RESEARCH ARTICLE

Characterisation of paediatric brain tumours by their MRS metabolite profiles

Simrandip K. Gill[1,2] | Heather E. L. Rose[1,2] | Martin Wilson[1,2] |
Daniel Rodriguez Gutierrez[3] | Lara Worthington[1,2,4] | Nigel P. Davies[1,2,4] |
Lesley MacPherson[2] | Darren R. Hargrave[5] | Dawn E. Saunders[5] |
Christopher A. Clark[6] | Geoffrey S. Payne[7] | Martin O. Leach[7] |
Franklyn A. Howe[8] | Dorothee P. Auer[9,10,11] | Tim Jaspan[9,11] |
Paul S. Morgan[3,9] | Richard G. Grundy[9] | Shivaram Avula[12] | Barry Pizer[13] |
Theodoros N. Arvanitis[2,14] | Andrew C. Peet[1,2]

[1]Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK

[2]Birmingham Children's Hospital NHS Foundation Trust, Birmingham, UK

[3]Medical Physics, Nottingham University Hospital, Queen's Medical Centre, Nottingham, UK

[4]Department of Imaging and Medical Physics, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK

[5]Paediatric Oncology Unit, Great Ormond Street Hospital For Sick Children, London, UK

[6]Developmental Imaging and Biophysics Section, UCL Great Ormond Street Institute of Child Health, University College London, London, UK

[7]CRUK Cancer Imaging Centre, The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust, London, UK

[8]Neurosciences Research Section, Molecular and Clinical Sciences Research Institute, St George's, University of London, London, UK

[9]The Children's Brain Tumour Research Centre, University of Nottingham, Nottingham, UK

[10]Radiological Sciences, Department of Clinical Neuroscience, University of Nottingham, Nottingham, UK

[11]Neuroradiology, Nottingham University Hospital, Queen's Medical Centre, Nottingham, UK

[12]Department of Radiology, Alder Hey Children's NHS Foundation Trust, Liverpool, UK

[13]Department of Paediatric Oncology, Alder Hey Children's NHS Foundation Trust, Liverpool, UK

[14]Department of Electronic, Electrical and Systems Engineering, University of Birmingham, Birmingham, UK

Correspondence
Andrew C. Peet, Institute of Child Health, University of Birmingham, Clinical Research Block, Whittall Street, Birmingham, B4 6NH, UK.
Email: a.peet@bham.ac.uk

Abstract

[1]H-magnetic resonance spectroscopy (MRS) has the potential to improve the noninvasive diagnostic accuracy for paediatric brain tumours. However, studies analysing large, comprehensive, multicentre datasets are lacking, hindering translation to widespread clinical practice. Single-voxel MRS (point-resolved single-voxel spectroscopy sequence, 1.5 T: echo time [TE] 23–37 ms/135–144 ms, repetition time [TR] 1500 ms; 3 T: TE 37–41 ms/135–144 ms, TR 2000 ms) was performed from 2003 to 2012 during routine magnetic resonance imaging for a suspected brain tumour on 340 children from five hospitals with 464 spectra being available for analysis and 281 meeting

---

Abbreviations: Ala, alanine; Cit, citrate; Cr, creatine; GABA, γ-aminobutyric acid; Glc, glucose; Gln, glutamine; Glu, glutamate; Glx, glutamine + glutamate; Gly, glycine; Gua, guanidinoacetate; Lac, lactate; LMM, lipids and macromolecules; mIns, myo-inositol; sIns, scyllo-inositol; Tau, taurine; tCho, total choline; tNAA, total N-acetylaspartate.

---

quality control. Mean spectra were generated for 13 tumour types. Mann–Whitney U-tests and Kruskal–Wallis tests were used to compare mean metabolite concentrations. Receiver operator characteristic curves were used to determine the potential for individual metabolites to discriminate between specific tumour types. Principal component analysis followed by linear discriminant analysis was used to construct a classifier to discriminate the three main central nervous system tumour types in paediatrics. Mean concentrations of metabolites were shown to differ significantly between tumour types. Large variability existed across each tumour type, but individual metabolites were able to aid discrimination between some tumour types of importance. Complete metabolite profiles were found to be strongly characteristic of tumour type and, when combined with the machine learning methods, demonstrated a diagnostic accuracy of 93% for distinguishing between the three main tumour groups (medulloblastoma, pilocytic astrocytoma and ependymoma). The accuracy of this approach was similar even when data of marginal quality were included, greatly reducing the proportion of MRS excluded for poor quality. Children's brain tumours are strongly characterised by MRS metabolite profiles readily acquired during routine clinical practice, and this information can be used to support noninvasive diagnosis. This study provides both key evidence and an important resource for the future use of MRS in the diagnosis of children's brain tumours.

KEYWORDS

[1]H-magnetic resonance spectroscopy, brain tumours, classification, metabolites, paediatric

1 | INTRODUCTION

Diagnosis of paediatric brain tumours is routinely made by multidisciplinary teams including radiologists, oncologists and histopathologists, with histopathology and molecular characterisation after biopsy or resection considered the gold standard.[1] However, magnetic resonance imaging (MRI) is the method of choice for the noninvasive diagnosis and characterisation of brain tumours in adults and children.[2–4] While this modality illustrates exceptional anatomical and spatial information, conventional sequences provide little information on tumour properties and have limited diagnostic accuracy.[5] The addition of advanced methods, such as diffusion-weighted imaging (DWI), perfusion imaging and magnetic resonance spectroscopy (MRS), can improve diagnostic accuracy and may play an important role in clinical management.[6–8] DWI, in particular, has become an essential technique for image-based diagnosis.[9] Coupling DWI with machine learning has demonstrated its ability to diagnose between three major paediatric brain tumours, medulloblastoma (MB), pilocytic astrocytoma (PA) and ependymomas (EPs), across a multicentre cohort.[10] Preliminary studies of single-voxel proton MRS have shown how it can aid in the characterisation of adult brain tumours,[11,12] leading to an interest in its use in paediatric neuro-oncology. Early spectroscopy studies were performed at long echo time (TE) with limited metabolite information. However, it became evident that short TE, providing more metabolite information and greater signal-to-noise ratio (SNR), was better for the characterisation of some tumours.[13] Analysis is now preferentially performed using automated software for metabolite fitting and water referencing.[14,15] As a result, subsequent studies have been able to discriminate between closely related metabolite profiles of brain tumours and have aided the detection of tumour-specific markers.[13,16] Visual interpretation of index cases against mean MR spectra and mean metabolite profiles of different tumour types have been shown to add clinical value,[17] and the combination of these metabolite profiles with machine learning techniques for brain tumour classification has become popular. The most commonly used multivariate techniques include principal component analysis (PCA) and linear discriminant analysis (LDA),[18] and classifiers developed from these often give excellent accuracy for diagnosis.[19,20]

While there are many studies demonstrating the diagnostic accuracy of MRS, reports are also published that show the added value of MRS compared with conventional MRI. Most commonly, simple viewing of MRS has been evaluated and shown to improve noninvasive diagnosis when compared with radiological review of conventional MRI,[7] including DWI.[17] It has also been shown that radiologists' diagnostic accuracy improves sequentially when conventional MRI is viewed followed by MRS and then a decision support system giving the results of a machine learning classifier.[21] Currently, such decision support systems are not available in a routine clinical setting.

With the advent of molecular testing, the diagnostic categories for children's brain tumours have changed considerably over the past decade.[22] This has led to some categories that were poorly defined on histology, such as supratentorial primitive neuroectodermal tumours (PNETs), being reclassified into several different diagnoses, some previously recognised, such as glioblastoma multiforme (GBM), and others that are new, such as embryonal tumour with multilayered rosettes. At the same time, some tumour types, such as MB, are composed of molecular subtypes, which only partially match the histological subtypes previously established.[23] This evolving picture provides many challenges for studies of noninvasive diagnosis, particularly as children's brain tumours are not common and large datasets take many years to accrue. Increasing sub-categorisation also reduces the numbers in each diagnostic group, making it difficult to obtain sufficient data to build reliable machine learning classifiers. Although there are definite challenges posed by this change in diagnosis, MRS has been shown to improve noninvasive diagnosis and clinical decision-making for children with brain tumours.[17,24]

Many spectroscopy studies are single-centre with small cohort numbers.[25–27] This makes it difficult to demonstrate the robustness and bene-fits of using the technique.[21] Progress has been made with small, multicentre studies showing the diagnostic potential of MRS across multiple cen-tres, at both 1.5 and 3 T,[24,28] and understanding of MRS as a multicentre diagnostic technique is continuing to develop. Inclusion of automated diagnosis,[21] as well as the integration of MRS interpretation in the clinical workflow,[29] has the potential to improve the accuracy of clinical decision-making. Nevertheless, large comprehensive, multicentre datasets for MRS are still unavailable, hampering its clinical use and impact.

The aim of the current study was to determine metabolite profiles of a cohort of paediatric brain tumours using short TE and, where available, long TE spectroscopy at 1.5 and 3 T, thus establishing distinguishing features from a large multiscanner cohort. Demonstrating the effectiveness of the compiled cohort in classifying the three main types of paediatric brain tumours with a view to aiding preoperative diagnosis, and could be used as an adjunct to conventional and advanced MRI.

2 | MATERIALS AND METHODS

2.1 | Patients

The current study received ethical approval and parental consent was obtained (5512M[B] and 04/MRE04/41). Single-voxel ^{1}H MRS was acquired during routine MRI for a suspected brain tumour, prior to treatment, from July 2003 to October 2012. In total, 340 children were scanned, with 464 spectra acquired for analysis. Where surgery was undertaken as part of the patient's routine clinical management, tumour diagnosis and grad-ing were obtained from histopathology using the diagnostic criteria relevant at the time.[1] For those cases where surgery was not performed or a tissue diagnosis was inconclusive, diagnosis was made on clinical and imaging findings, after review by the local multidisciplinary team.

2.2 | MRS acquisition

Scans were acquired at both 1.5 and 3 T at five principal treatment centres, then collected and analysed retrospectively. At 1.5 T, spectroscopy was performed on five different scanner types at five centres. The scanners used were the Siemens Symphony, Siemens Avanto, GE Signa Excite, Philips Intera and Philips Achieva. At 3 T, all spectroscopy was performed on the Philips Achieva from three centres. The acquisition protocol was agreed at the start of the project, but with sufficient flexibility to allow implementation on all scanners (Table S1). MRS was acquired with a point-resolved single-voxel spectroscopy (PRESS) sequence after conventional MRI including T1- and T2-weighted and T1-weighted postcontrast sequences. These images were used to guide the placement of the MRS voxel within the solid-appearing component of the tumour, avoiding areas of cyst, necrosis or normal-appearing brain tissue. The voxel was placed to avoid close proximity to scalp or other fatty tissue to minimise the risk of lipid contamination of the MRS signal. If the tumour was sufficiently large to allow for some flexibility in voxel placement, contrast enhance-ment and low apparent diffusion coefficient (ADC) were used to guide optimal position. The MRS parameters at 1.5 T were short TE (23–37 ms), long TE (135–144 ms), TR 1500 ms, 128–256 signal averages, spectral bandwidth 2000 or 2500 Hz and vector length 2048. The parameters at 3 T were short TE (37–41 ms), long TE (135–144 ms), TR 2000 ms, 64 to 196 signal averages, spectral bandwidth 2000 Hz and vector length 2048. Voxel size ranged from 3.38 to 8 cm^3 at 1.5 T and from 2.2 to 8 cm^3 at 3 T. B_0 shimming and water suppression were undertaken according to the automated methods available on the scanners. Free induction decays made available routinely on each scanner were exported for analysis and are referred to here as the raw data. The signal from each receiver coil is usually combined by the scanner software prior to export.

2.3 | MRS quantification

Raw data were analysed using TARQUIN (version 4.2 with extended basis set), fitting to a linear combination of 18 metabolite basis functions generated at the correct field strength and TE: alanine (Ala), aspartate (Asp), citrate (Cit), creatine (Cr), γ-aminobutyric acid (GABA),

glycerophosphocholine (GPC), glucose (Glc), glutamine (Gln), glutamate (Glu), glycine (Gly), guanidoacetate (Gua), myo-inositol (mIns), lactate (Lac), N-acetylaspartate (NAA), N-acetyl-aspartyl-glutamate (NAAG), phosphocholine (PCh), scyllo-inositol (sIns) and taurine (Tau) with an additional nine lipid (Lip) and macromolecular (MM) components: Lip0.9, Lip1.3a, Lip1.3b, Lip2.0, MM0.9, MM1.2, MM1.4, MM1.7 and MM2.0 (denoted by their chemical shift, e.g., Lip0.9 is a lipid peak at 0.9 ppm).[15,30,31] The metabolite basis sets were simulated using PRESS, with ideal pulses and the correct TE and TE1. Where TE1 was not available, it was set as a default to 12 ms. The following combinations of metabolites and lipid macromolecular components were reported: GPC + PCh as total choline (tCho), Glu + Gln as Glx, NAA + NAAG as total N-acetylaspartate (tNAA), Lip0.9 + MM0.9 as LMM0.9, Lip1.3a + Lip1.3b + MM1.2 + MM1.4 as LMM1.3 and Lip2.0 + MM2.0 as LMM 2.0. Quantitation was performed relative to a water reference spectrum with an assumed water content of 35,880 mM, the assumed average water content of white matter and default value in TARQUIN. Corrections for T2 relaxation times of metabolites and water were applied using the default values for TARQUIN (water 80 ms, metabolites 400 ms). No correction for T1 relaxation times was applied in the analysis because the TR values were selected to mitigate the differences in T1 at the two field strengths.

2.3.1 | Quality control criteria

A pipeline showing the quality control (QC) steps and the resulting cohort numbers is summarised in Figure 1; 363 1.5-T and 101 3-T MRS scans were assessed using the following QC criteria. Spectra failed QC if they did not have TARQUIN values for full-width half-maximum (FWHM) at less than 0.150 ppm, SNR more than 5 and a measure of fit quality (Q) less than 2.5, or if on visual inspection they exhibited baseline abnormalities and artefacts. The FWHM cut-off was chosen to give reasonable quantification of the tCho and total Cr peaks around 3.2 and 3.0 ppm, respectively, which is important clinically. In all, 140/363 failed the QC checks at 1.5 T, and 43/101 failed at 3 T. Spectra were then excluded if the diagnostic groups had less than three cases, the diagnostic group was unknown or insufficient diagnostic information was available for inclusion, or finally if the spectrum was a replication of same spectrum for the same patient (1.5 T, $n = 56$; 3 T, $n = 31$). Tumour types with three or more cases ($n = 13$) were included in the analysis with 142 short TE and 25 long TE for 1.5 T and 17 short TE and 10 long TE for 3 T, and henceforth patient numbers refer to that cohort. For short TE MRS data at 1.5 T, the diagnostic categories are given in Table 1. There were 25 cases with long TE data at 1.5 T; 10 PAs, 10 MBs and five diffuse intrinsic pontine gliomas (DIPGs). At 3 T, a total of 17 cases had short TE MRS, eight of which were PA and nine MB, and 10 cases had long TE MRS, five of which were PA and five MB.

A relatively high proportion of spectra were excluded for not meeting the QC criteria. The main reasons for this were poor SNR and poor FWHM. These were often a consequence of the tumour itself, with low cellularity leading to poor SNR and calcification, or haemorrhage leading to poor FWHM. Some childhood brain tumours are in locations where it is difficult to perform spectroscopy because of the proximity of bone, such as the suprasellar region or the inferior brain stem.
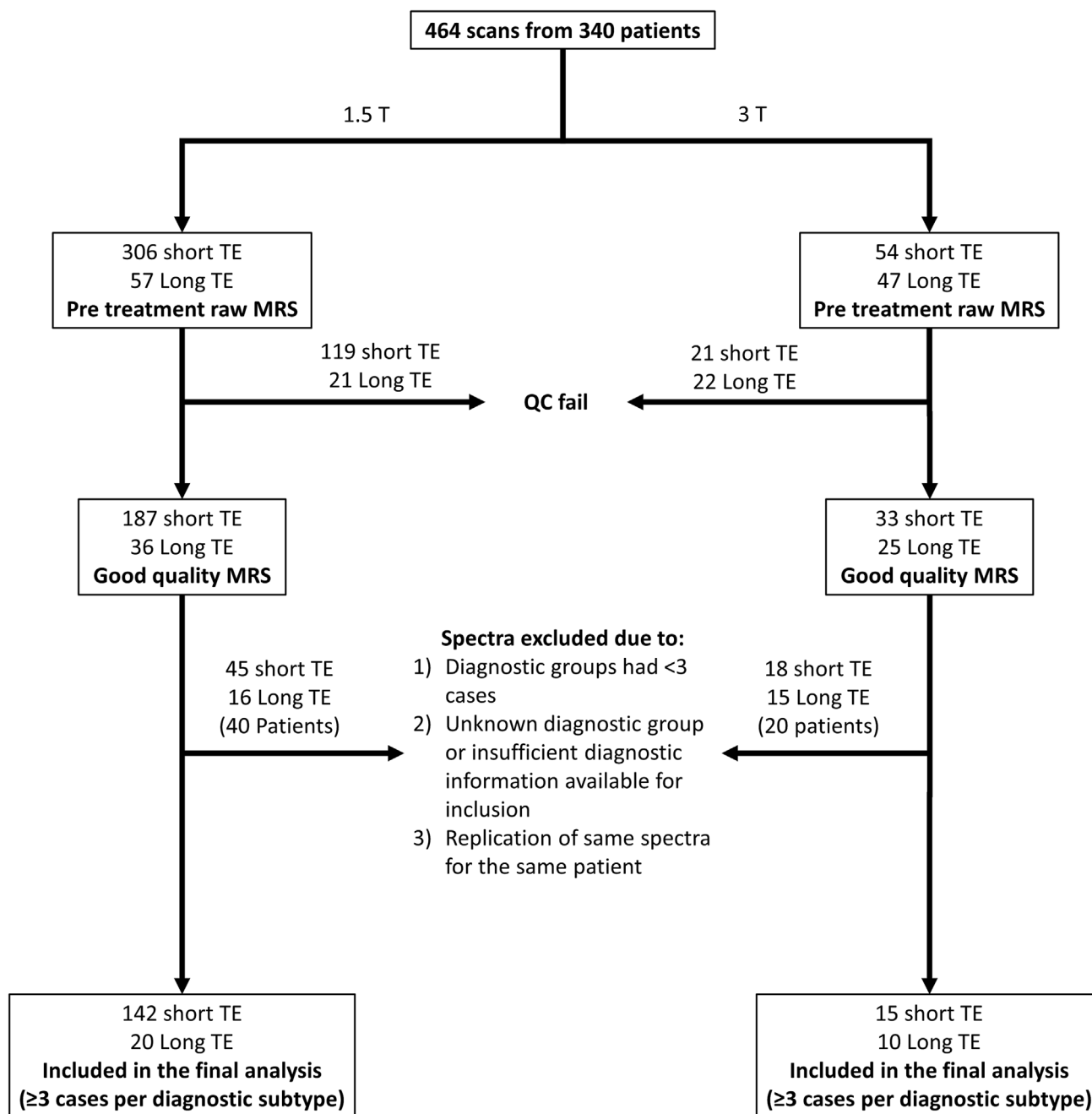
2.4 | Statistical analysis

Statistical analysis was performed using R Statistical Software (version 2.13.0, 2011)[32] on metabolite concentrations (mM). Metabolites with a Cramer–Rao lower bound (CRLB) of less than 30% in at least two subjects were included in the analysis. These criteria resulted in the exclusion of the metabolites Asp and GABA from the final analysis. Lipids and macromolecules at 0.9 ppm (LMM0.9), 1.3 ppm (LMM1.3) and 2.0 ppm (LMM2.0) are reported as combined signals.

Mean metabolite values were compared between different groups using Mann–Whitney U-tests. For MRS acquired using short TE at 1.5 T, each tumour type was compared with all the other tumour types combined into a single group (All Other). MB and PA were compared at 1.5-T long TE and at 3-T short and long TE. A comparison was also made between tumour types that are often included in a differential diagnosis on conventional MRI sequences, 1.5-T short TE (Table S2). A formal correction for multiple comparisons was not undertaken, but a reduced cut-off of p less than 0.01 was used to recognise the challenge and to provide consistency with previous publications.[13]

Receiver operator characteristic (ROC) curves were constructed for pairs of tumour groups and the area under the curve (AUC) was determined, with AUCs of more than 0.8 reported. Cut-off values (maximum sensitivity plus specificity) were determined to optimise discrimination. Cut-off value accuracy was determined using leave-one-out cross-validation (LOOCV).[33]

A diagnostic classifier was constructed to discriminate between the three most common paediatric tumour types, MB, PA and EP, by using metabolite profiles from short TE 1.5-T MRS. Principal components accounting for 95% of the variance were used in the LDA. Classifier accuracy was determined using LOOCV and is reported as the balanced accuracy rate (BAR).[34] There were insufficient 3-T data to build a classifier and test it using LOOCV; however, we performed an exploratory study of field strength dependence, in which the 3-T cases were used as a test set in the 1.5-T classifier. To evaluate interscanner MRS variability, spectra acquired at 1.5 T and short TE from MB on four of the scanners (Siemens Symphony, Siemens Avanto, GE Signa Excite and Philips Intera) were compared using a Kruskal–Wallis test.

```
                        ┌─────────────────────────────┐
                        │  464 scans from 340 patients │
                        └─────────────────────────────┘
                    1.5 T                            3 T

┌─────────────────────────┐              ┌─────────────────────────┐
│ 306 short TE             │              │ 54 short TE             │
│ 57 Long TE              │              │ 47 Long TE              │
│ Pre treatment raw MRS   │              │ Pre treatment raw MRS   │
└─────────────────────────┘              └─────────────────────────┘

        119 short TE                          21 short TE
        21 Long TE          QC fail           22 Long TE

┌─────────────────────────┐              ┌─────────────────────────┐
│ 187 short TE             │              │ 33 short TE             │
│ 36 Long TE              │              │ 25 Long TE              │
│ Good quality MRS        │              │ Good quality MRS        │
└─────────────────────────┘              └─────────────────────────┘
```

Spectra excluded due to:

45 short TE   1) Diagnostic groups had <3   18 short TE
16 Long TE   cases   15 Long TE
(40 Patients) 2) Unknown diagnostic group   (20 patients)
                  or insufficient diagnostic
                  information available for
                  inclusion
               3) Replication of same spectra
                  for the same patient

```
┌────────────────────────────────┐     ┌────────────────────────────────┐
│ 142 short TE                    │     │ 15 short TE                     │
│ 20 Long TE                     │     │ 10 Long TE                     │
│ Included in the final analysis │     │ Included in the final analysis │
│ (≥3 cases per diagnostic subtype)│   │ (≥3 cases per diagnostic subtype)│
└────────────────────────────────┘     └────────────────────────────────┘
```

FIGURE 1 Flow diagram of patients studied at 1.5 and 3 T. MRS, magnetic resonance spectroscopy; QC, quality control; TE, echo time.

To test the appropriateness of the QC criteria used in this study, with relation to effect on classification accuracy, all cases of MB, PA and EP acquired at 1.5 T that failed were rereviewed and placed into three categories: large, intermediate and borderline QC fail. The definitions were:

1. Large QC fail: A spectrum where major metabolite peaks were not discernible.
2. Intermediate QC fail: A spectrum where major peaks were discernible but artefacts or broad peaks were seen in more than one region of the spectrum.
3. Borderline QC fail: A spectrum where the voxel contained a small proportion of normal brain, or the voxel was over a metastatic lesion and/or artefacts were seen around the water peak.

These cases were incorporated into the original classifier as test cases.

TABLE 1 Patient demographics of the paediatric brain tumours in the cohort with short echo time ^{1}H MRS at 1.5 T.

	Subjects	WHO grade	Age (years) (mean ± SD)	Age range (years)	Sex (F/M)
All cases	142	I–IV	7.4 ± 4.5	0–17	58/84
Tumour type					
Pilocytic astrocytoma (PA)[a]	35	I/II	7.2 ± 4.3	0–16	17/18
Medulloblastoma (MB)[b]	42	IV	6.9 ± 3.6	1–15	12/30
Ependymoma (EP)[c]	10	II/III	5.1 ± 5.7	1–16	4/6
Diffuse pontine glioma (DIPG)	12	N/A	6.8 ± 3.4	3–14	7/5
Anaplastic astrocytoma (AA)	3	III	9.5 ± 5.8	4–15	2/1
Atypical teratoid rhabdoid tumour (ATRT)	5	IV	1.6 ± 1.6	0–4	2/3
Choroid plexus papilloma (CPP)	3	I	3.0 ± 2.6	1–6	0/3
CNS primitive neuroectodermal tumour (CNS PNET)	3	IV	8.3 ± 5.5	2–13	2/1
Craniopharyngioma (CRANIO)[d]	3	I	7.4 ± 2.3	5–10	2/1
Diffuse astrocytoma (DA)[e]	5	II/III	11.0 ± 6.6	2–17	4/1
Ganglioglioma (GG)	5	I	11.8 ± 4.6	5–16	3/2
Germinoma (GER)	7	N/A	12.9 ± 2.2	8–16	1/6
Glioblastoma multiforme (GBM)	9	IV	8.4 ± 4.5	0–14	2/7

Abbreviations: ^{1}H MRS, proton magnetic resonance spectroscopy; F, female; M, male; N/A, not diagnosed on the basis of histopathology so no WHO grading available; SD, standard deviation; WHO, World Health Organisation.
[a]Includes pilomyxoid.
[b]Includes classic, desmoplastic and large cell.
[c]Includes tanycytic and anaplastic.
[d]Includes classic and adamantinomatous.
[e]Includes classic, fibrillary and gemistocytic.

3 | RESULTS

3.1 | Metabolic features

Mean metabolite concentrations (mM) for diagnostic groups are given in Table 2 with mean spectra in Figures 2 to 6. All these figures show 1.5-T spectra, except for Figure 3, which shows 3-T spectra. Notable features are reported below; comments apply to short TE MRS at 1.5 T and the comparison is with all other cases in the cohort unless specified.

3.1.1 | Pilocytic astrocytoma

PA had lower Cr, tCho and lipids + MMs, particularly LMM1.3 and LMM2.0 ($p < 0.0001$) (Figure 2A). At long TE, less Cr, tCho and Gly ($p < 0.01$) was observed in PA than in MB (Figure S1). At 3-T short TE, PAs had lower Cr, tCho, Gly and LMM1.3 ($p < 0.01$) and higher tNAA ($p < 0.001$) compared with MBs (Figure 3A,B and Table S3). At 3-T long TE, MRS demonstrated lower Cr ($p < 0.01$) in PAs than in MBs.

3.1.2 | Medulloblastoma

MB had higher Cr, tCho, Gly, LMMs and Tau and significantly lower mIns ($p < 0.0001$) (Figure 4A). The most discriminatory feature of the MB mean spectra was high Tau. Gly makes a greater contribution than mIns in MBs (Figure 3B,D).

3.1.3 | Ependymoma

EP spectra showed higher mIns and Cr ($p < 0.01$) (Figure 5). Grade III had a lower Cr ($p < 0.01$) than grade II (Figure 5B,C). The higher LMM1.3 and lower mIns seen in the mean spectra of grade III relative to grade II were not significant.

TABLE 2 Mean absolute concentrations (mM) ± SD of short echo time pretreatment paediatric brain tumours at 1.5 T compared using the Mann–Whitney U-test.

Metabolites (mM ± SD)	AA (n = 3)	ATRT (n = 5)	CPP (n = 3)	CNS PNET (n = 3)	CRANIO[a] (n = 3)	DA[b] (n = 5)	DIPG (n = 12)	EP[c] (n = 10)	GG (n = 5)	GER (n = 7)	GBM (n = 9)	MB[d] (n = 42)	PA[e] (n = 35)	All (n = 142)
Ala	0.1 ± 0.2	0.2 ± 0.5	1.0 ± 1.3	0.4 ± 0.5	0.1 ± 0.2	0.1 ± 0.2	0.3 ± 0.4	0.2 ± 0.4	0.5 ± 0.3	0.04 ± 0.1	0.3 ± 0.5	0.5 ± 0.9	0.4 ± 0.7	0.4 ± 0.7
Cit	2.1 ± 1.6	0.5 ± 0.3	0.3 ± 0.3	0.3 ± 0.3	0.1 ± 0.1	0.6 ± 0.6	**1.0 ± 0.6****	0.7 ± 0.4	0.5 ± 0.5	0.8 ± 0.4	0.6 ± 0.5	0.6 ± 0.4	**0.4 ± 0.3****	0.6 ± 0.5
Cr	4.2 ± 1.8	1.5 ± 2.2	0.8 ± 1.3	0.6 ± 0.5	**0.2 ± 0.1****	2.2 ± 1.5	2.9 ± 1.4	**4.4 ± 2.4****	2.4 ± 0.7	1.3 ± 0.4	1.2 ± 0.9	**2.9 ± 1.3******	**0.9 ± 0.8******	2.1 ± 1.7
tCho	1.2 ± 0.4	1.9 ± 1.7	2.2 ± 2.3	1.6 ± 1.5	**0.3 ± 0.3****	1.0 ± 0.4	1.2 ± 0.6	2.0 ± 1.1	1.3 ± 0.5	1.2 ± 0.6	1.8 ± 1.3	**3.6 ± 1.5******	**1.1 ± 0.4******	2.0 ± 1.5
Glc	3.5 ± 1.0	2.7 ± 4.4	3.3 ± 0.8	0.01 ± 0.02	0.9 ± 0.9	1.4 ± 1.8	2.7 ± 1.5	1.1 ± 1.2	1.3 ± 1.3	2.6 ± 1.6	1.4 ± 1.5	**0.6 ± 1.0******	**2.8 ± 1.9******	1.7 ± 1.9
Gln	3.0 ± 1.0	2.0 ± 2.4	1.4 ± 1.8	2.2 ± 0.1	2.2 ± 1.5	3.2 ± 1.1	3.7 ± 4.2	3.2 ± 1.9	4.0 ± 1.8	3.1 ± 2.1	4.1 ± 3.0	2.9 ± 2.6	3.5 ± 1.9	3.2 ± 2.4
Glu	4.1 ± 2.5	2.4 ± 2.1	2.7 ± 1.7	2.1 ± 2.6	0.4 ± 0.7	1.8 ± 1.4	2.0 ± 1.3	2.8 ± 1.8	3.1 ± 1.1	**4.9 ± 1.4*****	1.2 ± 1.6	3.0 ± 2.2	1.7 ± 1.4	2.4 ± 2.0
Glx	7.1 ± 1.7	4.5 ± 2.7	4.1 ± 2.6	4.3 ± 2.5	2.6 ± 0.9	5.0 ± 1.7	5.6 ± 3.8	5.9 ± 3.4	7.2 ± 2.6	8.1 ± 2.5	5.3 ± 3.2	5.9 ± 3.4	5.2 ± 2.2	5.6 ± 2.9
Gly	0	0.9 ± 1.2	3.3 ± 3.3	1.6 ± 2.8	0.5 ± 0.6	0.04 ± 0.1	1.2 ± 2.1	3.4 ± 3.3	1.6 ± 1.1	1.4 ± 1.6	1.3 ± 1.0	**3.5 ± 2.3******	**0.6 ± 1.1*****	1.9 ± 2.2
Gua	0.4 ± 0.4	1.4 ± 2.7	0.4 ± 0.7	0.5 ± 0.9	1.0 ± 0.9	0.4 ± 0.5	0.7 ± 2.1	0.6 ± 1.0	0.5 ± 0.4	0.5 ± 1.3	0.4 ± 0.7	**1.1 ± 1.1******	0.2 ± 0.5	0.7 ± 1.1
mIns	**11.0 ± 2.0****	1.1 ± 1.8	8.1 ± 9.7	1.9 ± 2.8	0	2.8 ± 2.3	**6.0 ± 3.3******	**6.1 ± 4.9****	1.3 ± 1.9	1.2 ± 1.3	0.7 ± 0.9	**0.8 ± 2.0******	0.7 ± 0.9	2.1 ± 3.4
Lac	0.6 ± 0.6	3.4 ± 2.3	2.7 ± 3.6	4.8 ± 3.0	3.7 ± 1.5	1.7 ± 1.4	**0.7 ± 0.9******	1.6 ± 1.0	1.9 ± 0.7	2.8 ± 4.6	**5.2 ± 4.0****	2.7 ± 1.8	2.3 ± 1.9	2.5 ± 2.3
LMM0.9	3.0 ± 1.7	10.4 ± 4.4	7.5 ± 8.3	6.3 ± 2.3	6.2 ± 3.0	3.7 ± 2.8	**2.2 ± 1.7******	5.0 ± 3.5	2.9 ± 1.0	5.0 ± 4.2	9.1 ± 5.1	**6.9 ± 3.3******	**4.3 ± 2.3***	5.6 ± 3.7
LMM1.3	3.6 ± 2.1	**41.0 ± 25****	12.6 ± 18	15.0 ± 14	9.4 ± 7.0	5.4 ± 6.2	**3.0 ± 2.3******	15.0 ± 16	2.9 ± 2.1	18.0 ± 22	14.0 ± 6.1	**17.0 ± 13******	**6.1 ± 4.3******	12.0 ± 13
LMM2.0	7.9 ± 1.3	12.0 ± 5.7	5.8 ± 3.0	7.2 ± 2.6	5.7 ± 2.0	9.9 ± 6.4	**6.0 ± 3.0*****	12.0 ± 7.5	8.4 ± 2.6	9.4 ± 3.2	12.0 ± 2.4	**11.0 ± 3.0******	**6.5 ± 2.8******	9.1 ± 4.2
tNAA	1.5 ± 1.7	1.1 ± 0.9	0.3 ± 0.5	1.2 ± 0.4	1.0 ± 0.7	1.8 ± 0.5	**2.5 ± 1.1*****	1.1 ± 0.7	1.8 ± 0.9	1.2 ± 0.7	1.0 ± 0.9	1.3 ± 0.8	1.6 ± 0.8	1.4 ± 0.9
sIns	0.2 ± 0.2	0.1 ± 0.2	0.4 ± 0.5	0.2 ± 0.2	0	0.08 ± 0.08	0.3 ± 0.3	0.5 ± 0.4	0.1 ± 0.2	0.3 ± 0.4	0.1 ± 0.1	**0.5 ± 0.5******	**0.05 ± 0.1******	0.3 ± 0.4
Tau	1.4 ± 1.4	1.4 ± 2.1	1.0 ± 0.9	0.4 ± 0.7	0.4 ± 0.4	0.2 ± 0.4	0.4 ± 0.5	0.9 ± 1.3	0.4 ± 0.4	1.9 ± 1.5	**0.1 ± 0.3****	**2.7 ± 2.3******	0.6 ± 1.0	1.3 ± 1.79

Note: While the feature around 2 ppm is fitted to NAA in all spectra, macromolecules are known to be the major contributor to this feature in PA and CRANIO. Bold type indicates that there is a significant difference.

Abbreviations: AA, anaplastic astrocytoma; Ala, alanine; ATRT, atypical teratoid rhabdoid tumour; Cit, citrate; CNS PNET, CNS primitive neuroectodermal tumour; CPP, choroid plexus papilloma; Cr, creatine; CRANIO, craniopharyngioma; DA, diffuse astrocytoma; DIPG, diffuse intrinsic pontine glioma; EP, ependymoma; GBM, glioblastoma multiforme; GER, germinoma; GG, ganglioglioma; Glc, glucose; Gln, glutamine; Glu, glutamate; Glx, glutamate + glutamate; Gly, glycine; Gua, guanidinoacetate; Lac, lactate; LMM, lipids and macromolecules; MB, medulloblastoma; mIns, myo-inositol; NAA, N-acetylaspartate; PA, pilocytic astrocytoma; SD, standard deviation; sIns, scyllo-inositol; Tau, taurine; tCho, total choline; tNAA, total N-acetylaspartate.

[a]Includes classic and adamantinomatous.

[b]Includes classic, fibrillary and gemistocytic.

[c]Includes tanycytic and anaplastic.

[d]Includes classic, desmoplastic and large cell.

[e]Includes pilomyxoid.

*$p < 0.05$. **$p < 0.01$, versus all the other tumour types combined into a single group. ***$p < 0.001$, versus all the other tumour types combined into a single group. ****$p < 0.0001$, versus all the other tumour types combined into a single group.
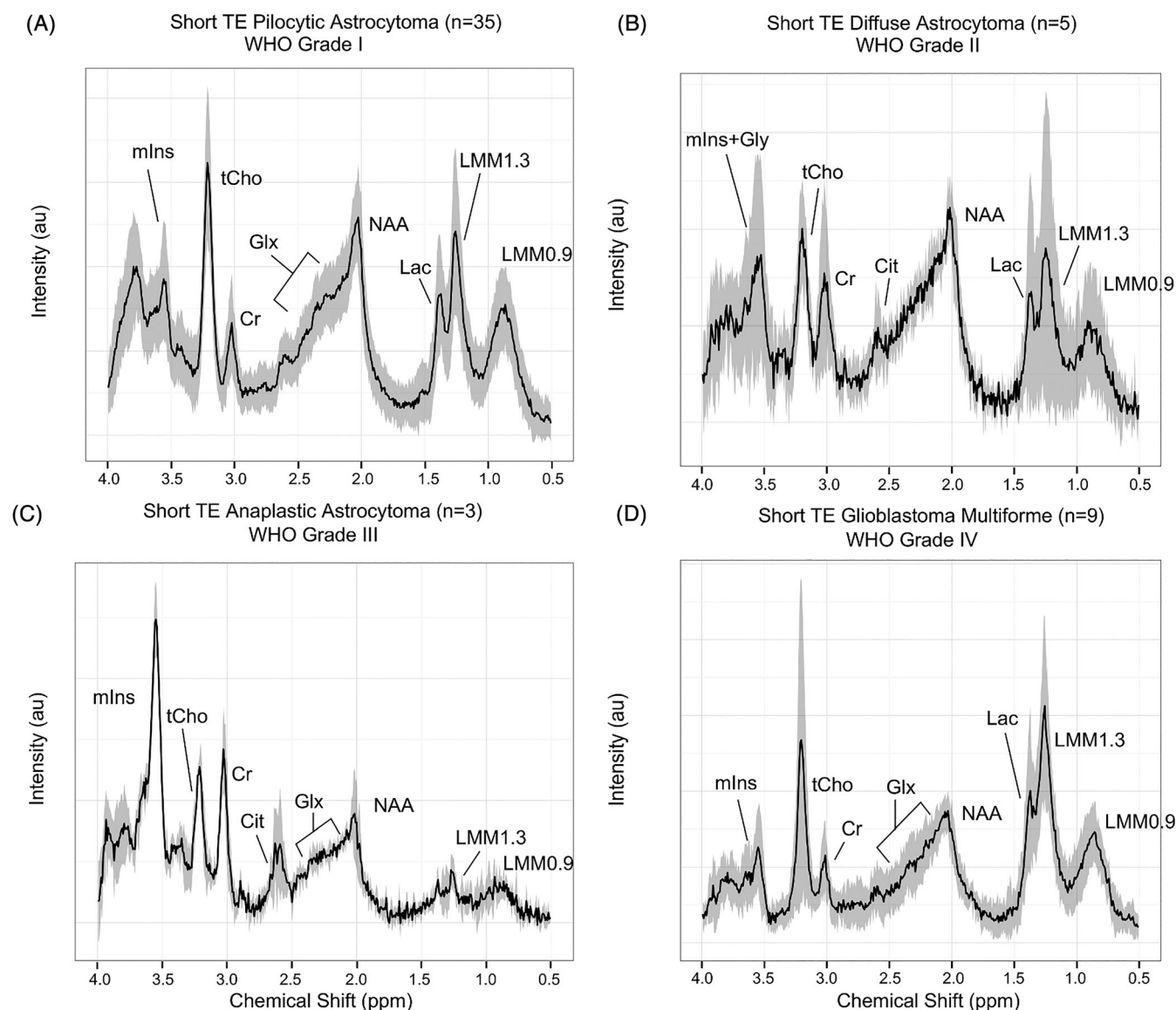
FIGURE 2 Mean short echo time (TE) 1.5-T spectra for (A) Pilocytic astrocytomas ($n = 35$), (B) Diffuse astrocytomas ($n = 5$), (C) Anaplastic astrocytomas ($n = 3$), and (D) Glioblastoma multiforme ($n = 9$), with a solid black line indicating the mean spectra and SD indicated by the shaded region. Note that in pilocytic astrocytoma, the peak labelled NAA is known to be consistent with an assignment to macromolecules. Cit, citrate; Cr, creatine; Glx, glutamine + glutamate; Gly, glycine; Lac, lactate; LMM, lipids and macromolecules; mIns, myo-inositol; NAA, N-acetylaspartate; SD, standard deviation; tCho, total choline; WHO, World Health Organisation.

3.1.4 | Diffuse pontine glioma

DIPG spectra possessed significantly higher Cit ($p < 0.01$), mIns ($p < 0.0001$) and tNAA ($p < 0.001$) and lower Lac, LMM0.9 and LMM1.3 ($p < 0.0001$) (Figure S2A). Long TE DIPG spectra are demonstrated in Figure S2B.

3.1.5 | Anaplastic astrocytoma

Anaplastic astrocytoma (AA) showed elevated mIns ($p < 0.01$) (Figure 2C).

3.1.6 | Atypical teratoid rhabdoid tumour

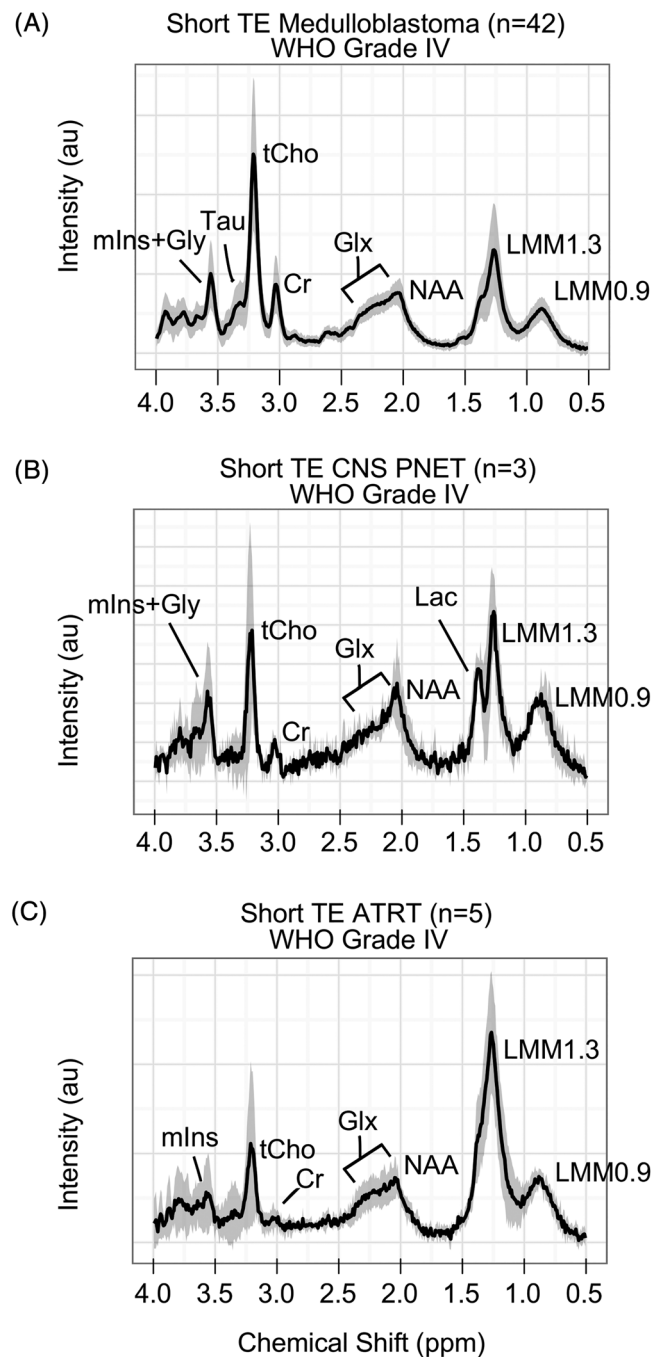Atypical teratoid rhabdoid tumour (ATRT) exhibited elevated LMM1.3 ($p < 0.01$) (Figure 4C).
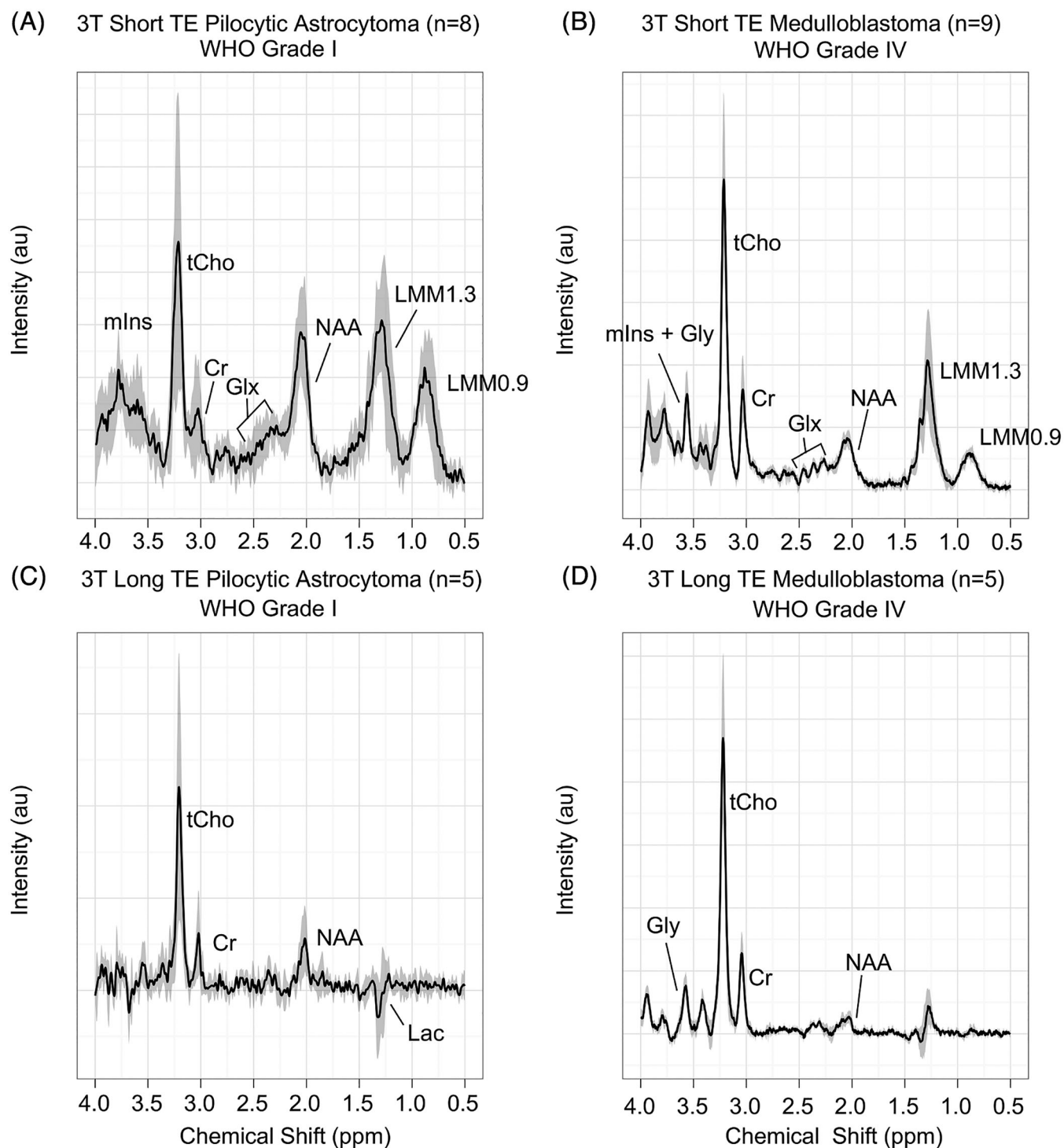
FIGURE 3 Mean 3-T spectra for (A) Pilocytic astrocytomas at short echo time (TE) ($n = 8$), (B) Medulloblastoma at short TE ($n = 9$), (C) Pilocytic astrocytoma at long TE ($n = 5$), and (D) Medulloblastoma at long TE ($n = 5$), with a solid black line indicating the mean spectra and SD indicated by the shaded region. Note that in pilocytic astrocytoma, the peak labelled NAA is known to be consistent with an assignment to macromolecules. Cr, creatine; Glx, glutamine + glutamate; Gly, glycine; Lac, lactate; LMM, lipids and macromolecules; mIns, myo-inositol; NAA, N-acetylaspartate; SD, standard deviation; tCho, total choline; WHO, World Health Organisation.

3.1.7 | Central nervous system PNET

No significantly different variables ($p > 0.01$) were identified when comparing central nervous system (CNS) PNETs with All Other (Figure 4B).
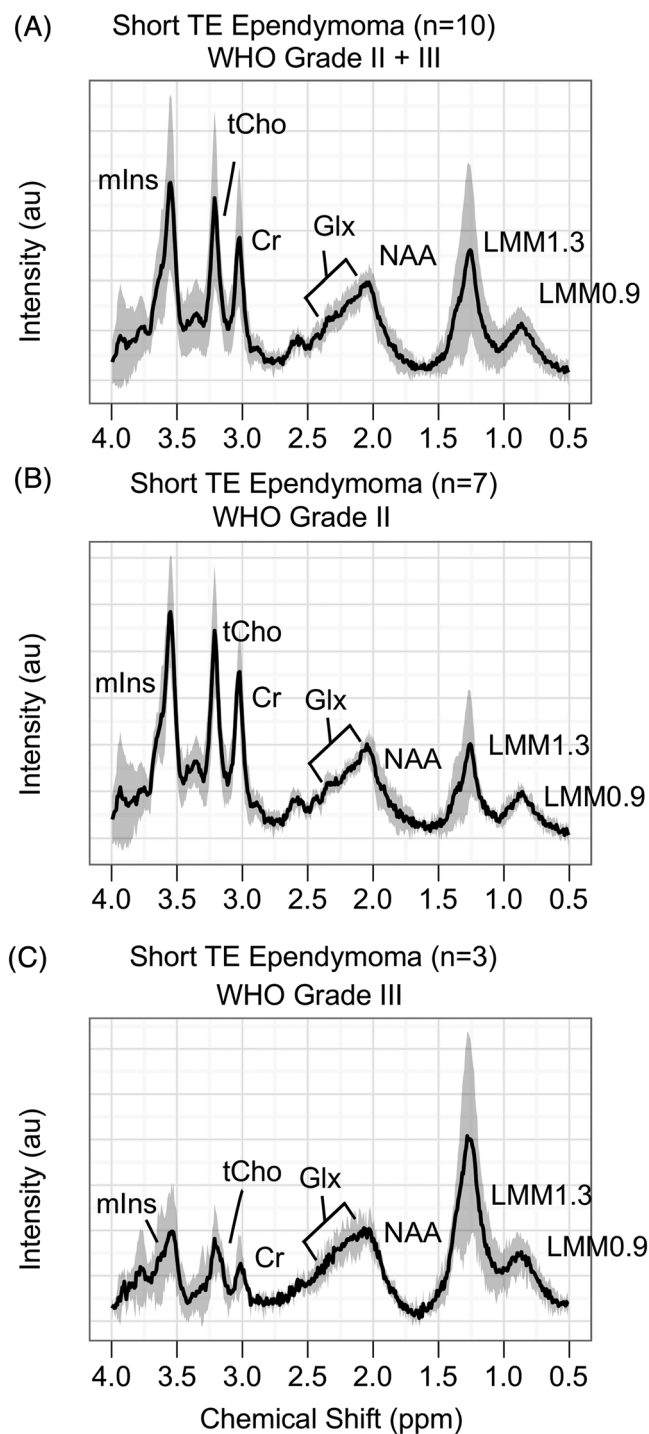
FIGURE 4 Mean short echo time (TE) 1.5-T spectra for (A) Medulloblastomas (*n* = 42), (B) Central nervous system primitive neuroectodermal tumours (CNS PNET) (*n* = 3), and (C) Atypical teratoid rhabdoid tumours (ATRT) (*n* = 5), with a solid black line indicating the mean spectra and SD indicated by the shaded region. Cr, creatine; Glx, glutamine + glutamate; Gly, glycine; Lac, lactate; LMM, lipids and macromolecules; mIns, myo-inositol; NAA, N-acetylaspartate; SD, standard deviation; Tau, taurine; tCho, total choline; WHO, World Health Organisation.

3.1.8 | Craniopharyngioma

Craniopharyngioma had lower Cr and tCho ($p < 0.01$) (Figure 6B).

3.1.9 | Glioblastoma multiforme

GBM had significantly higher Lac ($p < 0.01$) and lower Tau ($p < 0.01$) (Figure 2D).

FIGURE 5 Mean short echo time (TE) 1.5-T spectra for (A) All ependymomas ($n = 10$), (B) World Health Organisation (WHO) grade II ependymomas ($n = 7$), and (C) WHO grade III ependymomas ($n = 3$), with a solid black line indicating the mean spectra and SD indicated by the shaded region. Cr, creatine; Glx, glutamine + glutamate; LMM, lipids and macromolecules; mIns, myo-inositol; NAA, N-acetylaspartate; SD, standard deviation; tCho, total choline.

3.2 | Comparison of mean metabolite values between tumours commonly found in a differential diagnosis

Results comparing specific tumours are reported in Table 3. The following comparisons are of particular clinical relevance. MB demonstrated higher Gly compared with ATRT. MB demonstrated higher tCho and Tau and lower mIns when compared with EP. PA demonstrated lower LMM1.3 and LMM2.0 compared with GBM.
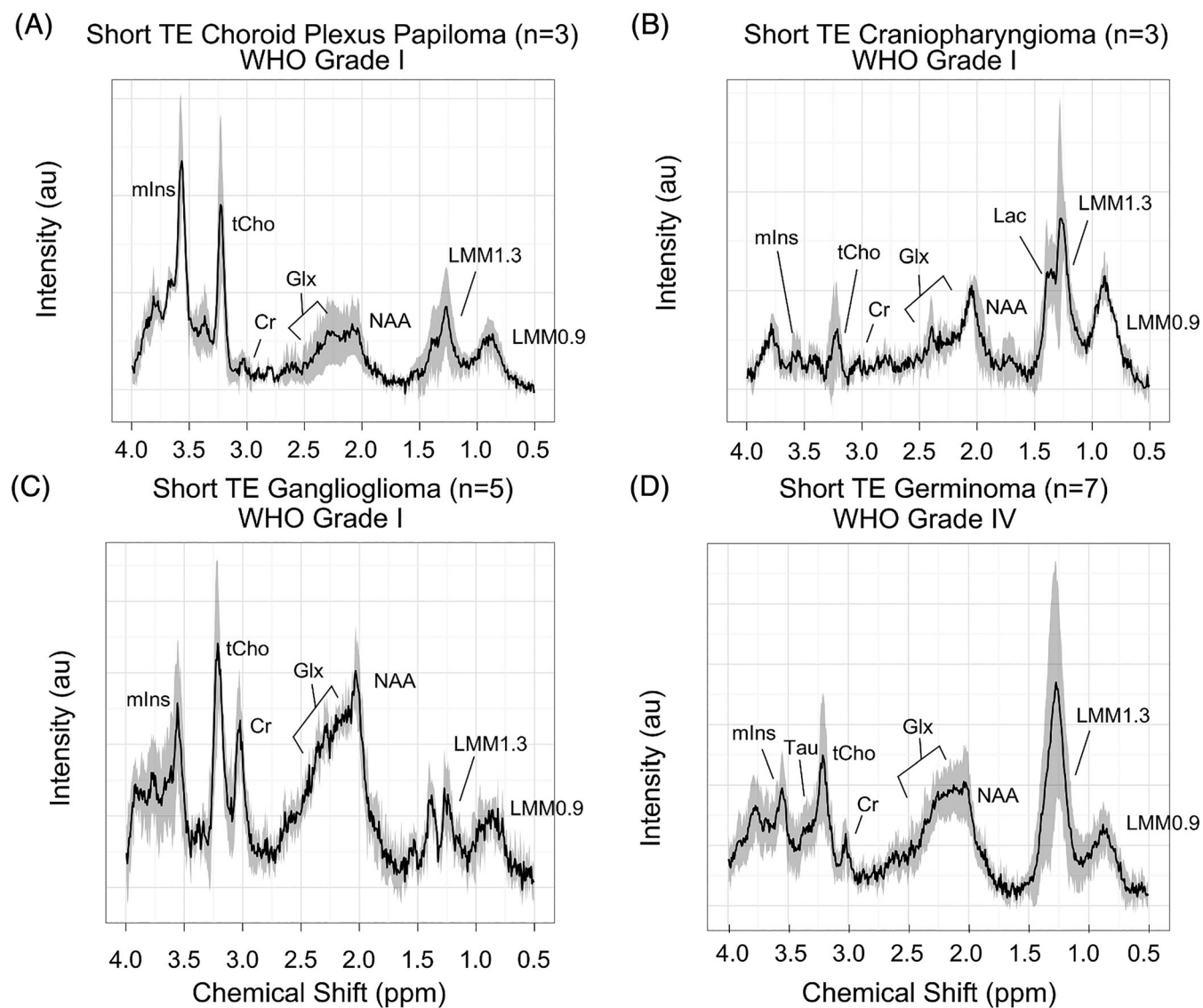
FIGURE 6 Mean short echo time (TE) 1.5-T spectra for (A) Choroid plexus papilloma ($n = 3$), (B) Craniopharyngioma ($n = 3$), (C) Ganglioglioma ($n = 5$), and (D) Germinoma ($n = 7$), with a solid black line indicating the mean spectra and SD indicated by the shaded region. Note that in craniopharyngioma, the peak labelled NAA is known to be consistent with an assignment to macromolecules. Cr, creatine; Glx, glutamine + glutamate; Lac, lactate; LMM, lipids and macromolecules; mIns, myo-inositol; NAA, N-acetylaspartate; SD, standard deviation; Tau, taurine; tCho, total choline; WHO, World Health Organisation.

3.3 | Classification of differential diagnoses using individual metabolites

The accuracy of individual metabolites to discriminate between specific pairs of tumours is given in Table 4 and Figure 7. Myo-inositol was found to be the best discriminator between MB and EP. Creatine, mIns, tCho and LMM2.0 were found to discriminate between PA and EP. Lactate, LMM0.9, LMM1.3 and LMM2.0 were found to discriminate between DIPG and GBM, although all LMMs demonstrated a low specificity. The LOOCV error rate was more than 30% for each pair of comparisons.

3.4 | Classification of common brain tumours by metabolite profile

The diagnostic classifier built from short TE MRS at 1.5 T showed good separation between the three tumour types of MB, PA and EP (Figure 8A). MBs were distinguished primarily by discriminant function 1, with high tCho, sIns, Gua and Tau, and PAs with high tNAA and Lac

TABLE 3 Differentiation between tumour metabolite profiles at 1.5 T comparing those that are often included in differential diagnoses on conventional radiology using the Mann–Whitney U-test with p values reported. Metabolite changes are relative to the tumour type in the left column.

	Atypical teratoid rhabdoid tumour (n = 5)	Ependymoma[b] (n = 10)	Pilocytic astrocytoma[c] (n = 35)
Medulloblastoma[a] (n = 42)	↑ Gly**	↑ tCho**	↑ Cit**
		↓ mIns***	↑ Cr****
		↑ Tau**	↑ tCho****
			↓Glc****
			↑ Glu**
			↑ Gly****
			↑ Gua****
			↑ LMM0.9****
			↑ LMM1.3****
			↑ LMM2.0****
			↑ sIns****
			↑Tau****

	Ependymoma[b] (n = 10)	Glioblastoma multiforme (n = 9)
Pilocytic astrocytroma[c] (n = 35)	↓ Cr****	↓ LMM1.3****
	↓ tCho**	↓ LMM2.0****
	↓Glc**	
	↓ mIns**	
	↓ LMM1.3**	
	↓ LMM2.0***	
	↓ sIns**	

		Glioblastoma multiforme (n = 9)
Diffuse intrinsic pontine glioma (n = 12)		↑ Cr***
		↑ mIns***
		↓ Lac***
		↓ LMM0.9***
		↓ LMM1.3****
		↓ LMM2.0***
		↑ tNAA**

Abbreviations: Cit, citrate; Cr, creatine; Glc, glucose; Glu, glutamate; Gly, glycine; Gua, guanidinoacetate; Lac, lactate; LMM, lipids and macromolecules; mIns, myo-inositol; sIns, scyllo-inositol; Tau, taurine; tCho, total choline; tNAA, total N-acetylaspartate.
[a]Includes classic, desmoplastic and large cell.
[b]Includes tanycytic and anaplastic.
[c]Includes pilomyxoid.
$p < 0.01$.*$p < 0.001$.****$p < 0.0001$.

(Figure 8B). PA was distinguished from EP by high tNAA, with EP characterised by high mIns (Figure 8C). LOOCV demonstrated a success rate of 93% with the correct classification of 74 out of 80 cases, all with a posterior probability of more than 80%. The discrimination of the three tumour groups with a posterior probability of more than 80% obtained a BAR of 0.88. Figure S3 demonstrates the spread of scanner manufacturers in the classifier. The test set of 3-T cases demonstrated correct classification of eight out of 13, with a posterior probability of more than 80%. A plot depicting the location of each 3-T case in the 1.5-T classifier is shown in Figure S4. Of the 22 cases at 1.5 T rereviewed under the new QC criteria, there were three large QC fails, five intermediate QC fails and 14 borderline QC fails. The three large QC fails were deemed to have failed the new criteria. Two of these were due to artefacts noted on one particular scanner; the other had a voxel placed adjacent to scalp with very poor shimming. Of the remaining 19 cases, 17 (89%) had a posterior probability of more than 80%, of which all 17 were correctly classified. Incorporation of the rereviewed cases into the classifier demonstrated a success rate of 95% and obtained a BAR of 0.90.

TABLE 4 Binary classification of tumours with discriminatory biomarkers identified using the Mann–Whitney U-test. Optimal cut-offs for significant variables were determined using ROC curves where an area under the curve (AUC), sensitivity, specificity and leave-one-out cross-validation (LOOCV) error rate were calculated.

Binary classification	Biomarker	ROC AUC	Cut-off value (mM)	Sensitivity	Specificity	LOOCV error rate (%)
Medulloblastoma versus ependymoma	mIns	0.82	3.1	0.80	0.90	14
Pilocytic astrocytoma versus ependymoma	Cr	0.90	2.0	0.90	0.91	11
Pilocytic astrocytoma versus ependymoma	mIns	0.82	3.1	0.80	0.97	9
Pilocytic astrocytoma versus ependymoma	tCho	0.81	1.2	0.80	0.86	29
Pilocytic astrocytoma versus ependymoma	LMM2.0	0.85	7.4	1.00	0.83	27
Diffuse intrinsic pontine glioma versus glioblastoma multiforme	Lac	0.94	1.8	1.00	0.92	15
Diffuse intrinsic pontine glioma versus glioblastoma multiforme	LMM0.9	0.94	3.6	1.00	0.58	15
Diffuse intrinsic pontine glioma versus glioblastoma multiforme	LMM1.3	0.98	6.2	0.90	0.66	10
Diffuse intrinsic pontine glioma versus glioblastoma multiforme	LMM2.0	0.92	7.2	0.90	0.69	24

Abbreviations: Cr, creatine; Lac, lactate; LMM, lipids and macromolecules; mIns, myo-inositol; ROC, receiver operator characteristic; tCho, total choline.

3.5 | Scanner comparisons

Guanidinoacetate ($p < 0.01$) was the only metabolite that differed between the four scanner models measured in MBs at 1.5 T (Table S4). No significant difference in SNR was found between scanners (31 vs. 27 vs. 14 vs. 20, $p > 0.05$, Siemens Symphony, Siemens Avanto, GE Signa Excite and Philips Intera, respectively).

3.6 | Field strength comparisons

Creatine ($p < 0.01$) was found to be higher in 3-T MBs ($n = 9$) than 1.5-T MBs ($n = 42$) (table not shown).

3.7 | Clinical examples

Figure 9 demonstrates three cases where the suspected diagnosis given on the radiological report differed from that indicated by visual inspection and classification of the MRS. The diagnosis proposed by MRS was confirmed by biopsy in each of these cases.

4 | DISCUSSION

This large multicentre study reports extensive information on ^{1}H MRS of 13 different types of children's brain tumours and its accuracy as a diagnostic aid. Mean spectra and metabolite data are presented at two TEs, two field strengths and for rare as well as common tumours, providing an extensive resource for those using the technique. Mean metabolite concentrations were shown to differ between each tumour type and the combined group of all other tumours, showing that tumour metabolite profiles strongly reflect tumour type as determined by histological analysis. Individual metabolites were able to discriminate between some pairs of tumours where there is commonly a diagnostic dilemma on conventional MRI and may provide a useful diagnostic aid. However, the use of the complete set of metabolites provides greater promise in diagnosis than individual metabolites. This is demonstrated by the finding that pairs of the three most common tumour types can be discriminated by individual metabolites with an accuracy of 80% to 91% (Table 4), while a formal test of the diagnostic potential of the whole metabolite profile using a three-way PCA-LDA classification demonstrated a diagnostic accuracy of 95% and can be used on almost the whole range of data qualities acquired in routine practice.

The use of conventional MRI in the preoperative diagnosis of brain tumours is often challenging and can result in several proposed diagnoses.[13,35] The large number of metabolites showing differences in mean concentrations highlights the potential for these to be used in diagnosis. However, the large variability of values for specific cases about the mean hinders individual metabolites from discriminating between tumour types, confirming the results of smaller studies.[13] Previous studies have reported results using metabolite cut-offs to aid diagnosis, prognosis and treatment monitoring.[36–40] ROC curves demonstrated a number of biomarkers that were highly sensitive and specific for discriminating between
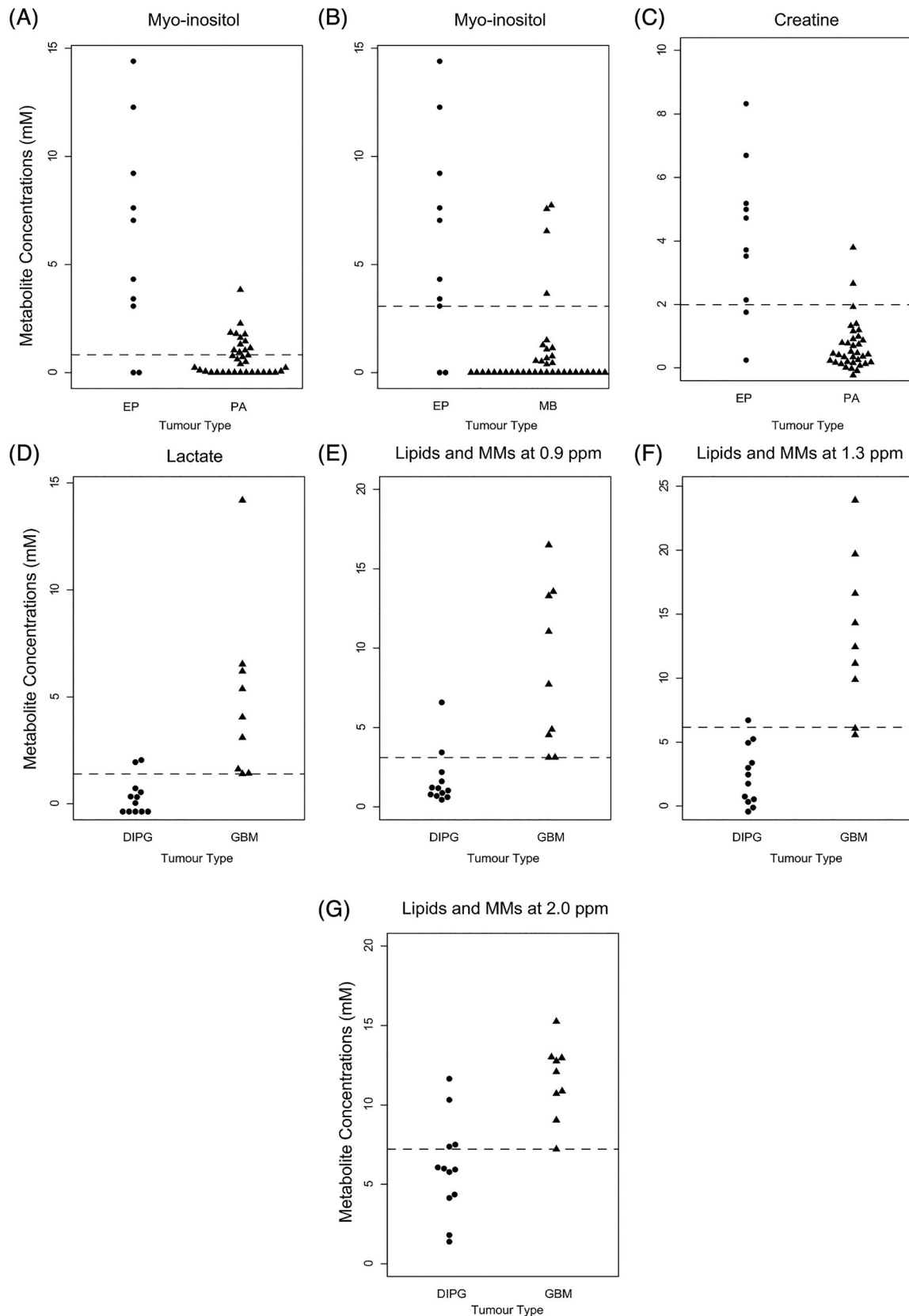
FIGURE 7 Dot density plots corresponding to discriminatory metabolites reported in Table 4. Cut-off values are shown with dotted lines. These include (A) Ependymoma (EP) versus pilocytic astrocytoma (PA)—myo-inositol, (B) EP versus medulloblastoma (MB)—myo-inositol, (C) EP versus PA—creatine, (D) Diffuse intrinsic pontine glioma (DIPG) versus glioblastoma multiforme (GBM)—lactate, (E) DIPG versus GBM—lipids and macromolecules (MMs) at 0.9 ppm, (F) DIPG versus GBM—lipids and MMs at 1.3 ppm, and (G) DIPG versus GBM—lipids and MMs at 2.0 ppm.
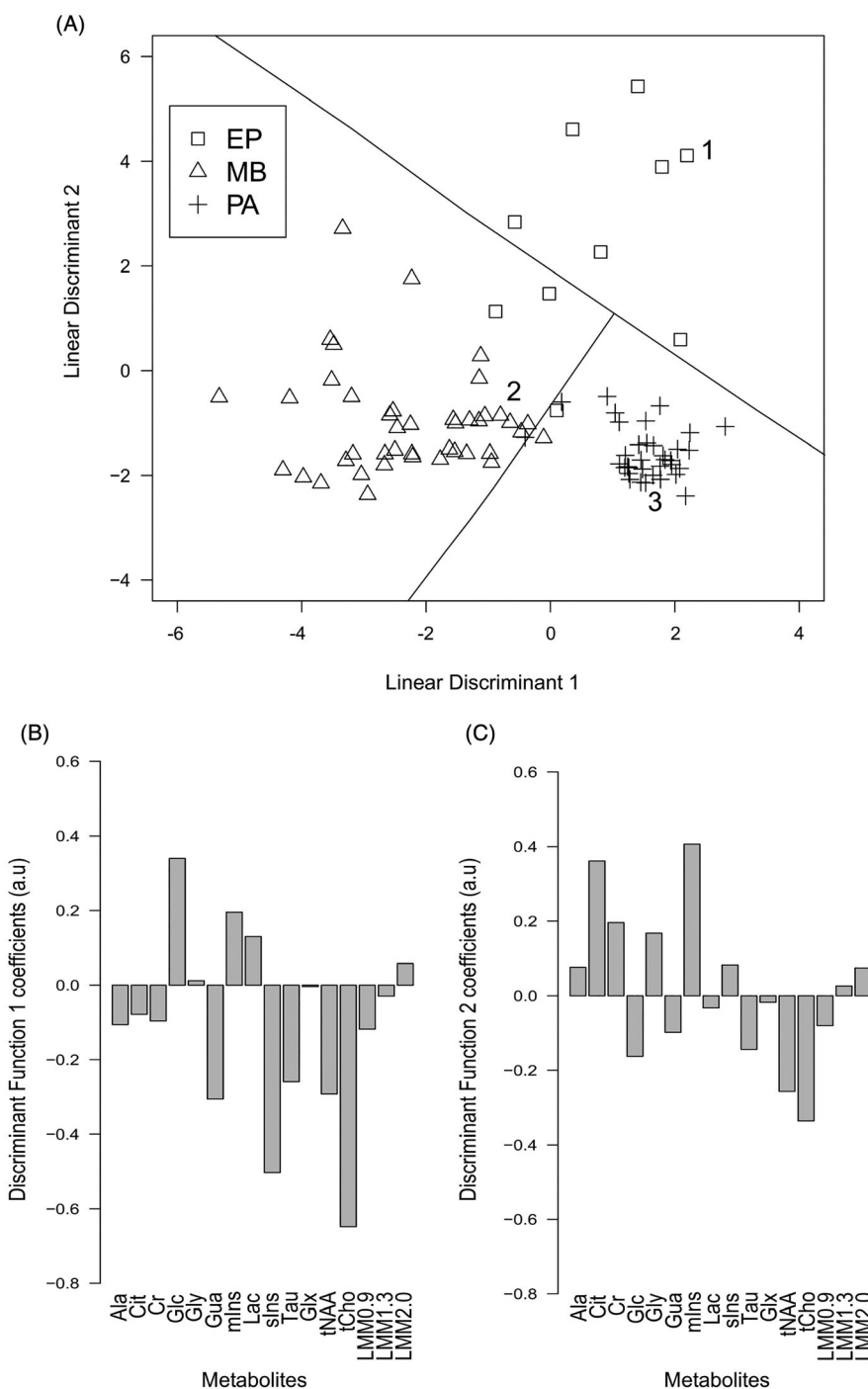
FIGURE 8 Results of the 1.5-T three-way classifier between medulloblastoma (MB) and pilocytic astrocytomas (PAs) and ependymomas (EPs). (A) A scatter plot of the discriminant function scores for each tumour type. (B) Discriminant function 1 coefficients, and (C) Discriminant function 2 coefficients of the standardised metabolite profile. Labelled cases correspond to examples in Figure 9. Ala, alanine; Cit, citrate; Cr, creatine; Glc, glucose; Glx, glutamine + glutamate; Gly, glycine; Gua, guanidinoacetate; Lac, lactate; LMM, lipids and macromolecules; mIns, myo-inositol; sIns, scyllo-inositol; Tau, taurine; tCho, total choline; tNAA, total N-acetylaspartate.

specific tumour types as validated using LOOCV. A future study should be undertaken to validate the findings prospectively and determine clinical impact.

Using whole metabolite profiles to discriminate between tumours is a logical extension to individual metabolites. Visual comparison of MRS with mean spectra from tumour groups is one approach that has been previously reported,[7,17,21] and the mean spectra provide a useful diagnostic aid. The accuracy of this approach needs to be assessed formally and its added value to conventional MRI evaluated. The use of pattern
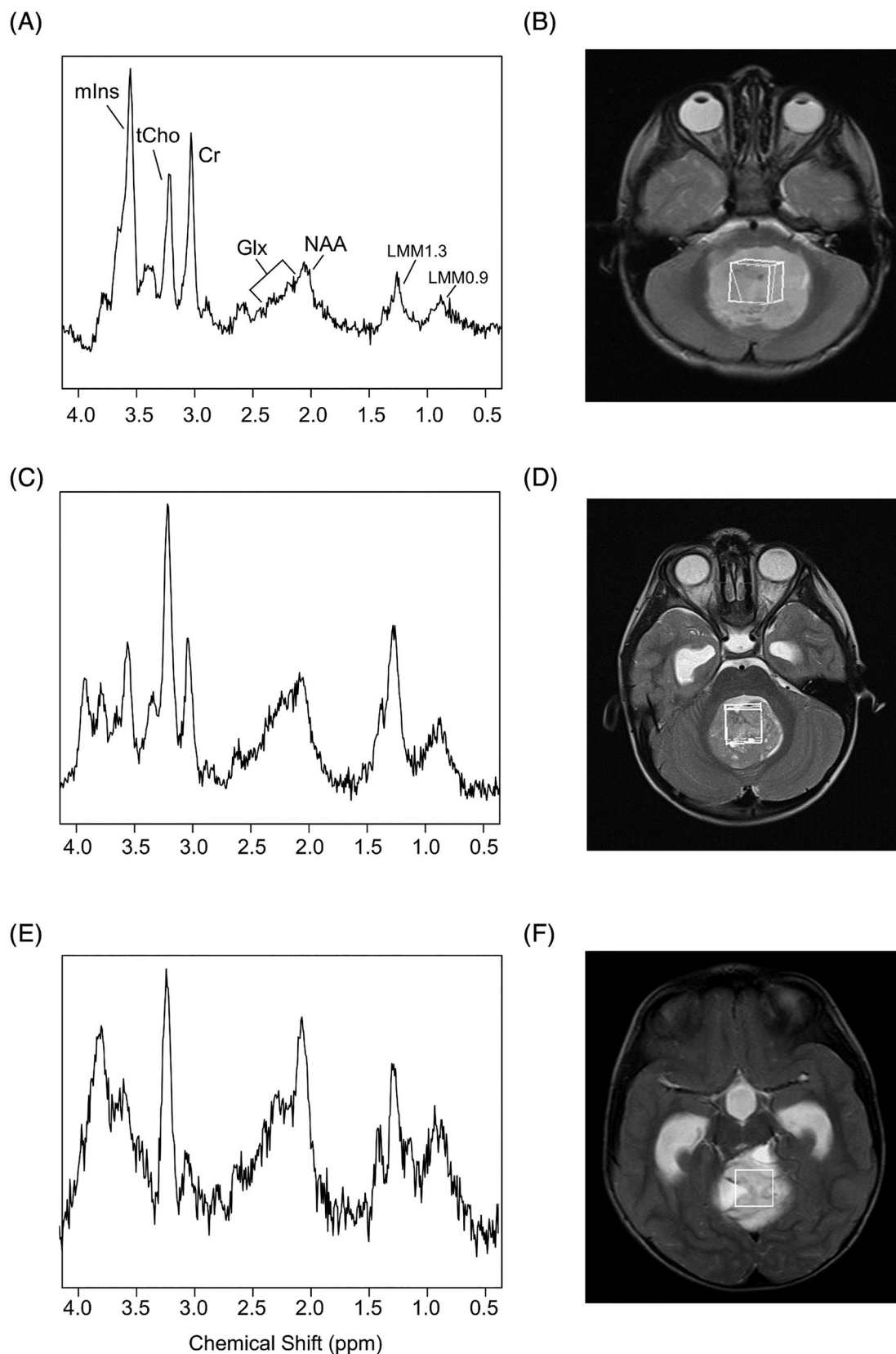
FIGURE 9 (A) Example case 1 spectra, and (B) Axial T2-weighted image with voxel placement over an ependymoma suspected to be a medulloblastoma on radiology reporting (2.5 years, male); (C) Example case 2 spectra, and (D) Axial T2-weighted image with voxel placement over a medulloblastoma suspected to be an ependymoma (4 years, male); and (E) Example case 3 spectra, and (F) Axial T2-weighted image with voxel placement over a pilocytic astrocytoma that was suspected to be an ependymoma (4.8 years, male) (as labelled in Figure 8). Cr, creatine; Glx, glutamine + glutamate; LMM, lipids and macromolecules; mIns, myo-inositol; NAA, N-acetylaspartate; tCho, total choline.

recognition methods for classification of tumours based on MRS has been reported previously.[19,20,41] High levels of accuracy for diagnosing PA, EP and MB in children have been achieved using these techniques from short TE MRS in a single-centre study,[20] but short and long TEs needed to be combined to provide high accuracy in a multicentre study.[19] The current multicentre study reproduces the high accuracy of the single-centre data at short TE, giving a saving in acquisition time. Further improvements should occur with MRS data acquired at 3 T but there are insufficient cases to explore this fully at present. While there is mounting evidence for the accuracy of MRS as a diagnostic aid when used in combination with pattern recognition techniques, the availability of such tools for clinical use is still lacking, and commercial, regulatory approved decision support systems need to be made available.

4.1 | Scanner comparisons

No cases had MRS acquired on more than one scanner, but comparing MRS on different scanners showed that for one specific tumour type (MB) the mean concentrations of Gua and Cr were different ($p < 0.05$). Both of these metabolites have an MRS peak close to water, making them subject to variation with level of shimming and water suppression. Modelling of Cr in TARQUIN has a correction term to allow for this known effect. Cr is also known to vary widely in MB and the differences could be due to population differences.[23]

4.2 | Quality control

Although MRS data fail stringent QC measures in approximately one-third of cases, spectra failing QC do not necessarily fail as a limitation of the technique, because the characteristics of the tumour also play an important role. A total of 39% of 1.5-T short TE spectra and 48% of 3-T short TE spectra that failed QC were acquired from Pas. The low cellularity of low-grade tumours and their cystic nature result in reduced SNR,[42,43] and this in itself can be diagnostic information. Similarly, it is often not possible to shim effectively over tumours containing regions of haemorrhage, leading to grossly broadened peaks. Importantly, spectra in the intermediate and borderline QC fail categories perform well on the metabolite profile classifier, with classification success rates being 95% with and 93% without these cases included. This finding demonstrates that less stringent QC criteria could be used, allowing inclusion of almost all cases. Furthermore, all three cases with very poor quality MRS could have been avoided; two were performed on a scanner known to produce artefacts, while the third had a poor voxel placement. Intermediate and borderline QC failures could be avoided in many cases by care with voxel placement, the addition of saturation bands and setting an appropriate number of averages. Applying these measures should enable clinically interpretable MRS to be obtained for the vast majority of children's brain tumours.

All metabolites that had a %CRLB of less than 30% in at least two cases were included in the analysis and reported in the results. Whether a particular metabolite has low %CRLB in a specific case depends on the concentration of that metabolite and the quality of the MRS, in particular the SNR. In cases where these conditions were present, some metabolites that are challenging to quantify, for example, Glu, Gln and Lac, were determined with acceptable %CRLB.

4.3 | Study limitations

The study does suffer from significant limitations both technical and related to study design. From a technical perspective, a fixed MRS processing protocol has been used with a lack of accurately known T1, T2s of metabolites and water for relaxation time correction. T2 values for water and metabolites will vary with field strength and even between tumours, leading to some inaccuracy in the estimated concentrations produced by the default values assumed in TARQUIN. Particular care should be taken when comparing metabolite values between the different field strengths. In addition, tumour heterogeneity within the voxel has not been taken account of, with partial volume effects of necrosis and high cellularity regions creating variability in water reference values due to proton density and T1 relaxation effects. Not taking account of these parameters could lead to a loss of useful diagnostic information and create variability that hinders classification. However, determining these parameters for each tumour is not currently feasible in a routine clinical environment. Metabolite ratios obviate the need to make assumptions about water concentrations and water relaxation rates and are an alternative approach.

We have elected to use a fixed predetermined metabolite basis set because, at the time of MRS acquisition and processing, the diagnosis would not be known and an adaptive basis set would be challenging to implement clinically. However, there are situations where the basis set may not contain key metabolites present in the tumour and mis-assignments occur. A pertinent example of this is the assignment of a peak around 2 ppm to NAA in PAs, a peak for which there is evidence that a macromolecular component exists.[44] While the metabolite concentrations are not all strictly correct in these circumstances, the use to provide a noninvasive diagnosis should not be significantly impaired.

Taurine levels in the brain do decrease with age but the main tumour types occur throughout childhood and so this does not explain the discriminatory nature of the metabolite in MBs.

Undertaking MRI in children can be challenging. All scans were performed in specialist centres with access to sedation and general anaesthetic where deemed clinically appropriate. Scans in which significant movement artefact was detected on conventional imaging did not proceed to acquisition of spectroscopy and so it is unlikely that movement artefact is a reason for degradation of spectroscopy data.

The relative lack of 3-T data is a limitation of this study and further acquisition and analysis at this field strength should be undertaken. There were few 3-T scanners available in most centres at the start of the study period, and even towards the end, 1.5-T scanners were often used in preference, most likely because of the acute presentation of the cases leading to the need to find an available scanner overriding the advantages of 3 T together with the requirement for spinal imaging, which was often better at the lower field strength.

Metabolite profiles are likely to vary across each tumour to some extent, making the MRS dependent on voxel placement. Many tumours were sufficiently small that there was little flexibility in voxel placement, reducing the variability due to intratumoural heterogeneity. Ideally, multiple voxels would be placed in large heterogenous tumours, but this was not practical within the clinical setting.

Chemical shift displacement effects could lead to certain metabolites having their values altered; however, the use of relatively small voxels and their placement well within the solid component of the tumour should mitigate against this. At higher field strengths, the effect is more pronounced, but smaller voxels tended to be used. The availability of mean spectra for a variety of tumour types as a comparator for new cases is an important facilitator for clinical practice and has been shown to aid decision-making.[17,21,29] In the future, we envisage that machine learning classifiers will become an increasingly important diagnostic aid. These should be incorporated into a computerised clinical decision support system, which would enable them to be incorporated into clinical workflows. However, this requires appropriate evaluation and can be subject to regulatory approval.

From a study design perspective, there are challenges and limitations posed by the length of time that the data were acquired over and the time that has elapsed since data collection was completed. Children's brain tumours are relatively rare, and large datasets can only be acquired from multiple sites over a prolonged period, during which diagnostic criteria and MRS acquisition techniques inevitably change. With the advent of molecular tests, some diagnoses based on histopathology are now recognised to be incorrect, while whole diagnostic categories such as supratentorial PNET are no longer recognised. Higher field strength scanners are increasingly used clinically with subsequent improvements in MRS quality, particularly when combined with newer pulse sequences such as semi-LASER. However, from a pragmatic viewpoint, diagnoses have only changed for a few cases, many children with brain tumours have MRS acquired at 1.5 T with very few above 3 T, and new pulse sequences have their biggest impact for MRS imaging and higher field strengths. The reporting of such a large multisite dataset of MRS for children's brain tumours is therefore still important scientifically and of clinical relevance.

5 | CONCLUSION

This study reports MRS metabolite profiles for a large multicentre cohort of childhood brain tumours. Mean spectra, metabolite concentrations and multivariate classifiers aid discrimination between tumour types. Pattern classification of metabolite profiles provides high levels of accuracy in discriminating between the major children's brain tumour types, even for data that do not meet stringent QC measures. Overall, the study should greatly aid clinical implementation of MRS across multiple centres.

ORCID
Heather E. L. Rose https://orcid.org/0000-0002-0346-1334
Geoffrey S. Payne https://orcid.org/0000-0001-8029-1765
Theodoros N. Arvanitis https://orcid.org/0000-0001-5473-135X
Andrew C. Peet https://orcid.org/0000-0002-4846-5152

REFERENCES
1. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol.* 2016;131(6):803-820. doi:10.1007/s00401-016-1545-1
2. Poussaint TY. Magnetic resonance imaging of pediatric brain tumors: state of the art. *Top Magn Reson Imaging.* 2001;12(6):411-433. doi:10.1097/00002142-200112000-00004
3. Kuperman V. *Magnetic Resonance Imaging, Physical Principles and Applications.* Academic Press; 2000.
4. Bryant SO, Cha S, Barkovich AJ. Modern neuroimaging of pediatric brain tumors. In: Gupta N, Banerjee A, Haas-Kogan D, eds. *Pediatric CNS Tumors.* Springer; 2010:247-272. doi:10.1007/978-3-540-87979-4_13

5. Rao P. Role of MRI in paediatric neurooncology. *Eur J Radiol.* 2008;68(2):259-270. doi:10.1016/j.ejrad.2008.06.033

6. Koral K, Zhang S, Gargan L, et al. Diffusion MRI improves the accuracy of preoperative diagnosis of common pediatric cerebellar tumors among reviewers with different experience levels. *Am J Neuroradiol.* 2013;34(12):2360-2365. doi:10.3174/ajnr.A3596

7. Shiroishi MS, Panigrahy A, Moore KR, et al. Combined MRI and MRS improves pre-therapeutic diagnoses of pediatric brain tumors over MRI alone. *Neuroradiology.* 2015;57(9):951-956. doi:10.1007/s00234-015-1553-1

8. Peet AC, Arvanitis TN, Leach MO, Waldman AD. Functional imaging in adult and paediatric brain tumours. *Nat Rev Clin Oncol.* 2012;9(12):700-711. doi:10.1038/nrclinonc.2012.187

9. Manias KA, Gill SK, MacPherson L, Foster K, Oates A, Peet AC. Magnetic resonance imaging based functional imaging in paediatric oncology. *Eur J Cancer.* 2017;72:251-265. doi:10.1016/j.ejca.2016.10.037

10. Novak J, Zarinabad N, Rose H, et al. Classification of paediatric brain tumours by diffusion weighted imaging and machine learning. *Sci Rep.* 2021; 11(1):2987. doi:10.1038/s41598-021-82214-3

11. Bruhn H, Frahm J, Gyngell ML, et al. 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. doi:10.1148/radiology.172.2.2748837

12. Langkowski JH, Wieland J, Bomsdorf H, et al. Pre-operative localized in vivo proton spectroscopy in cerebral tumors at 4.0 Tesla—first results. *Magn Reson Imaging.* 1989;7(5):547-555. doi:10.1016/0730-725X(89)90409-8

13. Panigrahy A, Krieger MD, Gonzalez-Gomez I, et al. Quantitative short echo time 1H-MR spectroscopy of untreated pediatric brain tumors: preoperative diagnosis and characterization. *Am J Neuroradiol.* 2006;27(3):560-572.

14. Provencher SW. Automatic quantitation of localized in vivo 1H spectra with LCModel. *NMR Biomed.* 2001;14(4):260-264. doi:10.1002/nbm.698

15. Wilson M, Reynolds G, Kauppinen RA, Arvanitis TN, Peet AC. A constrained least-squares approach to the automated quantitation of in vivo 1H magnetic resonance spectroscopy data. *Magn Reson Med.* 2011;65(1):1-12. doi:10.1002/mrm.22579

16. Kovanlikaya A, Panigrahy A, Krieger MD, et al. Untreated pediatric primitive neuroectodermal tumor in vivo: quantitation of taurine with MR spectroscopy. *Radiology.* 2005;236(3):1020-1025. doi:10.1148/radiol.2363040856

17. Manias KA, Gill SK, MacPherson L, et al. Diagnostic accuracy and added value of qualitative radiological review of H-1-magnetic resonance spectroscopy in evaluation of childhood brain tumors. *Neurooncol Pract.* 2019;6(6):428-437. doi:10.1093/nop/npz010

18. Opstad KS, Ladroue C, Bell BA, Griffiths JR, Howe FA. Linear discriminant analysis of brain tumour 1H MR spectra: a comparison of classification using whole spectra versus metabolite quantification. *NMR Biomed.* 2007;20(8):763-770. doi:10.1002/nbm.1147

19. Vicente J, Fuster-Garcia E, Tortajada S, et al. Accurate classification of childhood brain tumours by in vivo 1H MRS-A multi-centre study. *Eur J Cancer.* 2013;49(3):658-667. doi:10.1016/j.ejca.2012.09.003

20. Davies NP, Wilson M, Harris LM, et al. Identification and characterisation of childhood cerebellar tumours by in vivo proton MRS. *NMR Biomed.* 2008; 21(8):908-918. doi:10.1002/nbm.1283

21. Davies NP, Rose HEL, Manias KA, et al. Added value of magnetic resonance spectroscopy for diagnosing childhood cerebellar tumours. *NMR Biomed.* 2022;35(2):e4630. doi:10.1002/nbm.4630

22. Louis DN, Perry A, Wesseling P, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro Oncol.* 2021;23(8): 1231-1251. doi:10.1093/neuonc/noab106

23. Blüml S, Margol AS, Sposto R, et al. Molecular subgroups of medulloblastoma identification using noninvasive magnetic resonance spectroscopy. *Neuro Oncol.* 2016;18(1):126-131. doi:10.1093/neuonc/nov097

24. Manias KA, Harris LM, Davies NP, et al. Prospective multicentre evaluation and refinement of an analysis tool for magnetic resonance spectroscopy of childhood cerebellar tumours. *Pediatr Radiol.* 2018;48(11):1630-1641. doi:10.1007/s00247-018-4182-0

25. Horská A, Uluǧ AM, Melhem ER, et al. Proton magnetic resonance spectroscopy of choroid plexus tumors in children. *J Magn Reson Imaging.* 2001; 14(1):78-82. doi:10.1002/jmri.1154

26. Porto L, Kieslich M, Franz K, Lehrbecher T, Pilatus U, Hattingen E. Proton magnetic resonance spectroscopic imaging in pediatric low-grade gliomas. *Brain Tumor Pathol.* 2010;27(2):65-70. doi:10.1007/s10014-010-0268-6

27. Chang YW, Yoon HK, Shin HJ, Roh HG, Cho JM. MR imaging of glioblastoma in children: usefulness of diffusion/perfusion-weighted MRI and MR spectroscopy. *Pediatr Radiol.* 2003;33(12):836-842. doi:10.1007/s00247-003-0968-8

28. Zarinabad N, Abernethy LJ, Avula S, et al. Application of pattern recognition techniques for classification of pediatric brain tumors by in vivo 3T 1 H-MR spectroscopy—a multi-center study. *Magn Reson Med.* 2018;79(4):2359-2366. doi:10.1002/mrm.26837

29. Manias K, Gill SK, Zarinabad N, et al. Evaluation of the added value of 1H-magnetic resonance spectroscopy for the diagnosis of pediatric brain lesions in clinical practice. *Neurooncol Pract.* 2018;5(1):18-27. doi:10.1093/nop/npx005

30. Seymour ZA, Panigrahy A, Finlay JL, Nelson, MD Jr, Blüml S. Citrate in pediatric CNS tumors? *Am J Neuroradiol.* 2008;29:1006-1011. doi:10.3174/ajnr.A1018

31. Davies NP, Wilson M, Natarajan K, et al. Non-invasive detection of glycine as a biomarker of malignancy in childhood brain tumours using in-vivo 1H MRS at 1.5 Tesla confirmed by ex-vivo high-resolution magic-angle spinning NMR. *NMR Biomed.* 2010;23(1):80-87. doi:10.1002/nbm.1432

32. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Accessed April 2011. https://www.r-project.org/

33. Geisser S. The predictive sample reuse method with applications. *J Am Stat Assoc.* 1975;70(350):320-328. doi:10.1080/01621459.1975.10479865

34. Guyon I, Alamdari ARSA, Dror G, Buhmann JM. Performance prediction challenge. In: *IEEE International Conference on Neural Networks—Conference Proceedings.* IEEE; 2006. doi:10.1109/ijcnn.2006.246632

35. Panigrahy A, Blüml S. Neuroimaging of pediatric brain tumors: From basic to advanced magnetic resonance imaging (MRI). *J Child Neurol.* 2009;24: 1343-1365. doi:10.1177/0883073809342129

36. Harris LM, Davies N, MacPherson L, et al. The use of short-echo-time 1H MRS for childhood cerebellar tumours prior to histopathological diagnosis. *Pediatr Radiol.* 2007;37(11):1101-1109. doi:10.1007/s00247-007-0571-5

37. Orphanidou-Vlachou E, Auer D, Brundler MA, et al. 1H magnetic resonance spectroscopy in the diagnosis of paediatric low grade brain tumours. *Eur J Radiol.* 2013;82(6):E295-E301. doi:10.1016/j.ejrad.2013.01.030

38. Smith EA, Carlos RC, Junck LR, Tsien CI, Elias A, Sundgren PC. Developing a clinical decision model: MR spectroscopy to differentiate between recurrent tumor and radiation change in patients with new contrast-enhancing lesions. *Am J Roentgenol.* 2009;192(2):W45-W52. doi:10.2214/AJR.07.3934

39. Astrakas LG, Zurakowski D, Tzika AA, et al. Noninvasive magnetic resonance spectroscopic imaging biomarkers to predict the clinical grade of pediatric brain tumors. *Clin Cancer Res.* 2004;10(24):8220-8228. doi:10.1158/1078-0432.CCR-04-0603

40. Marcus KJ, Astrakas LG, Zurakowski D, et al. Predicting survival of children with CNS tumors using proton magnetic resonance spectroscopic imaging biomarkers. *Int J Oncol.* 2007;30(3):651-657. doi:10.3892/ijo.30.3.651

41. Tate AR, Underwood J, Acosta DM, et al. 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. doi:10.1002/nbm.1016

42. Gupta RK, Cloughesy TF, Sinha U, et al. Relationships between choline magnetic resonance spectroscopy, apparent diffusion coefficient and quantitative histopathology in human glioma. *J Neurooncol.* 2000;50(3):215-226. doi:10.1023/A:1006431120031

43. Harris LM, Davies NP, MacPherson L, et al. Magnetic resonance spectroscopy in the assessment of pilocytic astrocytomas. *Eur J Cancer.* 2008;44(17): 2640-2647. doi:10.1016/j.ejca.2008.08.012

44. Tamrazi B, Nelson MD Jr, Blüml S. MRS of pilocytic astrocytoma: the peak at 2 ppm may not be NAA. *Magn Reson Med.* 2017;78(2):452-456. doi:10.1002/mrm.26374

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Gill SK, Rose HEL, Wilson M, et al. Characterisation of paediatric brain tumours by their MRS metabolite profiles. *NMR in Biomedicine.* 2024;1-21. doi:10.1002/nbm.5101