Received: 12 April 2023 | Revised: 7 January 2024 | Accepted: 3 February 2024

DOI: 10.1002/nbm.5129

RESEARCH ARTICLE

NMR IN BIOMEDICINE WILEY

Noise suppression of proton magnetic resonance spectroscopy improves paediatric brain tumour classification

Teddy Zhao[1,2] | James T. Grist[1] | Dorothee P. Auer[3] | Shivaram Avula[4] | Simon Bailey[5] | Nigel P. Davies[6] | Richard G. Grundy[3] | Omar Khan[7] | Lesley MacPherson[8] | Paul S. Morgan[3,9,10] | Barry Pizer[11] | Heather E. L. Rose[1,2] | Yu Sun[1,2] | Martin Wilson[12] | Lara Worthington[1,2,13] | Theodoros N. Arvanitis[1,2,7,14] | Andrew C. Peet[1,2]

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

[2]Oncology, Birmingham Children's Hospital, Birmingham, UK

[3]Clinical Neuroscience, University of Nottingham, Nottingham, UK

[4]Radiology, Alder Hey Children's NHS Foundation Trust, Liverpool, UK

[5]Paediatric Oncology, Great North Children's Hospital, Newcastle upon Tyne, UK

[6]Imaging and Medical Physics, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK

[7]Digital Healthcare, WMG, University of Warwick, Coventry, UK

[8]Radiology, Birmingham Children's Hospital, Birmingham, UK

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

[10]Medical Physics, Nottingham University Hospitals NHS Trust, Nottingham, UK

[11]University of Liverpool, Liverpool, UK

[12]Centre for Human Brain Health, University of Birmingham, Birmingham, UK

[13]RRPPS, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK

[14]Engineering, University of Birmingham, Birmingham, UK

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

Abstract

Proton magnetic resonance spectroscopy (^{1}H-MRS) is increasingly used for clinical brain tumour diagnosis, but suffers from limited spectral quality. This retrospective and comparative study aims at improving paediatric brain tumour classification by performing noise suppression on clinical ^{1}H-MRS. Eighty-three/forty-two children with either an ependymoma (ages 4.6 ± 5.3/9.3 ± 5.4), a medulloblastoma (ages 6.9 ± 3.5/6.5 ± 4.4), or a pilocytic astrocytoma (8.0 ± 3.6/6.3 ± 5.0), recruited from four centres across England, were scanned with 1.5T/3T short-echo-time point-resolved spectroscopy. The acquired raw ^{1}H-MRS was quantified by using Totally Automatic Robust Quantitation in NMR (TARQUIN), assessed by experienced spectroscopists, and processed with adaptive wavelet noise suppression (AWNS). Metabolite concentrations were extracted as features, selected based on multiclass receiver operating characteristics, and finally used for identifying brain tumour types with supervised machine learning. The minority class was oversampled through the synthetic minority oversampling technique for comparison purposes. Post-noise-suppression ^{1}H-MRS showed significantly elevated signal-to-noise ratios (P < .05, Wilcoxon signed-rank test), stable full width at half-maximum (P > .05, Wilcoxon signed-rank test), and significantly higher classification accuracy (P < .05, Wilcoxon signed-rank test). Specifically, the cross-validated overall and balanced classification accuracies can be improved from 81% to 88% overall and 76% to 86% balanced for the 1.5T cohort, whilst for the 3T cohort they can be improved from 62% to 76% overall and 46% to 56%, by applying Naïve Bayes on the oversampled ^{1}H-MRS. The study shows that fitting-based signal-to-noise ratios of clinical ^{1}H-MRS can be significantly improved by using AWNS with insignificantly altered line width, and the post-

Abbreviations: ^{1}H-MRS, proton magnetic resonance spectroscopy; AUC, area under the curve; AWNS, adaptive wavelet noise suppression; CRLB, Cramér–Rao lower bound; DF, discriminant functions; fSNR, fitting-based SNR; FWHM, full width at half-maximum; mAUC, multiclass area under the curve; NS, noise suppression; PBTs, paediatric brain tumours; ROC, receiver operating characteristics; SMOTE, synthetic minority oversampling technique; SNR, signal-to-noise ratio; wSNR, whole-spectrum SNR.

Funding information
Health Data Research UK; CHILDREN with CANCER UK; Action Medical Research and the Brain Tumour Charity; Birmingham Women's and Children's Hospital Charities; The Children's Cancer and Leukaemia Group Little Princess Trust; Cancer Research UK and NIHR Experimental Cancer Medicine Centre Paediatric Network; Children's Research Fund; NIHR Nottingham Biomedical Research Centre; Cancer Research UK and EPSRC Cancer Imaging Programme at the Children's Cancer and Leukaemia Group (CCLG) in association with the MRC and Department of Health (England); NIHR Research Professorship; Help Harry Help Others

noise-suppression ^{1}H-MRS may have better diagnostic performance for paediatric brain tumours.

KEYWORDS
machine learning, metabolite concentration, noise suppression, paediatric brain tumour, proton magnetic resonance spectroscopy, wavelet

1 | INTRODUCTION

Paediatric brain tumours (PBTs) remain amongst the most lethal cancers in childhood.[1] Half of PBTs arise in the posterior fossa,[2] among which the most frequent tumour types are pilocytic astrocytomas, medulloblastomas, and ependymomas.[3] Pilocytic astrocytomas, a subset of gliomas, are the most common PBTs in the posterior fossa, and they are classified as World Health Organization (WHO) grade I tumours.[4] Medulloblastomas are the next most common tumours, which are classified as WHO grade IV. They have four main molecular groups that are associated with diverse prognoses.[5] Ependymomas[6] are less common and are classified as grade II or III,[7] although grading is challenging and not closely linked to their malignant behaviour. The three tumour types require different treatment strategies that usually involve surgical resection, with adjuvant treatment, including radiotherapy and chemotherapy, being dependent on diagnosis and a set of clinical, pathological, and radiological factors.[8] For pilocytic astrocytomas, surgical resection alone is usually curative and, even if resection is not complete, patients often require no further treatment. Surgical strategy is therefore aimed at minimising morbidity whilst aiming for as complete a resection as possible.[9] For medulloblastomas, surgical resection is an important part of their treatment but small residuals can be successfully treated with adjuvant radiotherapy and chemotherapy. Many medulloblastomas have metastases at diagnosis and the prognostic value of extensive resection of the primary is less certain in these cases, although often still attempted.[10] For ependymomas, a complete resection is the key to maximising the chances of survival and second-look surgery is advocated where the tumour has not been completely resected and the residual is thought amenable to further resection.[11] Given the different surgical strategies required for these three tumour types, a pre-operative diagnosis is a significant contribution to their clinical care.[12] In addition, an early diagnosis can allow timely planning of adjuvant treatment and more informed discussions with the family.[13] An initial noninvasive diagnosis is made from clinical and imaging information, with this being confirmed or refined by histology and molecular analysis after surgical resection.[14]

Conventional MRI is the standard diagnostic imaging modality, and it can present structural details of tissue and is normally involved in clinical diagnostic determination.[8] It allows limited differential classification between the three PBTs with conventional machine learning and requires heavy computational analysis to achieve better performance.[15] In contrast, proton magnetic resonance spectroscopy (^{1}H-MRS)[16] can reveal metabolite profiles of human tissue in vivo[17] by observing the T_2 relaxation time variations of metabolites. Metabolite profiles were observed to be cancer-specific[18] and could reflect malignancy,[19] supporting their use in the clinical diagnosis of PBTs.[20] However, clinical ^{1}H-MRS is challenging to acquire at high quality and in particular suffers from limited signal-to-noise ratio (SNR).[21] This is due to the relatively low concentration of protons from metabolites in comparison with bulk water and fat protons and the limitation on voxel size required for tumour localisation. Even for well-designed MR systems, noise may still exist due to thermal motion of charged particles and electrons in the receiver coil.[22] Spectral noise[23] can introduce errors in estimating metabolite concentrations, which may lead to inaccurate metabolite- or spectrum-based PBT classification. In clinical practice, the ^{1}H-MRS SNR may not always be of acceptable quality, and it is affected by various factors such as the number of signals averaged and voxel size.[24] Apodisation is the commonly used method for clinical ^{1}H-MRS postprocessing.[25] Despite being able to increase the SNR to some degree, apodisation decreases the spectral resolution, which shows as increased full width at half-maximum (FWHM). Consequently, apodisation makes overlapping spectra more difficult to separate and metabolite quantification no longer reliable.[24] Wavelet analysis is a unified multiresolution processing technique for nonstationary signals.[26] Wavelets were reported to be useful in ^{1}H-MRS for quantification[27], analysis,[28] noise suppression (NS),[29] and clinical imaging biomarker identification.[30] To reduce the detrimental effects of noise in clinical ^{1}H-MRS, wavelet analysis could potentially improve the accuracy of metabolite concentration estimation.

This article hypothesises that the metabolite concentrations that are estimated from post-noise-suppression (postNS) ^{1}H-MRS have improved diagnostic accuracy for PBTs. The aim is to investigate the potential of ^{1}H-MRS NS for improving PBT classification.

2 | MATERIALS AND METHODS

This retrospective study (Figure 1) was approved by the local research ethics committee (ethics number: 04/MRE04/41). Informed consent was obtained from parents or guardians of all patients.

2.1 | Data acquisition

Structural imaging and ^{1}H-MRS data of PBT cases, with a diagnosis of ependymoma (posterior fossa or supratentorial), medulloblastoma, and pilocytic astrocytoma (posterior fossa), were collected from our four hospitals nationally between October 2004 and December 2019. Each patient underwent MRI and ^{1}H-MRS before surgical resection. Histological diagnosis was reviewed by local tumour boards. Histological subtypes of the medulloblastomas and ependymomas were grouped together. Multisite imaging data were acquired by using scanners including Siemens Symphony 1.5T, GE Signa LX 1.5T, Philips Ingenia 1.5T, Philips Intera 1.5T, Philips Achieva 3T, Philips X-series 3T, and Siemens Verio 3T (Tables D1–D3). Structural images were acquired by using T_1-weighted, T_2-weighted, T_1-weighted post-contrast and diffusion-weighted MR imaging sequences.

Following conventional imaging that included gadolinium administration, ^{1}H-MRS with water reference was acquired by using the point-resolved spectroscopy sequence (field strength 1.5T or 3T, head coils or head and neck coils 8–32 channels, sampling frequency 1000–2500 Hz, chemical shift displacement < 4% per ppm, echo time 30–41 ms, number of complex points 512, 1024, or 2048, number of signals averaged 128, pulse repetition time 1500–2200 ms, voxel size 13 × 13 × 13–20 × 20 × 20 mm^3: Tables D1–D3). Water suppression was performed through chemical shift selective saturation pulses without outer-volume suppression. Volumes of interest were placed within the tumour by clinical radiologists and MR technicians with reference to structural images. Contrast enhancement and low apparent diffusion coefficient were used as guides where tumours exhibited some heterogeneity.

2.2 | Spectroscopy quantification and quality control

^{1}H-MRS were quantified by using Totally Automatic Robust Quantitation in NMR (TARQUIN) (version 4.3.11) with the ^{1}H brain full basis that includes the basis of lipids and macromolecules.[31] Quality-control parameters, which were obtained through quantification, included fitting-based SNR (fSNR), whole-spectrum SNR (wSNR), FWHM (Appendix A.1), and Cramér–Rao lower bound (CRLB).[24,32] ^{1}H-MRS filtering was subsequently performed according to the aforementioned quality-control parameters and assessed visually by experienced spectroscopists for general quality features, namely phasing, fitting, baseline drifting, and the presence of artefacts.

The following exclusion criteria were applied for raw patient ^{1}H-MRS screening (Table D4): (1) missing histological diagnosis; (2) missing water suppressed signals, water reference signals, or structural MR images that indicated the ^{1}H-MRS voxel location; (3) ^{1}H-MRS voxels only partially containing tumours, as determined by visual inspection of the voxel location images produced on the scanner aided by reference to the available image set; (4) very poor FWHM (> 0.15 ppm) of the spectrum; (5) very poor fSNR (< 4) of the spectrum. All the cases that passed quality-control filtering described above were included in the final cohort for the following analysis in this study.

2.3 | Spectroscopy noise suppression

^{1}H-MRS passing quality screening were processed by using a designed framework, adaptive wavelet noise suppression (AWNS), that combines wavelets and a data-driven approach (Figures 1 and E1, Algorithm 1). Firstly, raw ^{1}H-MRS was initially processed by a series of wavelet variations (Table D5) in the frequency domain.[33] Secondly, quantification of both original and postNS ^{1}H-MRS was performed by using TARQUIN as described above. Finally, quality metrics were used as the criteria for selecting the final postNS results. Among the multiple spectra produced from the initial input spectrum, the spectrum was selected based on the fSNR and FWHM for the rest of the study. Apart from overall comparison, ^{1}H-MRS were additionally divided into three groups according to the noise level in each cohort to perform postNS evaluation, providing a poor-quality group (4 < fSNR ≤ 10), a medium-quality group (10 < fSNR < 20), and a good-quality group (fSNR > 20). ^{1}H-MRS NS was conducted by using MATLAB (version 2020a, MathWorks, Natick, Massachusetts, United States).

2.4 | Metabolite analysis

For each individual case, metabolite concentrations were normalised based on the sum of all metabolite concentrations, including lipids and macromolecules, as the features for tumour classification. Metabolites whose CRLB percentage values, according to the raw spectrum, were greater

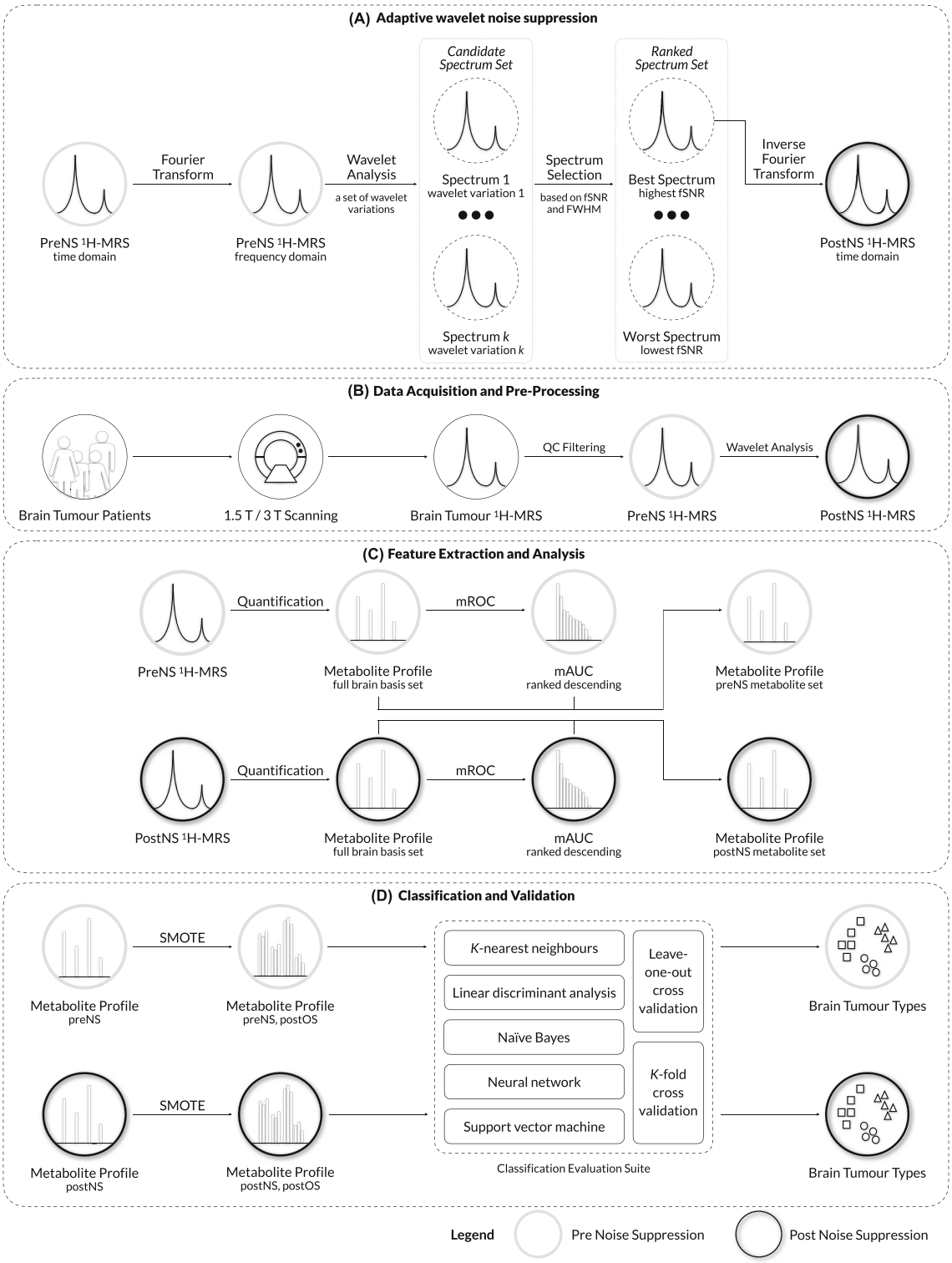

FIGURE 1 Flowchart showing the methods used for clinical magnetic resonance spectroscopy in this project. *Abbreviations:* ^{1}H-MRS, proton magnetic resonance spectroscopy; AWNS, adaptive wavelet noise suppression; FID, free induction decay; NS, noise suppression; SMOTE, synthetic minority oversampling technique; fSNR, fitting-based signal-to-noise ratio; FWHM, full width at half-maximum; mAUC, multiclass area under the curve; PreNS, pre-noise suppression; PostNS, post-noise suppression; PreOS, pre oversampling; PostOS, post oversampling.

than 50% in all cases were excluded (Table D6). The diagnostic ability of a given metabolite to discriminate between the three tumour types was calculated through a multiclass area under the curve (mAUC). The area under the curve (AUC) of the metabolite combinations that consist of highly overlapped metabolites in chemical shifts was further evaluated. Such combinations include (1) creatine and phosphocreatine, at 3.9 ppm;

(2) glucose and glutathione, at 3.8 ppm; (3) glycine and *myo*-Inositol, at 3.6 ppm; (4) *scyllo*-Inositol and taurine, at 3.3 ppm; (5) glycerophosphocholine, phosphocholine, and free choline, at 3.2 ppm; (6) citrate, glutamate, and glutamine, at 2.35 ppm; (7) *N*-acetylaspartate, *N*-acetylaspartylglutamate, lipids at 2.0 ppm, and macromolecules at 2.0 ppm; (8) lipids and macromolecules at 1.3 ppm and lactate; (9) lipids and macromolecules at 0.9 ppm. Between the three PBT types, bootstrapped multiclass receiver operating characteristics (ROC)[34] were used to evaluate the diagnostic ability of individual metabolites and metabolite combinations for discriminating between the three tumour types, and bootstrapped multivariate and multiclass ROC with the multinomial logistic regression were used to evaluate the diagnostic ability for all the metabolites.

2.5 | Tumour classification

Normalised metabolite concentrations were used as the potential features for tumour classification. Metabolite concentrations were ranked and selected according to their mAUC, which demonstrates their diagnostic value across the three tumour types.[35] Features used for classification were set as no more than the number in the minority group minus one. Classification of tumours was performed by using linear and nonlinear classifiers, including linear discriminant analysis, *k*-nearest neighbours, Naïve Bayes, multinomial log–linear models via neural networks, and support vector machines with a linear kernel. Oversampling was performed with an oversampling rate of 100% or 200% for the minority class by using the adaptive synthetic minority oversampling technique (SMOTE) for classification only.[36] The decision tree[37] of tumour classification was assessed by using resubstitution on linear discriminant functions. Classification accuracy was determined by using both leave-one-out and *k*-fold cross-validation, and the statistical significance values of the difference between classification accuracies were adjusted by using Bonferroni correction (Appendix A.2). Test sets were generated through stratified sampling, and *k* was set based on the sample size of the minority class as 10 for the 1.5T cohort and 4 for the 3T cohort.

2.6 | Statistical analysis

The chi-squared test was performed to evaluate the association between patient sex and tumour types.

The Wilcoxon rank-sum test was performed to evaluate the patient age between tumour types and the tumour classification accuracy between preNS and postNS ^{1}H-MRS. The Kruskal–Wallis *H* test was performed to evaluate the noncategorical variables between the three tumour types. The Wilcoxon signed-rank test was performed to evaluate the variables between preNS and postNS, namely the quality-control parameters (fSNR, wSNR, FWHM, CRLB, and CRLB percentage values) of ^{1}H-MRS and the AUC of metabolites. Specifically, the multivariant AUC[38] was used to compare multiple metabolites for their diagnostic ability, where metabolites were combined by performing regression before being used for prediction and calculating ROC. Significance of the statistical analysis was determined when $P < .05$. All statistical analysis and machine-learning experiments were conducted by using R (version 4.2.2, The R Foundation, Vienna, Austria).

3 | RESULTS

3.1 | Demographics

In the final cohort, a total of 83 patients (Table 1) were scanned at 1.5T (57% were male) and 42 patients scanned at 3T (43% were male). Ependymomas were located in the posterior fossa except for three cases in the 1.5T cohort and one in the 3T cohort. Demographic statistics did not suggest significant differences in sex (Table D7) and age (Table D8) between groups ($P > .05$), with the exception of significantly younger ependymoma patients than the other two tumour types in the 1.5T cohort ($P < .05$).

3.2 | Spectral quality

PostNS ^{1}H-MRS showed improved spectral quality across the three tumour types and three quality levels in both 1.5T and 3T cohorts (Table D9). The fSNR was significantly improved from 19 ± 13 to 27 ± 17 ($V = 0$, $P < .001$) in the 1.5T cohort and 16 ± 11 to 21 ± 12 ($V = 0$, $P < .001$) in the 3T cohort (Figures 2,E2–E5). The wSNR was significantly improved from 27 ± 26 to 40 ± 27 ($V = 0$, $P < .001$) in the 1.5T cohort and 25 ± 23 to 35 ± 30 ($V = 0$, $P < .001$) in the 3T cohort (Figures 2 and E2–E5). FWHM was not altered significantly in either 1.5T ($V = 326$, $P = .17$) or 3T ($V = 136$, $P = .21$) postNS ^{1}H-MRS (Figures 2,E2–E5). An example showed reduced residual and slightly altered spectra (Figure 2). CRLB

TABLE 1 Demographic and clinical variables of patients.

	1.5T cohort	3T cohort
Sample size	83	42
Site of primary tumours		
Ependymomas	13†	4‡
Medulloblastomas	31	17
Pilocytic astrocytomas	39	21
Patient age (yrs)		
At diagnosis		
Ependymomas	4.6 ± 5.3	9.2 ± 5.3
Medulloblastomas	6.9 ± 3.5	6.5 ± 4.4
Pilocytic astrocytomas	8.0 ± 3.6	6.3 ± 5.1
At image acquisition		
Ependymomas	4.6 ± 5.3	9.3 ± 5.4
Medulloblastomas	6.9 ± 3.5	6.5 ± 4.4
Pilocytic astrocytomas	8.0 ± 3.6	6.3 ± 5.0
Patient sex (F:M)		
Ependymomas	8:5	3:1
Medulloblastomas	9:22	7:10
Pilocytic astrocytomas	19:20	13:8
Histological subtypes		
Ependymomas		
Grade II	7	2
Grade III, anapaestic	6	2
Medulloblastomas		
Grade IV, classic	26	17
Grade IV, demoplastic-nodular	5	0
Pilocytic astrocytomas		
Grade I	39	21

†Three of the 13 ependymomas were located supratentorially.
‡One of the four ependymomas was located supratentorially.

(Table D10) and CRLB percentage values (Table D11) were observed to be significantly reduced in postNS ^{1}H-MRS ($P < .05$), with the exception of insignificant change of *scyllo*-Inositol from the 1.5T cohort ($V = 244$, $P = 0.194$).

3.3 | Metabolite concentrations

Metabolite concentrations showed significant differences between different tumour types (Table 2) in both the 1.5T and the 3T postNS ^{1}H-MRS, some of which had their significance in postNS ^{1}H-MRS. Most metabolites were consistently significantly different between the three tumour types ($P < .05$) for either prenoise suppression (preNS, Table D12) or postNS ^{1}H-MRS. For the 1.5T cohort, such metabolites included citrate, total choline, total creatine, glutathione, glycine, total lipids and macromolecules at 0.9 ppm, 1.3 ppm, and 2.0 ppm, *myo*-Inositol, *scyllo*-Inositol, and taurine. For the 3T cohort, such metabolites included total choline, total creatine, glutathione, glycine, lactate, macromolecules at 0.9 ppm and 1.3 ppm, *scyllo*-Inositol, and taurine. However, some metabolites were originally significantly different ($P < .05$) but lost their significance ($P > .05$) in postNS ^{1}H-MRS between the three tumour types. Such metabolites included glucose and lactate for the 1.5T cohort and total *N*-acetylaspartate for the 3T cohort. Nevertheless, there were still some other metabolites that were significantly different ($P < .05$) in postNS ^{1}H-MRS and were originally not significantly different ($P > .05$) between the three tumour types, namely citrate, glucose, total lipids, and macromolecules at 0.9 ppm from the 3T cohort.

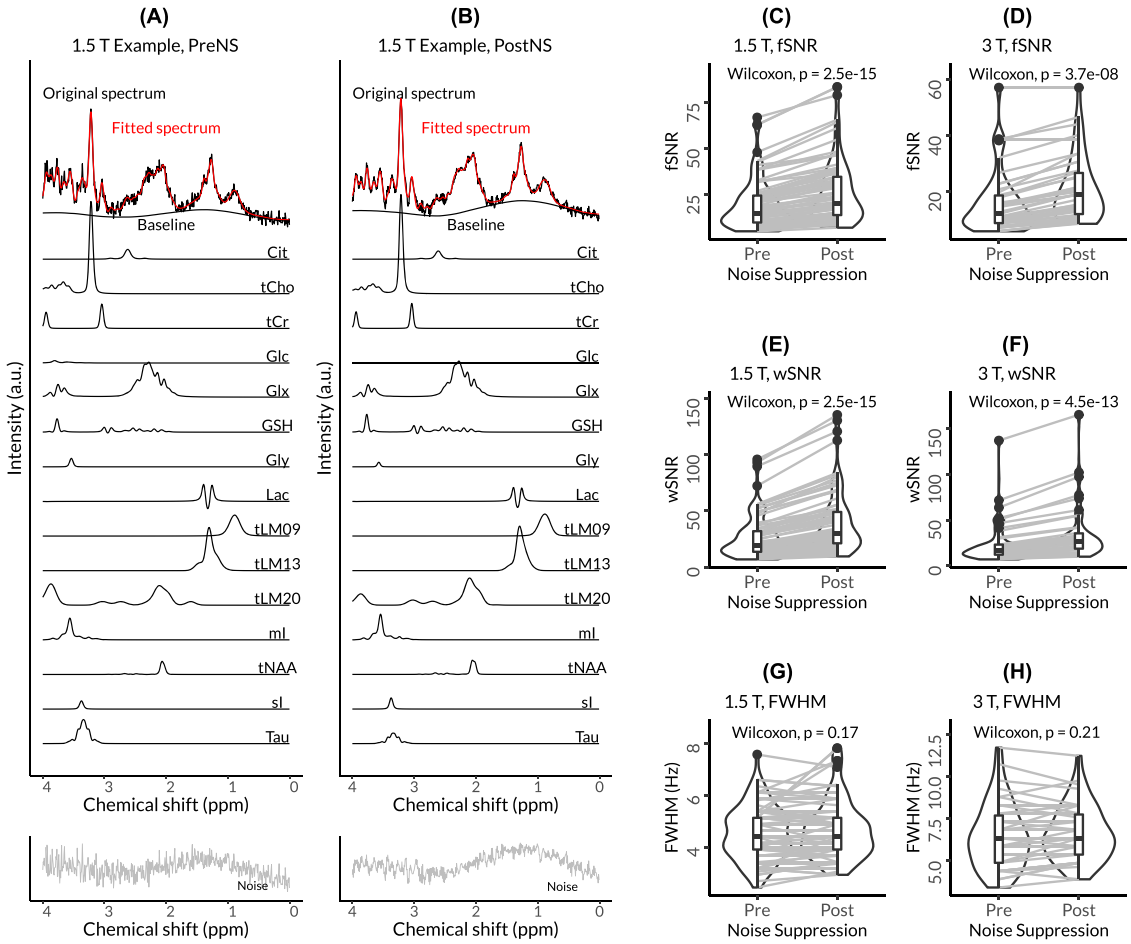

FIGURE 2 Plots showing an example of (A) pre- and (B) post-noise suppression ^{1}H-MRS spectrum for a clinical case, and violin plots comparatively showing the quality-control parameters for (C, E, G) 1.5T and (D, F, H) 3T pre- and post-noise suppression proton magnetic resonance spectroscopy (1.5T, $N = 83$; 3T, $N = 42$), including (C–D) fitting-based signal-to-noise ratio (fSNR) and (E–F) whole-spectrum ratios (wSNR) as well as (G–H) full width at half-maximum (FWHM)

3.4 | Diagnostic ability of metabolites

The mAUC was significantly improved for 10 and 9 out of 16 metabolites in the postNS ^{1}H-MRS for the 1.5T and 3T cohorts, respectively (Tables D13–D14). Among postNS ^{1}H-MRS, the 1.5T cohort (Figure 3) showed that lactate and total lipids and macromolecules at 1.3 ppm ($V = 0$, $P < .001$), combined glutamate, glutamine and citrate ($V = 3486$, $P < .001$), total choline ($V = 3486$, $P < .001$), *scyllo*-Inositol and taurine ($V = 818$, $P < .001$), glycine and *myo*-Inositol ($V = 1$, $P < .001$), and total creatine ($V = 149$, $P < .001$) had significantly improved mAUC, while the combination of *N*-acetylaspartate, *N*-acetylaspartylglutamate, and total lipids and macromolecules at 2.0 ppm ($V = 3486$, $P < .001$) and the combination of glucose and glutathione had significantly reduced mAUC ($V = 3486$, $P < .001$).

At the same time, the 3T cohort (Figure 4) showed significantly improved mAUC for the combination of *N*-acetylaspartate, *N*-acetylaspartylglutamate, and total lipids and macromolecules at 2.0 ppm ($V = 42$, $P < .001$), the combination of glutamate, glutamine, and citrate ($V = 0$, $P < .001$), *scyllo*-Inositol and taurine ($V = 0$, $P < .001$), glycine and *myo*-Inositol ($V = 0$, $P < .001$), and glucose and glutathione ($V = 0$, $P < .001$), whilst it showed significantly decreased mAUC for total choline ($V = 825$, $P < .001$) and total creatine ($V = 739$, $P < .001$).

The ranking of individual metabolites by mAUC differed from preNS to postNS ^{1}H-MRS (Figures 3 and 4). Taking the 3T pre-oversampling cohort as an example, the top four metabolites in preNS ^{1}H-MRS were total *N*-acetylaspartate (mAUC = 0.80 ± 0.01), glycine (mAUC = 0.79 ± 0.01), total choline (mAUC = 0.76 ± 0.01), and *scyllo*-Inositol (mAUC = 0.74 ± 0.01). However, the top four metabolites in postNS ^{1}H-MRS were glycine (mAUC = 0.85 ± 0.01), total lipids and macromolecules at 2.0 ppm (mAUC = 0.81 ± 0.01), glutamate and glutamine (mAUC = 0.79 ± 0.01), and total *N*-acetylaspartate (mAUC = 0.79 ± 0.01). Both the 1.5T and 3T postNS ^{1}H-MRS provides improved multivariate AUC in comparison with preNS ^{1}H-MRS (Figures 3 and 4). However, only the 1.5T cohort showed significant improvements ($W = 41$, $P < .01$), whilst the 3T cohort showed insignificant improvement ($W = 70$, $P = .13$).

TABLE 2 Estimated metabolite concentrations in mmol from 1.5T and 3T post-noise suppression proton magnetic resonance spectroscopy.

	Ependymomas		Medulloblastomas		Pilocytic astrocytomas		All		Kruskal–Wallis *H*		
Metabolites	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	***H***	***P***	
1.5T Cohort	*N* = 13		*N* = 31		*N* = 39		*N* = 83				
Citrate	1.1	0.5	0.7	0.5	0.4	0.3	0.6	0.5	22.4	< .000 1	****
total Choline	2.4	0.7	3.8	1.9	1.0	0.6	2.2	1.8	51.9	< .000 1	****
total Creatine	5.4	2.4	3.2	1.8	1.8	1.9	2.9	2.3	23.0	< .000 1	****
Glucose	0.8	1.2	0.7	1.1	1.0	1.1	0.8	1.1	3.9	.141 2	ns
Glutamate and Glutamine	7.0	1.6	6.3	2.4	8.0	4.9	7.2	3.8	1.6	.449 7	ns
Glutathione	0.7	0.5	0.6	0.6	0.2	0.4	0.4	0.6	11.6	.003 1	**
Glycine	0.7	1.3	3.1	2.8	0.0	0.1	1.3	2.3	54.5	< .000 1	****
Lactate	1.9	1.2	2.4	1.1	2.0	1.5	2.1	1.3	3.8	.146 7	ns
total LM at 0.9 ppm	3.9	1.6	8.8	4.7	4.4	1.5	6.0	3.8	30.0	< .000 1	****
total LM at 1.3 ppm	10.5	6.8	22.7	15.1	9.2	4.9	14.4	11.9	30.2	< .000 1	****
total LM at 2.0 ppm	8.8	2.6	11.4	4.8	6.5	3.3	8.7	4.4	25.3	< .000 1	****
myo-Inositol	12.4	7.0	1.6	2.3	3.4	2.6	4.1	5.1	30.5	< .000 1	****
total *N*-acetylaspartate	1.5	1.4	1.2	0.8	1.2	1.0	1.2	1.0	0.3	.858 3	ns
scyllo-Inositol	0.7	0.6	0.7	0.7	0.0	0.1	0.4	0.6	38.7	< .000 1	****
Taurine	2.7	1.8	2.8	2.8	1.3	1.3	2.1	2.1	7.2	.027 8	*
3T Cohort	*N* = 4		*N* = 17		*N* = 21		*N* = 42				
Citrate	0.5	0.6	0.9	1.1	0.2	0.3	0.5	0.8	11.4	.003 4	**
total Choline	1.3	0.5	5.0	4.2	1.6	1.0	2.9	3.2	20.0	< .000 1	***
total Creatine	2.0	1.2	5.1	3.8	3.2	2.1	3.8	3.0	8.7	.013 2	*
Glucose	0.8	0.6	2.4	2.6	0.7	1.2	1.4	2.0	10.0	.006 9	**
Glutamate and Glutamine	3.5	0.8	5.7	5.1	6.2	3.2	5.7	4.0	3.3	.194 3	ns
Glutathione	0.8	0.2	2.3	1.0	1.0	0.7	1.5	1.0	16.8	.000 2	***
Glycine	1.6	1.2	3.9	2.0	0.5	0.6	2.0	2.1	24.7	< .000 1	****
Lactate	1.8	1.1	2.2	1.6	0.9	0.6	1.5	1.3	7.0	.030 8	*
total LM at 0.9 ppm	3.2	1.2	6.5	5.0	5.2	5.9	5.6	5.3	8.7	.013 1	*
total LM at 1.3 ppm	12.3	11.3	23.4	19.9	9.9	9.9	15.6	16.0	14.4	.000 8	***
total LM at 2.0 ppm	6.5	2.3	8.3	4.7	6.5	4.8	7.2	4.6	3.8	.147 5	ns
myo-Inositol	2.5	1.2	3.5	7.0	2.9	4.1	3.1	5.3	0.6	.759 0	ns
total *N*-acetylaspartate	1.0	0.6	1.3	0.9	2.1	1.4	1.7	1.2	5.0	.082 6	ns
scyllo-Inositol	0.0	0.1	0.6	1.7	0.1	0.2	0.3	1.1	9.2	.010 2	*
Taurine	1.2	0.9	6.3	4.6	1.3	1.3	3.3	4.0	21.0	< .000 1	****

The concentrations of metabolites as well as lipids and molecules (LM) prior to normalisation. *P* values, which evaluated the difference between the three tumour types, were calculated through Kruskal–Wallis *H* tests and highlighted with significant levels of < .05 as *, < .01 as **, and < .001 as ***.

3.5 | Classification assessment

The decision tree of tumour classification showed misclassified cases in each group (Figure 5). Misclassified cases are dominated by borderline cases, most of which are ependymomas. In general, postNS ^{1}H-MRS with or without oversampling both showed fewer misclassified tumours compared with the preNS and pre-oversampling ^{1}H-MRS. By performing NS on the 1.5T cohort, misclassified tumours were reduced from 7 (8.4%) to 5 (6.0%) before and from 7 to 3 (3.6%) after performing oversampling. Among these misclassified tumours, 2 out of 7 were ependymomas, reduced to 1 out of 5 before performing oversampling. For post-oversampling classification, misclassified cases were reduced from 7 to 3 for postNS ^{1}H-MRS, among which there was still one misclassified ependymoma. As for the 3.0T cohort, such overall reduction only showed in the post-oversampling 3.0T cohort as 5 (11.9%) to 4 (9.5%) by performing ^{1}H-MRS NS and two ependymomas remained misclassified after performing NS.

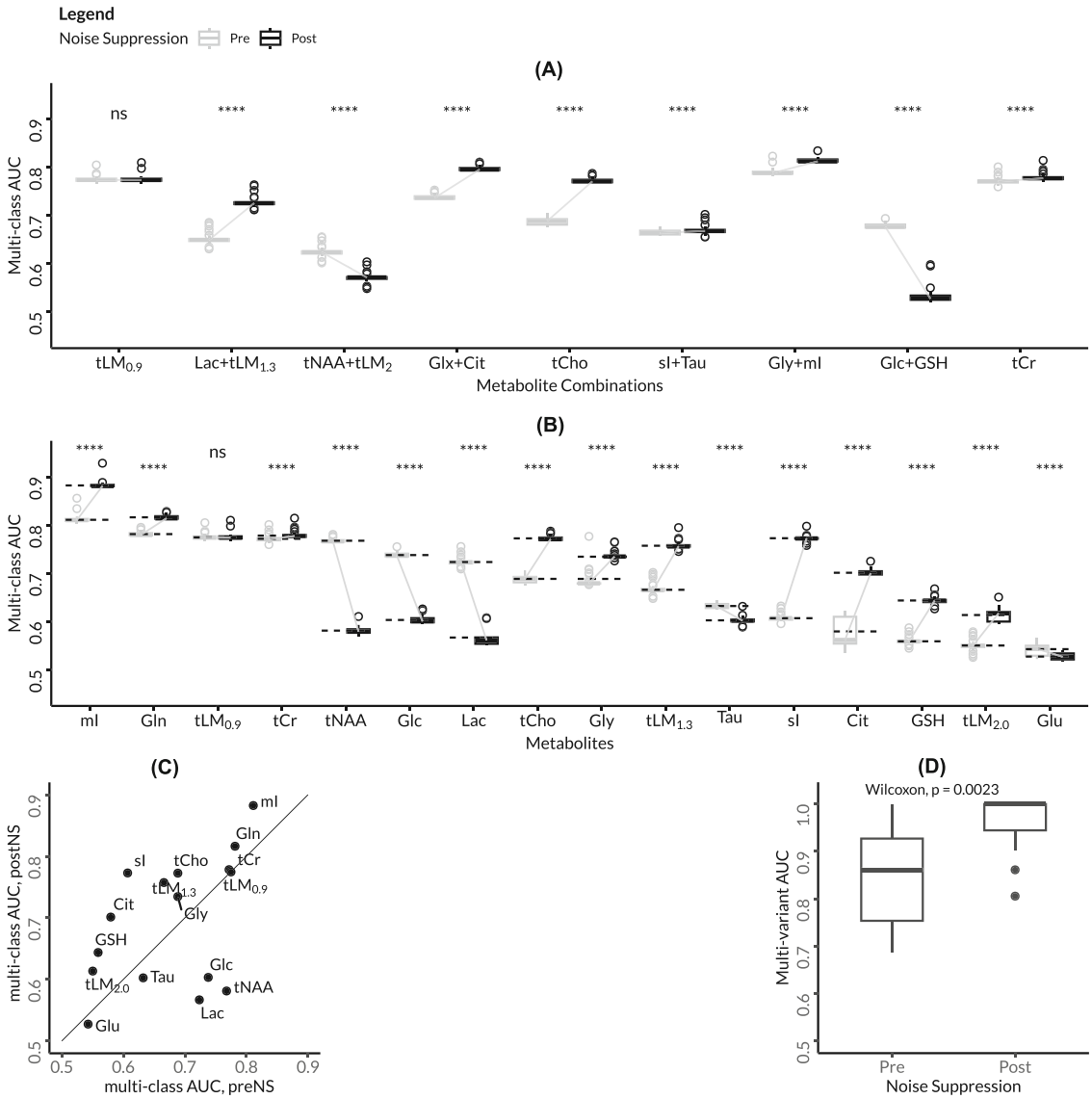

FIGURE 3 Box and scatter plots describing the diagnostic ability alteration of metabolites as (A) combinations or (B) individuals through multiclass (C) univariant or (D) multivariant receiver operating characteristics in the 1.5T cohort ($N = 83$), where the difference was compared between pre (left) and post (right) noise suppression.

Overall and balanced classification accuracies showed significant improvement after NS generally ($P < .05$, Figures 6 and E6–E9, Tables D15–D18). Prior to oversampling, NS had significantly improved tumour classification for both the 1.5T and 3T cohorts ($P < .05$). For the 3T cohort in particular, oversampling improved classification further ($P < .05$).

Optimised cross-validated classification for the 1.5T and 3T cohorts was achieved through postNS ^{1}H-MRS with Naïve Bayes. The balanced classification accuracy of the 1.5T cohort was improved to 86% from 76% through postNS ^{1}H-MRS. After oversampling, the optimal classification accuracy was achieved as 88% overall and 86% balanced. For the 3T cohort, the overall classification accuracy was improved to 74% from 69% and balanced accuracy to 55% from 51%. After performing oversampling, the overall and balanced classification accuracy was improved to 76% from 62% and 56% from 46% through postNS ^{1}H-MRS, respectively (Table 3, Figure 6). The results of *k*-fold cross-validation also showed significant improvement of classification accuracy after NS and oversampling (Figures 6,E6–E9).

According to classification measures, postNS ^{1}H-MRS showed optimal overall classification accuracy when combined with oversampling for both the 1.5T and 3T cohorts (Table 3). Ependymoma identification was improved after performing NS in postNS ^{1}H-MRS, showing as the improved F_1 score from 0.67 to 0.76 for the 1.5T cohort and 0 to 0.29 for the 3T cohort. After performing oversampling, optimal ependymoma identification was achieved for the 1.5T cohort, showing as an F_1 score improved from 0.69 ± 0.02 to 0.80 ± 0.03. For the 3T cohort, ependymoma identification was not improved clearly through oversampled ^{1}H-MRS.

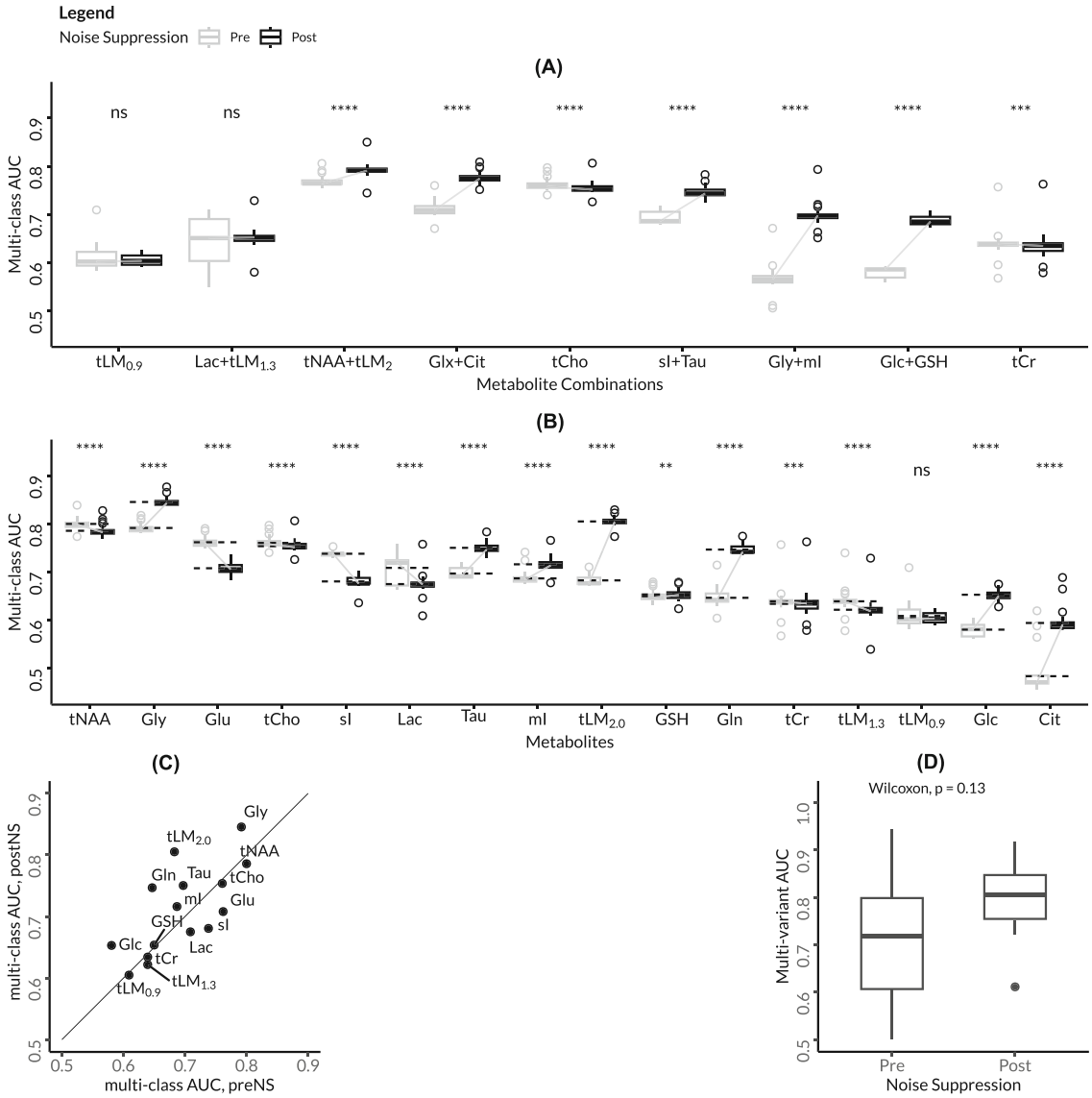

FIGURE 4 Box and scatter plots describing the diagnostic ability alteration of metabolites as (A) combinations or (B) individuals through multiclass (C) univariant or (D) multivariant receiver operating characteristics in the 3T cohort ($N = 42$), where the difference was compared between pre (left) and post (right) noise suppression.

4 | DISCUSSION

This study aimed to address the issue of ^{1}H-MRS spectral quality in PBT classification through performing NS on a multisite dataset. PostNS ^{1}H-MRS showed significantly improved fSNR and insignificantly altered FWHM across all three tumour types and quality levels. Machine-learning based classification performance was significantly improved through postNS ^{1}H-MRS. Although some metabolites showed decreased mAUC, most metabolites showed increased mAUC in postNS ^{1}H-MRS. This finding corresponds to the final improved classification accuracy through all the classifiers and indicates the potential utility of postNS ^{1}H-MRS for PBT diagnosis. The improved classification performance of postNS ^{1}H-MRS might indicate better metabolite concentrations for tumour diagnosis. Wavelets are a useful tool and expected to enhance the clinical value of ^{1}H-MRS. Clinical ^{1}H-MRS often has a limited fSNR (< 10) due to inherent lack of sensitivity and limitations on acquisition time. The metabolite concentrations estimated from such ^{1}H-MRS spectra may be inaccurate, which will lead to poor clinical performance of ^{1}H-MRS. Higher field strength with an optimised scanning protocol may be able to increase the resolution of MR spectra and the certainty of metabolite concentration determination. However, the spectral noise of ^{1}H-MRS may still be a problem in clinical practice and be a barrier to adoption. For example, the size of voxels is restricted for a small tumour to avoid partial volume effects, and such a small voxel will inevitably lead to noisy ^{1}H-MRS. Wavelets can provide a flexible solution for ^{1}H-MRS with a wide range of noise levels. The frequency-uniform filter used by wavelets can preserve the probability distribution of noise (Appendix C),[39] indicating the valid use of CRLB for postNS ^{1}H-MRS spectra.[25]

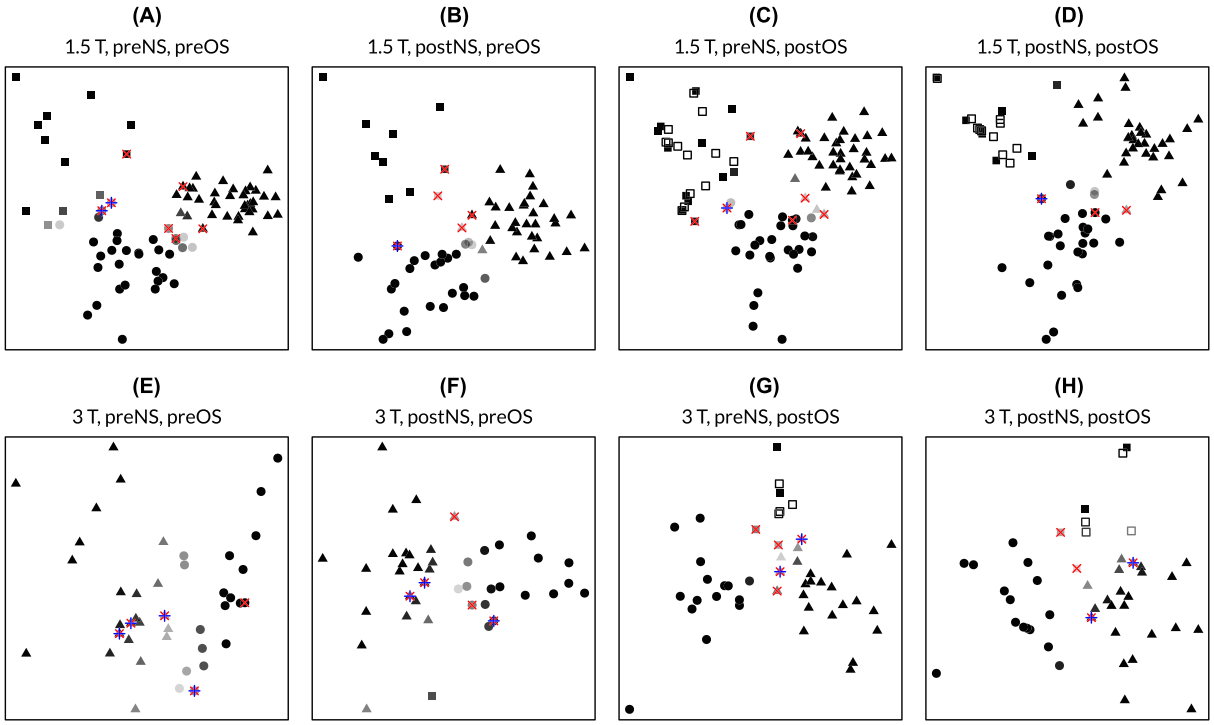

FIGURE 5 Two-dimensional scatter plots showing the classification performance of common paediatric brain tumours with discriminant functions (DF) through resubstitution and eligible metabolites. Classification was comparatively evaluated between (A, C, E, G) pre- and (B, D, F, H) post-noise suppression (NS) and (A–B, E–F) pre- and (C–D, G–H) post-oversampling (OS) 1.5T (A–D, preOS $N = 83$, postOS $N = 96$) and 3T (E–H, preOS $N = 42$, postOS $N = 46$) proton magnetic resonance spectroscopy. Investigated tumour types include ependymomas as square markers, medulloblastomas as circular markers, and pilocytic astrocytomas as triangle markers, among which ependymomas were oversampled, shown by white square markers, to 200% as comparison (C–D, G–H). Uncertainly classified cases are shown as transparent, with the contrast indicating the probability of classification. Misclassified cases are marked with a cross mark, and misclassified ependymomas are additionally marked with a plus mark.

AWNS achieves a robust performance by combining wavelet analysis with a data-driven approach. The use of wavelets has been considered for ^{1}H-MRS NS, but previous studies are limited to exploring the performance of specific wavelet bases and variations.[29] However, the unpredictable noise in clinical ^{1}H-MRS makes it challenging to select a universal wavelet variation that is optimal in all situations. This question remains unanswered in previous wavelet-related approaches. AWNS is designed to address this issue, where wavelet variations are selected adaptively through a data-driven approach, where the selection is guided by quality-control parameters. In such a way, AWNS can theoretically preserve the metabolite-related spectral components whilst largely removing signals originating from other sources. In addition, the improvement of fSNR brought by AWNS remains robust for high-SNR spectra, and the CRLB for most metabolites can be reduced as well. This is because AWNS considers the existence of noise and tries to suppress even low-level noise instead of only showing efficiency for low-SNR spectra.

Previous approaches to ^{1}H-MRS NS evaluated their performance by observing the SNRs of experimental data, where the SNRs are defined by using regional noise.[29,40] Although recommended by the recent consensus,[24] regional SNRs are unlikely to be able to reflect the accuracy of quantified metabolite concentrations, because the fitting error that is relevant to ^{1}H-MRS quantification has not been considered. In contrast, AWNS uses fSNR (Equation A.1) to guide NS, thus the fitting performance of metabolites has been considered. Furthermore, the ^{1}H-MRS obtained from simulated, phantom, or healthy human brain in previous studies may contain only a few metabolites that have relatively simpler line shapes.[27,29,40] This leaves the question about how useful such approaches can be in clinical studies. To address this question, AWNS is designed by considering the need of clinical questions and assessed by evaluating the case-by-case fSNR alteration and targeting the diagnostic performance of ^{1}H-MRS. Given the complexity of the brain tumour ^{1}H-MRS used in this study, AWNS could be a more robust, powerful, and practical solution for real-world ^{1}H-MRS.

AWNS might be able to improve the ^{1}H-MRS quantification performance. A naive wavelet decomposition and reconstruction process on a time-domain signal simply removes the signal components according to their locations in the frequency domain, and in most cases the signal components with higher frequency are considered as the noise to remove. Considering ^{1}H-MRS, the self-repeated spectral components, which contain little metabolite information, are considered as noise. To make sure the removed spectral components are dominated by noise and the preserved ones are mainly contributed to by metabolites, AWNS optimises wavelet variations by maximising the fSNR and keeping the FWHM within an acceptable range. As a result, postNS ^{1}H-MRS showed significantly lower median CRLB for all metabolites and CRLB percentage values

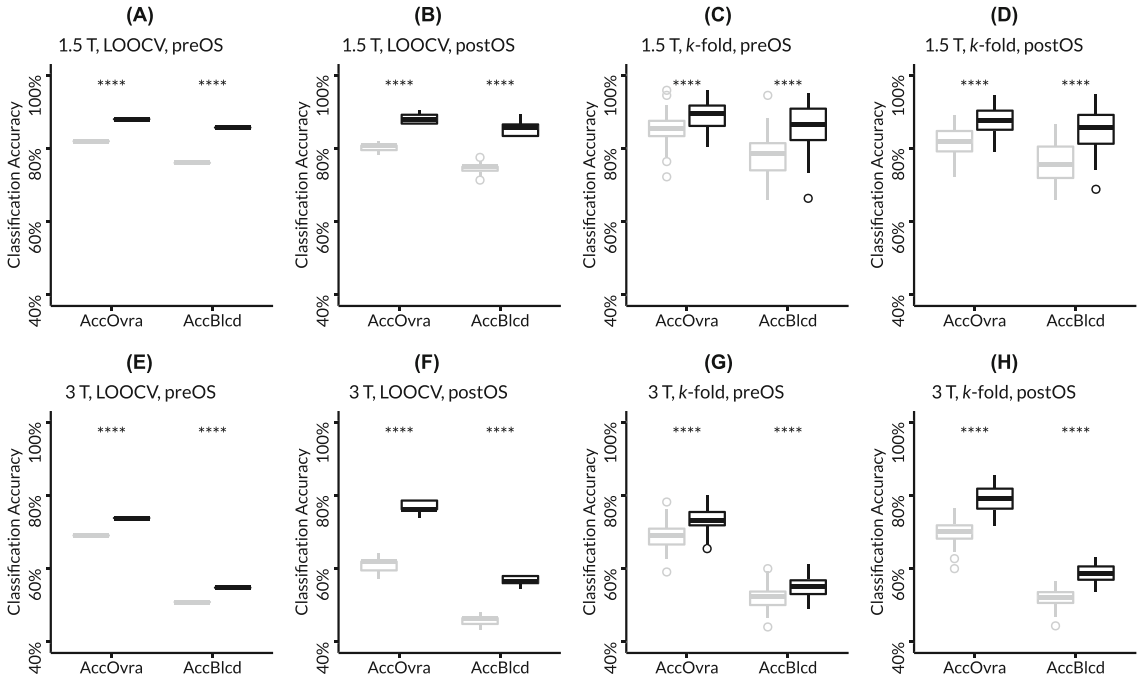

FIGURE 6 Box plots showing the significantly improved overall and balanced classification accuracy (AccOvra and AccBlcd) of the three brain tumour types, ependymomas, medulloblastomas, and pilocytic astrocytomas, determined through (A–B, E–F) leave-one-out (LOO) and (C–D, G–H) *k*-fold cross-validation (CV) and Naïve Bayes for (A, C, E, G) pre- and (B, D, F, H) post-oversampling (OS) 1.5T (A–D, preOS $N = 83$, postOS $N = 96$) and 3T (E–H, preOS $N = 42$, postOS $N = 46$) pre- (grey) and post- (black) noise suppression (NS) proton magnetic resonance spectroscopy, where oversampling was performed for ependymomas with an oversampling rate of 100%. Level of significance: ****, $P < .0001$

for most metabolites. This is because CRLB[41] assesses whether the observation has unknown probability distributions by providing precision estimators, and the removal of the irrelevant information makes the distribution clearer. In this process, AWNS does not add any new information to spectra, but it learns from the prior knowledge of existing metabolites in the acquired ^{1}H-MRS and attempts to remove the spectral components that are irrelevant to such prior knowledge as much as possible. Since quantification[24] is undertaken by fitting spectra with prior knowledge, AWNS may thereby be able to improve the quantification performance. Compared with apodisation, a technique that increases the line width theoretically and therefore is only recommended for visualisation purposes,[42] AWNS can keep the line width stable. Some metabolites are highly overlapping in spectra, which makes it a challenge to increase the fitting performance for these overlapping metabolites and a NS method that increases line width is undesirable.

The diagnostic ability of some metabolites has been different in postNS ^{1}H-MRS. The results show that the alteration of metabolites' diagnostic ability, determined through mAUC, varies after NS, even though most metabolites, as well as most metabolite combinations, showed significantly improved mAUC. The potential reasons for this phenomenon include the vulnerability of some metabolite spectra to noise, overlapping metabolite spectra, and the nature of metabolites itself. Firstly, metabolites that have slightly higher intensity than noise can be fitted successfully, but the results are affected by noise, thus these metabolites can potentially be estimated more accurately in postNS ^{1}H-MRS. For instance, the combination of glutamate, glutamine, and taurine showed stably better diagnostic ability in both 1.5T and 3T postNS ^{1}H-MRS. Secondly, some metabolites overlap to such an extent that they are difficult to identify separately even if the noise has been suppressed and so little may be gained by suppressing spectral noise. A typical example is the refined glycine in the 3T results that showed high mAUC, whilst it was mixed with *myo*-Inositol in 1.5T spectra and therefore showed relatively lower mAUC. Lastly, it is assumed that more accurate metabolite concentrations can provide improved classification performance of tumours, but some metabolites may then have less powerful diagnostic ability.

PostNS ^{1}H-MRS improves the classification performance for these PBTs. The results of classification accuracy through metabolite selection showed a significant improvement by postNS ^{1}H-MRS rather than preNS ^{1}H-MRS. Metabolite selection is a newly proposed method of feature extraction for machine learning in PBT classification, which showed advantages over PCA.[35] Our previous results showed improved classification accuracy by using PCA on postNS ^{1}H-MRS as well,[33] indicating the robustness of noise suppression for ^{1}H-MRS based tumour classification. It remains unknown whether postNS ^{1}H-MRS will provide more accurate metabolite concentration determination, but the classification performance is improved, showing the advantage of ^{1}H-MRS noise suppression in clinical decision making.

Decision evaluation through resubstitution showed that postNS ^{1}H-MRS provides fewer misclassified cases, which indicates that postNS ^{1}H-MRS has the potential of improving clinical diagnosis of PBTs. However, misclassified cases were still presenting and also mostly borderline cases.

TABLE 3 Classification measures comparing pre- and post-noise suppression proton magnetic resonance spectroscopy with Naive Bayes.

Oversampling	Pre		Post	
Noise suppression	Pre	Post	Pre	Post
1.5T Cohort ($N = 83$)				
All tumour types				
Accuracy, overall	0.81	0.88	0.81 ± 0.01	0.88 ± 0.01
Accuracy, balanced	0.76	0.86	0.74 ± 0.01	0.86 ± 0.02
Matthew's correlation coefficient	0.71	0.81	0.67 ± 0.02	0.79 ± 0.02
Ependymomas				
F_1 score	0.52	0.74	0.52 ± 0.02	0.69 ± 0.03
Precision	0.50	0.71	0.48 ± 0.02	0.66 ± 0.02
Recall	0.54	0.77	0.54 ± 0.02	0.77 ± 0.02
Sensitivity	0.54	0.77	0.54 ± 0.02	0.80 ± 0.02
Specificity	0.90	0.94	0.89 ± 0.01	0.93 ± 0.01
3T Cohort ($N = 42$)				
All tumour types				
Accuracy, overall	0.69	0.74	0.62 ± 0.03	0.76 ± 0.01
Accuracy, balanced	0.51	0.55	0.46 ± 0.02	0.56 ± 0.02
Matthew's correlation coefficient	0.48	0.54	0.37 ± 0.04	0.58 ± 0.02
Ependymomas				
F_1 score	0.00	0.00	0.00	0.00
Precision	0.00	0.00	0.00	0.00
Recall	0.00	0.00	0.00	0.00
Sensitivity	0.00	0.00	0.00	0.00
Specificity	0.84	0.95	0.82 ± 0.02	0.89 ± 0.01

It is noteworthy that resubstitution indicates ependymomas were classified well, but classification accuracy estimated through cross-validation indicates ependymomas were poorly classified. This indicates that ependymomas are diverse and often different from other groups, and such ependymomas are hard to classify, since prior knowledge of them does not exist in the training set when performing cross-validation, even when oversampling has been performed. At the same time, these results also indicate that postNS ^{1}H-MRS may have limited ability for classifying these diverse ependymomas but can improve the classification for the rest tumours.

Oversampling through SMOTE was performed in this study because the ependymomas had a much smaller sample size than the other two tumour classes and the balanced classification accuracy was consequently limited. The improvement of classification accuracy contributed by oversampling ependymomas depended on the machine-learning classifiers used. Whether postNS ^{1}H-MRS provides more sustainably improved classification performance than oversampling is unclear. The results have not suggested whether the combination of NS and oversampling always outperforms only NS or oversampling, although either of them can be helpful. NS aims to improve the classification accuracy through providing more accurately estimated metabolite concentrations, while oversampling aims to rebalance the group size by creating artificial cases for the minority group. Therefore, whether NS or oversampling is more helpful will be dependent on the cohort size and the ^{1}H-MRS spectral quality.

The clinical utility of ^{1}H-MRS NS for the tumours in this study should also be considered. The tumours all have surgical resection as their preferred initial treatment. However, the extent of surgical resection required depends greatly on the tumour type. A complete macroscopic resection is crucial if long-term survival is to be achieved in ependymomas, with even small fragments of residual tumour being difficult to control with adjuvant radiotherapy and chemotherapy. Conversely, in cases of pilocytic astrocytoma, residual tumour post-surgery often requires no further intervention, since many cases do not experience further tumour growth. In medulloblastomas, small tumour residuals post-surgery can often but not always be successfully treated with radiotherapy and chemotherapy. Prior knowledge of the tumour type is therefore important in surgical decision making. Furthermore, complex planning of radiotherapy and in particular proton therapy, often at a centre remote to the surgery, benefits from early initiation rather than when a final histological or molecular diagnosis is available several days later. At the same time, families find the time between the initial diagnostic MRI and definitive diagnosis particularly challenging, and an early noninvasive diagnosis can improve the quality of discussions with the family. In addition to the direct clinical improvement that could result for the tumour types used in this study,

the method should be readily applicable to other tumour types for which surgical resection is not the preferred initial treatment and an accurate noninvasive diagnosis is particularly important.

Although more accurate determination of metabolite concentrations could be obtained through optimising MR sequences, increasing scanning durations, or applying higher field strength, limitations imposed by clinical practice determine that ^{1}H-MRS still suffers from low SNR and improvements in postprocessing remain important. Current clinical MR investigations are often undertaken on 1.5T scanners, particularly where spinal imaging is required. Where ^{1}H-MRS is undertaken at higher field strength, smaller voxel sizes may be used, negating the advantages. Clinically applicable sequences are also limited in acquisition time, especially for children, as longer scanning duration may not be tolerated. Instead, NS as postprocessing can assist in improving the diagnostic accuracy of metabolite concentrations, which can be applied widely across different scanners and protocols. In addition to improving the diagnostic performance of metabolites for clinical ^{1}H-MRS, NS may also make it possible to use some ^{1}H-MRS data that fail to meet the SNR quality control screening.

The study has a series of limitations. For the method itself, the selection of wavelets is unsupervised and only determined according to the fSNRs that are dominated by fitting residual, thus the risk of introducing signal-dependent noise variance[43] has not been addressed. Regarding the cohort, there is a limited sample size, particularly in the 3T cohort for its relatively smaller voxel size that leads to more noise in spectra. Meanwhile, the ependymoma patients are relatively younger than the remaining two groups in the 1.5T cohort, which could be due to higher frequency of ependymomas in younger population.[44] The ground truth of metabolite concentrations of in vivo brain tissues is not known and so there is no definitive proof that the concentrations are determined more accurately in postNS ^{1}H-MRS. Therefore, this study was undertaken with the hypothesis that the metabolite concentrations determined from postNS ^{1}H-MRS might provide better classification accuracy. Meanwhile, the study of PBT classification is limited by the methodology and sample size. It is challenging to know the ground truth of metabolite concentrations of in vivo brain tissues. Consequently, this study is under the hypothesis that the metabolite concentrations determined from postNS ^{1}H-MRS might provide better classification accuracy. Cell-signalling pathways in PBTs suggest association between metabolites[45], whilst metabolites were generally considered as independent during this classification. The limited classification accuracy suggested by the 3T cohort could be due to the small ependymoma group, which meant that the range of ^{1}H-MRS for this diagnosis would not be fully represented but also limited the number of features we allowed in the classification. Higher accuracy would be expected in a dataset containing more ependymoma cases.

Considering noise suppression for clinical ^{1}H-MRS, this article only presents a method that can improve the diagnostic ability of ^{1}H-MRS by suppressing the noise as an initial step. As following work, further optimisation for wavelet computing[46] and selection[28] is required, prior to making the method available as a software package for being translated into clinical practice. The observed improvement of fSNR by AWNS seems to be related to metabolite concentration levels, which means the performance of AWNS could be related to the spectral line shape. Therefore, further assessment of AWNS will address not only simulated and phantom ^{1}H-MRS, but also in vivo ^{1}H-MRS acquired from multiple types of tissues.

5 | CONCLUSIONS

Noise suppression for clinical ^{1}H-MRS can provide significantly improved spectral quality, metabolite concentrations with increased diagnostic ability, and better classification performance for paediatric brain tumours.

AUTHOR CONTRIBUTIONS

Conceptualisation: Teddy Zhao and Andrew C. Peet. *Methodology*, *software*, *validation*, *formal analysis*, *investigation*, *data curation*, *writing—original draft*, *visualisation*: Teddy Zhao. *Resources*: all authors. *Writing—review and editing*: Teddy Zhao, Andrew C. Peet, Barry Pizer, Dorothee P. Auer, Heather E. L. Rose, Lesley MacPherson, James T. Grist, Martin Wilson, Nigel P. Davies, and Theodoros N. Arvanitis. *Supervision*: James T. Grist and Andrew C. Peet. *Project administration*, *Funding acquisition*: Andrew C. Peet.

ACKNOWLEDGEMENTS

NIHR Research Professorship, Grant/Award Numbers: NIHR-RP-R2-12-019; Help Harry Help Others; UK National Institute of Health Research's Nottingham Biomedical Research Centre; Cancer Research UK and EPSRC Cancer Imaging Programme at the Children's Cancer and Leukaemia Group (CCLG) in association with the MRC and Department of Health (England), Grant/Award Numbers: C7809/A10342; Cancer Research UK and NIHR Experimental Cancer Medicine Centre Paediatric Network, Grant/Award Numbers: C8232/A25261; Children's Research Fund; Birmingham Women's and Children's Hospital Charities; The Children's Cancer and Leukaemia Group Little Princess Trust, Grant/Award Numbers: 2017/15 and 2019/01; Children with Cancer, Grant/Award Numbers: 15/118; Action Medical Research and the Brain Tumour Charity, Grant/Award Numbers: GN2181; Health Data Research UK (HDR UK). The reviewers are genuinely appreciated for their valuable and constructive comments to this article.

CONFLICTS OF INTEREST STATEMENT

The authors declare no potential conflicts of interest.

FINANCIAL DISCLOSURE

None reported.

CODE AVAILABILITY

The code of Adaptive Wavelet Noise Suppression is available at https://doi.org/10.25500/edata.bham.00000981 or can be requested by contacting the first author at dxz709@alumni.bham.ac.uk.

ORCID

Teddy Zhao https://orcid.org/0000-0001-7727-3638
James T. Grist https://orcid.org/0000-0001-7223-4031
Dorothee P. Auer https://orcid.org/0000-0002-4745-3635
Shivaram Avula https://orcid.org/0000-0003-1742-7206
Simon Bailey https://orcid.org/0000-0003-4763-4329
Nigel P. Davies https://orcid.org/0000-0002-7978-5511
Richard G. Grundy https://orcid.org/0000-0002-2585-2539
Omar Khan https://orcid.org/0000-0001-6851-3804
Lesley MacPherson https://orcid.org/0000-0001-5017-3135
Paul S. Morgan https://orcid.org/0000-0002-5870-1446
Heather E. L. Rose https://orcid.org/0000-0002-0346-1334
Yu Sun https://orcid.org/0000-0001-6729-139X
Martin Wilson https://orcid.org/0000-0002-2089-3956
Theodoros N. Arvanitis https://orcid.org/0000-0001-5473-135X
Andrew C. Peet https://orcid.org/0000-0002-4846-5152

REFERENCES

1. Aldape K, Brindle KM, Chesler L, et al. Challenges to curing primary brain tumours. *Nat Rev Clin Oncol*. 2019;16(8):509-520.
2. Steliarova-Foucher E, Colombet M, Ries LAG, et al. International incidence of childhood cancer, 2001-10: a population-based registry study. *Lancet Oncol*. 2017;18(6):719-731.
3. Rickert CH, Paulus W. Epidemiology of central nervous system tumors in childhood and adolescence based on the new who classification. *Child's Nerv Syst*. 2001;17(9):503-511.
4. Collins VP, Jones DT, Giannini C. Pilocytic astrocytoma: pathology, molecular mechanisms and markers. *Acta Neuropathol*. 2015;129(6):775-788.
5. Northcott PA, Robinson GW, Kratz CP, et al. Medulloblastoma. *Nat Rev Dis Primers*. 2019;5(1–20):11.
6. Kilday JP, Rahman R, Dyer S, et al. Pediatric ependymoma: Biological perspectives. *Mol Cancer Res*. 2009;7(6):765-786.
7. 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.
8. Pollack IF, Jakacki RI. Childhood brain tumors: epidemiology, current management and future directions. *Nat Rev Neurol*. 2011;7(9):495-506.
9. Parsons MW, Whipple NS, Poppe MM, Mendez JS, Cannon DM, Burt LM. The use and efficacy of chemotherapy and radiotherapy in children and adults with pilocytic astrocytoma. *J Neurooncol*. 2021;151(2):93-101.
10. Dressler EV, Dolecek TA, Liu M, Villano JL. Demographics, patterns of care, and survival in pediatric medulloblastoma. *J Neurooncol*. 2017;132(3):497-506.
11. Swanson EL, Amdur RJ, Morris CG, et al. Intracranial ependymomas treated with radiotherapy: long-term results from a single institution. *J Neurooncol*. 2011;102(3):451-457.
12. Wilne SH, Dineen RA, Dommett RM, Chu TP, Walker DA. Identifying brain tumours in children and young adults. *BMJ*. 2013;347:f5844.
13. 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.
14. Regis J, Bouillot P, Rouby-Volot F, Figarella-Branger D, Dufour H, Peragut JC. Pineal region tumors and the role of stereotactic biopsy: Review of the mortality, morbidity, and diagnostic rates in 370 cases. *Neurosurg*. 1996;39(5):907-914.
15. Orphanidou-Vlachou E, Vlachos N, Davies NP, Arvanitis TN, Grundy RG, Peet AC. Texture analysis of t1- and t2-weighted mr images and use of probabilistic neural network to discriminate posterior fossa tumours in children. *NMR Biomed*. 2014;27(6):632-639.
16. Jeener J, Meier BH, Bachmann P, Ernst RR. Investigation of exchange processes by two-dimensional nmr spectroscopy. *J Chem Phys*. 1979;71(11):4546-4553.
17. Barker PB, Lin DDM. In vivo proton mr spectroscopy of the human brain. *Prog Nucl Magn Reson Spectrosc*. 2006;49(2):99-128.
18. Faubert B, Solmonson A, DeBerardinis RJ. Metabolic reprogramming and cancer progression. *Science*. 2020;368(6487):eaaw5473.
19. Haddadin IS, McIntosh A, Meisamy S, et al. Metabolite quantification and high-field mrs in breast cancer. *NMR Biomed*. 2009;22(1):65-76.
20. Preul MC, Caramanos Z, Collins DL, et al. Accurate, noninvasive diagnosis of human brain tumors by using proton magnetic resonance spectroscopy. *Nat Med*. 1996;2(3):323-325.
21. Keevil SF. Spatial localization in nuclear magnetic resonance spectroscopy. *Phys Med Biol*. 2006;51(16):R579-636.
22. Levitt MH. Fourier transform nmr. In: Levitt MH, ed. *Spin dynamics*. Second Edition. Chichester, West Sussex, England: John Wiley & Sons Ltd; 2007:86.

23. Kreis R. Issues of spectral quality in clinical 1h-magnetic resonance spectroscopy and a gallery of artifacts. *NMR Biomed*. 2004;17(6):361-381.
24. Kreis R, Boer V, Choi IY, et al. Terminology and concepts for the characterization of in vivo MR spectroscopy methods and MR spectra: Background and experts' consensus recommendations. *NMR Biomed*. 2020;34(5):e4347.
25. Wilson M, Andronesi O, Barker PB, et al. Methodological consensus on clinical proton mrs of the brain: Review and recommendations. *Magn Reson Med*. 2019;82(2):527-550.
26. Rioul O, Vetterli M. Wavelets and signal processing. *IEEE Signal Process Mag*. 1991;8(4):14-38.
27. Serrai H, Senhadji L, Jacques D, Coatrieux JL. Time-domain quantification of amplitude, chemical shift, apparent relaxation time t*2, and phase by wavelet-transform analysis: Application to biomedical magnetic resonance spectroscopy. *J Magn Reson*. 1997;124(1):20-34.
28. Antoine JP, Chauvin C, Coron A. Wavelets and related time-frequency techniques in magnetic resonance spectroscopy. *NMR Biomed*. 2001;14(4):265-270.
29. Cancino-De-Greiff HF, Ramos-Garcia R, Lorenzo-Ginori JV. Signal de-noising in magnetic resonance spectroscopy using wavelet transforms. *Concepts Magn Reson*. 2002;14(6):388-401.
30. Stanwell P, Siddall P, Keshava N, et al. Neuro magnetic resonance spectroscopy using wavelet decomposition and statistical testing identifies biochemical changes in people with spinal cord injury and pain. *Neuroimage*. 2010;53(2):544-552.
31. 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.
32. Cavassila S, Deval S, Huegen C, van Ormondt D, Graveron-Demilly D. Cramér-Rao bounds: an evaluation tool for quantitation. *NMR Biomed*. 2001;14(4):278-283.
33. Zhao D, Grist JT, Sun Y, Peet AC. Impact of wavelets and apodisation in magnetic resonance spectroscopy quality for paediatric brain tumours. *Proc Intl Soc Magn Reson Med*. 2019;27:4243.
34. Hand DJ, Till RJ. A simple generalisation of the area under the ROC curve for multiple class classification problems. *Mach Learn*. 2001;45(2):171-186.
35. Zhao D, Grist JT, Rose HEL, et al. Metabolite selection for machine learning in childhood brain tumour classification. *NMR Biomed*. 2022;35(6):e4673.
36. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. *J Artif Intell Res*. 2002;16:321-357.
37. Quinlan JR. Induction of decision trees. *Mach Learn*. 1986;1:81-106.
38. Robin X, Turck N, Hainard A, et al. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform*. 2011;12:77.
39. Stéphane M. Denoising. In: Stéphane M, ed. *A wavelet tour of signal processing, the sparse way*. Third Edition. Boston: Academic Press; 2009:535-610.
40. Ji B, Hosseini Z, Wang L, Zhou L, Tu X, Mao H. Spectral wavelet-feature analysis and classification assisted denoising for enhancing magnetic resonance spectroscopy. *NMR Biomed*. 2021;34(6):e4497.
41. Cramér H. A contribution to the theory of statistical estimation. *Scand Actuar J*. 1946;1946(1):85-94.
42. Near J, Harris AD, Juchem C, et al. Preprocessing, analysis and quantification in single-voxel magnetic resonance spectroscopy: Experts' consensus recommendations. *NMR Biomed*. 2021;34(5):e4257.
43. Clarke WT, Chiew M. Uncertainty in denoising of MRSI using low-rank methods. *Magn Reson Med*. 2022;87(2):574-588.
44. Arora RS, Alston RD, Eden TO, Estlin EJ, Moran A, Birch JM. Age-incidence patterns of primary cns tumors in children, adolescents, and adults in england. *Neuro Oncol*. 2009;11(4):403-413.
45. Sullivan LB, Gui DY, Heiden MGV. Altered metabolite levels in cancer: Implications for tumour biology and cancer therapy. *Nat Rev Cancer*. 2016;16(11):680-693.
46. Arts LPA, van den Broek EL. The fast continuous wavelet transformation (fcwt) for real-time, high-quality, noise-resistant time-frequency analysis. *Nat Comput Sci*. 2022;2(1):47-58.

How to cite this article: Zhao T, Grist JT, Auer DP, et al. Noise suppression of proton magnetic resonance spectroscopy improves paediatric brain tumour classification. *NMR in Biomedicine*. 2024;37(6):e5129. doi:10.1002/nbm.5129

APPENDIX A: DEFINITIONS

A.1 | Quality-control parameters

A.1.1 | fSNR

The fitting-based signal-to-noise ratio (fSNR) was calculated based on the amplitude of the highest peak L of the spectrum and the fitting residual σ_1 between 0.2 and 4 ppm as

$$\mathrm{fSNR} = \frac{L}{\sigma_1},$$

where σ_1 denotes the fitting residual, derived as the difference between the original spectrum and the combination of fitting bases finally determined, at the range of 0.2–4 ppm, and L is the height of the highest peak of the spectrum.

A.1.2 | wSNR

The whole-spectrum signal-to-noise ratio (wSNR) was calculated based on the amplitude of the highest peak L of the spectrum and the fitting residual σ_0 of the whole spectrum as

$$\mathrm{wSNR} = \frac{L}{\sigma_0},$$

where σ_0 denotes the fitting residual derived as the difference between the original spectrum and the combination of fitting bases finally determined, throughout the whole spectrum, and L is the height of the highest peak of the spectrum.

A.1.3 | FWHM

The full width at half-maximum (FWHM) was calculated based on the width of the unsuppressed water peak at half its full height.[24]

A.2 | Classification accuracy parameters

$$\begin{aligned}
\alpha_{\mathrm{LOOCV}} &= \frac{\mathrm{CP_{EP}} + \mathrm{CP_{MB}} + \mathrm{CP_{PA}}}{\mathrm{TP_{EP}} + \mathrm{TP_{MB}} + \mathrm{TP_{PA}}}, \\
\beta_{\mathrm{LOOCV}} &= \frac{1}{3}\left(\frac{\mathrm{CP_{EP}}}{\mathrm{TP_{EP}}} + \frac{\mathrm{CP_{MB}}}{\mathrm{TP_{MB}}} + \frac{\mathrm{CP_{PA}}}{\mathrm{TP_{PA}}}\right), \\
\alpha_{\text{k-fold}} &= \frac{1}{k}\sum^{k}\frac{\mathrm{CPF_{EP}} + \mathrm{CPF_{MB}} + \mathrm{CPF_{PA}}}{\mathrm{TPF_{EP}} + \mathrm{TPF_{MB}} + \mathrm{TPF_{PA}}}, \\
\beta_{\text{k-fold}} &= \frac{1}{3k}\sum^{k}\left(\frac{\mathrm{CP_{EP}}}{\mathrm{TP_{EP}}} + \frac{\mathrm{CP_{MB}}}{\mathrm{TP_{MB}}} + \frac{\mathrm{CP_{PA}}}{\mathrm{TP_{PA}}}\right),
\end{aligned}$$

where α_{LOOCV} denotes the leave-one-out cross-validated overall classification accuracy, β_{LOOCV} the leave-one-out cross-validated balanced classification accuracy, $\alpha_{\text{k-fold}}$ the k-fold cross-validated overall classification accuracy, $\beta_{\text{k-fold}}$ the k-fold cross-validated balanced classification accuracy, $\mathrm{CP_X}$ the correct predictions for tumour type X, $\mathrm{TP_X}$ the total predictions for tumour type X, $\mathrm{CPF_X}$ the correct predictions of the fold for tumour type X, and $\mathrm{TPF_X}$ the total predictions of the fold for tumour type X.

APPENDIX B: ALGORITHM

Algorithm 1 Adaptive Wavelet Noise Suppression

Require: a noisy ^{1}H-MRS signal $\mathbf{x}(n)$, wavelet transform variations (including wavelet basis functions, wavelet transform methods, wavelet decomposition methods, thresholding methods, decomposition levels) Ψ.

Ensure: the noise suppressed ^{1}H-MRS signal $\mathbf{y}(n)$.

1: **Preparation** derive the ^{1}H-MRS spectrum $\hat{\mathbf{x}}(s)$ from the ^{1}H-MRS signal $\mathbf{x}(n)$ with discrete Fourier transform: $\hat{\mathbf{x}}(s) = \mathcal{F}\mathbf{x}(n)$

2: **for** any ψ_i in Ψ **do**

3: **Spectrum decomposition** derive spectral components from the obtained ^{1}H-MRS spectrum $\hat{\mathbf{x}}(s)$ with the wavelet specification ψ_i as: $\hat{\mathbf{x}}(s) \rightarrow \{\hat{\mathbf{c}}_{\psi_i,\tau}\}$

4: **Spectrum reconstruction** reconstruct ^{1}H-MRS spectrum by using thresholded spectral components as $\{\hat{\mathbf{c}}_{\psi_i,\tau}\}|_{\tau<\tau_0} \rightarrow \hat{\mathbf{x}}_i'(s)$

5: **Spectrum evaluation** derive the quality-control parameters of the original and reconstructed ^{1}H-MRS spectrum as $\hat{\mathbf{x}}'(s) \rightarrow \mathbf{Q}_0$ (original) and $\hat{\mathbf{x}}_i'(s) \rightarrow \mathbf{Q}_i$ (reconstructed).

6: **end for**

7: **Spectrum selection** Select the outperforming spectrum to be $\hat{\mathbf{y}}(s)$ among the obtained reconstructed spectra $\hat{\mathbf{x}}_i'(s)$ according to their corresponding quality-control metric $\mathbf{Q}_i$ with the conditions (a) $\mathrm{FWHM}_i < 1.5 \times \mathrm{FWHM}_0$ and (b) max fSNR_i as $\hat{\mathbf{x}}_i'(s) \rightarrow \hat{\mathbf{y}}(s)$

8: **Finalisation** derive the time-domain signal of the outperforming spectrum to be the noise suppressed signal $\mathbf{y}(n)$ as $\hat{\mathbf{y}}(s) \rightarrow \mathbf{y}(n)$ by using inverse discrete Fourier transform.

Notation

$\mathbf{x}(n)$	A ^{1}H-MRS signal defined in the time domain, where n denotes the time points.
Ψ	A set of wavelet variations, including wavelet basis functions, wavelet transform methods, wavelet decomposition methods, thresholding methods, and decomposition levels, where a wavelet variation is denoted by ψ_i.
$\mathcal{F}$	Performing discrete Fourier transform on a time-domain signal.
$\hat{\mathbf{x}}(s)$	A ^{1}H-MRS spectrum defined in the frequency domain, where s denotes the chemical shift.
$\{\hat{\mathbf{c}}_{\psi_i,\tau}\}$	A set of spectral components derived by using the wavelet ψ_i and indexed by τ.
$\hat{\mathbf{x}}_i'(s)$	The reconstructed spectrum from spectral components $\{\hat{\mathbf{c}}_{\psi_i,\tau}\}$ with the cut-off τ_0.
$\mathbf{Q}$	The quality-control metric, i.e. $\mathbf{Q}_0$ for that of $\hat{\mathbf{x}}(s)$ and $\mathbf{Q}_i$ for that of $\hat{\mathbf{x}}_i'(s)$.
$\hat{\mathbf{y}}(s)$	The finally chosen spectrum with optimal $\mathbf{Q}$ from all $\hat{\mathbf{x}}_i'(s)$.
$\mathbf{y}(n)$	The post-noise suppression ^{1}H-MRS signal that is defined in the time domain.

APPENDIX C: PROOF

Consider that we have the noise $\omega(n)$ presenting in the free induction decay signal $x(n)$:

$$x(n) = x_0(n) + \omega(n), \text{where } n = 1,\ldots,N.$$

Assuming the original noise follows the Gaussian distribution with the mean of zero, then

$$\mathcal{E}\omega[n_1]\omega[n_2] = \sigma^2\delta[n_1 - n_2].$$

Considering that the noise will be decomposed by using the orthonormal basis $\mathcal{B} = \{b_k\}_{0 \le k < N}$, the decomposed noise will be calculated as

$$\hat{\omega}_{\mathcal{B}}[k] = \sum_{n=1}^{N} \omega[n] b_k^*[n].$$

Such decomposed noise will also satisfy

$$\begin{aligned}\mathcal{E}\{\hat{\omega}_{\mathcal{B}}[k_1]\hat{\omega}_{\mathcal{B}}[k_2]\} &= \mathcal{E}\sum_{n_1=1}^{N_1}\omega[n_1]b_{k_1}^{*}[n_1]\sum_{n_2=1}^{N_2}\omega[n_2]b_{k_2}^{*}[n_2]\\ &= \sum_{n_1=1}^{N}\sum_{n_2=1}^{N}b_{k_1}(n_1)b_{k_2}(n_2)\mathcal{E}\{\omega[n_1]\omega[n_2]\}\\ &= \sigma^2 < b_{k_1}, b_{k_2} >\\ &= \sigma^2\delta[k_1 - k_2],\end{aligned}$$

indicating that the suppressed noise in the processed ^{1}H-MRS signals also meets the Gaussian distribution.

TABLE D1 The selected clinical variables, scanning parameters, and quality-control parameters of the 1.5T cohort, Part I.

Case	Site	Tumour type	Scanner	NCC	FS (Hz)	TE (ms)	TR (ms)	NCP	Voxel size (mm^3)	fSNR preNS	fSNR postNS
#01	BCH	EP	Siemens Avanto	12	2,000	30	1,500	2,048	15×15×15	9.2	12.7
#02	BCH	EP	Siemens Avanto	12	2,000	30	1,500	2,048	15×15×15	9.7	15.3
#03	BCH	EP	GE Signa	8	2,500	30	1,500	2,048	15×15×15	11.1	14.6
#04	QMC	EP	Philips Intera	16	2,500	30	2,000	1,024	15×15×15	14.5	18.9
#05	BCH	EP	Siemens Symphony	16	2,000	30	1,500	2,048	15×15×15	19.0	30.6
#06	BCH	EP	Siemens Aera	20	1,000	30	2,200	1,024	15×15×15	22.7	34.6
#07	BCH	EP	Siemens Symphony	16	1,000	30	1,500	1,024	20×20×20	24.0	31.4
#08	BCH	EP	Siemens Symphony	16	2,000	30	1,500	2,048	15×15×15	24.5	39.3
#09	BCH	EP	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	24.6	33.1
#10	AHCH	EP	Philips Ingenia	8	1,000	31	1,500	512	20×20×20	24.7	38.9
#11	BCH	EP	Siemens Symphony	16	1,000	30	1,500	1,024	20×20×20	25.5	32.4
#12	BCH	EP	Siemens Aera	20	1,000	30	2,200	1,024	15×15×15	25.7	38.3
#13	BCH	EP	Siemens Avanto	12	2,000	30	1,500	2,048	20×20×20	32.2	39.4
#14	BCH	MB	GE Signa	8	2,500	30	1,500	2,048	15×15×15	8.5	9.8
#15	BCH	MB	GE Signa	8	2,500	30	1,500	2,048	15×15×15	8.9	10.4
#16	BCH	MB	Siemens Avanto	12	2,000	30	1,500	2,048	15×15×15	13.4	19.6
#17	BCH	MB	Siemens Avanto	12	2,000	30	1,500	2,048	20×20×20	14.2	20.3
#18	BCH	MB	Siemens Symphony	16	1,000	30	1,500	1,024	15×15×15	14.8	23.5
#19	BCH	MB	Siemens Avanto	12	1,500	30	1,500	2,048	15×15×15	16.4	26.2
#20	BCH	MB	Siemens Symphony	16	1,000	30	1,500	1,024	15×15×15	16.5	22.6
#21	BCH	MB	Siemens Symphony	16	1,000	30	1,500	1,024	15×15×15	16.7	25.7
#22	BCH	MB	Siemens Symphony	16	1,000	30	1,500	1,024	15×15×15	18.8	24.1
#23	BCH	MB	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	21.1	26.2
#24	BCH	MB	Siemens Avanto	12	2,000	30	1,500	2,048	20×20×20	21.2	28.7
#25	BCH	MB	Siemens Symphony	16	2,500	30	1,500	2,048	15×15×15	21.8	30.8
#26	BCH	MB	Siemens Avanto	12	2,000	30	1,500	2,048	20×20×20	21.9	32.5
#27	BCH	MB	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	21.9	32.8
#28	BCH	MB	Siemens Avanto	12	2,000	30	1,500	2,048	15×15×15	24.2	36.8
#29	BCH	MB	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	25.7	34.9
#30	BCH	MB	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	25.7	37.3
#31	BCH	MB	Siemens Avanto	12	2,000	30	1,500	2,048	15×15×15	27.1	39.7
#32	BCH	MB	GE Signa	8	2,500	30	1,500	2,048	20×20×20	31.7	41.4
#33	BCH	MB	Siemens Avanto	12	2,000	30	1,500	2,048	20×20×20	32.2	36.5
#34	BCH	MB	Siemens Symphony	16	2,000	30	1,500	2,048	15×15×15	32.7	47.5

(Continues)

TABLE D1 (Continued)

Case	Site	Tumour type	Scanner	NCC	FS (Hz)	TE (ms)	TR (ms)	NCP	Voxel size (mm^3)	fSNR preNS	fSNR postNS
#35	BCH	MB	Siemens Avanto	12	2,000	30	1,500	2,048	20×20×20	38.5	46.4
#36	BCH	MB	GE Signa	8	2,000	30	1,500	2,048	20×20×20	38.6	55.0
#37	BCH	MB	Siemens Symphony	16	1,000	30	1,500	1,024	20×20×20	38.7	48.5
#38	BCH	MB	Siemens Avanto	12	2,000	30	1,500	2,048	20×20×20	41.2	46.3
#39	BCH	MB	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	42.7	61.2
#40	BCH	MB	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	43.3	64.5
#41	BCH	MB	Siemens Symphony	16	1,000	30	1,500	1,024	20×20×20	48.1	65.8
#42	BCH	MB	Siemens Symphony	16	1,000	30	1,500	1,024	20×20×20	62.9	83.5
#43	BCH	MB	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	63.0	83.6
#44	BCH	MB	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	67.0	79.1

Abbreviations: AHCH, Alder Hey Children's NHS Foundation Trust, Liverpool; BCH, Birmingham Children's Hospital, Birmingham Women's and Children's NHS Foundation Trust; QMC, Queen's Medical Centre, Nottingham University Hospitals NHS Trust; RVI, Royal Victoria Infirmary, The Newcastle upon Tyne Hospitals NHS Foundation Trust. NCC, number of coil channels; FS, sampling frequency; TE, echo time; TR, repetition time; NCP, number of complex points; fSNR, fitting-based signal-to-noise ratio; preNS, pre-noise suppression; postNS, post-noise suppression.

TABLE D2 The selected clinical variables, scanning parameters, and quality-control parameters of the 1.5T cohort, Part II.

Case	Site	Tumour type	Scanner	NCC	FS (Hz)	TE (ms)	TR (ms)	NCP	Voxel size (mm^3)	fSNR preNS	fSNR postNS
#45	BCH	PA	Siemens Symphony	16	2,500	30	1,500	2,048	15×15×15	5.1	7.1
#46	BCH	PA	Siemens Avanto	12	2,000	30	1,500	2,048	15×15×15	5.8	8.7
#47	BCH	PA	Siemens Symphony	16	1,000	30	1,500	1,024	15×15×15	6.3	8.8
#48	BCH	PA	Siemens Avanto	12	2,000	30	1,500	2,048	15×15×15	6.4	9.8
#49	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	15×15×15	6.5	11.2
#50	BCH	PA	Siemens Symphony	16	1,000	30	1,500	1,024	15×15×15	6.6	10.1
#51	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	6.7	9.9
#52	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	15×15×15	6.9	9.6
#53	BCH	PA	Siemens Avanto	12	2,000	30	1,500	2,048	15×15×15	6.9	10.6
#54	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	15×15×15	7.6	12.3
#55	BCH	PA	Siemens Symphony	16	1,000	30	1,500	1,024	15×15×15	7.7	14.3
#56	BCH	PA	GE Signa	8	2,500	30	1,500	2,048	15×15×15	7.9	8.7
#57	BCH	PA	Siemens Avanto	12	2,000	30	1,500	2,048	15×15×15	8.2	12.5
#58	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	15×15×15	8.4	13.5
#59	BCH	PA	Siemens Avanto	12	2,000	30	1,500	2,048	20×20×20	8.5	10.6
#60	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	15×15×15	8.8	12.8
#61	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	15×15×15	9.7	12.1
#62	BCH	PA	Siemens Avanto	12	2,000	30	1,500	2,048	20×20×20	10.1	14.3
#63	BCH	PA	Siemens Symphony	16	2,000	30	1,500	1,024	20×20×20	10.1	14.7
#64	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	10.5	17.3
#65	BCH	PA	Siemens Avanto	12	2,000	30	1,500	2,048	20×20×20	10.6	14.2
#66	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	10.6	15.3
#67	BCH	PA	GE Signa	8	2,500	30	1,500	2,048	20×20×20	10.7	14.9
#68	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	11.0	14.5
#69	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	11.1	13.6
#70	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	15×15×15	11.1	13.7
#71	BCH	PA	Siemens Avanto	12	2,000	35	1,500	2,048	13×13×13	12.3	17.4

TABLE D2 (Continued)

Case	Site	Tumour type	Scanner	NCC	FS (Hz)	TE (ms)	TR (ms)	NCP	Voxel size (mm^3)	fSNR preNS	fSNR postNS
#72	BCH	PA	Siemens Avanto	12	2,000	30	1,500	2,048	15×15×15	12.6	17.0
#73	BCH	PA	Siemens Avanto	12	2,000	30	1,500	2,048	15×15×15	13.5	18.3
#74	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	13.6	19.9
#75	BCH	PA	Siemens Avanto	12	2,000	30	1,500	2,048	20×20×20	13.8	14.9
#76	BCH	PA	Siemens Aera	20	1,000	30	1,500	1,024	15×15×15	14.8	18.9
#77	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	14.9	19.6
#78	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	15.1	23.2
#79	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	15.2	19.0
#80	BCH	PA	Siemens Avanto	12	2,000	30	1,500	2,048	20×20×20	16.3	21.0
#81	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	15×15×15	17.2	26.4
#82	BCH	PA	Siemens Symphony	16	2,000	30	1,500	2,048	20×20×20	18.1	23.6
#83	BCH	PA	Siemens Symphony	16	1,000	30	1,500	2,048	20×20×20	20.6	29.7

Abbreviations: AHCH, Alder Hey Children's NHS Foundation Trust, Liverpool; BCH, Birmingham Children's Hospital, Birmingham Women's and Children's NHS Foundation Trust; QMC, Queen's Medical Centre, Nottingham University Hospitals NHS Trust; RVI, Royal Victoria Infirmary, The Newcastle upon Tyne Hospitals NHS Foundation Trust. NCC, number of coil channels; FS, sampling frequency; TE, echo time; TR, repetition time; NCP, number of complex points; fSNR, fitting-based signal-to-noise ratio; preNS, pre-noise suppression; postNS, post-noise suppression.

TABLE D3 The selected clinical variables, scanning parameters, and quality-control parameters of the 3T cohort.

Case	Site	Tumour type	Scanner	NCC	FS (Hz)	TE (ms)	TR (ms)	NCP	Voxel size (mm^3)	fSNR preNS	fSNR postNS
#01	BCH	EP	Philips Achieva	16	2,000	41	2,000	1,024	13×13×13	5.9	8.8
#02	BCH	EP	Philips Achieva	16	2,000	36	2,000	1,024	13×13×13	7.5	9.1
#03	AHCH	EP	Philips Achieva	8	2,000	30	2,000	1,024	15×15×15	10.0	13.3
#04	AHCH	EP	Philips Achieva	8	2,000	40	2,000	1,024	15×15×15	15.5	20.6
#05	AHCH	MB	Philips Ingenia	8	2,000	39	2,000	1,024	15×15×15	8.9	13.0
#06	AHCH	MB	Philips Achieva	8	2,000	39	2,000	1,024	15×15×15	11.1	17.3
#07	BCH	MB	Siemens Avanto	16	2,000	38	2,000	1,024	13×13×13	14.9	21.8
#08	RVI	MB	Siemens Verio	32	2,000	30	2,120	2,048	20×20×20	15.5	18.9
#09	BCH	MB	Philips Achieva	32	2,000	36	2,000	1,024	13×13×13	17.9	25.6
#10	BCH	MB	Philips Achieva	16	2,000	36	2,000	1,024	15×15×15	18.3	27.1
#11	BCH	MB	Philips Achieva	16	2,000	36	2,000	1,024	15×15×15	18.9	25.2
#12	AHCH	MB	Philips Achieva	8	2,000	37	2,000	1,024	20×20×20	19.8	33.9
#13	AHCH	MB	Philips Achieva	8	2,000	41	2,000	1,024	15×15×15	20.9	27.7
#14	BCH	MB	Philips Achieva	16	2,000	36	2,000	1,024	20×20×20	22.8	30.8
#15	BCH	MB	Philips Achieva	16	2,000	38	2,000	1,024	20×20×20	23.0	30.8
#16	BCH	MB	Philips Achieva	16	2,000	37	2,000	1,024	15×15×15	26.7	33.7
#17	AHCH	MB	Philips Achieva	8	2,000	37	2,000	1,024	15×15×15	32.1	40.4
#18	BCH	MB	Philips Achieva	16	2,000	37	2,000	512	15×15×15	38.3	46.8
#19	AHCH	MB	Philips Achieva	8	2,000	37	2,000	1,024	20×20×20	38.6	47.0
#20	AHCH	MB	Philips Ingenia	8	2,000	40	2,000	1,024	20×20×20	38.9	44.2
#21	AHCH	MB	Philips Achieva	8	2,000	39	2,000	1,024	20×20×20	49.9	57.2
#22	BCH	PA	Philips Achieva	16	2,000	41	2,000	1,024	13×13×13	5.9	9.4
#23	BCH	PA	Philips Achieva	16	2,000	37	2,000	1,024	13×13×13	6.1	8.6
#24	AHCH	PA	Philips Achieva	8	2,000	37	2,000	1,024	15×15×15	7.1	10.2
#25	AHCH	PA	Philips Achieva	8	2,000	40	2,000	1,024	15×15×15	7.2	9.8

(Continues)

TABLE D3 (Continued)

Case	Site	Tumour type	Scanner	NCC	FS (Hz)	TE (ms)	TR (ms)	NCP	Voxel size (mm^3)	fSNR preNS	fSNR postNS
#26	QMC	PA	Philips Achieva	8	2,000	37	2,000	1,024	15×15×15	7.7	10.8
#27	AHCH	PA	Philips Achieva	8	2,000	37	2,000	1,024	15×15×15	7.7	11.1
#28	AHCH	PA	Philips Ingenia	8	1,000	38	2,000	512	15×15×15	8.4	13.9
#29	AHCH	PA	Philips Achieva	8	2,000	41	2,000	1,024	15×15×15	8.7	11.7
#30	AHCH	PA	Philips Achieva	8	2,000	40	2,000	1,024	15×15×15	8.8	11.8
#31	RVI	PA	Siemens Verio	32	2,000	30	2,120	2,048	15×20×19	9.1	10.0
#32	AHCH	PA	Philips Achieva	8	2,000	37	2,000	1,024	15×15×15	9.2	14.5
#33	AHCH	PA	Philips Ingenia	8	2,000	41	2,000	1,024	15×15×15	9.8	14.1
#34	BCH	PA	Philips Achieva	16	2,000	38	2,000	1,024	13×13×13	10.2	15.6
#35	BCH	PA	Philips Achieva	32	2,000	37	2,000	1,024	13×13×13	10.4	15.0
#36	BCH	PA	Philips Achieva	32	2,000	38	2,000	1,024	15×15×15	11.0	15.3
#37	BCH	PA	Philips Achieva	16	2,000	37	2,000	1,024	15×15×15	12.1	19.3
#38	BCH	PA	Philips Achieva	16	2,000	37	2,000	1,024	15×15×15	12.8	19.3
#39	RVI	PA	Siemens Verio	32	2,000	30	2,120	2,048	20×20×20	14.6	19.0
#40	BCH	PA	Siemens Avanto	16	2,000	37	2,000	1,024	15×15×15	15.1	21.0
#41	BCH	PA	Philips Achieva	16	2,000	40	2,000	1,024	20×20×20	15.6	21.6
#42	BCH	PA	Philips Achieva	16	2,000	37	2,000	1,024	20×20×20	16.1	20.1

Abbreviations: AHCH, Alder Hey Children's NHS Foundation Trust, Liverpool; BCH, Birmingham Children's Hospital, Birmingham Women's and Children's NHS Foundation Trust; QMC, Queen's Medical Centre, Nottingham University Hospitals NHS Trust; RVI, Royal Victoria Infirmary, The Newcastle upon Tyne Hospitals NHS Foundation Trust. NCC, number of coil channels; FS, sampling frequency; TE, echo time; TR, repetition time; NCP, number of complex points; fSNR, fitting-based signal-to-noise ratio; preNS, pre-noise suppression; postNS, post-noise suppression.

TABLE D4 Details of quality filtering for the acquired brain tumour proton magnetic resonance spectroscopy.

	Ependymomas	Medulloblastomas	Pilocytic astrocytomas	Total
1.5T cohort				
Initial sample	23	39	54	116
Removed sample	10	8	15	33
Missing histology	0	1	0	1
Incomplete imaging data	2	4	3	9
Partial volume effects	1	1	3	5
Poor line width	3	0	1	4
Poor fSNR	2	0	2	4
Artefacts	2	2	6	10
Final sample	†13	31	39	83
Birmingham (96%)	10	31	39	80
Liverpool (1%)	1	0	0	1
Nottingham (3%)	2	0	0	2
3T cohort				
Initial sample	7	25	41	73
Removed sample	3	8	20	31
Missing histology	0	2	1	3
Incomplete imaging data	2	4	6	12
Partial volume effects	1	0	4	5
Poor line width	0	0	0	0
Poor fSNR	0	0	1	1

TABLE D4 (Continued)

	Ependymomas	Medulloblastomas	Pilocytic astrocytomas	Total
Artefacts	0	2	8	10
Final sample	‡4	17	21	42
Birmingham (53%)	3	8	11	22
Liverpool (38%)	1	8	7	16
Newcastle (7%)	0	1	2	3
Nottingham (2%)	0	0	1	1

†Three of the 13 ependymomas were located supratentorially.
‡One of the four ependymomas was located supratentorially.

TABLE D5 List of wavelet variations used.

	Vanish moments in synthetic wavelets	Vanish moments in analytic wavelets
Bi-orthogonal	1	1, 3, 5
	2	2, 4, 6, 8
	3	1, 3, 5, 7, 9
	4	4
	5	5
	6	6
Reversed bi-orthogonal	1	1, 3, 5
	2	2, 4, 6, 8
	3	1, 3, 5, 7, 9
	4	4
	5	5
	6	6
Coiflets	1, 2, 3, 4, 5	
Daubechies	1, 2, 3, 4, 5, 6, 10, 15, 20, 25	
Fejer–Korovki	4, 6, 8, 14, 22	
Symlets	2, 3, 4, 5, 10, 15, 20, 25	
Discrete approximation of Meyer	/	
Farras nearly symmetric	first-stage filters	
Haar	/	

TABLE D6 Excluding metabolites according to duplication and Cramér–Rao lower bound percentage values.

Metabolite	Condition	Reason
−CrCH2	Excluded	CRLB% is greater than 50% in all cases.
Alanine (Ala)	Excluded	CRLB% is greater than 50% in all cases.
Aspartate (Asp)	Excluded	CRLB% is greater than 50% in all cases.
Citrate (Cit)	Included	/
Creatine (Cr)	Excluded	Total creatine is included instead.
γ-aminobutyric acid (GABA)	Excluded	CRLB% is greater than 50% in all cases.
Glycerophosphocholine (GPC)	Excluded	Total choline is included instead.
Glucose (Glc)	Included	/
Glutamine (Gln)	Excluded	Combined glutamate and glutamine is included instead.
Glutathione (Glth or GSH)	Included	/
Glutamate (Glu)	Excluded	Combined glutamate and glutamine is included instead.
Glycine (Gly)	Included	/
myo-Inositol (mI or Ins)	Included	/
Lactate (Lac)	Included	/
Lipids at 0.9 ppm (Lip09)	Excluded	Total lipids and macromolecules at 0.9 ppm is included instead.
Lipids at 1.3 ppm, a (Lip13a)	Excluded	Total lipids and macromolecules at 1.3 ppm is included instead.
Lipids at 1.3 ppm, b (Lip13b)	Excluded	Total lipids and macromolecules at 1.3 ppm is included instead.
Lipids at 2.0 ppm (Lip20)	Excluded	Total lipids and macromolecules at 2.0 ppm is included instead.
MacroMolecules at 0.9 ppm (MM09)	Excluded	Total lipids and macromolecules at 0.9 ppm is included instead.
MacroMolecules at 1.2 ppm (MM12)	Excluded	Total lipids and macromolecules at 1.3 ppm is included instead.
MacroMolecules at 1.4 ppm (MM14)	Excluded	Total lipids and macromolecules at 1.3 ppm is included instead.
MacroMolecules at 1.7 ppm (MM17)	Excluded	Total lipids and macromolecules at 2.0 ppm is included instead.
MacroMolecules at 2.0 ppm (MM20)	Excluded	Total lipids and macromolecules at 2.0 ppm is included instead.
MacroMolecules at 3.8 ppm (MM38)	Excluded	Total lipids and macromolecules at 2.0 ppm is included instead.
N-acetylaspartate (NAA)	Excluded	Total *N*-acetylaspartate is included instead.
N-acetylaspartylglutamate (NAAG)	Excluded	Total *N*-acetylaspartate is included instead.
Phosphocholine (PCh)	Excluded	Total choline is included instead.
Phosphocreatine (PCr)	Excluded	Total creatine is included instead.
Phosphoethanolamine (PEth)	Excluded	CRLB% is greater than 50% in all cases.
scyllo-Inositol (sI or Scyllo)	Included	/
Taurine (Tau)	Included	/
total *N*-acetylaspartate (tNAA)	Included	/
total Choline (tCho)	Included	/
total Creatine (tCr)	Included	/
combined Glutamate and Glutamine (Glx)	Included	/
total Lipids and Macromolecules at 0.9 ppm (tLM09)	Included	/
total Lipids and Macromolecules at 1.3 ppm (tLM13)	Included	/
total Lipids and Macromolecules at 2.0 ppm (tLM20)	Included	/

TABLE D7 Statistics of patient sex.

	1.5T				3T			
	Sex (F:M)		Chi-squared test		Sex (F:M)		Chi-squared test	
	Type A	Type B	χ^2	*P*	Type A	Type B	χ^2	*P*
Ependymoma | Medulloblastoma	8: 5	9: 22	1.114 6	.291 1	3: 1	7: 10	0.312 5	.576 2
Ependymoma | Pilocytic astrocytoma	8: 5	19: 20	0.231 1	.630 7	3: 1	13: 8	< 0.000 1	.999 9
Medulloblastoma | Pilocytic astrocytoma	9: 22	19: 20	0.354 1	.551 7	7: 10	13: 8	0.652 3	.419 3

Table showing the statistics for patient sex between two tumour types. Significant levels are determined by the conditions as $P < .05$ (*), $P < .01$ (**), $P < .001$ (***), and $P < .0001$ (****).

TABLE D8 Statistics of patient age.

	1.5T					3T				
	Age (Mean ± SD)		Wilcoxon rank-sum			Age (Mean ± SD)		Wilcoxon rank-sum		
	Type A	Type B	*W*	*P*		Type A	Type B	*W*	*P*	
Ependymoma | Medulloblastoma	4.6 ± 5.3	6.9 ± 3.5	74	.001 0	**	9.2 ± 5.3	6.5 ± 4.4	37	.655 7	ns
Ependymoma | Pilocytic astrocytoma	4.6 ± 5.3	8.0 ± 3.6	84	.000 2	***	9.2 ± 5.3	6.3 ± 5.0	44	.653 8	ns
Medulloblastoma | Pilocytic astrocytoma	6.9 ± 3.5	8.0 ± 3.6	529	.377 5	ns	6.5 ± 4.4	6.3 ± 5.0	162	.763 0	ns

Table showing the statistics for patient age between two tumour types. Significant levels are determined by the conditions as $P < .05$ (*), $P < .01$ (**), $P < .001$ (***), and $P < .0001$ (****).

TABLE D9 Statistics of quality-control parameters.

	1.5T							3T						
	PreNS		PostNS		Wilcoxon signed-rank			PreNS		PostNS		Wilcoxon signed-rank		
	Mean	SD	Mean	SD	*V*	*P*		Mean	SD	Mean	SD	*V*	*P*	
fSNR														
All	19.3	13.4	26.6	17.5	0	< .000 1	****	16.1	11.0	21.1	11.6	0	< .000 1	****
Groups by tumour type														
EP	20.6	7.2	29.2	10.1	0	.000 2	***	9.7	4.2	13.0	5.5	0	.125 0	ns
MB	29.7	15.6	40.0	19.9	0	< .000 1	****	24.9	12.4	31.3	11.5	0	< .000 1	****
PA	10.6	3.8	15.0	5.1	0	< .000 1	****	10.2	3.1	14.4	4.2	0	< .000 1	****
Groups by spectral quality														
PQ	7.6	1.3	11.0	2.1	0	< .000 1	****	8.0	1.3	11.2	2.0	0	< .000 1	****
MQ	13.7	2.7	19.3	4.5	0	< .000 1	****	14.7	3.0	21.0	4.8	0	< .000 1	****
GQ	32.8	13.0	44.4	16.1	0	< .000 1	****	33.2	11.6	38.9	9.4	0	.022 5	*
wSNR														
All	26.4	19.3	39.6	27.3	0	< .000 1	****	24.6	23.5	35.4	30.0	0	< .000 1	****
Groups by tumour type														
EP	29.3	10.5	44.3	16.7	0	.000 2	***	12.0	4.6	17.8	7.5	0	.125 0	ns
MB	41.0	22.6	60.5	30.9	0	< .000 1	****	40.7	30.3	57.3	37.2	0	< .000 1	****
PA	13.8	5.1	21.5	7.8	0	< .000 1	****	13.9	4.7	21.0	6.5	0	< .000 1	****
Groups by spectral quality														
PQ	11.2	5.2	17.2	7.3	0	< .000 1	****	10.8	2.5	16.4	4.0	0	< .000 1	****
MQ	17.8	4.2	27.8	6.6	0	< .000 1	****	19.8	5.7	29.9	7.7	0	< .000 1	****
GQ	45.2	19.5	66.6	26.6	0	< .000 1	****	58.1	32.5	79.5	38.7	0	.003 9	**

(Continues)

TABLE D9 (Continued)

	1.5T							3T						
	PreNS		PostNS		Wilcoxon signed-rank			PreNS		PostNS		Wilcoxon signed-rank		
	Mean	SD	Mean	SD	*V*	*P*		Mean	SD	Mean	SD	*V*	*P*	
FWHM														
All	4.5	1.0	4.6	1.1	326	.174 6	ns	6.5	1.9	6.6	1.6	136	.206 6	ns
Groups by tumour type														
EP	4.8	0.9	5.0	0.9	3	.280 7	ns	6.5	2.5	6.8	1.7	2	.375 0	ns
MB	4.3	1.2	4.3	1.1	83	.930 3	ns	6.4	1.8	6.6	1.6	12	.111 5	ns
PA	4.6	0.8	4.8	1.1	62	.315 4	ns	6.5	2.0	6.6	1.8	44	.916 5	ns
Groups by spectral quality														
PQ	4.9	1.1	5.3	1.5	35	.052 1	ns	6.7	2.3	7.0	1.9	18	.197 1	ns
MQ	4.6	0.9	4.5	0.8	33	.605 6	ns	6.7	1.6	6.7	1.4	32	.681 3	ns
GQ	4.3	1.0	4.3	0.9	47	.728 9	ns	5.5	1.5	5.9	1.6	0	.031 0	ns

Table showing the statistical results of quality-control parameters (QCPs), namely fitting-based signal-to-noise ratio (fSNR), whole-spectrum signal-to-noise ratio (wSNR), and full width at half-maximum (FWHM), from pre-noise suppression (preNS) and post-noise suppression (postNS) clinical brain tumour proton magnetic resonance spectroscopy. The evaluation was presented for specific cases based on tumour type or spectral quality, namely ependymomas (EP), medulloblastomas (MB), and pilocytic astrocytomas (PA) for tumour type-based groups, and poor quality (PQ), medium quality (MQ), and good quality (GQ) for spectral quality-based groups. The difference between preNS and postNS QCPs was compared by performing Wilcoxon signed-rank tests, where the significant levels are determined by the conditions as $P < .05$ (*), $P < .01$ (**), $P < .001$ (***), and $P < .0001$ (****).

TABLE D10 Statistics of Cramér–Rao lower bound values.

	Median			Mean			SD			Wilcoxon signed-rank		
	PreNS	PostNS	δ%	PreNS	PostNS	δ%	PreNS	PostNS	δ%	*V*	*P*	
1.5T Cohort												
Citrate	0.49	0.03	−94%	0.71	0.40	−44%	1.00	0.71	−29%	2,034	.003 1	**
total Choline	0.36	0.03	−92%	0.49	0.25	−49%	0.62	0.51	−18%	2,469	< .000 1	****
total Creatine	0.51	0.03	−94%	0.69	0.43	−38%	0.68	0.74	9%	2,314	< .000 1	****
Glucose	2.89	0.20	−93%	3.91	2.26	−42%	3.52	4.55	29%	1,155	.009 2	**
Glutathione	1.01	0.09	−91%	1.12	0.93	−17%	0.50	1.51	202%	539	.038 2	*
Glutamate and Glutamine	1.88	0.13	−93%	2.54	1.47	−42%	2.37	2.83	19%	2,607	< .000 1	****
Glycine	2.38	0.18	−92%	3.25	2.31	−29%	2.55	5.39	111%	606	.008 6	**
Lactate	5.19	0.41	−92%	7.57	4.25	−44%	8.09	8.41	4%	2,244	< .000 1	****
total LM at 0.9 ppm	2.18	0.20	−91%	2.54	1.20	−53%	2.20	1.97	−10%	2,699	< .000 1	****
total LM at 1.3 ppm	9.50	0.77	−92%	11.93	6.16	−48%	13.68	10.23	−25%	2,627	< .000 1	****
total LM at 2.0 ppm	3.78	0.25	−93%	4.78	3.05	−36%	4.21	8.09	92%	2,625	< .000 1	****
myo-Inositol	2.37	0.15	−94%	3.52	2.27	−36%	4.08	5.14	26%	1,829	.001 6	***
total *N*-acetylaspartate	1.13	0.09	−92%	1.70	0.99	−42%	2.09	1.77	−15%	2,310	.002 2	**
scyllo-Inositol	0.60	0.05	−92%	0.80	1.05	31%	0.57	2.16	279%	244	.190 4	ns
Taurine	1.39	0.08	−94%	2.29	1.73	−24%	2.77	5.21	88%	1,290	.005 8	**
3T Cohort												
Citrate	0.62	0.10	−84%	0.89	0.25	−72%	0.88	0.31	−65%	443	.000 1	***
total Choline	0.40	0.07	−83%	0.67	0.18	−73%	1.16	0.26	−78%	739	< .000 1	****
total Creatine	0.74	0.12	−84%	1.03	0.27	−74%	1.01	0.36	−64%	785	< .000 1	****
Glucose	2.96	0.54	−82%	5.34	1.07	−80%	8.77	1.46	−83%	245	.001 2	**
Glutathione	1.29	0.19	−85%	2.00	0.60	−70%	3.28	0.97	−70%	690	.000 2	***
Glutamate and Glutamine	3.04	0.47	−85%	4.09	1.15	−72%	3.63	1.69	−53%	710	< .000 1	****
Glycine	0.92	0.23	−75%	1.33	0.48	−64%	1.16	0.67	−42%	480	.000 4	***

TABLE D10 (Continued)

	Median			Mean			SD			Wilcoxon signed-rank		
	PreNS	PostNS	δ%	PreNS	PostNS	δ%	PreNS	PostNS	δ%	*V*	*P*	
Lactate	2.52	0.64	−75%	2.92	1.44	−51%	2.21	1.96	−11%	332	.013 7	*
total LM at 0.9 ppm	3.39	0.42	−88%	4.79	1.39	−71%	6.08	2.35	−61%	750	< .000 1	****
total LM at 1.3 ppm	14.57	1.70	−88%	29.35	5.14	−82%	75.46	8.02	−89%	816	< .000 1	****
total LM at 2.0 ppm	5.57	0.82	−85%	6.43	1.88	−71%	7.02	2.89	−59%	741	< .000 1	****
myo-Inositol	2.63	0.48	−82%	3.50	1.10	−69%	4.59	1.60	−65%	412	.001 4	**
N-acetylaspartate	1.02	0.14	−86%	1.39	0.42	−70%	1.50	0.64	−57%	743	< .000 1	****
scyllo-Inositol	0.39	0.06	−85%	1.36	0.08	−94%	3.86	0.09	−98%	151	.000 5	***
Taurine	2.21	0.51	−77%	2.49	1.25	−50%	1.25	1.47	18%	281	.007 7	**

Table comparing the uncertainty of metabolites as well as lipids and macromolecules (LM) measured from 1.5T and 3T pre- and post-noise suppression proton magnetic resonance spectroscopy that is given by Cramér–Rao lower bound (CRLB) values. Significant levels are determined by the conditions as $P < .05$ (*), $P < .01$ (**), $P < .001$ (***), and $P < .0001$ (****).

TABLE D11 Statistics of Cramér–Rao lower bound percentage values.

	Median			Mean			SD			Wilcoxon signed-rank		
	PreNS	PostNS	δ%	PreNS	PostNS	δ%	PreNS	PostNS	δ%	*V*	*P*	
1.5T Cohort												
Citrate	127%	9%	−93%	163%	51%	−69%	160%	81%	−49%	2,493	< .000 1	****
total Choline	18%	2%	−92%	32%	11%	−66%	37%	20%	−47%	2,898	< .000 1	****
total Creatine	39%	4%	−89%	76%	23%	−70%	163%	58%	−64%	2,770	< .000 1	****
Glucose	296%	22%	−93%	451%	396%	−12%	511%	2,188%	328%	1,430	< .000 1	****
Glutathione	145%	11%	−92%	190%	240%	26%	169%	646%	282%	540	.037 0	*
Glutamate and Glutamine	32%	2%	−93%	47%	22%	−52%	48%	55%	15%	2,878	< .000 1	****
Glycine	111%	11%	−90%	196%	81%	−59%	213%	243%	14%	708	< .000 1	****
Lactate	299%	23%	−92%	520%	165%	−68%	647%	282%	−56%	2,574	< .000 1	****
total LM at 0.9 ppm	45%	3%	−93%	57%	21%	−64%	56%	30%	−46%	3,108	< .000 1	****
total LM at 1.3 ppm	97%	8%	−92%	160%	53%	−67%	313%	87%	−72%	3,018	< .000 1	****
total LM at 2.0 ppm	54%	4%	−93%	76%	37%	−51%	61%	89%	45%	2,941	< .000 1	****
myo-Inositol	154%	12%	−93%	426%	70%	−84%	1400%	155%	−89%	2,234	< .000 1	****
total *N*-acetylaspartate	100%	11%	−89%	174%	66%	−62%	173%	116%	−33%	2,736	< .000 1	****
scyllo-Inositol	143%	19%	−87%	471%	178%	−62%	844%	318%	−62%	274	.042 3	*
Taurine	87%	8%	−90%	233%	251%	8%	408%	1,612%	295%	1,524	< .000 1	****
3T Cohort												
Citrate	124%	19%	−85%	157%	34%	−79%	102%	37%	−64%	483	< .000 1	****
total Choline	21%	5%	−78%	35%	8%	−77%	41%	10%	−75%	806	< .000 1	****
total Creatine	27%	5%	−80%	31%	7%	−77%	20%	7%	−64%	834	< .000 1	****
Glucose	226%	43%	−81%	415%	60%	−86%	475%	61%	−87%	607	.000 2	***
Glutathione	141%	24%	−83%	205%	34%	−83%	311%	31%	−90%	795	< .000 1	****
Glutamate and Glutamine	79%	12%	−85%	100%	22%	−78%	65%	28%	−58%	754	< .000 1	****
Glycine	61%	11%	−82%	602%	26%	−96%	2691%	36%	−99%	530	< .000 1	****
Lactate	167%	34%	−80%	239%	100%	−58%	189%	180%	−5%	373	.000 8	***
total LM at 0.9 ppm	85%	10%	−89%	108%	25%	−77%	87%	38%	−57%	816	< .000 1	****
total LM at 1.3 ppm	128%	18%	−86%	183%	40%	−78%	191%	60%	−69%	847	< .000 1	****
total LM at 2.0 ppm	76%	10%	−87%	88%	20%	−77%	57%	26%	−55%	799	< .000 1	****
myo-Inositol	87%	15%	−83%	131%	45%	−65%	140%	69%	−51%	440	.000 2	***

(Continues)

TABLE D11 (Continued)

	Median			Mean			SD			Wilcoxon signed-rank		
	PreNS	PostNS	δ%	PreNS	PostNS	δ%	PreNS	PostNS	δ%	*V*	*P*	
total *N*-acetylaspartate	93%	14%	−85%	126%	27%	−79%	151%	36%	−76%	813	< .000 1	****
scyllo-Inositol	247%	22%	−91%	437%	54%	−88%	886%	73%	−92%	145	.001 3	**
Taurine	89%	22%	−75%	131%	40%	−70%	147%	50%	−66%	312	.000 6	***

Table comparing the uncertainty of metabolites as well as lipids and macromolecules (LM) measured from 1.5T and 3T pre- and post-noise suppression proton magnetic resonance spectroscopy that is given by Cramér–Rao lower bound (CRLB) percentage values. Significant levels are determined by the conditions as $P < .05$ (*), $P < .01$ (**), $P < .001$ (***), and $P < .0001$ (****).

TABLE D12 Estimated metabolite concentrations in mmol from 1.5T and 3T pre-noise suppression proton magnetic resonance spectroscopy.

	Ependymomas		Medulloblastomas		Pilocytic astrocytomas		All		Kruskal–Wallis *H*		
Metabolites	Mean	SD	Mean	SD	Mean	SD	Mean	SD	*H*	*P*	
1.5T Cohort	*N* = 13		*N* = 31		*N* = 39		*N* = 83				
Citrate	0.7	0.3	0.6	0.3	0.4	0.3	0.5	0.3	15.7	.000 4	***
total Choline	1.9	1.0	3.3	1.6	1.1	0.4	2.0	1.5	46.3	< .000 1	****
total Creatine	3.9	1.8	2.9	1.7	0.9	1.2	2.1	1.9	35.1	< .000 1	****
Glucose	1.2	0.9	0.6	0.9	1.0	0.8	0.9	0.9	10.4	.005 4	**
Glutamate and Glutamine	6.9	2.4	6.4	2.5	5.7	2.2	6.1	2.4	3.8	.146 9	ns
Glutathione	0.4	0.5	0.7	0.5	0.3	0.3	0.4	0.4	9.3	.009 5	**
Glycine	1.9	3.5	3.1	2.5	0.2	0.3	1.5	2.5	49.4	< .000 1	****
Lactate	1.7	1.7	2.6	1.6	1.9	1.1	2.1	1.4	6.1	.048 1	*
total LM at 0.9 ppm	4.3	2.6	7.4	4.1	4.0	1.5	5.3	3.3	24.6	< .000 1	****
total LM at 1.3 ppm	14.9	16.8	20.4	15.2	7.3	4.7	13.4	13.1	29.6	< .000 1	****
total LM at 2.0 ppm	7.9	2.1	9.6	3.1	5.2	1.9	7.2	3.2	40.4	< .000 1	****
myo-Inositol	9.6	5.7	1.7	1.7	2.0	1.8	3.1	3.9	20.7	< .000 1	****
total *N*-acetylaspartate	1.2	0.9	0.9	0.7	1.3	0.8	1.2	0.8	5.4	.067 2	ns
scyllo-Inositol	0.2	0.3	0.3	0.4	0.1	0.1	0.2	0.3	27.8	< .000 1	****
Taurine	1.8	1.6	3.1	2.7	0.8	1.1	1.8	2.2	16.8	.000 2	***
3T Cohort	*N* = 4		*N* = 17		*N* = 21		*N* = 42				
Citrate	0.5	0.6	0.7	0.3	0.4	0.3	0.5	0.4	5.0	.080 1	ns
total Choline	1.3	0.4	4.5	3.7	1.7	0.9	2.8	2.8	18.9	< .000 1	***
total Creatine	1.9	1.1	4.4	2.8	3.1	2.2	3.5	2.5	7.2	.027 3	*
Glucose	0.9	0.3	1.5	1.2	0.6	0.7	1.0	0.1	4.5	.105 0	ns
Glutamate and Glutamine	3.5	0.1	4.8	3.8	5.1	3.6	4.8	3.5	1.1	.589 6	ns
Glutathione	0.9	0.3	2.2	2.1	0.9	0.8	1.4	1.5	12.4	.002 0	**
Glycine	1.2	0.8	3.9	2.3	0.9	1.5	2.2	2.3	21.1	< .000 1	****
Lactate	1.7	0.8	2.2	1.9	0.9	1.1	1.5	1.5	8.3	.015 6	*
total LM at 0.9 ppm	3.6	1.3	6.4	6.2	5.0	5.7	5.4	5.6	6.0	.050 3	ns
total LM at 1.3 ppm	13.3	10.8	23.0	18.3	10.3	10.9	15.7	15.3	14.7	.000 6	***
total LM at 2.0 ppm	6.2	2.2	8.7	4.1	6.8	3.7	7.5	3.8	5.2	.075 7	ns
myo-Inositol	2.9	2.4	2.7	2.6	3.0	3.9	2.9	3.2	0.1	.965 9	ns
total *N*-acetylaspartate	1.0	0.5	1.2	1.1	1.9	1.1	1.6	1.1	9.7	.008 0	**
scyllo-Inositol	0.0		0.4	0.7	0.1	0.3	0.2	0.5	8.8	.012 3	*
Taurine	1.0	0.7	5.7	4.6	1.1	0.3	3.0	0.5	18.5	< .000 1	****

The concentrations of metabolites as well as lipids and molecules (LM) prior to normalisation. *P* values, which evaluated the difference between the three tumour types, were calculated through Kruskal–Wallis *H* tests. Significant levels are determined by the conditions as $P < .05$ (*), $P < .01$ (**), $P < .001$ (***), and $P < .0001$ (****).

TABLE D13 Statistics of multiclass areas under the curve for metabolites as well as lipids and macromolecules.

	PreNS		PostNS		Wilcoxon signed-rank		
	Mean	SD	Mean	SD	*V*	*P*	
1.5T Cohort							
Citrate	0.579 5	0.028 7	0.701 3	0.005 0	0	< .000 1	****
total Choline	0.688 6	0.008 3	0.772 6	0.005 0	0	< .000 1	****
total Creatine	0.771 7	0.005 8	0.778 3	0.006 2	149	< .000 1	****
Glucose	0.738 0	0.004 5	0.603 3	0.009 0	3,486	< .000 1	****
Glutathione	0.559 0	0.007 3	0.643 7	0.005 8	0	< .000 1	****
Glutamate and Glutamine	0.738 6	0.004 8	0.791 9	0.004 7	0	< .000 1	****
Glycine	0.688 5	0.027 5	0.734 5	0.005 9	48	< .000 1	****
Lactate	0.723 5	0.007 0	0.566 8	0.019 5	3,486	< .000 1	****
total LM at 0.9 ppm	0.774 9	0.005 2	0.774 5	0.005 7	1,789	.836 3	ns
total LM at 1.3 ppm	0.665 9	0.009 2	0.757 3	0.007 6	0	< .000 1	****
total LM at 2.0 ppm	0.550 3	0.008 2	0.613 5	0.012 1	0	< .000 1	****
myo-Inositol	0.811 1	0.006 5	0.882 4	0.005 6	0	< .000 1	****
total *N*-acetylaspartate	0.767 7	0.004 4	0.581 2	0.006 5	3,486	< .000 1	****
scyllo-Inositol	0.607 0	0.006 7	0.772 9	0.006 1	0	< .000 1	****
Taurine	0.632 3	0.005 2	0.602 6	0.007 2	3,486	< .000 1	****
3T Cohort							
Citrate	0.484 0	0.034 6	0.594 4	0.020 5	1	< .000 1	****
total Choline	0.760 5	0.010 6	0.753 7	0.012 1	825	< .000 1	****
total Creatine	0.639 4	0.023 7	0.634 1	0.025 8	739	.000 3	***
Glucose	0.580 6	0.015 1	0.653 0	0.011 9	0	< .000 1	****
Glutathione	0.649 9	0.010 4	0.653 6	0.010 9	230	.009 5	**
Glutamate and Glutamine	0.668 6	0.019 7	0.790 6	0.009 8	0	< .000 1	****
Glycine	0.791 6	0.009 3	0.845 5	0.008 0	0	< .000 1	****
Lactate	0.708 8	0.026 0	0.675 1	0.018 3	858	< .000 1	****
total LM at 0.9 ppm	0.609 0	0.022 0	0.604 8	0.012 2	459	.930 2	ns
total LM at 1.3 ppm	0.639 4	0.020 6	0.621 8	0.022 5	901	< .000 1	****
total LM at 2.0 ppm	0.682 8	0.010 2	0.805 1	0.008 1	0	< .000 1	****
myo-Inositol	0.686 9	0.012 3	0.716 2	0.012 6	3	< .000 1	****
total *N*-acetylaspartate	0.800 0	0.009 3	0.785 7	0.010 8	901	< .000 1	****
scyllo-Inositol	0.738 1	0.006 5	0.680 6	0.011 7	903	< .000 1	****
Taurine	0.697 0	0.010 6	0.750 4	0.010 9	0	< .000 1	****

Table showing the multiclass area under the curve of metabolites as well as lipids and macromolecules (LM), comparatively between pre- and post-noise suppression proton magnetic resonance spectroscopy. Significant levels are determined by the conditions as $P < .05$ (*), $P < .01$ (**), $P < .001$ (***), and $P < .0001$ (****).

TABLE D14 Statistics of multiclass areas under the curve for the combinations of metabolites as well as lipids and macromolecules.

		PreNS		PostNS		Wilcoxon signed-rank		
	Main peak	Mean	SD	Mean	SD	*V*	*P*	
1.5T Cohort								
total LM at 0.9 ppm	~0.9 ppm	0.774 9	0.005 2	0.774 5	0.005 7	1,789	.836 3	ns
Lactate and total LM at 1.3 ppm	~1.3 ppm	0.649 5	0.009 3	0.726 2	0.008 0	0	< .000 1	****
total LM at 2.0 ppm and total *N*-acetylaspartate	~2.0 ppm	0.624 4	0.008 2	0.571 5	0.007 4	3,486	< .000 1	****
Citrate, glutamate, and glutamine	~2.35 ppm	0.738 6	0.004 9	0.797 3	0.004 8	0	< .000 1	****
total Choline	~3.2 ppm	0.688 6	0.008 3	0.772 6	0.005 0	0	< .000 1	****
scyllo-Inositol and taurine	~3.3 ppm	0.665 0	0.005 1	0.668 7	0.007 2	818	< .000 1	****
Glycine and *myo*-Inositol	~3.6 ppm	0.789 4	0.005 6	0.814 2	0.003 6	1	< .000 1	****
Glucose and glutathione	~3.8 ppm	0.678 6	0.004 7	0.533 9	0.019 7	3,486	< .000 1	****
total Creatine	~3.9 ppm	0.771 7	0.005 8	0.778 3	0.006 2	149	< .000 1	****
3T Cohort								
total LM at 0.9 ppm	~0.9 ppm	0.609 0	0.022 0	0.604 8	0.012 2	459	.930 2	ns
Lactate and total LM at 1.3 ppm	~1.3 ppm	0.646 5	0.046 4	0.652 4	0.018 3	312	.082 1	ns
total LM at 2.0 ppm and total *N*-acetylaspartate	~2.0 ppm	0.768 2	0.009 9	0.792 0	0.012 5	42	< .000 1	****
Citrate, glutamate, and glutamine	~2.35 ppm	0.710 8	0.013 4	0.776 1	0.009 8	0	< .000 1	****
total Choline	~3.2 ppm	0.760 5	0.010 6	0.753 7	0.012 1	825	< .000 1	****
scyllo-Inositol and taurine	~3.3 ppm	0.692 1	0.011 0	0.747 4	0.115 7	0	< .000 1	****
Glycine and *myo*-Inositol	~3.6 ppm	0.567 1	0.022 7	0.697 7	0.019 1	0	< .000 1	****
Glucose and glutathione	~3.8 ppm	0.579 2	0.011 2	0.686 5	0.009 7	0	< .000 1	****
total Creatine	~3.9 ppm	0.639 4	0.023 7	0.634 1	0.025 8	739	.000 3	***

Table showing the multiclass area under the curve of metabolites as well as lipids and macromolecule (LM) combinations, comparatively between pre- and post-noise suppression proton magnetic resonance spectroscopy. Significant levels are determined by the conditions as $P < .05$ (*), $P < .01$ (**), $P < .001$ (***), and $P < .0001$ (****).

TABLE D15 Statistics of the overall classification accuracy for the 1.5T cohort.

	ACC, preNS		ACC, postNS		Wilcoxon rank-sum		
Type	Mean	SD	Mean	SD	*W*	*P*	
LOOCV							
preOS							
*k*NN	82%	1%	88%	0%	0	< .000 1	****
LDA	80%	0%	87%	0%	0	< .000 1	****
NB	82%	0%	88%	0%	0	< .000 1	****
NN	81%	0%	86%	0%	0	< .000 1	****
SVM	83%	0%	92%	0%	0	< .000 1	****
postOS							
*k*NN	84%	1%	91%	1%	0	< .000 1	****
LDA	78%	1%	90%	0%	0	< .000 1	****
NB	80%	1%	88%	1%	0	< .000 1	****
NN	81%	1%	87%	1%	0	< .000 1	****
SVM	88%	1%	92%	0%	0	< .000 1	****
k-fold							
preOS							
*k*NN	84%	4%	87%	3%	673	< .000 1	****
LDA	83%	4%	88%	3%	366	< .000 1	****
NB	85%	5%	89%	4%	521	< .000 1	****

TABLE D15 (Continued)

	ACC, preNS		ACC, postNS		Wilcoxon rank-sum		
Type	Mean	SD	Mean	SD	*W*	*P*	
NN	84%	4%	87%	3%	799	< .000 1	****
SVM	82%	4%	90%	3%	213	< .000 1	****
postOS							
*k*NN	85%	4%	90%	3%	424	< .000 1	****
LDA	80%	5%	90%	3%	84	< .000 1	****
NB	82%	4%	88%	4%	331	< .000 1	****
NN	83%	4%	87%	3%	487	< .000 1	****
SVM	86%	4%	91%	4%	445	< .000 1	****

Table comparatively showing the overall classification accuracy (ACC) for the 1.5T cohort between pre- and post-noise suppression (NS) in either pre- or post-oversampling (OS) conditions. The ACC was generated by using either leave-one-out cross-validation (LOOCV) or *k*-fold cross-validation ($k = 4$), and the machine-learning classifiers considered include *k*-nearest neighbours (*k*NN), linear discriminant analysis (LDA), naïve Bayes (NB), single-layer neural network (NN), and support vector machine (SVM). Significant levels are determined by the conditions as $P < .05$ (*), $P < .01$ (**), $P < .001$ (***), and $P < .0001$ (****).

TABLE D16 Statistics of the balanced classification accuracy for the 1.5T cohort.

	ACC, preNS		ACC, postNS		Wilcoxon rank-sum		
Type	Mean	SD	Mean	SD	*W*	*P*	
LOOCV							
preOS							
*k*NN	75%	1%	85%	0%	0	< .000 1	****
LDA	72%	0%	84%	0%	0	< .000 1	****
NB	76%	0%	86%	0%	0	< .000 1	****
NN	76%	0%	85%	0%	0	< .000 1	****
SVM	80%	0%	90%	0%	0	< .000 1	****
postOS							
*k*NN	79%	2%	89%	2%	0	< .000 1	****
LDA	71%	1%	89%	1%	0	< .000 1	****
NB	74%	1%	85%	2%	0	< .000 1	****
NN	77%	1%	86%	1%	0	< .000 1	****
SVM	84%	1%	90%	0%	0	< .000 1	****
k-fold							
preOS							
*k*NN	75%	6%	84%	5%	299	< .000 1	****
LDA	74%	6%	87%	4%	78	< .000 1	****
NB	78%	6%	86%	6%	408	< .000 1	****
NN	80%	4%	86%	5%	464	< .000 1	****
SVM	78%	5%	89%	5%	152	< .000 1	****
postOS							
*k*NN	80%	6%	89%	5%	368	< .000 1	****
LDA	72%	6%	88%	4%	36	< .000 1	****
NB	76%	5%	85%	6%	327	< .000 1	****
NN	79%	6%	86%	4%	354	< .000 1	****
SVM	82%	6%	89%	5%	428	< .000 1	****

Table comparatively showing the balanced classification accuracy (ACC) for the 1.5T cohort between pre- and post-noise suppression (NS) in either pre- or post-oversampling (OS) conditions. The ACC was generated by using either leave-one-out cross-validation (LOOCV) or *k*-fold cross-validation ($k = 4$), and the machine-learning classifiers considered include *k*-nearest neighbours (*k*NN), linear discriminant analysis (LDA), naïve Bayes (NB), single-layer neural network (NN), and support vector machine (SVM). Significant levels are determined by the conditions as $P < .05$ (*), $P < .01$ (**), $P < .001$ (***), and $P < .0001$ (****).

TABLE D17 Statistics of the overall classification accuracy for the 3T cohort.

	ACC, preNS		ACC, postNS		Wilcoxon rank-sum		
Type	Mean	SD	Mean	SD	*W*	*P*	
LOOCV							
preOS							
*k*NN	83%	0%	73%	2%	2,500	< .000 1	****
LDA	86%	0%	69%	0%	2,500	< .000 1	****
NB	69%	0%	74%	0%	0	< .000 1	****
NN	79%	0%	76%	0%	2,500	< .000 1	****
SVM	81%	0%	71%	0%	2,500	< .000 1	****
postOS							
*k*NN	63%	3%	70%	3%	126	< .000 1	****
LDA	71%	1%	78%	2%	5	< .000 1	****
NB	61%	2%	76%	2%	0	< .000 1	****
NN	69%	2%	79%	2%	0	< .000 1	****
SVM	72%	2%	76%	2%	356	< .000 1	****
k-fold							
preOS							
*k*NN	76%	4%	71%	3%	2,182	< .000 1	****
LDA	77%	3%	75%	3%	1,639	.007 2	**
NB	69%	4%	73%	3%	450	< .000 1	****
NN	74%	3%	74%	3%	1,228	.881 6	ns
SVM	77%	3%	75%	3%	1,841	< .000 1	****
postOS							
*k*NN	68%	5%	68%	3%	1,292	.774 0	ns
LDA	72%	4%	75%	4%	730	< .000 1	****
NB	70%	4%	79%	3%	66	< .000 1	****
NN	71%	4%	75%	4%	457	< .000 1	****
SVM	72%	4%	77%	3%	452	< .000 1	****

Table comparatively showing the overall classification accuracy (ACC) for the 3T cohort between pre- and post-noise suppression (NS) in either pre- or post-oversampling (OS) conditions. The ACC was generated by using either leave-one-out cross-validation (LOOCV) or *k*-fold cross-validation ($k = 4$), and the machine-learning classifiers considered include *k*-nearest neighbours (*k*NN), linear discriminant analysis (LDA), naïve Bayes (NB), single-layer neural network (NN), and support vector machine (SVM). Significant levels are determined by the conditions as $P < .05$ (*), $P < .01$ (**), $P < .001$ (***), and $P < .0001$ (****).

TABLE D18 Statistics of the balanced classification accuracy for the 3T cohort.

	ACC, preNS		ACC, postNS		Wilcoxon rank-sum		
Type	Mean	SD	Mean	SD	*W*	*P*	
LOOCV							
preOS							
*k*NN	62%	0%	54%	1%	2,500	< .000 1	****
LDA	63%	0%	51%	0%	2,500	< .000 1	****
NB	51%	0%	55%	0%	0	< .000 1	****
NN	58%	0%	63%	0%	0	< .000 1	****
SVM	59%	0%	53%	0%	2,500	< .000 1	****
postOS							
*k*NN	47%	2%	52%	2%	124	< .000 1	****
LDA	53%	1%	61%	4%	9	< .000 1	****
NB	46%	1%	56%	1%	0	< .000 1	****
NN	54%	4%	59%	3%	554	< .000 1	****
SVM	53%	2%	56%	2%	340	< .000 1	****
k-fold							
preOS							
*k*NN	57%	3%	53%	3%	2,174	< .000 1	****
LDA	57%	2%	56%	3%	1,497	.089 8	ns
NB	52%	3%	55%	3%	619	< .000 1	****
NN	55%	3%	57%	3%	723	.000 2	***
SVM	57%	2%	56%	3%	1,618	.011 4	*
postOS							
*k*NN	53%	4%	52%	3%	1,470	.130 2	ns
LDA	56%	4%	60%	5%	690	.000 1	***
NB	52%	3%	59%	2%	52	< .000 1	****
NN	54%	3%	61%	5%	333	< .000 1	****
SVM	54%	3%	60%	4%	317	< .000 1	****

Table comparatively showing the balanced classification accuracy (ACC) for the 3T cohort between pre- and post-noise suppression (NS) in either pre- or post-oversampling (OS) conditions. The ACC was generated by using either leave-one-out cross-validation (LOOCV) or *k*-fold cross-validation ($k = 4$), and the machine-learning classifiers considered include *k*-nearest neighbours (*k*NN), linear discriminant analysis (LDA), naïve Bayes (NB), single-layer neural network (NN), and support vector machine (SVM). Significant levels are determined by the conditions as $P < .05$ (*), $P < .01$ (**), $P < .001$ (***), and $P < .0001$ (****).

APPENDIX D: TABLES

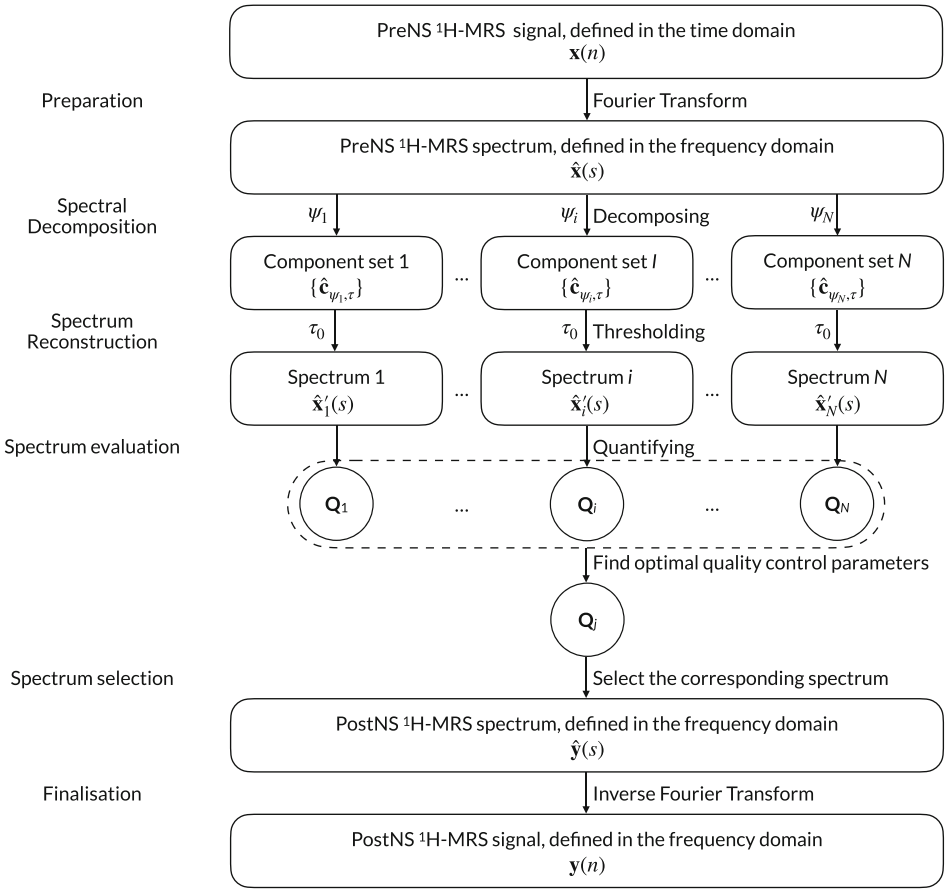

FIGURE E1 The detailed flowchart of adaptive wavelet noise suppression (AWNS).

APPENDIX E: FIGURES

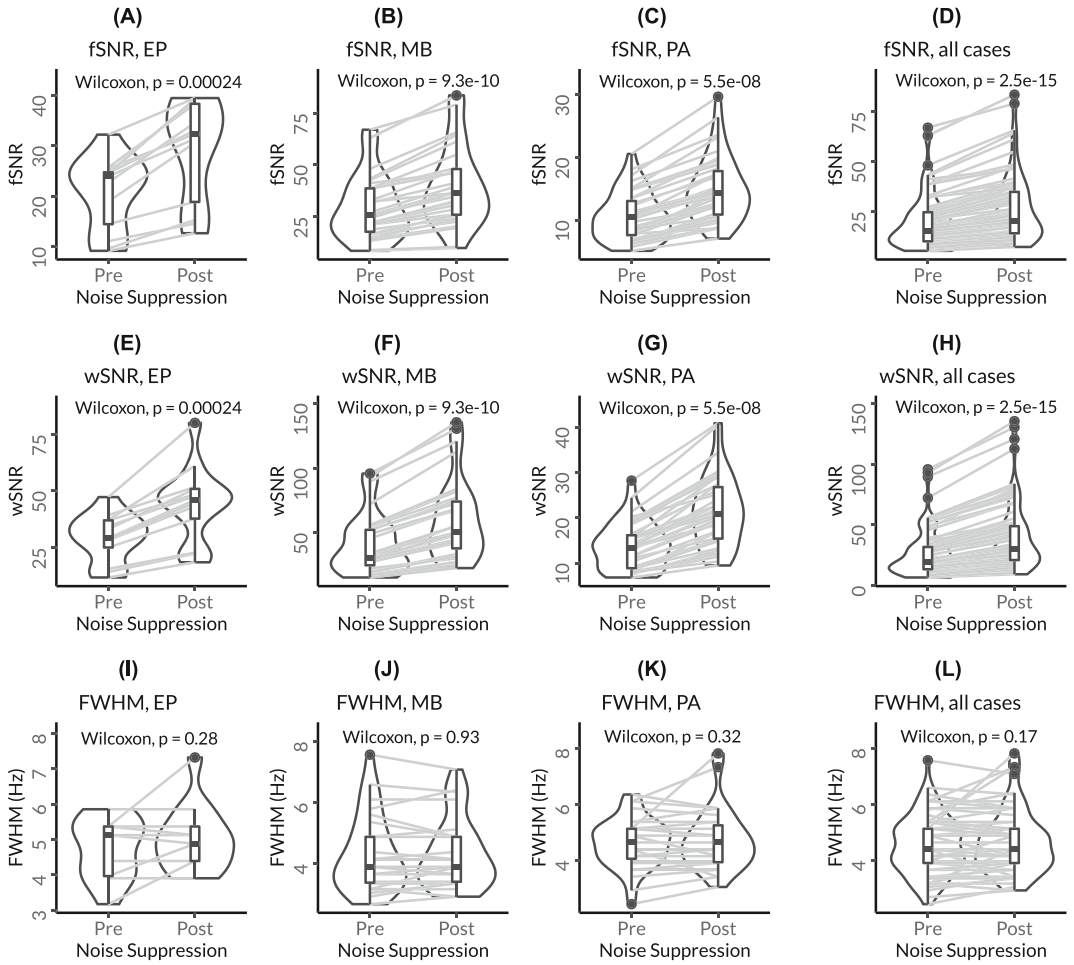

FIGURE E2 Diagrams showing the quality-control parameters of the 1.5T proton magnetic resonance spectroscopy data, including (A–D) fitting-based signal-to-noise ratio (fSNR), (E–H) whole-spectrum signal-to-noise ratio (wSNR), and (I–L) full width at half-maximum (FWHM), for each brain tumour type, including (A, E, I) ependymoma (EP), (B, F, J) medulloblastoma (MB), and (C, G, K) pilocytic astrocytoma (PA), with (D, H, L) all the cases as the reference.

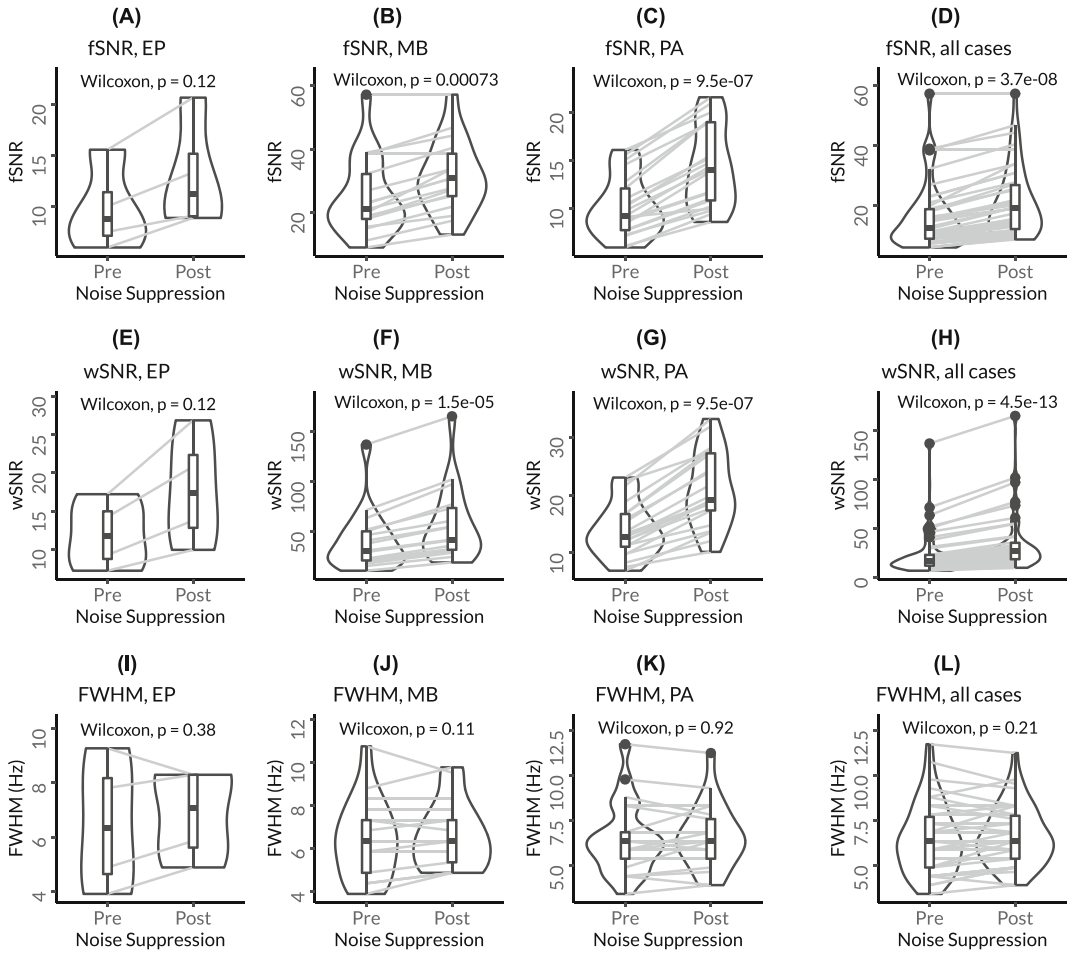

FIGURE E3 Diagrams showing the quality-control parameters of the 3T proton magnetic resonance spectroscopy data, including (A–D) fitting-based signal-to-noise ratio (fSNR), (E–H) whole-spectrum signal-to-noise ratio (wSNR), and (I–L) full width at half-maximum (FWHM), for each brain tumour type, including (A, E, I) ependymoma (EP), (B, F, J) medulloblastoma (MB), and (C, G, K) pilocytic astrocytoma (PA), with (D, H, L) all the cases as the reference.

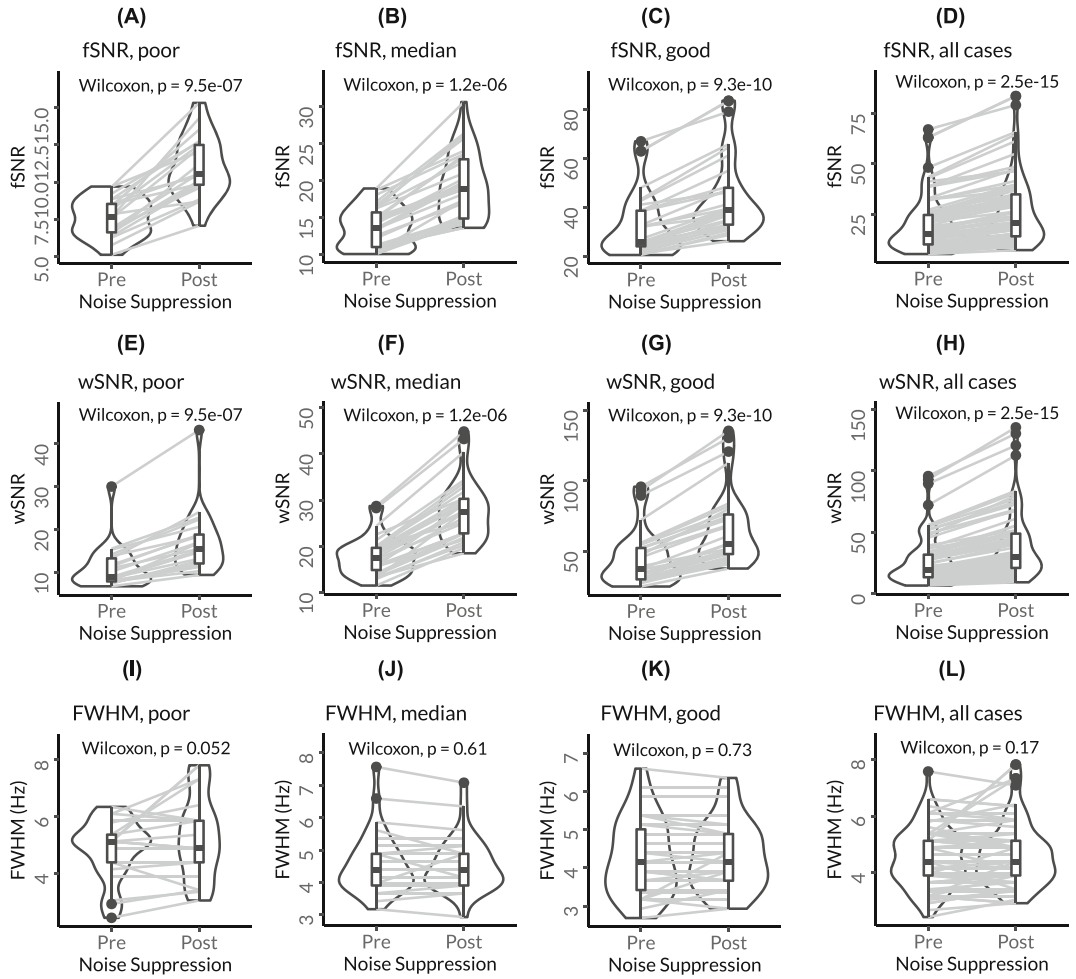

FIGURE E4 Diagrams showing the quality-control parameters of the 1.5T proton magnetic resonance spectroscopy data, including (A–D) fitting-based signal-to-noise ratio (fSNR), (E–H) whole-spectrum signal-to-noise ratio (wSNR), and (I–L) full width at half-maximum (FWHM), for each level of spectral quality, including (A, E, I) poor, (B, F, J) median, and (C, G, K) good quality, with (D, H, L) all the cases as the reference.

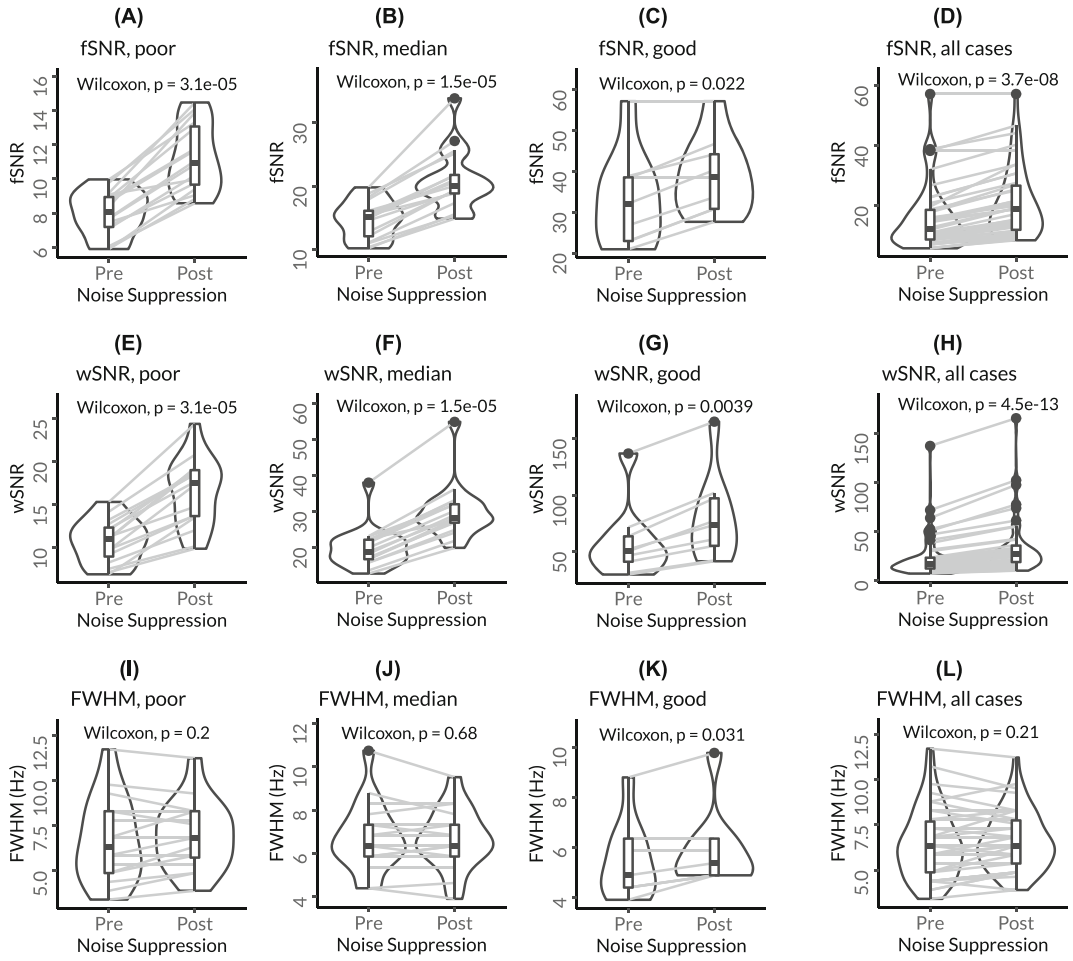

FIGURE E5 Diagrams showing the quality-control parameters of the 3T proton magnetic resonance spectroscopy data, including (A–D) fitting-based signal-to-noise ratio (fSNR), (E–H) whole-spectrum signal-to-noise ratio (wSNR), (I–L) and full width at half-maximum (FWHM), for each level of spectral quality, including (A, E, I) poor, (B, F, J) median, and (C, G, K) good quality, with (D, H, L) all the cases as the reference.

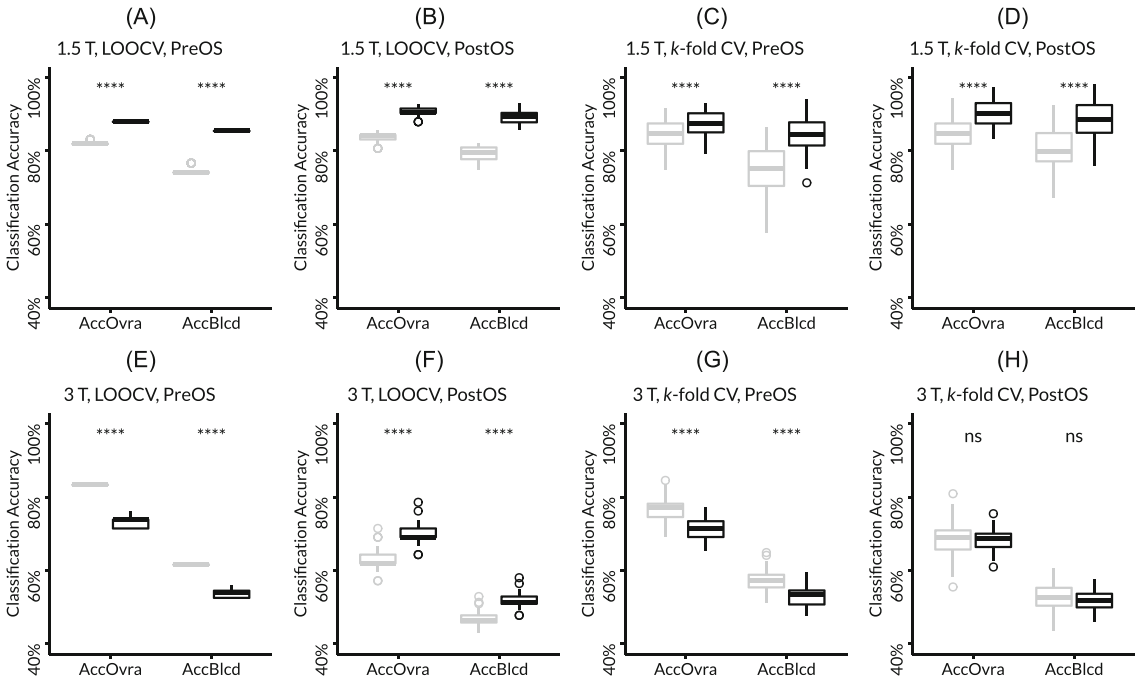

FIGURE E6 Box plots showing the significantly improved overall and balanced classification accuracy (AccOvra and AccBlcd) of the three brain tumour types, ependymomas, medulloblastomas, and pilocytic astrocytomas, determined through (A–B, E–F) leave-one-out (LOO) and (C–D, G–H) *k*-fold ($k = 10$ for the 1.5T cohort and $k = 4$ for the 3T cohort) cross-validation (CV) and *k*-nearest neighbours for (A, C, E, G) pre- and (B, D, F, H) post-oversampling (OS) (A–D) 1.5T and (E–H) 3T pre- (grey) and post- (black) noise suppression proton magnetic resonance spectroscopy data, where oversampling was performed for ependymomas with an oversampling rate of 100%. Level of significance: ns, $P > 0.05$; **, $P < 0.01$; ****, $P < 0.0001$.

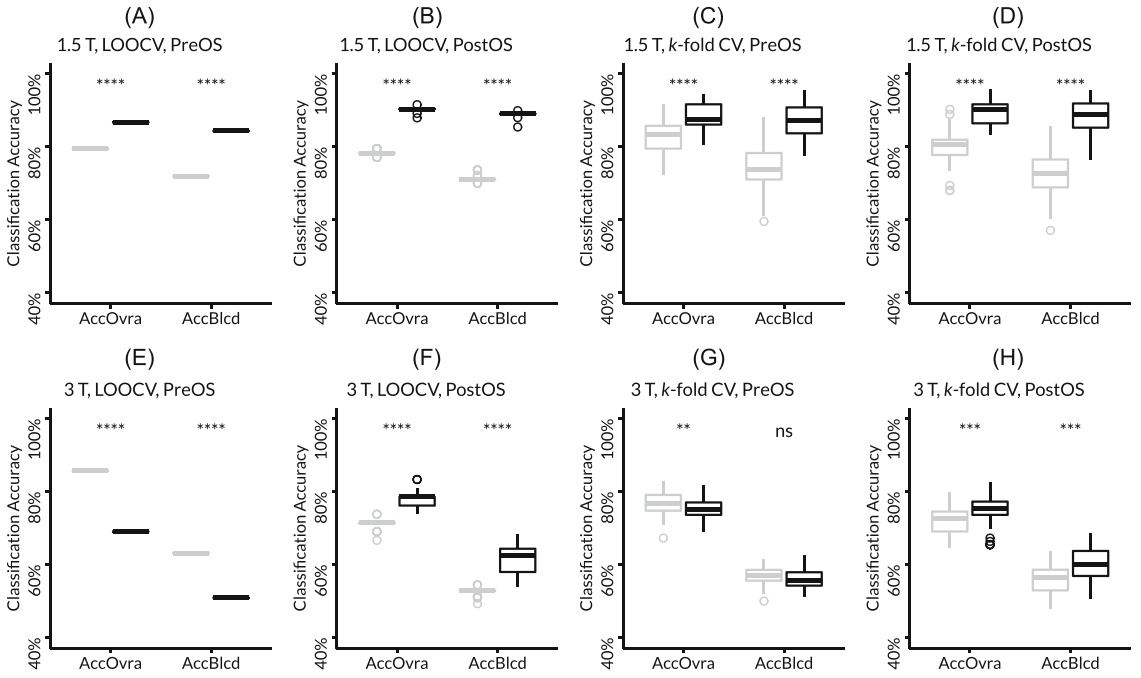

FIGURE E7 Box plots showing the significantly improved overall and balanced classification accuracy (AccOvra and AccBlcd) of the three brain tumour types, ependymomas, medulloblastomas, and pilocytic astrocytomas, determined through (A–B, E–F) leave-one-out (LOO) and (C–D, G–H) *k*-fold ($k = 10$ for the 1.5T cohort and $k = 4$ for the 3T cohort) cross-validation (CV) and linear discriminant analysis for (A, C, E, G) pre- and (B, D, F, H) post-oversampling (A–D) 1.5T and (E–H) 3T pre- (grey) and post- (black) noise suppression proton magnetic resonance spectroscopy data, where oversampling was performed for ependymomas with an oversampling rate of 100%. Level of significance: ns, $P > 0.05$; **, $P < .01$; ***, $P < .001$; ****, $P < .0001$.

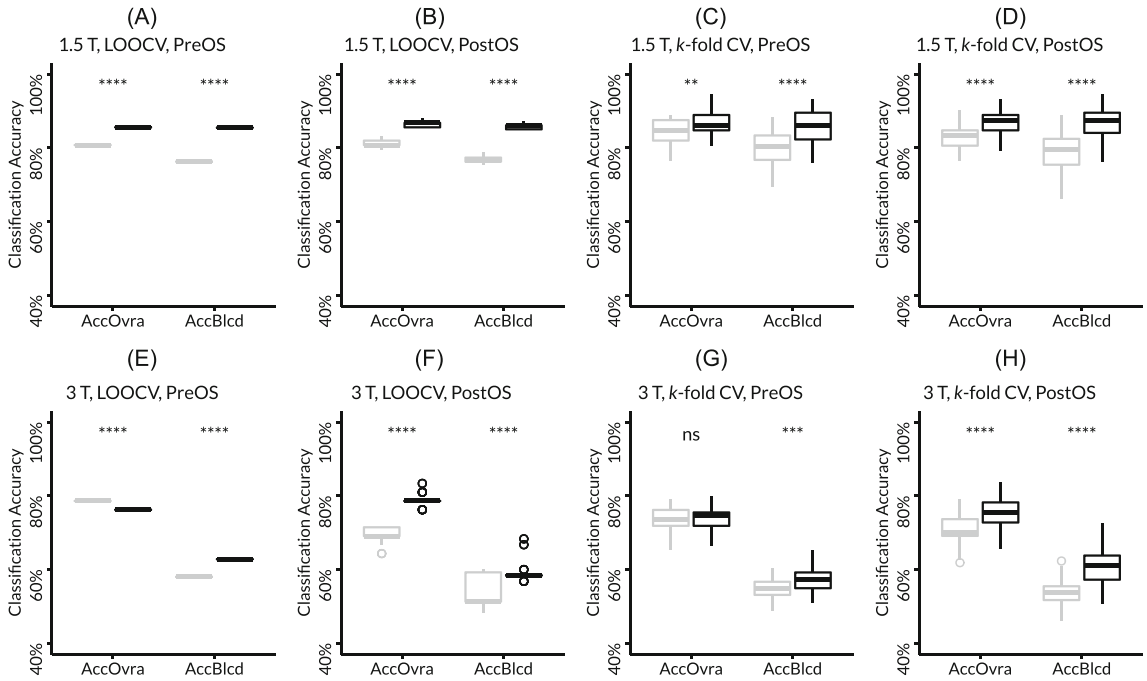

FIGURE E8 Box plots showing the significantly improved overall and balanced classification accuracy (AccOvra and AccBlcd) of the three brain tumour types, ependymomas, medulloblastomas, and pilocytic astrocytomas, determined through (A–B, E–F) leave-one-out (LOO) and (C–D, G–H) *k*-fold ($k = 10$ for the 1.5T cohort and $k = 4$ for the 3T cohort) cross-validation (CV) and single-layer neural network for (A, C, E, G) pre- and (B, D, F, H) post-oversampling (A–D) 1.5T and (E–H) 3T pre- (grey) and post- (black) noise suppression proton magnetic resonance spectroscopy data, where oversampling was performed for ependymomas with an oversampling rate of 100%. Level of significance: ns, $P > 0.05$; **, $P < .01$; ***, $P < .001$; ****, $P < .0001$.

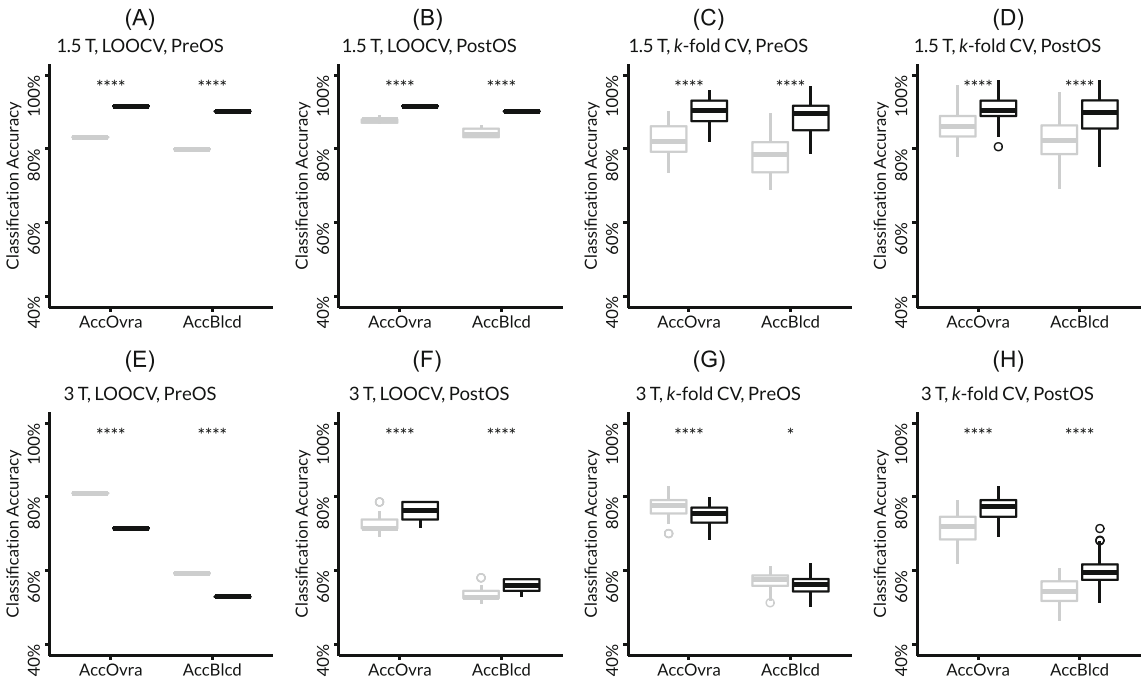

FIGURE E9 Box plots showing the significantly improved overall and balanced classification accuracy (AccOvra and AccBlcd) of the three brain tumour types, ependymomas, medulloblastomas, and pilocytic astrocytomas, determined through (A–B, E–F) leave-one-out (LOO) and (C–D, G–H) *k*-fold ($k = 10$ for the 1.5T cohort and $k = 4$ for the 3T cohort) cross-validation (CV) and support vector machine with a linear kernel for (A, C, E, G) pre- and (B, D, F, H) post-oversampling (A–D) 1.5T and (E–H) 3T pre- (grey) and post- (black) noise suppression proton magnetic resonance spectroscopy data, where oversampling was performed for ependymomas with an oversampling rate of 100%. Level of significance: *, $P < .05$; ****, $P < .0001$.