-invasive prognostic biomarkers, especially for the minority class, provided by 3T MRS is required.

In conclusion, this study demonstrates the ability and high diagnostic accuracy of 3T MRS in detecting key differences in the metabolite profiles for the main types of childhood tumors. Classification performance of 3T MRS compares favorably with previously reported multi-center data sets at 1.5T. 3T MRS with automated processing and pattern recognition providing a useful technique for accurate, non-invasive diagnosis and classification of childhood brain tumors and thereby a powerful diagnostic tool for clinical practice.

ACKNOWLEDGMENTS

Data collection was supported by the Paediatric network ECMC.

REFERENCES

15. Davies NP, Wilson M, Natarajan K, Sun Y, Macpherson L, Brundler MA, Arvanitis TN, Grundy RG, Peet AC. Non-invasive detection of


