Original Article

Diagnostic accuracy of qualitative MRI in 550 paediatric brain tumours: evaluating current practice in the computational era

Luke Dixon[1]^, Gurpreet Kaur Jandu[2]^, Jai Sidpra[3,4]^, Kshitij Mankad[3,4]^

[1]Department of Neuroradiology, Imperial University Healthcare NHS Foundation Trust, London, UK; [2]Cardiff University Medical School, Cardiff, UK; [3]Developmental Biology and Cancer Section, University College London Great Ormond Street Institute of Child Health, London, UK; [4]Department of Neuroradiology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK

Contributions: (I) Conception and design: L Dixon, K Mankad; (II) Administrative support: None; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: L Dixon, GK Jandu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Dr. Kshitij Mankad, FRCR. Consultant Paediatric Neuroradiologist, Department of Neuroradiology, Great Ormond Street Hospital for Children NHS Foundation Trust, Great Ormond Street, London, WC1N 3JH, UK. Email: drmankad@gmail.com.

Background: To investigate the accuracy of qualitative reporting of conventional magnetic resonance imaging (MRI) in the classification of paediatric brain tumours.

Methods: Preoperative MRI reports of 608 children prior to resection or biopsy of an intracranial lesion were retrospectively reviewed. A total of 550 children had complete radiological and histopathological notes, thereby reaching our inclusion criteria. Concordance between MRI report and final histopathological diagnosis was assessed using an established lexicon derived from the WHO 2016 classification of CNS tumours. Levels of agreement based on cellular origin, tumour type, and tumour grade were evaluated. Diagnostic accuracy, sensitivity, specificity, confidence intervals, and positive and negative predictive values were calculated.

Results: Diagnostic accuracy differed significantly between tumour types and tumour grades. Sensitivities were highest for ependymomas and sellar, pituitary, pineal, and cranial and/or paraspinal nerve tumours (range 80.65–100%). Sensitivity was slightly lower for astrocytic gliomas, oligodendrogliomas, and choroid plexus, neuronal, mixed neuronal-glial, embryonal, and histiocytic tumours (range 63.33–79.59%). Low sensitivities were noted for meningiomas and mesenchymal non-meningothelial, melanocytic, and germ cell tumours (range 0–56.25%). The most correct tumour type predictions were made in the posterior fossa whilst the most incorrect predictions were made in the lobar regions, pineal/tectal plate area, and the supratentorial ventricles.

Conclusions: This is the largest published series investigating the predictive accuracy of MRI in paediatric brain tumours. We show that diagnostic accuracy varies greatly by tumour type and location. Looking forward, we should develop and leverage computational methods to improve accuracy in the tumour types and anatomical locations where qualitative diagnostic accuracy is lower.

Keywords: Brain; brain tumour; diagnostic imaging; magnetic resonance imaging (MRI); neoplasm; neuro-oncology

Submitted Dec 23, 2020. Accepted for publication Jun 16, 2021.
doi: 10.21037/qims-20-1388
View this article at: https://dx.doi.org/10.21037/qims-20-1388

^ ORCID: Luke Dixon, 0000-0003-4060-7897; Gurpreet Kaur Jandu, 0000-0001-5815-3390; Jai Sidpra, 0000-0003-1985-5503; Kshitij Mankad, 0000-0001-5979-9337.

Introduction

Paediatric brain tumours are the most common solid cancers of childhood, with an approximate prevalence of 5.6 per 100,000 (1-3). Over the past two decades, advances in surgical management, chemotherapy, and radiotherapy have improved the survival rates of children with certain brain tumours, such as medulloblastoma and low grade glioma (4,5). Despite this progress, brain tumours remain the leading cause of cancer-related deaths in children and there continues to be a poor prognosis for certain tumour types, such as diffuse midline gliomas (DMGs) (1,3,6). Moreover, there is an increasing concern that curative therapy can itself exact a heavy price in terms of late neurodevelopmental morbidity (4). In this light, attention has shifted in recent years to more individualised treatment regimens that are tailored to specific tumour types and clinical phenotypes. This has been greatly bolstered by an improved understanding of the molecular and genetic underpinnings of brain tumours (7).

Magnetic resonance imaging (MRI) has played a pivotal role in tumour detection, preliminary diagnosis, treatment guidance, and follow-up monitoring (8). Historically, this was largely based on qualitative assessment of MRI by radiologists and neurosurgeons but recent advances in computing power have greatly facilitated the development of quantitative image analysis techniques and machine learning approaches, which are now hoped to augment and improve tumour diagnosis (9-11). Furthermore, via correlation with molecular profiling, it is hoped that these and related computational techniques may permit the accurate, image-based, non-invasive phenotyping and predictive genotyping of brain tumours. This would be of particular benefit in guiding therapy in the 'non-operable tumours', such as in certain DMGs (12,13). There are, however, many challenges with these techniques, including problems with external generalisation and reproducibility. Another key issue is redundancy: where computational methods are being developed and proposed for situations where there is little pressing need for a new solution.

In this era of computational imaging, it is crucial to look back at our past experience in qualitative image analysis. In this study, we aim to retrospectively assess the diagnostic accuracy of preoperative MRI reports performed during the course of routine clinical practice. Through this process, we hope to better understand our strengths and shortcomings in conventional image interpretation of paediatric brain tumours such that we can target our computational focus onto areas where such assistance may be of most benefit.

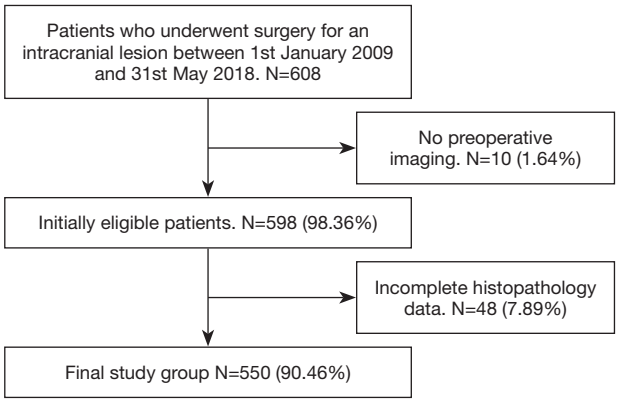

Figure 1 Patient enrolment flow diagram.

Methods

This clinical audit was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and approved by The Great Ormond Street Hospital for Children NHS Foundation Trust institutional review board prior to commencement, with individual consent waived for this retrospective analysis of imaging data.

Patient population and inclusion criteria

We performed a retrospective observational study of paediatric patients who underwent biopsy and/or resection of a brain tumour at our institution, a specialist tertiary and university-affiliated children's hospital. To reduce bias, the study was designed according to the Standards for The Reporting of Diagnostic Accuracy Studies (STARD) Criteria (14). Patients were identified for inclusion via retrospective review of a prospectively maintained histopathology database of all brain tumours evaluated at our institution between 1st January 2009 and 1st May 2018. The inclusion criteria for this study were: (I) all patients with preoperative imaging and an associated radiology report and (II) all patients with available and complete histopathology results. Our initial search identified 608 patients. Patients with no preoperative MRI or incomplete postoperative histopathology data were excluded, resulting in a total of 550 eligible patients in the final study group (*Figure 1*).

Pathological sampling and findings

Pathology samples were acquired by surgical resection

or targeted surgical biopsy. Specimens were analysed by experienced board-certified consultant paediatric neuropathologists and classed as either neoplastic or non-neoplastic (e.g., tumefactive demyelination). Neoplastic lesions were then classified and graded according to the contemporaneous World Health Organisation (WHO) central nervous system (CNS) tumour classification at the time of sampling (15). The final recorded pathological diagnostic report was retrospectively reviewed and recorded in a secure, anonymised database.

MRI protocol

All patients underwent clinically indicated preoperative MRI as per our standardised institutional brain tumour protocol: sagittal T1- and axial T2-weighted imaging, coronal fluid attenuated inversion recovery, and diffusion weighted imaging/apparent diffusion coefficient of the brain and post-contrast T1-weighted imaging of the brain and whole spine. Images were stored in a Picture Archiving and Communication System (PACS). All MRI scans were examined and formally reported by a board-certified consultant paediatric neuroradiologist. Clinical and demographic information available at the time of diagnosis was considered and factored into the final radiological diagnosis.

Review of MRI reports and ontological classification of MRI prediction and pathological diagnosis

The MRI and histopathology reports were reviewed by two blinded observers. The MRI prediction and histopathology diagnosis were coded based on a simplified form of the current WHO 2016 classification of CNS tumours. A lexicon of different terms was formulated to translate into an agreed standardised ontology to account for synonymous nosology and inevitable heterogeneity in tumour naming due to the study period covering both the WHO 2007 and WHO 2016 classification of CNS tumours (15,16). A consensus agreement between the two neuroradiologists was reached for classification. Where relevant, a simplified grade was defined as an attribute of the ontology with WHO grade I and II tumours classified as 'low grade' whilst WHO grade III and IV tumours were classified as 'high grade'. Where the radiology report offered no explicit prediction of grade, then tumours was classified as 'not otherwise specified'. Where reports offered a hierarchy of differential diagnoses, the favoured top diagnosis was taken as the preoperative MRI prediction. *Table 1* depicts the lexicon of terms followed and their standardised ontological definitions.

Statistical analysis

MRI report prediction and the final reference standard histopathological diagnosis were dichotomised according to the ontological definitions. The MRI reports' sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and corresponding 95% confidence intervals (CIs) using Clopper-Pearson score method were estimated for broad tumour type, grade, and the overall patient population. MRI report predictive accuracy for tumour type was also assessed based on tumour location. Categorical values were expressed as numbers with percentages. All data was analysed on MedCalc version 19.5.3 (MedCalc Software Limited, Ostend, Belgium).

Results

A total of 550 patients were included. The most common tumour types were astrocytic and oligodendrogliomas, neuronal and mixed neuronal-glial tumours, and embryonal tumours. These three tumour categories collectively made up 74.5% of all cases reviewed. A total of 399 (72.55%) cases had the broad tumour type correctly predicted on MRI. A summary of correct MRI predictions and test performance results for each tumour type are presented in *Table 2*. As expected, the specificity, NPV, and overall accuracy increased with the rarer tumour types as the relative number of true negatives greatly increased—thereby limiting interpretation of this measure in these groups. Predictive performance of MRI varied amongst tumour types. Sensitivity greatly varied by tumour type, from 0–100%, whilst specificity was generally high ranging from 93.6–100%. Sensitivities were highest for ependymomas, tumours of sellar region, pituitary, pineal, and cranial and/or paraspinal nerve tumours—range 80.65–100%—whilst there were slightly lower sensitivities for astrocytic gliomas and oligodendrogliomas in addition to choroid plexus, neuronal, mixed neuronal-glial, embryonal, and histiocytic tumours—range 63.33–79.59%. Low sensitivities were noted for meningiomas, mesenchymal non-meningothelial, melanocytic, and germ cell tumours—range 0–56.25%. A cross-table of MRI predictions versus final histopathological diagnosis is shown in *Figure 2*. Regarding astrocytic gliomas and oligodendrogliomas, when not predicted accurately by

Table 1 Ontological definitions of different tumour definitions as per the WHO 2016 classification

WHO classification	Grade	Tumour diagnosis
Astrocytic gliomas and oligodendrogliomas	High	"High grade astrocytoma", "Anaplastic astrocytoma", "Glioblastoma", "GBM", "Diffuse midline glioma", "DIPG", "Diffuse intrinsic pontine glioma", "Anaplastic oligodendroglioma", "Astrocytoma grade III", "Astrocytoma grade IV"
	Low	"Low grade astrocytoma", "Oligodendroglioma", "Astrocytoma grade I", "Astrocytoma grade II", "Low grade glial", "Diffuse astrocytoma",
	Not specified	"Astrocytoma", "diffuse glioma"
Other astrocytic gliomas	High	"Anaplastic pleomorphic xanthoastrocytoma"
	Low	"Pilocytic astrocytoma", "Subependymal giant cell astrocytoma", "SEGA", "Pleomorphic xanthoastrocytoma", "PXA", "Pilomyxoid astrocytoma"
Ependymoma	High	"Ependymoma RELA Positive", "High grade ependymoma", Grade III ependymoma", "Anaplastic ependymoma"
	Low	"Subependymoma", "Ependymoma"
Other gliomas	Low	"Angiocentric glioma", "Chordoid glioma of the third ventricle"
Choroid plexus tumour	High	"Choroid plexus carcinoma"
	Low	"Choroid plexus papilloma"
Neuronal and mixed neuronal-glial tumours	High	"Anaplastic ganglioglioma"
	Low	"Dysembryoplastic neuroepithelial tumour", "DNET", "Gangliоglioma", "Gangliocytoma", "Dysplastic gangliocytoma of the cerebellum", "Lhermitte Duclos", "Desmoplastic infantile astrocytoma and ganglioglioma", "Papillary glioneuronal tumour", "Rosette-forming glioneuronal tumour", "Central neurocytoma", "Extraventricular neurocytoma", "Cerebellar liponeurocytoma", "Multinodular and vacuolating tumour"
Pineal tumours	High	"Pineoblastoma", "High grade pineal tumour"
	Low	"Pineocytoma", "Low grade pineal tumour"
Embryonal tumours	High	"Medulloblastoma", "Embryonal tumour with multi-layered rosettes", "ETMR", "Medulloepithelioma", "Embryonal tumour NOS/not specified", "Atypical teratoid/rhabdoid tumour", "ATRT", "CNS embryonal tumour with rhabdoid features", "Neuroblastoma", "PNET", "Small round cell"

the MRI report, they were either labelled as a non-specific tumour or predicted to be ependymomas, embryonal, or neuronal and mixed neuronal-glial tumours.

In some categories of tumour type, two tiers of analysis for both tumour type and grade was possible. If grade was also assessed, there was a marked drop in sensitivity for certain tumour types, specifically in high grade astrocytic gliomas and oligodendrogliomas, ependymomas, low grade choroid plexus tumours, high grade neuronal and mixed neuronal-glial tumours, low grade pineal tumours and high grade meningiomas (*Table 2*)—range 0–40%. *Figure 3* is a bar chart demonstrating sensitivity for more common tumour types and grades. For gliomas and neuronal and mixed neuronal-glial classes, a drop in sensitivity and accuracy when predicting grade was significantly related to reduced precision of reporting. 7% of astrocytic glioma and oligodendroglioma and 10% of neuronal and mixed neuronal-glial tumour reports only stated the broad tumour class without any reference to suspected grade or tumour subtype.

The locations of the most common different tumour types and grades are summarised in *Figure 4*. The percentage of discordant predictions varied depending on tumour location. Tumours in the thalamus, posterior fossa, pituitary, and suprasellar regions had the most correct predictions. Tumours in lobar locations, the pineal region, and in the supratentorial ventricles had the greatest percentage of incorrect predictions (*Figure 5*). Prediction of

Table 2 Test performance of MRI predictions by broad tumour type compared to final pathological diagnosis. In aggregate, 399/500 (72.55%) of tumours were correctly diagnosed from MRI

Tumour type	Grade	Number of cases					Test performance measures (LCB-UCB)				
		N	TP	FP	FN	Prevalence	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
Astrocytic and oligodendroglioma	All	222	161	21	61	40.36	72.52 (66.15–78.28)	93.6 (90.38–95.99)	88.46 (83.41–92.12)	83.42 (80.22–86.19)	85.09 (81.84–87.96)
	High	57	19	6	38	10.36	33.33 (21.40–47.06)	98.78 (97.37–99.55)	76.00 (56.87–88.38)	92.76 (91.43–93.90)	92.00 (89.41–94.13)
	Low	165	98	21	67	30.00	59.39 (51.48–66.96)	94.55 (91.78–96.59)	82.35 (75.14–87.82)	84.45 (81.85–86.74)	84.00 (80.66–86.97)
Ependymoma	All	31	25	12	6	5.64	80.65 (62.53–92.55)	97.69 (96.00–98.80)	67.57 (53.71–78.09)	98.83 (97.63–99.43)	96.73 (94.88–98.05)
	High	27	6	1	21	4.91	22.22 (8.62–42.26)	99.81 (98.94–100.00)	85.71 (42.81–97.96)	96.13 (95.31–96.82)	96.00 (94.01–97.48)
	Low	4	0	0	4	0.73	0.00 (0.00–60.24)	100.00 (99.33–100.00)	–	99.27 (99.27–99.27)	99.27 (98.15–99.80)
Other gliomas	Low	2	0	0	2	0.36	0.00 (0.00–84.19)	100.00 (99.33–100.00)	–	99.64 (99.64–99.64)	99.64 (98.69–99.96)
Choroid plexus tumour	All	12	9	0	3	2.18	75.00 (42.81–94.51)	100.00 (99.32–100.00)	100.00	99.45 (98.54–99.79)	99.45 (98.41–99.89)
	High	2	1	0	1	0.54	66.67 (9.43–99.16)	100.00 (99.33–100.00)	100.00	99.82 (99.10–99.96)	99.82 (98.99–100.00)
	Low	10	4	0	6	1.82	40.00 (12.16–73.76)	100.00 (99.32–100.00)	100.00	98.90 (98.19–99.33)	98.91 (97.64–99.60)
Neuronal and mixed Neuronal-glial tumours	All	90	57	9	33	16.36	63.33 (52.51–73.25)	98.04 (96.32–99.10)	86.36 (76.50–92.49)	93.18 (91.24–94.72)	92.36 (89.82–94.44)
	High	2	0	0	2	0.36	0.00 (0.00–84.19)	100.00 (99.33–100.00)	–	99.64 (99.64–99.64)	99.64 (98.69–99.96)
	Low	88	57	7	31	16.00	64.77 (53.86–74.66)	98.48 (96.90–99.39)	89.06 (79.35–94.52)	93.62 (91.70–95.12)	93.09 (90.64–95.06)
Primary pineal gland tumours	All	9	5	2	3	99.17	95.00 (86.08–98.96)	99.63 (98.67–99.96)	96.61 (87.71–99.13)	99.45 (98.35–99.82)	99.17 (98.07–99.73)
	High	8	5	3	3	1.46	62.50 (24.49–91.48)	99.63 (98.67–99.96)	71.43 (36.18–91.68)	99.45 (98.66–99.77)	99.09 (97.89–99.70)
	Low	1	0	0	1	0.18	0.00 (0.00–97.50)	100.00 (99.33–100.00)	–	99.82 (99.82–99.82)	99.82 (98.99–100.00)
Embryonal tumours	High	98	78	14	20	17.82	79.59 (70.26–87.07)	96.90 (94.86–98.03)	84.78 (76.72–90.40)	95.63 (93.67–97.01)	93.28 (91.47–95.68)
Tumours of the cranial and/or paraspinal nerves	Low	2	2	1	0	0.36	100.00 (15.81–100.00)	99.82 (98.99–100.00)	66.67 (22.01–93.41)	100.00	99.82 (98.99–100.00)
Meningiomas	All	8	2	2	6	1.45	25.00 (3.19–65.09)	99.63 (98.67–99.96)	50.00 (13.81–86.19)	98.90 (98.37–99.26)	98.55 (97.15–99.37)
	High	5	0	0	5	0.91	0.00 (0.00–52.18)	100.00 (99.33–100.00)	–	99.09 (99.09–99.09)	99.09 (97.89–99.70)
	Low	3	1	0	2	0.55	33.33 (0.84–90.57)	100.00 (99.33–100.00)	100.00	99.64 (99.19–99.84)	99.64 (98.69–99.96)
Mesenchymal non-meningothelial	All	3	1	0	2	0.55	33.33 (0.84–90.57)	100.00 (99.33–100.00)	100.00	99.64 (99.19–99.84)	99.64 (98.69–99.96)
	High	1	0	0	1	0.18	0.00 (0.00–97.50)	100.00 (99.33–100.00)	–	99.82 (99.82–99.82)	99.82 (98.99–100.00)
	Low	2	1	0	1	0.54	66.67 (9.43–99.16)	100.00 (99.33–100.00)	100.00	99.82 (99.10–99.96)	99.82 (98.99–100.00)
Melanocytic	High	1	0	0	1	0.18	0.00 (0.00–97.50)	100.00 (99.33–100.00)	–	99.82 (99.82–99.82)	99.82 (98.99–100.00)

Table 2 (*continued*)

Table 2 (*continued*)

Tumour type	Grade	Number of cases					Test performance measures (LCB-UCB)				
		N	TP	FP	FN	Prevalence	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
Histiocytic	Nos	3	2	0	1	0.55	66.67 (9.43–99.16)	100.00 (99.33–100.00)	100.00	99.82 (99.10–99.96)	99.82 (98.99–100.00)
Germ cell tumours	Nos	16	9	4	7	2.91	56.25 (29.88–80.25)	99.25 (98.09–99.80)	69.23 (43.62–86.75)	98.70 (97.75–99.25)	98.00 (96.45–99.00)
Suprasellar tumours	Low	35	35	3	0	6.36	100 (90.00–100.00)	99.42 (98.31–99.88)	92.11 (79.06–97.30)	100.00	99.45 (98.41–99.89)
Metastatic tumours	Nos	4	3	0	1	6.19	97.22 (85.47–99.93)	100.00 (99.33–100.00)	100.00	99.82 (98.75–99.97)	99.83 (99.05–100.00)
Pituitary adenomas	Nos	11	9	0	2	2.00	81.82 (48.22–97.72)	100.00 (99.32–100.00)	100.00	99.63 (98.72–99.89)	99.64 (98.69–99.96)
Tumour uncategorised	All	2	1	0	1	0.54	66.67 (9.43–99.16)	100.00 (99.33–100.00)	100.00	99.82 (99.10–99.96)	99.82 (98.99–100.00)
	Low	2	1	0	1	0.54	66.67 (9.43–99.16)	100.00 (99.33–100.00)	100.00	99.82 (99.10–99.96)	99.82 (98.99–100.00)
Non-neoplastic	NA	1	0	17	1	0.18	0.00 (0.00–97.50)	96.90 (95.09–98.19)	0.00	99.81 (99.81–99.82)	96.73 (94.88–98.05)

MRI, magnetic resonance imaging; TP, true positives; FP, false positives; FN, false negatives; PPV, positive predictive value; NPV, negative predictive value; LCB/UCB, lower/upper 95% confidence interval calculated using Clopper-Pearson score method.

tumour grade distinct from tumour type also varied across location with tumours in the basal ganglia and thalamus often being under-graded whilst tumours in the pineal region and supratentorial ventricles were often over-graded (*Figure 6*).

Notably, diagnostic accuracy did not change significantly over the course of our study and did not change following the release of the WHO 2016 classification of CNS tumours.

Discussion

This is the largest published series investigating the predictive accuracy of qualitative MRI assessment of paediatric brain tumours. Overall accuracy of MRI reports for broad tumour type was good, with 72.55% of cases reaching the correct prediction. However, the diagnostic error rate varied greatly between different tumour types and anatomical locations. This emphasises clear strengths and limitations in current qualitative MRI interpretation of paediatric brain tumours whilst permitting an insight into where future quantitative methods should be targeted.

Variation in diagnostic accuracy between different tumour types is understandable. The broad tumour groups with the highest diagnostic accuracy often have well-recognised, characteristic radiological features and, for certain classes, typical locations where there is a limited differential—for instance, pineal and pituitary tumours. However, even in some of these groups, this high sensitivity was substantially reduced when looking at a subclass of tumour and grade, for instance, in ependymomas and pineal tumours. This reduced grading accuracy is, in part, likely related to limitations in interpretation and poor imaging differentiators. In the case of certain tumours, it is also important to recognise that even histopathological assessment is not necessarily decisive in differentiating grade (4,17,18). For instance, in ependymomas, there is well-recognised inter-observer variation in histological classification between grades II and III (17,18). This sometimes uncertain differentiation of grade II from grade III has limited clinical utility in guiding treatment where, in most cases, the mainstay of initial therapy is total surgical resection irrespective of predicted grade (17,18). In this light, the reporting radiologist may understandably and consciously omit a predicted grade as it has little impact on treatment decision.

There were, however, certain tumour groups where accurate prediction of tumour type and grade would

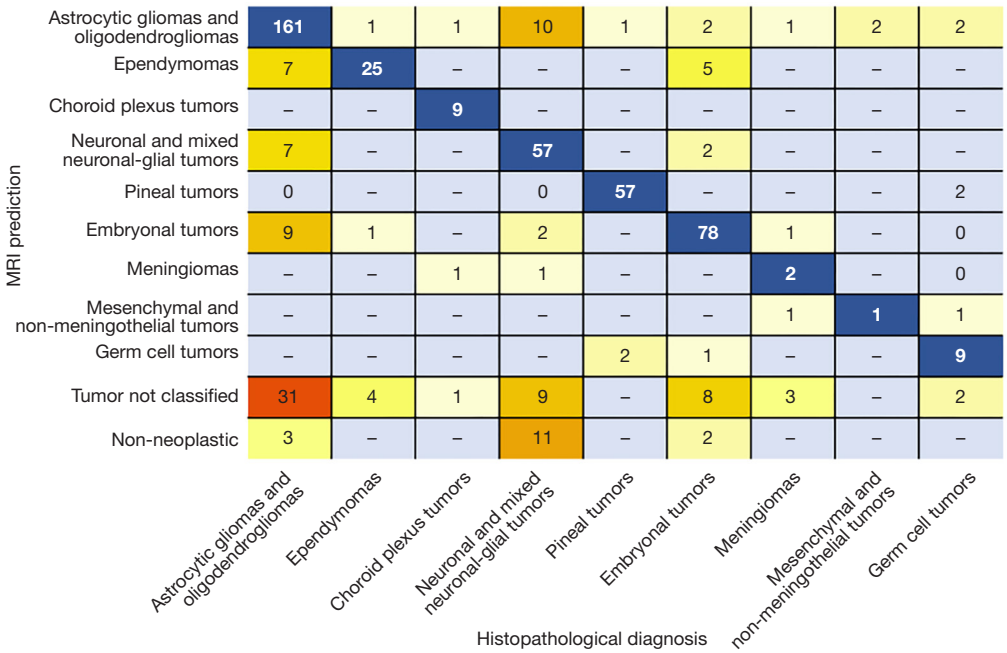

Figure 2 Cross-table of MRI predictions of different tumour types compared to final histopathological diagnosis. The horizontal axis corresponds to the gold-standard final histopathological diagnosis of the different tumour types, whereas the vertical axis corresponds to the MRI prediction. The numbers represent the counts of that specific tumour type. Dark blue reflects concordant correct predictions whilst the heatmap of yellow to red reflects the most frequently incorrectly predicted tumour types. MRI, magnetic resonance imaging.

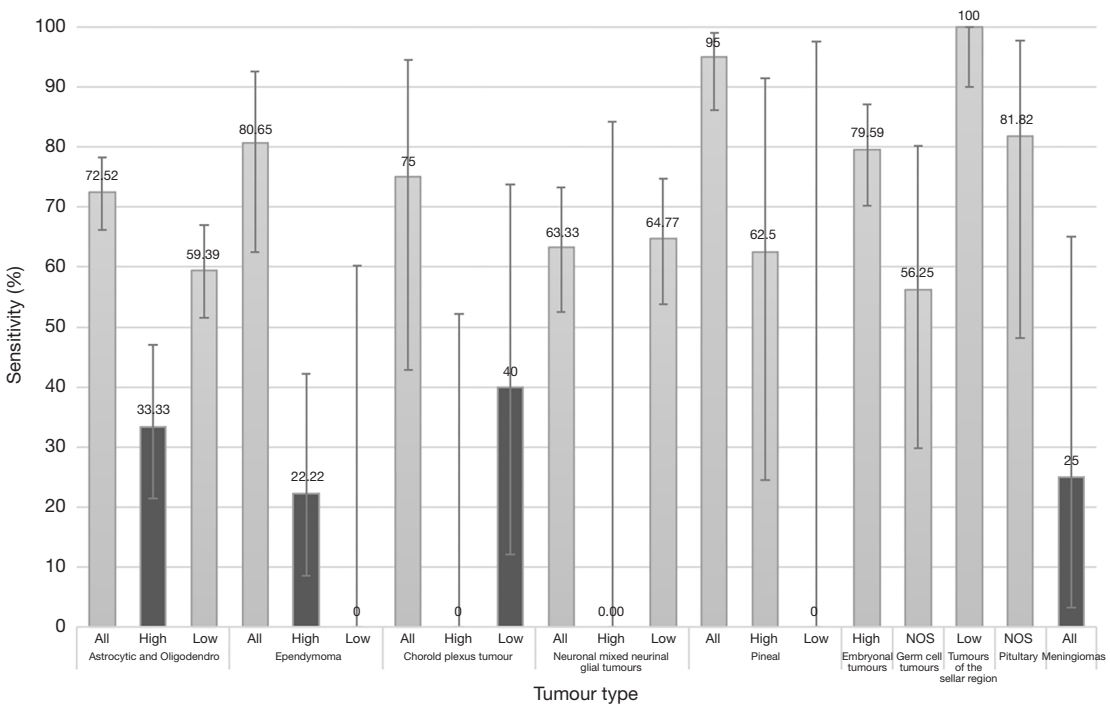

Figure 3 Bar chart of sensitivity with confidence intervals for different tumour types and grades. Dark grey bars reflect tumour types and grades which had a sensitivity of less than 50%.

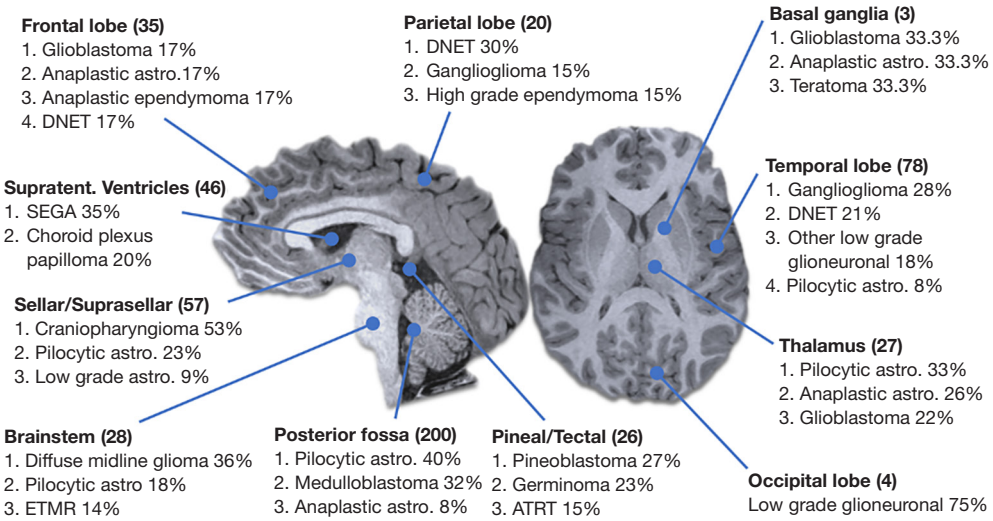

Figure 4 Schematic summarising the regional distribution of the most common tumour types. The number in brackets next to the intracranial location is the true number of tumours in that region. The most common tumour types and their percentage frequency in that region is listed below each.

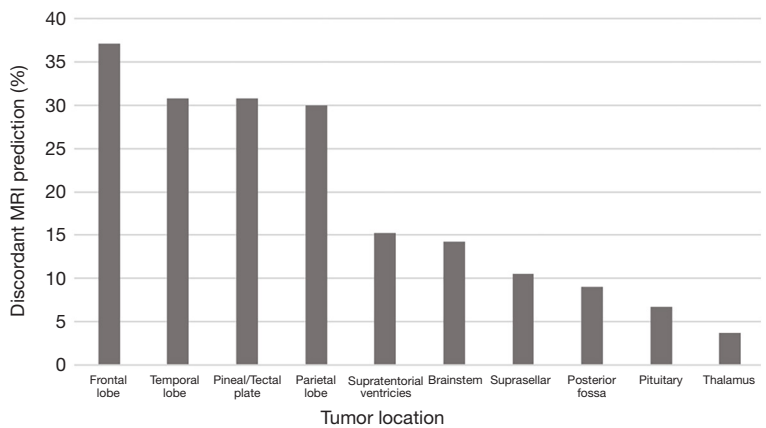

Figure 5 Percentage of discordant tumour type predictions on MRI compared to pathological diagnosis by different intracranial locations. MRI, magnetic resonance imaging.

be of greater benefit in guiding initial therapy. In both the astrocytic glioma and oligodendroglioma group and the neuronal and mixed neuronal-glial tumour group, there was a slightly poorer overall sensitivity for tumour type (72.52% and 63.33%, respectively) which dropped considerably when looking at sensitivity for the high-grade subclass (33.33% and 0%, respectively). *Figure* 7 describes one such instance in which a neuronal tumour was initially diagnosed as a focal cortical dysplasia. The poor sensitivity to grade in these groups could have potential implications on initial treatment, where conservative surveillance may be initially favoured over primary resection in cases of incorrectly presumed low-grade. In the neuronal and mixed neuronal-glial group, one possible reason for a low sensitivity for the high-grade class is that this is a relatively rare subgroup compared to the low-grade class, with only 2.22% of tumours being high-grade. Reduced sensitivity in rare paediatric brain tumour types and subclasses was also noticed in meningiomas, mesenchymal non-meningothelial, melanocytic, and germ cell tumours. This

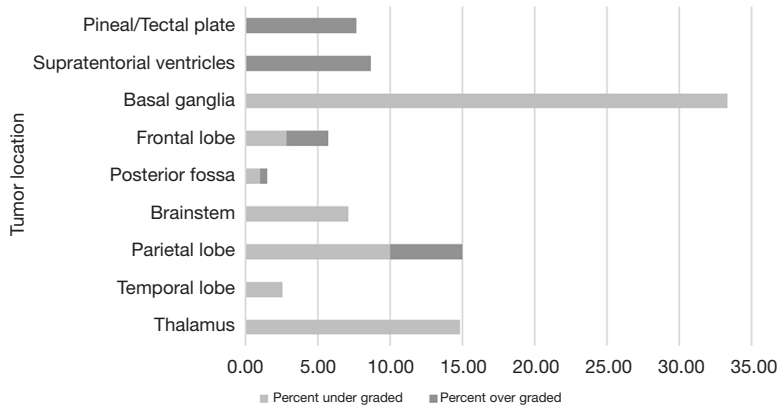

Figure 6 Percentage of incorrect tumour grading irrespective of tumour type by intracranial location.

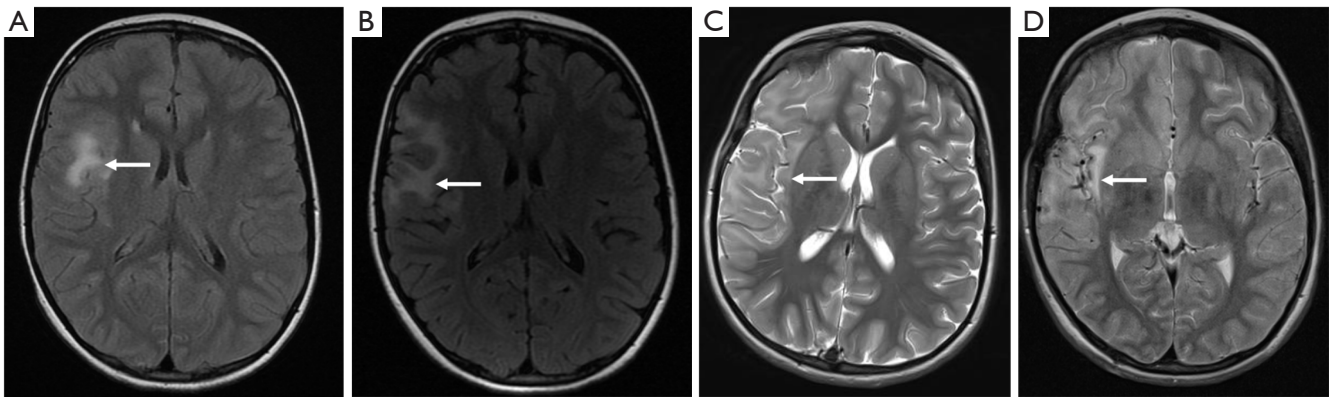

Figure 7 A 7.5-year-old girl presented with a recent history of explosive epilepsy, cognitive decline, hypersalivation, left-sided facial twitching, and frequent vomiting. Magnetic resonance imaging of the brain at presentation shows a region of non-enhancing cortical-subcortical signal abnormality with some local mass effect involving the right opercular region and extending into the mesial temporal region [A: axial T2-weighted fluid attenuated inversion recovery (FLAIR)]. The radiological diagnosis based on these appearances favoured that of a malformation of cortical development. Follow-up axial T2 FLAIR at 2 months (B) showed an increase in the extent of signal abnormality and a greater involvement of the right frontal lobe. Later axial T2-weighted imaging at 6 months follow-up (C) revealed interval extension of the lesion into the orbitofrontal cortex, temporal pole, and parietal lobe. This expanding imaging phenotype favoured that of gliomatosis cerebri though subsequent brain biopsy showed non-specific inflammatory changes and the patient was returned to surveillance imaging. Axial T2-weighted imaging at 12-month follow-up (D) shows peri-Sylvian parenchymal volume loss and less conspicuous signal abnormality. A second cortical brain biopsy at this time showed reactive changes with prominent inflammation and the patient was ultimately diagnosed with Rasmussen's encephalitis. Arrows: leisions.

reduced sensitivity for rarer tumours is understandable as radiologists are less experienced with these tumour groups and there is inevitably less published data on their imaging phenotype. These shortcomings in the rare tumour groups arguably emphasises the need for greater sharing of imaging data across institutions. This would help generate larger case series of a particular tumour type, thereby facilitating pooled data analysis to investigate imaging patterns and discriminators. Pooling of rare tumours would also be beneficial in creating training data sets for machine learning techniques, the accuracy of which is often dependent upon the quality and quantity of input data. Furthermore, in the case of diagnostic image classification, this creates a natural bias to more common diseases where there is readily available data. In the context of paediatric brain tumours, this could potentially mean developing diagnostic

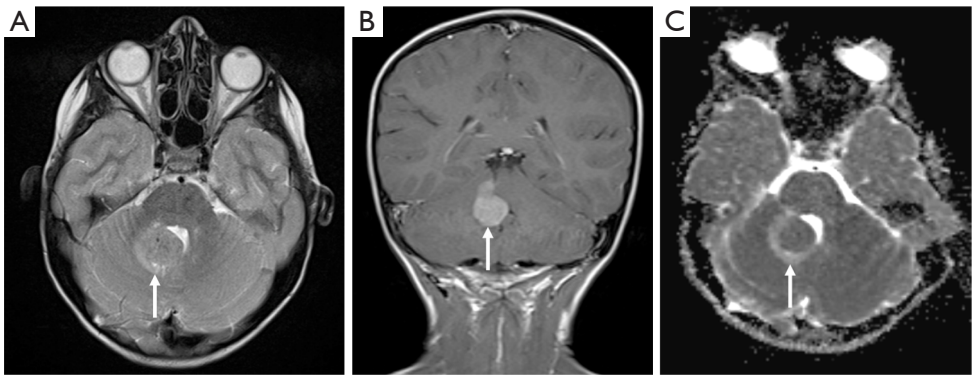

Figure 8 A 6-year-old girl presented with a five-month history of diplopia. Magnetic resonance imaging of the brain at presentation showed a mass in the medial aspect of the right cerebellar hemisphere protruding into the fourth ventricle (A: axial T2-weighted imaging). Coronal post-contrast T1-weighted imaging (B) showed enhancement and extension upwards to the right superior cerebellar peduncle and to the inferior quadrigeminal plate, with the apparent diffusion coefficient map (C) showing restriction. The lesion was initially reported as a probable medulloblastoma but was later confirmed as a meningioma following histopathological examination. Arrows: leisions.

algorithms which target the more common tumours that an experienced radiologist can already diagnose with a fair degree of accuracy, whilst neglecting the rarer tumours which may benefit more from simultaneously run computational diagnostics.

Currently, machine learning techniques in paediatric brain tumours have focused on the classification of posterior fossa tumours as they have a high incidence in the paediatric population (10,11). In our series, only 18 of 200 (9%) posterior fossa paediatric brain tumours had an incorrectly predicted tumour type on the MRI report. In contrast, the number of discordant predictions for tumour type and grade was much higher in the supratentorial lobar regions and in the pineal/tectal plate area (30–37.76%). Furthermore, correct grading of tumour irrespective of tumour type was best in the posterior fossa and worst in the basal ganglia, thalamus, and supratentorial ventricles. It is, again, arguably expected that the radiological interpretation of these tumours will have a higher degree of accuracy in the posterior fossa as it is a relatively common anatomical location for tumours in children—with 80% of the tumours in this region being either pilocytic astrocytomas, medulloblastomas, or anaplastic astrocytomas.

This skew towards a particular anatomical distribution of pathology may also, however, skew predictions towards the more common imaging phenotype. Indeed, *Figure 8* portrays the case of a 6-year-old girl who was initially diagnosed with a medulloblastoma, but which was later histologically confirmed to be a meningioma. In contrast, in the lobar region, pineal/tectal plate area, and deep grey structures, there was a much broader range of tumours with no dominant subtype (*Figure 4*). In particular, erroneous diagnostic predictions of lesions of the deep grey nuclei often occurred when a tumour was misdiagnosed as an inflammatory, metabolic, or vasculitic process (*Figure 9*). In the case of the intraventricular tumours, specifically the rare choroid plexus tumours, prediction of grade was particularly poor, again showing that predictions are often skewed towards the more common pathology (*Figure 10*).

Recent advances in the genetic underpinnings of paediatric CNS tumours have greatly increased the diagnostic power of conventional MRI. Though not always routinely available in clinical practice, the implementation of advanced MRI techniques such as magnetic resonance spectroscopy (MRS), arterial spin labelling (ASL) and amide proton transfer (APT) has been shown to further aid the diagnosis and grading of paediatric brain tumours (19). Indeed, the combination of this richer functional imaging with an ever-increasing understanding of tumour biology may provide a deeper understanding of imaging phenotype and heterogenous variations such as location, outcome, and response to therapy. The combination of advanced imaging with machine learning techniques has also shown to be an effective method of classifying paediatric brain tumours (20).

This study has several limitations. Firstly, this is a retrospective review which only included cases where there was complete preoperative MRI data and histology. Secondly, only patients who underwent surgical intervention were enrolled in this study. This would have excluded many

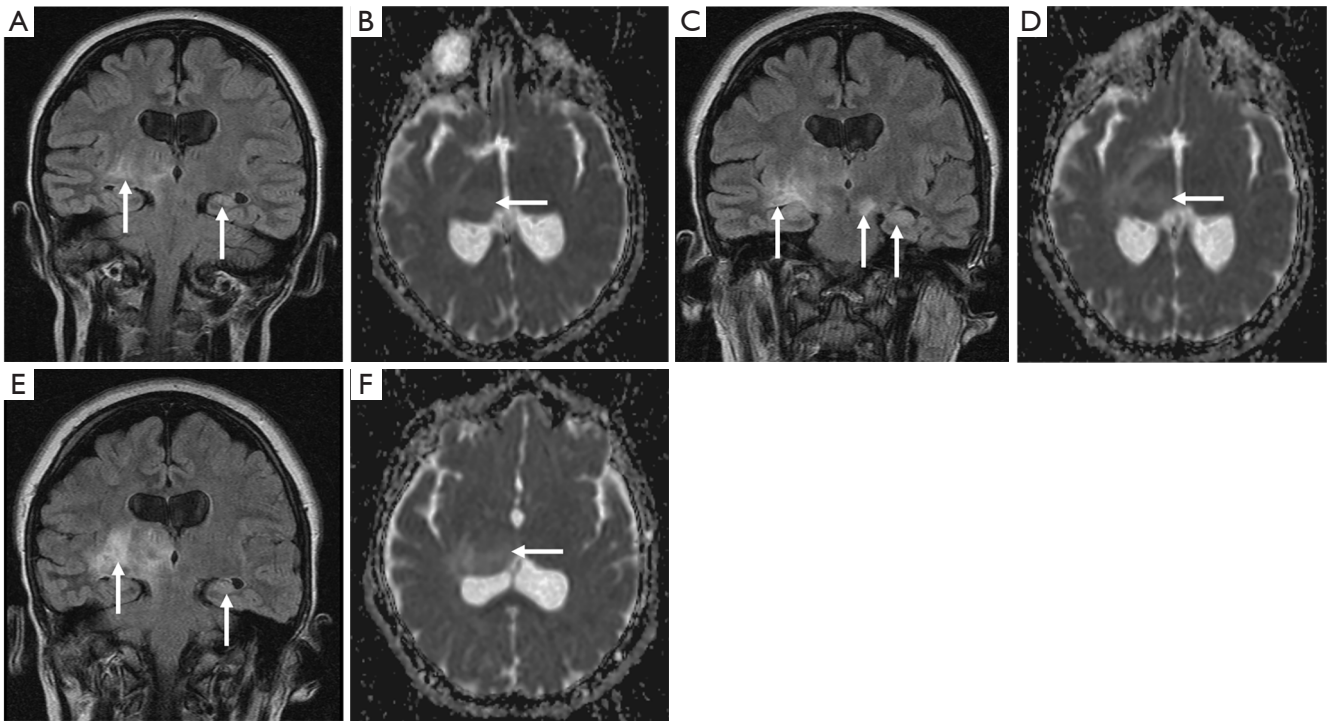

Figure 9 A 16-year-old girl presented with severe headaches. Magnetic resonance imaging of the brain at presentation showed a non-enhancing and ill-defined lesion centred on the right thalamus with further lesions in the contralateral hippocampal formation (A: coronal T2 fluid attenuate inversion recovery and B: apparent diffusion coefficient) and cerebellum (not shown). These appearances were initially reported as acute disseminated encephalomyelitis. Annual surveillance of the lesion at 12 months (C: coronal T2 fluid attenuated inversion recovery and D: apparent diffusion coefficient) showed an increase in size with more oedema. Given this time course and progression, a vasculitic process was favoured. 24-month follow-up (E: coronal T2 fluid attenuated inversion recovery and F: apparent diffusion coefficient) confirmed a further interval increase in the extent of the lesion such that it extended laterally into the medial temporal lobe and involved the right periventricular white matter posteriorly. There was also secondary obstructive hydrocephalus. With these late appearances, the lesion was biopsied, and histopathology confirmed it to be a diffusively infiltrative glioma. Arrows: leisions.

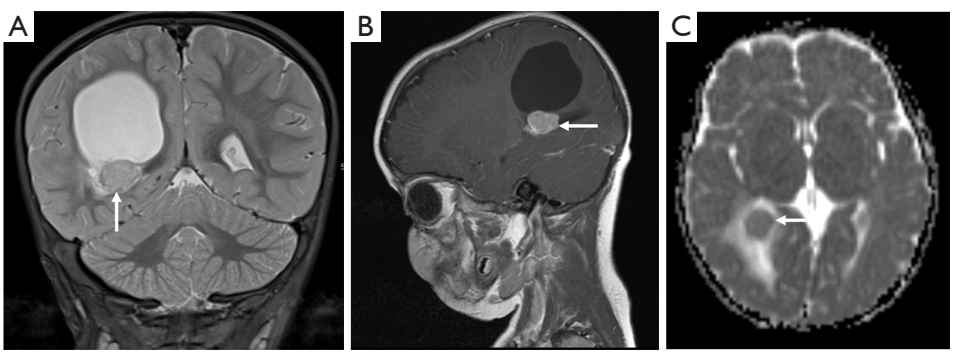

Figure 10 A 17-month-old boy presented with a sudden episode of fainting. Magnetic resonance imaging of the brain at presentation showed a well-defined tumour associated with the right choroid plexus and concomitant cyst-like dilatation of the right lateral ventricle with some periventricular parenchymal changes (A: coronal T2-weighted imaging). Sagittal post-contrast T1-weighted imaging (B) showed a uniformly enhancing lesion with restriction on the apparent diffusion coefficient map (C). With these imaging appearances, the initial radiological diagnosis was that of a meningioma. However, later histopathological examination favoured a choroid plexus carcinoma. Arrows: leisions.

patients with radiologically classified benign lesions or lesions that were unsuitable for surgery. In consequence, the assessment of diagnostic accuracy in certain tumour types (e.g., glioneuronal tumours) is unlikely to be reflective of real-world accuracy and should be interpreted with some caution. Thirdly, assessment of accuracy in rare brain tumour groups is limited due to the relatively small number of cases and the high relative number of true negatives. Finally, this study assessed the accuracy of the written radiological report and not necessarily the overall accuracy of the radiologist's interpretation. Presently, at our institution, there is no mandate to provide a predicted tumour grade, and this is likely to only be reported in cases of high surety. Given the innate difficulty of tumour grading with conventional imaging and the drop in sensitivity for certain tumour types, we anticipate this to be an ideal target for the application of computational methods which can identify and weigh complex variables when little is known about their relationship or underlying distribution, such as those imaging features indicative of tumour grade. Additionally, when a differential list of diagnoses was given, only the first and most probable diagnosis was assessed. This is not reflective of real clinical practice and overlooks the valuable insights that a concise differential list can offer the clinician. Furthermore, in some instances, the written report may have been limited in its detail due to a concomitant discussion in the setting of the multidisciplinary tumour board where a more detailed interpretation is relayed and recorded elsewhere in the patients' clinical record.

Conclusions

This study highlights broad areas of strength and weakness in the current qualitative interpretation of paediatric brain tumour MRIs. Diagnostic accuracy greatly varies between tumour types and intracranial location. Highly accurate classification was noted for posterior fossa tumours whilst accuracy fell in the lobar, basal ganglia, thalamic and intraventricular tumours. As we move into the exciting and promising era of computer-aided diagnosis, these areas of paediatric neuro-oncologic imaging should perhaps be prioritised for the greatest potential improvements in diagnostic accuracy.

Acknowledgments

Funding: None.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/qims-20-1388). KM serves as an unpaid associate editor of *Quantitative Imaging in Medicine and Surgery*. The other authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This clinical audit was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and approved by The Great Ormond Street Hospital for Children NHS Foundation Trust institutional review board prior to commencement, with individual consent waived for this retrospective analysis of imaging data.

References

1. Ostrom QT, Gittleman H, Truitt G, Boscia A, Kruchko C, Barnholtz-Sloan JS. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011-2015. Neuro Oncol 2018;20:iv1-86.
2. Pollack IF. Brain tumors in children. N Engl J Med 1994;331:1500-7.
3. Ostrom QT, de Blank PM, Kruchko C, Petersen CM, Liao P, Finlay JL, Stearns DS, Wolff JE, Wolinsky Y, Letterio JJ, Barnholtz-Sloan JS. Alex's Lemonade Stand Foundation Infant and Childhood Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2007-2011. Neuro Oncol 2015;16 Suppl 10:x1-36.
4. Pollack IF, Agnihotri S, Broniscer A. Childhood brain tumors: current management, biological insights, and

future directions. J Neurosurg Pediatr 2019;23:261-73.
5. Millard NE, De Braganca KC. Medulloblastoma. J Child Neurol 2016;31:1341-53.
6. El-Khouly FE, Veldhuijzen van Zanten SEM, Santa-Maria Lopez V, Hendrikse NH, Kaspers GJL, Loizos G, et al. Diagnostics and treatment of diffuse intrinsic pontine glioma: where do we stand? J Neurooncol 2019;145:177-84.
7. Ramaswamy V, Remke M, Bouffet E, Bailey S, Clifford SC, Doz F, et al. Risk stratification of childhood medulloblastoma in the molecular era: the current consensus. Acta Neuropathol 2016;131:821-31.
8. Medina LS, Kuntz KM, Pomeroy S. Children with headache suspected of having a brain tumor: a cost-effectiveness analysis of diagnostic strategies. Pediatrics 2001;108:255-63.
9. Zhou M, Scott J, Chaudhury B, Hall L, Goldgof D, Yeom KW, Iv M, Ou Y, Kalpathy-Cramer J, Napel S, Gillies R, Gevaert O, Gatenby R. Radiomics in Brain Tumor: Image Assessment, Quantitative Feature Descriptors, and Machine-Learning Approaches. AJNR Am J Neuroradiol 2018;39:208-16.
10. Kapoor N, Lacson R, Khorasani R. Workflow Applications of Artificial Intelligence in Radiology and an Overview of Available Tools. J Am Coll Radiol 2020;17:1363-70.
11. Quon JL, Bala W, Chen LC, Wright J, Kim LH, Han M, et al. Deep Learning for Pediatric Posterior Fossa Tumor Detection and Classification: A Multi-Institutional Study. AJNR Am J Neuroradiol 2020;41:1718-25.
12. Thust S, Micallef C, Okuchi S, Brandner S, Kumar A, Mankad K, Wastling S, Mancini L, Jäger HR, Shankar A. Imaging characteristics of H3 K27M histone-mutant diffuse midline glioma in teenagers and adults. Quant Imaging Med Surg 2021;11:43-56.
13. Tam LT, Yeom KW, Wright JN, Jaju A, Radmanesh A, Han M, et al. MRI-based radiomics for prognosis of pediatric diffuse intrinsic pontine glioma: an international study. Neurooncol Adv 2021;3:vdab042.
14. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar DA, Cohen JF; STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;351:h5527.
15. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol 2016;131:803-20.
16. Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A, Scheithauer BW, Kleihues P. The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol 2007;114:97-109.
17. Hübner JM, Kool M, Pfister SM, Pajtler KW. Epidemiology, molecular classification and WHO grading of ependymoma. J Neurosurg Sci 2018;62:46-50.
18. Merchant TE. Current Clinical Challenges in Childhood Ependymoma: A Focused Review. J Clin Oncol 2017;35:2364-9.
19. Lequin M, Hendrikse J. Advanced MR Imaging in Pediatric Brain Tumors, Clinical Applications. Neuroimaging Clin N Am 2017;27:167-90.
20. Novak J, Zarinabad N, Rose H, Arvanitis T, MacPherson L, Pinkey B, et al. Classification of paediatric brain tumours by diffusion weighted imaging and machine learning. Sci Rep 2021;11:2987.

Cite this article as: Dixon L, Jandu GK, Sidpra J, Mankad K. Diagnostic accuracy of qualitative MRI in 550 paediatric brain tumours: evaluating current practice in the computational era. Quant Imaging Med Surg 2022;12(1):131-143. doi: 10.21037/qims-20-1388

