

Multicentre prospective classification of childhood brain tumours based on metabolite profiles derived from ^1H MRS

N. P. Davies^{1,2}, S. Gill^{2,3}, T. N. Arvanitis^{3,4}, D. Auer⁵, R. Grundy^{6,7}, F. A. Howe⁸, D. Hargrave⁹, T. Jaspan⁷, L. MacPherson³, K. Natarajan^{1,3}, G. Payne^{9,10}, D. Saunders¹¹, Y. Sun^{2,3}, M. Wilson^{2,3}, and A. C. Peet^{2,3}

¹Medical Physics, University Hospitals Birmingham NHS Foundation Trust, Birmingham, United Kingdom, ²Cancer Sciences, University of Birmingham, Birmingham, United Kingdom, ³Birmingham Children's Hospital NHS Foundation Trust, Birmingham, United Kingdom, ⁴Department of Electrical, Electronic, and Computer Engineering, University of Birmingham, Birmingham, United Kingdom, ⁵Academic Radiology, University of Nottingham, Nottingham, United Kingdom, ⁶Children's Brain Tumour Research Centre, University of Nottingham, Nottingham, United Kingdom, ⁷University Hospital Nottingham, Nottingham, United Kingdom, ⁸St. George's University of London, London, United Kingdom, ⁹Royal Marsden Hospital, London, United Kingdom, ¹⁰Institute of Cancer Research, London, United Kingdom, ¹¹Great Ormond Street Hospital, London, United Kingdom

Introduction: ^1H MRS provides non-invasive metabolite profiles of childhood brain tumours aiding diagnosis and potentially improving characterisation. Although ^1H MRS classifiers have been developed for adult brain tumours, child specific classifiers are required, since childhood brain tumours differ in the prevalence of tumour types and in their biology. Previous studies of ^1H MRS for classifying childhood brain tumours have been limited by small numbers of cases and by their retrospective, single-centre design. The aim of this study was to perform a large prospective multicentre evaluation of ^1H MRS as a diagnostic tool for grading childhood brain tumours.

Method: A low grade vs high grade classifier for childhood brain tumours was trained using single-voxel (SV) ^1H MRS acquired using a standard protocol (PRESS, TE/TR 30/1500 ms) on two 1.5 T scanners in a single centre (Centre 1) over a 5 year period up to May 2008. A total of 123 cases met the inclusion criteria of having a pre-treatment SV MRS with subsequent brain tumour diagnosis and grading confirmed by histopathology according to WHO criteria (N = 97) or by radiological review in the absence of a biopsy (N = 26). Of these, 81 were diagnosed as low grade (LG; WHO grade I or II) and 42 as high grade (HG; WHO grade III or IV). MRS data were processed using TARQUIN [1] to determine metabolite concentrations. TARQUIN was used in preference to LCModel, due to its effective and convenient use of simulated basis sets, to account for differences in MRS data acquisition protocols that are difficult to avoid in multicentre studies. Cases were screened for artefacts and quality control (QC) criteria were applied (SNR > 5 and full-width-at-half maximum (FWHM) of water peak < 10 Hz). Classifier training consisted of principal components analysis (PCA) of the standardised metabolite profile, followed by linear discriminant analysis (LDA) with subsequent model tuning and cross-validation [2]. The classifier was then tested in a prospective manner using short-TE SV MRS data acquired on 6 different scanners in 4 centres. The test dataset consisted of 55 cases from Centre 1, acquired between June 2008 and September 2010, and 55 cases from Centres 2 - 4, of which 10 were acquired on a 3 T scanner and 28 were acquired with TE between 23 - 40 ms.

Results: Table 1 shows the classifier performance. The high classifier accuracy, obtained using for the training set, is very similar using TARQUIN to derive the metabolite profile, compared to that previously obtained for the same cohort using LCModel [3]. Prospective testing compares well with the training set accuracy with an overall accuracy rate of 86%. Accuracy for cases from centres 2-4 was lower (80%) than that of cases from centre 1 (92%). However, for centres 2-4, the classification rate for cases collected using the same acquisition protocol as the training set (1.5 T, TE 30ms) was 83%, which is comparable to the 78% found for those that did not adhere to the acquisition protocol.

Conclusions: High classification accuracy for tumour grade has been shown in a prospective multi-centre evaluation of a childhood brain tumour classifier, based on multivariate analysis of metabolite profiles derived from ^1H MRS. Using a simulated basis set allows for variations in echo time and field strength and as a consequence provides comparable classification accuracy for data collected using different protocols, a property which is important for multi-centre application of the approach.

REFERENCES:

[1] Wilson M, et al. Magn Reson Med. 2010; Sep 27 [Published online ahead of print] PMID: 20878762.

[2] Davies NP, et al. NMR in Biomed. 2008; 21:908-918.

[3] Davies NP, et al. MAGMA, 2008; 21(Suppl. 1):18.

Table 1: Performance of the LG vs HG classifier based on ^1H MRS metabolite profiles derived by TARQUIN.

| Dataset | N (Before QC) | | | N (After QC) | | | Accuracy | N (misclassified) | |
|-----------------------|---------------|----|-------|--------------|----|-------|----------|-------------------|----|
| | LG | HG | Total | LG | HG | Total | | LG | HG |
| Training (Centre 1) | 81 | 42 | 123 | 65 | 38 | 103 | 93% | 3 | 4 |
| Testing (Centre 1) | 39 | 16 | 55 | 35 | 15 | 50 | 92% | 1 | 3 |
| Testing (Centres 2-4) | 34 | 21 | 55 | 29 | 17 | 46 | 80% | 3 | 6 |
| Testing (All Centres) | 73 | 37 | 110 | 64 | 32 | 96 | 86% | 4 | 9 |

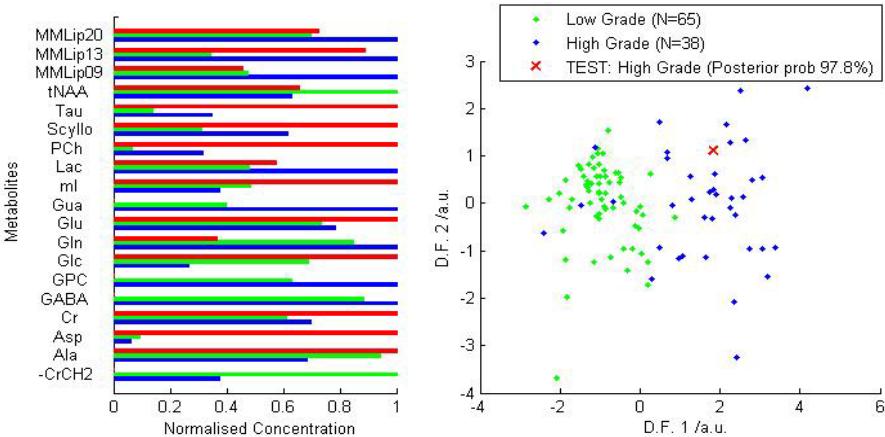


Figure 1: Classifier output for example test case of a medulloblastoma (HG) from an external centre, showing normalised metabolite concentrations and discriminant function (D.F.) scores plot for LG (green), HG (blue) and test case (red).