WHO grade II/III glioma molecular status: Prediction by MRI morphological features, apparent diffusion coefficient and age

Manuscript type

Original research

Summary statement

An algorithm based on standard MRI sequences and age predicted isocitrate dehydrogenase status in lower grade gliomas with comparable accuracy to advanced MRI sequences and computational methods.

Key results

- Apparent diffusion coefficient (ADC) measurements (minimum, mean) and their ratios to normal appearing white matter were reproducible (intraclass correlation coefficient 0.83-0.96) and distinguished three lower grade glioma subtypes: isocitrate dehydrogenase (IDH) wild-type, IDH mutant/1p19q intact, and IDH mutant/1p19q co-deleted ($p<0.001$).
- A negative association (β 0.09, Pseudo R^2 0.34) was identified between age and IDH mutations ($p<0.001$). Glioma location, enhancement characteristics, calcification, and cyst formation were univariable and multivariable predictors of IDH genotype.
- Two predictive models incorporating ADC, age and morphology defined IDH genotype with accuracies of 92% and 91% (AUC 0.94-0.96).

Commented [BE1]: Thank you for your revision. Most of the reviewers' comments have been addressed. There are a few remaining issues, though. Please see the reviewers' comments and the attached document with changes tracked. If you agree with these changes, please accept them. If you disagree, please annotate your responses.

Deleted: clinically relevant

Commented [BE2]: Or did you really mean other modalities, e.g. PET?

Deleted: modalities

Commented [WU3]: Please include numerical data in addition to p values.

Abbreviations

ADC	Apparent diffusion coefficient
AUC	Area under the curve
ICC	Intraclass correlation coefficient
IDH	Isocitrate dehydrogenase
NAWM	Normal-appearing white matter
OR	Odds ratio
ROI	Region of interest
1p19q	Short arm of chromosome 1 and long arm of chromosome 19

Abstract

Background: A readily implemented MRI biomarker for glioma genotyping is currently lacking.

Purpose: To evaluate the accuracy of routine clinical MRI parameters to predict isocitrate dehydrogenase (IDH) status in patients with glioma.

Materials and Methods: In this retrospective study (07/08-02/19), untreated World Health Organization (WHO) grade II/III gliomas were analyzed by 3 neuroradiologist readers blinded to tissue results. ADC minimum (ADC_{min}) and mean (ADC_{mean}) regions of interest were defined in tumor and normal appearing white matter (ADC_{NAWM}). A visual rating of anatomical (T1, T1+contrast, T2, FLAIR) features was performed. Interobserver comparison (intraclass correlation coefficient (ICC), Cohen's kappa (κ)) was followed by non-parametric (Kruskal-Wallis ANOVA) and effect size (η^2) testing of associations between ADC metrics and glioma genotypes. Descriptors with sufficient concordance (ICC>0.8, κ>0.6) underwent univariable analysis. Predictive variables ($p<0.05$) were entered into a multivariable logistic regression and tested in a new glioma sample.

Results: The study included 290 patients (median 40 (IQR 19.25) years, 169 male) with 82 IDH wild-type, 107 IDH mutant/1p19q intact and 101 IDH mutant/1p19q co-deleted gliomas. ADC_{min}, ADC_{mean}, ADC_{min}:ADC_{NAWM} and ADC_{mean}:ADC_{NAWM} were reproducible (ICC 0.83-0.96), enabling the distinction of glioma subtypes ($p<0.001$, η^2 0.28-0.38). A negative association (Pseudo R^2 0.34) was identified between age and IDH mutations ($p<0.001$). Glioma location, enhancement characteristics, calcification and cyst formation were univariable (p<0.001-0.045) and multivariable predictors of genotype. Two predictive models A) mandating calcification result and B) incorporating ADC_{mean}:ADC_{NAWM} ratio, age and morphology classified tumor type with accuracies of 91.7% (266/290) and 90.9% (264/290) (area under the curve 0.94-0.96). In the test sample of 49 gliomas (9 IDH wild-type, 21 IDH mutant/1p19q intact and 19 IDH mutant/1p19q co-deleted), the classification accuracy was 81.6% (40/49) for model A and 85.7% (42/49) for model B.

Conclusion: Two proposed algorithms predicted isocitrate dehydrogenase status in WHO grade II/III gliomas based on standard clinical MRI sequences alone.

Deleted: glioma genotypes, specifically

Deleted: .

Commented [BE4]: Please comment on correction for multiple testing.

Commented [BE5]: This is confusing (how do you get a range from <0.001-0.045?) – please rephrase

Commented [BE6]: I did not quite understand your parentheses here – please check whether this is what you meant.

Deleted: (

Deleted: (

Deleted:)

Deleted: lower grade gliomas

Introduction

A subgroup of lower grade gliomas is characterized by genetic overlap with primary glioblastoma and exhibits similarly rapid disease progression **(1,2)**. Such malignant neoplasms are indistinguishable from indolent astrocytomas by assessing proliferative indices and cell morphology **(3)**.

Mutations in the isocitrate dehydrogenase gene (IDH^{mut}), most commonly IDH1 (R132H), define most slow growing gliomas (>70%) within the World Health Organization (WHO) histological grades II/III **(4)**. IDH mutations are absent (IDH wild-type, IDH^{wt}) in lower grade tumors of the primary glioblastoma spectrum, which further differ by genetic hallmarks of combined chromosome 7 gain and chromosome 10 loss, epidermal growth factor receptor (EGFR) amplification and telomerase reverse transcriptase (TERT) promoter mutations **(2)**. Amongst IDH^{mut} gliomas, synchronous deletion of the short arm of chromosome 1 and long arm of chromosome 19 ($IDH^{mut}/1p19q^{del}$) constitutes a specific feature of oligodendrogliomas, whereas IDH^{mut} astrocytomas are mostly 1p19q intact ($IDH^{mut}/1p19q^{int}$) **(5)**. This genetic grouping serves an important clinical purpose of stratifying tumors with differential susceptibility to adjuvant treatment, for example $IDH^{mut}/1p19q^{del}$ gliomas have greater sensitivity to alkylating chemotherapy **(6)**.

Glioblastoma outcomes are improved with gross total gadolinium-based contrast agent enhancing lesion resection **(7)** and potentially beyond this for T2/FLAIR component removal **(8)**. The similarity between the biology of 'low grade' IDH^{wt} glioma and glioblastoma makes it crucial to identify glioblastoma early and separate it from the more favorable IDH^{mut} entities.

Physiological imaging techniques and computational algorithms have shown potential for IDH status prediction**(9)**, but a lack of transferable thresholds and, in some instances, technical complexity, impede their clinical translation for this purpose.

Diffusion-weighted MRI (DWI) is routinely used in cancer imaging. It functions on the assumption that free water motion in tissues diminishes with growing tumor cellularity **(10)**. Three-direction DWI is widely performed and integrated into clinical glioma imaging protocols, with quantitative results available immediately during reporting **(11)**. Diffusion-based methods can support grading and have shown capability for IDH typing **(12–14)**, including for gliomas lacking contrast enhancement **(15)**.

Prior studies suggest that lesion properties such as location, internal architecture and enhancement patterns differ between glioma genetic subtypes **(16)**. Additionally, consideration of patient age may aid diagnosis, as it has been shown that IDH^{wt} gliomas more commonly arise in older patients **(17)**. This study investigated i) the accuracy and reproducibility of four ADC parameters to distinguish the WHO 2016 glioma subtypes, ii)

the contribution of age and anatomical lesion features for the prediction of glioma IDH status, using routinely available imaging information without machine learning.

Commented [BE7]: Please rephrase as a clear **purpose** statement (not just stating what you investigated, but what you set out to do) and adapt it to precisely reflect the purpose statement in your abstract.

Materials and Methods

Ethics review board approval was obtained and written informed consent waived for this retrospective study.

Patients

All patients consecutively diagnosed with WHO grade II/III glioma at our national brain tumour referral institution between July 2008 to January 2018 were eligible for the study. This reflects the time period during which molecular genetic testing has been available. Inclusion criteria were a proven histological diagnosis of WHO grade II/III glioma, available IDH and 1p19q genetic test results and MRI prior to treatment.

Exclusion criteria included previous glioma treatment, a tumor other than WHO grade II/III glioma, missing, inconclusive or ambiguous molecular results (e.g. IDHwt/1p19qdel), prolonged (≥ 1 year) interval from MRI to surgery, or missing images. In 44 of the included 290 patients, ADC_{mean} values were reported in a previous study, which compared volumetric and regional ADC_{mean} measurements (15). In the current study, multiple region derived ADC metrics and new morphological descriptors were analysed (by different observers) in these patients. The results were validated using a previously unseen test sample (n=49), newly referred since the main study (January 2018 – February 2019).

Deleted: not

Deleted: ,

Deleted: .

MRI Acquisition

All MRI examinations included T2-weighted, T2-FLAIR, and T1-weighted sequence pre and post administration of a gadolinium-based contrast agent as well as DWI sequences (n=211 at 1.5 T, n=79 at 3T). Our institution is a quaternary center, therefore the MRI examinations originated from multiple sites and scanners (57 GE, 206 Siemens, 26 Phillips and 1 Toshiba). No scanner model contributed more than 14 % gliomas of one molecular subtype. The range of MRI parameters is provided in **Supplementary Table S1.**

Deleted: and

Histopathology

All tissue samples were fixed as paraffin blocks and analyzed at our institution's neuropathology department, using the latest methodology consistent with the WHO 2016 guidance on histopathology and immunohistochemistry (18). For IDH R132H negative tumors, multiple gene Sanger sequencing was performed to identify alternative IDH

mutations. A quantitative polymerase chain reaction-based copy number assay was employed to determine 1p/19q status.

ADC quantification

The ADC measurements were blinded to tissue diagnosis (reference standard) and age. Three independent observers (2 board certified neuroradiologists, 1 resident (MK 6 years, WM 3 years experience, SO in training)) placed 3 different 30-40 mm^2 regions of interest (ROIs) into the visually perceived lowest ADC portions of each glioma. From these, the mean value of the numerically lowest ADC ROI measurement was designated ADC_{min} as in (12). Subsequently, one large ROI (ADC_{mean}) was drawn to cover the largest axial tumor cross-section, excluding tumor margins, necrosis, macroscopic hemorrhage and calcifications. A comparative ROI was sited in normal appearing centrum semiovale white matter (ADC_{NAWM}), following (15), amounting to 5 ROI measurements per patient. Multifocal tumors were measured as one glioma. Observer 1 analyzed all (n=290) gliomas, observer 2 re-analyzed n=75 gliomas and observer 3 re-analyzed the remaining n=215 gliomas, totaling 2900 ADC measurements. From these, ADC_{min}:ADC_{NAWM} and ADC_{mean}:ADC_{NAWM} ratios were calculated, resulting in 4 ADC parameters (ADC_{min}, ADC_{min}:ADC_{NAWM}, ADC_{mean} and ADC_{mean}:ADC_{NAWM}) per patient. For the test sample (n=49), one researcher newly trained in the ADC method (AAB, board certified, 3 years experience) obtained all ADC values blinded as above. **Figure 1** shows examples of the region placements.

Morphology assessment

Three observers (2 board certified neuroradiologists, 1 resident (ST 8 years experience, AAB, SO)) independently reviewed 290 MRI datasets and were blinded to diagnosis. Morphology readings were performed at a separate (≥2 weeks) time point to ADC measurements. Feature categories were based on previous publications (17,19), adapted to be practicable within clinical time constraints. Tumor location was specified by epicenter. Multifocality was marked positive, if >1 discrete tumor deposit was visible, or if ≥3 lobes were involved. The non-enhancing tumor margin was described using a visual rating scale from 1=able to clearly draw around the lesion on T_2- weighted images to 4=indistinct margin on T_2-weighted and FLAIR images). Hemorrhage and calcification were assessed on T1-weighted imaging, together with CT, T2* sequences and susceptibility weighted imaging as available. The option 'uncertain' was added for these categories to allow for variability in the diagnostic sequences. The single largest tumor diameter was measured on T_2-weighted images according to (20). Contrast uptake was categorized into

Commented [BE8]: Please check

Deleted: +

Deleted:

Deleted:

Deleted:

Deleted:

non-enhancing, patchy-solid or rim-enhancing. Rim enhancement surrounding central necrosis was distinguished from cysts, defined as exhibiting fluid signal isointense to cerebrospinal fluid, with absent or minimal rim enhancement. T_2/FLAIR mismatch was specified according to (21). Examples of different glioma morphologies are shown in **Figure 2.**

Statistical analysis

Statistical testing was performed in SPSS 25 (IBM, New York) and STATA 15 (Statacorp, College Station, TX). The concordance of ADC measurements was examined by intraclass correlation coefficient (ICC) analysis, using a 2-way random effects model. For each ADC region of interest, the mean of the observers' measurements was adopted as the final value. Cohen's kappa testing was used to evaluate the observer agreement for morphological categories, with the raters' majority opinion designated the final value. If 3 opinions differed, this was resolved in consensus.

The relation between ADC and glioma subtypes was analyzed using non-parametric testing (Kruskal-Wallis ANOVA), including Dunn's pairwise comparisons with Bonferroni correction. The strength of the association between glioma subtype and ADC metrics was probed using Eta^2 (η^2).

Univariable logistic regression was applied to test if ADC metrics, age or morphological criteria could predict IDH^{wt} status. Visual categories with $\kappa \geq 0.6$ were subjected to univariable analysis. If significant ($p<0.05$) in univariable analysis, features with substantial agreement were tested in a multivariable regression. A backwards elimination was performed to discard features that did not contribute significantly to the prediction. To assess model discrimination, we used a receiver operating characteristic (ROC) analysis.

Results

Patient demographics

On commencing the study, 515 patients were eligible for inclusion. Following removal of duplicates (n=42), 183 patients were excluded due to previous glioma treatment (n=60), tumor other than WHO grade II/III glioma (n=43 and n=1 cord tumor), ambiguous molecular result (n=29), no preoperative DWI (n=24 and n=15 ADC map not computable), missing histopathology report (n=2), prolonged (≥ 1 year) interval from MRI to surgery (n=3), or missing images (n=1). A total of 290 patients (median 40 (interquartile range (IQR) 19.25) years, 169 male) were included in the analysis. An overview of the case selection process is shown in **Figure 3.** An overview of patient demographics and

Deleted: not

Deleted: OP

molecular groups is shown in **Table 1**. The relation between glioma IDH status and age was found to be non-linear, with an exponential rise in the likelihood of IDH^{wt} status towards older age.

ADC quantification for glioma molecular subtyping

The interobserver reproducibility was good to excellent (ICC 0.83-0.96) for all ADC parameters. Consistency and absolute agreement were identical, indicating no systematic difference between the raters. Detailed ICC test results are shown in **Supplementary Table S2**. Each of the ADC parameters enabled $IDH^{mut}/1p19q^{del,}$ $IDH^{mut}/1p19q^{int}$ and IDH^{wt} glioma discrimination ($p<0.01$) as presented in **Supplementary Table S3**.

Eta^2 (η^2) testing revealed a strong association between ADC values and glioma subtype for non-gadolinium-enhancing and solidly enhancing tumors, but not for rim-enhancing masses, see **Supplementary Table S4**. For ADC_{mean}:ADC_{NAWM} ratio, the optimal threshold for IDH typing across all 290 gliomas was 1.8 (sensitivity 87.3% and specificity 59.6%. For non-enhancing gliomas, an ADC_{mean}:ADC_{NAWM} ratio threshold of 1.8 yielded a sensitivity of 84.8% and specificity of 66.4% for IDH^{wt} identification, compared to a sensitivity of 97% and specificity of 54.3% for a higher ADC_{mean}:ADC_{NAWM} ratio threshold of 1.9. Molecular group ADC differences are shown in **Figure 4.**

Morphology assessment

Observer comparison

For tumor location, the agreement between the 3 observers was good ($\kappa = 0.81 – 0.89$, $p<0.001$). Measurement of the single longest tumor diameter (<6 cm or ≥6 cm) demonstrated good agreement ($\kappa = 0.80 – 0.82$, $p=<0.001$). Defining calcification as present reached substantial agreement ($\kappa = 0.67 – 0.74$, $p=<0.001$) with uncertain results (e.g. missing sequences), excluded. In 63.4% (184/290) cases, one of 3 raters marked calcification as uncertain. In 11.7% (34/290) more than one rater specified calcification status as uncertain. The raters' opinion on tumor cysts showed substantial agreement ($\kappa = 0.66 – 0.70$, $p=<0.001$). The categorization of enhancement patterns yielded substantial agreement (weighted $\kappa = 0.69 – 0.77$, $p=<0.001$).

Moderate interobserver agreement was found for non-enhancing tumor margin (weighted $\kappa = 0.45 – 0.61$, $p=<0.001$) and for the T2-FLAIR mismatch sign ($\kappa = 0.44 – 0.62$, $p<0.001$). Fair agreement was observed for multifocality ($\kappa = 0.20 – 0.46$, $p<0.001$) and hemorrhage ($\kappa = 0.29 – 0.51$, $p<0.001$). Please see **Supplementary Table S5** for details on the kappa test results.

Deleted: a

Univariable analysis

The univariable logistic regression results are displayed in **Table 2**. Several features were significant predictors, including all four ADC metrics (negative association), age (negative association) and several morphological categories (enhancement pattern, non-enhancing margin, calcification and cysts). Locations were grouped according to whether <1/3, 1/3-2/3 or >2/3 of tumors represented IDH^{wt} gliomas to reduce the number of variables for statistical analysis. The presence of calcification was positively associated (odds ratio 2.18, $p<0.001$) with $1p19q^{del}$ status in IDH^{mut} gliomas (not tabulated). Tumor diameter and T2/FLAIR mismatch sign demonstrated no association with IDH status.

Commented [BE9]: Please state in the methods section

Multivariable logistic regression model

The multivariable regression results are listed in **Table 2 and Figure 5**. The best performing model (model A) consisted of ADC_{mean}:ADC_{NAWM} ratio, age in years + age^2 (joint term), enhancement pattern, tumor location category (3 groups: frontal/insula region, thalamus/brainstem or elsewhere) and absence of calcification. Based on a likelihood cut-off value of 0.5 (50%), model A correctly classified 91.7% (266/290) of the gliomas (area under the curve (AUC) 0.96). In developing this model, 11.7% (34/290) patients were excluded by the statistics software due to uncertain calcification status as per the raters' majority result.

Deleted: ,

An alternative model (model B), derived by the same backwards elimination method except for not considering calcification status, performed nearly as well achieving a correct classification of IDH status in 90.9% (264/290) of gliomas (AUC 0.94, sensitivity 75.9% and specificity 96.6%). Model B consisted of ADC_{mean}:ADC_{NAWM} ratio, age in years + age^2 (joint term), enhancement pattern, tumor location category and absence of tumor cyst(s). The diagnostic contribution from age and tumor morphology is highlighted in **Figure 6**.

The numerical results from the study sample were transcribed into a Microsoft Excel for Mac (Version 14.5.2) formula (please see Table 2 footnote) to calculate the IDH^{wt} status probability for individual glioma patients in the subsequent test sample. .

Commented [BE10]: Please state in the methods section

Test sample

In the sample of patients with newly diagnosed glioma (n=49 (9 IDH^{wt}, 21 IDH^{mut}/$1p19q^{int}$, 19 IDH^{mut}/$1p19q^{del}$)) the single blinded rater (AAB) replicated the methodology of the main study. In cases of uncertainty regarding calcification (n=5), 'no calcification' was specified to permit results calculation.

Model A correctly classified IDH mutational status in 81.6% (40/49) gliomas (sensitivity 88.9% and specificity 80%). Model B predicted IDH status in 85.7% (42/49) gliomas with a lower sensitivity of 66.7% but greater specificity of 90%.

Of the IDH^{mut} gliomas, which were erroneously diagnosed as IDH^{wt} (n=8 using model A and n=4 using model B), 75% (6/8 and 3/4 respectively) were $IDH^{mut}/1p19q^{del}$ with an average ADC_{mean}:ADC_{NAWM} ratio of 1.43 (1.21-1.76). One $IDH^{mut}/1p19q^{int}$ astrocytoma with an ADC_{mean}:ADC_{NAWM} ratio of 1.84 was misclassified by both models in an elderly patient (81 years), and one anaplastic $IDH^{mut}/1p19q^{int}$ astrocytoma with an ADC_{mean}:ADC_{NAWM} ratio of 1.46 was misclassified by model A alone. The IDH^{wt} gliomas erroneously predicted as IDH^{mut} tumors (n=1/9 model A, n=3/9 model B) had ADC_{mean}:ADC_{NAWM} ratio values of 1.73-1.87. On subsequent review, all misclassified IDH^{wt} tumors exhibited a gliomatosis growth pattern with diffusely T_2 hyperintense infiltration of ≥ 3 lobes. In one IDH^{wt} glioma, the comparison ADC_{NAWM} ROI was sited in ring artifact.

Discussion

Advanced imaging and computational methods for glioma molecular characterization are not yet widely integrated into clinical practice, therefore an unmet need exists for non-invasive tumor genotyping. This study has demonstrated that region derived ADC metrics and several visual descriptors are reproducible and valuable for molecular characterisation of lower grade gliomas. By combining ADC values, age and morphology, we developed a probability calculator to predict lower grade glioma IDH genotype using information derived from MRI sequences available in standard care (11).

Volumetric (13,15) and region derived minimum (12) and mean (15) ADC measurements have previously been used to estimate WHO grade II/III glioma IDH status. Our study confirms excellent interobserver agreement for ROI measurements, consistent with the reproducibility of ADC values described in other cancer research (22). Significant differences between $IDH^{mut}/1p19q^{del}$, $IDH^{mut}/1p19q^{int}$ and IDH^{wt} glioma subtypes were apparent for each ADC parameter, with the ADC_{mean}:ADC_{NAWM} ratio performing best for IDH status prediction. While ADC values are independent of hardware and field strength under fixed parameters (23), using a ratio offers the further advantage of being vendor neutral. Drawing one maximum size round ADC_{mean} ROI in the largest tumor cross-section is considered feasible on most clinical workstations. Good reproducibility was shown previously for 2 observers using ADC_{mean}:ADC_{NAWM} ratio regions of interest, representative of entire lesion volumetric measurements (15). In the current analysis, 3 new observers used the technique in the study sample, and 1 observer in the test sample, totalling 6

Commented [WU11]: At the beginning, please state your major finding and why it is important. Provide a succinct 1 paragraph of your results in the 1st paragraph. Instead of providing general statements that something was generally better or superior, provide specific metrics that indicate your results (when possible) Please limit this to 1-2 paragraphs.

different observers between the studies. It is hypothesized that most lower grade gliomas are sufficiently homogenous to make such ROI measurements reliable.

When testing ADC_{mean}:ADC_{NAWM} ratio alone for IDH typing, we confirmed a threshold in the region of 1.8 (15), applicable to solid tumors with or without contrast enhancement. ADC values appear unreliable for IDH typing in rim-enhancing, necrotic gliomas even when measured in macroscopically solid components, which mirrors a previous study in WHO grade IV glioblastoma (24).

The stand-alone accuracy of ADC_{mean}:ADC_{NAWM} ratio for predicting WHO II/III grade glioma IDH status in our research (AUC 0.83 across all tumor morphologies) exceeds that of published approaches using multi-shell diffusion (NODDI, AUC max 0.76) (25) and diffusion kurtosis (AUC max 0.72 (26) and 0.788 (27)). The ADC_{mean}:ADC_{NAWM} ratio was a highly significant predictor ($p<0.001$) in both multivariable models, indicating a strong inverse relationship between ADC and the likelihood of IDH^{wt} status.

Age at disease onset is known to improve multivariable prediction of IDH status (16,17). Remarkably, in a recent study by Zhou et al., age offered a higher predictive value than machine learned texture features (17). In this analysis, age was negatively associated with IDH genotype in univariable regression, and in the multivariable regression the consideration of age squared (joint term $p <0.001$) improved the final model.

Because the qualitative description of glioma features is subjective, we limited the statistical modelling to visual categories with substantial agreement, such as tumor location. Frontal glioma location has repeatedly been associated with IDH^{mut} status (28,29). Goze et al. found 100% of insula centred low grade gliomas to be IDH mutant (30). In our study, both locations were similarly associated with a greater likelihood of IDH^{mut} status, which is also consistent with a report by Xiong et al. (31). Conversely, we confirmed that thalamic or brainstem location is predominantly a feature of IDH^{wt}, which may variably be associated with malignant glioma mutations such as H3 K27M (32).

The presence versus absence of solid enhancement was not consistently associated with IDH status (multivariable p=0.41-0.58) in our study. However, 'glioblastoma morphology' featuring rim enhancement was a predictor of IDH^{wt} status. We did not test percentage enhancement, which in a study by Delfanti et al. failed to predict IDH type [32].

The absence of calcification strongly correlated with IDH^{wt} status and negatively with $1p19q^{del}$ in univariable analysis. In a study by Kanazawa et al., both calcification and cysts were significantly related to $IDH^{mut}/1p19q^{del}$ (19). We hypothesize that with many patients undergoing CT at diagnosis, consistent availability of this and/or SWI imaging could further increase observer certainty and concordance. In keeping with our observations (model B), absence of cysts has been proposed previously as an IDH^{wt} glioma feature (33).

When combining all variables, which were significant in univariable analysis, the best multivariable model using a stepwise backwards elimination strategy was model A (including *no calcification*). However, several issues were noted with generating model A as follows: Firstly, 34 gliomas with uncertain calcification status were automatically excluded by the statistical software and thus could not contribute to the modelling. Secondly, as a consequence *no cystic* becomes non-significant and was therefore eliminated from model A. In clinical reality, some patients may lack imaging that permits a definitive diagnosis of calcification status (i.e. no CT and/or T2* or SWI sequences).

For this reason, we tested the alternative model B, which was derived by the same backwards elimination strategy, except for not requiring knowledge of calcification status. Interestingly, in model B the variable *no cystic* was significant, meaning some of the 34 gliomas, which were no longer dropped in the analysis due to missing data, contributed to this significance. Model B does appear valuable, as evidenced by its greater accuracy in testing. In the test sample, the accuracy of model B (86%) outperformed that of model A (82%) for IDH status prediction. For this reason, both models are presented. Because of a tendency for multivariable model over-fitting, the performance of the algorithm in terms of sensitivity, specificity and accuracy is best represented by the test sample results.

In our study sample, no statistical association was identified between multifocality and IDH status, which was recently proposed as a feature predictive of IDH^{wt} in WHO II glioma (34). Our results support that the T2-FLAIR mismatch sign is a specific feature of $IDH^{mut}/1p19q^{int}$ status. However, the interobserver agreement was moderate ($\kappa = 0.44 – 0.62$), closer to the lower 95% CI bound ($\kappa = 0.53$) of Patel et al.'s original publication (21). The T2-FLAIR mismatch sign did not directly predict IDH status, given that all molecular glioma subtypes can lack this feature. For non-enhancing margin definition the agreement was moderate, meaning that although IDH^{wt} gliomas are less well demarcated (17,35), subjectivity and overlap with $IDH^{mut}/1p19q^{del}$ indistinct margins (36) limit the reproducibility of this feature.

Our study had some limitations, including retrospective study design and the lack of a definitive calcification result in a proportion of patients. Both models may have a misclassification risk for low diffusivity $IDH^{mut}/1p19q^{del}$ oligodendrogliomas and for IDH^{wt} tumors exhibiting a T2 and ADC hyperintense gliomatosis growth pattern. We have not tested the models on WHO grade IV gliomas. Our analysis contains data from multiple institutions, which could be perceived as a limitation but in fact underscores the generalizability of results. Differences in MRI systems did not contribute to particular molecular subtypes or imaging features; therefore, a systematic error appears unlikely.

Deleted: *no_calcification*

Formatted: Font: Italic

Deleted: _

Deleted: ,

Formatted: Font: Italic

Deleted: _

The regression analysis was derived from features with strong observer agreement. Furthermore, the good performance of the individual rater suggests that a predictive system as presented here may be valuable for clinical use.

In conclusion, the combination of ADC_{mean}:ADC_{NAWM} ratio, tumor morphology and age predicts the presence of IDH^{wt} glioma on clinical MRI with sufficiently high accuracy to be considered for probability estimates in practice. Such an algorithm could be implemented with ease, almost irrespective of the clinical setting.

Deleted: hospital

References

1. Eckel-Passow JE, Lachance DH, Molinaro AM, et al. Glioma Groups Based on 1p/19q, IDH, and TERT Promoter Mutations in Tumors. N Engl J Med. 2015;372(26):2499–2508.
2. Stichel D, Ebrahimi A, Reuss D, et al. Distribution of EGFR amplification, combined chromosome 7 gain and chromosome 10 loss, and TERT promoter mutation in brain tumors and their potential for the reclassification of IDHwt astrocytoma to glioblastoma. Acta Neuropathol (Berl). 2018;136(5):793–803.
3. Louis DN, Perry A, Burger P, et al. International Society Of Neuropathology--Haarlem consensus guidelines for nervous system tumor classification and grading. Brain Pathol Zurich Switz. 2014;24(5):429–435.
4. Yan H, Parsons DW, Jin G, et al. IDH1 and IDH2 mutations in gliomas. N Engl J Med. 2009;360(8):765–773.
5. Cancer Genome Atlas Research Network, Brat DJ, Verhaak RGW, et al. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. N Engl J Med. 2015;372(26):2481–2498.
6. van den Bent MJ, Brandes AA, Taphoorn MJB, et al. Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed anaplastic oligodendroglioma: long-term follow-up of EORTC brain tumor group study 26951. J Clin Oncol Off J Am Soc Clin Oncol. 2013;31(3):344–350.
7. Kreth F-W, Thon N, Simon M, et al. Gross total but not incomplete resection of glioblastoma prolongs survival in the era of radiochemotherapy. Ann Oncol. 2013;24(12):3117–3123.
8. Li YM, Suki D, Hess K, Sawaya R. The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection? J Neurosurg. 2016;124(4):977–988.

9. Andronesi OC, Rapalino O, Gerstner E, et al. Detection of oncogenic IDH1 mutations using magnetic resonance spectroscopy of 2-hydroxyglutarate. J Clin Invest. 2013;123(9):3659–3663.

10. Patterson DM, Padhani AR, Collins DJ. Technology insight: water diffusion MRI--a potential new biomarker of response to cancer therapy. Nat Clin Pract Oncol. 2008;5(4):220–233.

11. Thust SC, Heiland S, Falini A, et al. Glioma imaging in Europe: A survey of 220 centres and recommendations for best clinical practice. Eur Radiol. 2018;

12. Xing Z, Yang X, She D, Lin Y, Zhang Y, Cao D. Noninvasive Assessment of IDH Mutational Status in World Health Organization Grade II and III Astrocytomas Using DWI and DSC-PWI Combined with Conventional MR Imaging. AJNR Am J Neuroradiol. 2017;38(6):1138–1144.

13. Leu K, Ott GA, Lai A, et al. Perfusion and diffusion MRI signatures in histologic and genetic subtypes of WHO grade II-III diffuse gliomas. J Neurooncol. 2017;

14. Hempel J-M, Schittenhelm J, Brendle C, et al. Histogram analysis of diffusion kurtosis imaging estimates for in vivo assessment of 2016 WHO glioma grades: A cross-sectional observational study. Eur J Radiol. 2017;95:202–211.

15. Thust SC, Hassanein S, Bisdas S, et al. Apparent diffusion coefficient for molecular subtyping of non-gadolinium-enhancing WHO grade II/III glioma: volumetric segmentation versus two-dimensional region of interest analysis. Eur Radiol. 2018;28(9):3779–3788.

16. Juratli TA, Tummala SS, Riedl A, et al. Radiographic assessment of contrast enhancement and T2/FLAIR mismatch sign in lower grade gliomas: correlation with molecular groups. J Neurooncol. 2019;141(2):327–335.

17. Zhou H, Vallières M, Bai HX, et al. MRI features predict survival and molecular markers in diffuse lower-grade gliomas. Neuro-Oncol. 2017;19(6):862–870.

18. Reuss DE, Sahm F, Schrimpf D, et al. ATRX and IDH1-R132H immunohistochemistry with subsequent copy number analysis and IDH sequencing as a basis for an “integrated” diagnostic approach for adult astrocytoma, oligodendroglioma and glioblastoma. Acta Neuropathol (Berl). 2015;129(1):133–146.

19. Kanazawa T, Fujiwara H, Takahashi H, et al. Imaging scoring systems for preoperative molecular diagnoses of lower-grade gliomas. Neurosurg Rev. 2018;

20. Pignatti F, van den Bent M, Curran D, et al. Prognostic factors for survival in adult patients with cerebral low-grade glioma. J Clin Oncol Off J Am Soc Clin Oncol. 2002;20(8):2076–2084.

21. Patel SH, Poisson LM, Brat DJ, et al. T2-FLAIR Mismatch, an Imaging Biomarker for IDH and 1p/19q Status in Lower-grade Gliomas: A TCGA/TCIA Project. Clin Cancer Res Off J Am Assoc Cancer Res. 2017;23(20):6078–6085.

22. Newitt DC, Zhang Z, Gibbs JE, et al. Test-retest repeatability and reproducibility of ADC measures by breast DWI: Results from the ACRIN 6698 trial. J Magn Reson Imaging JMRI. 2018;

23. Ogura A, Tamura T, Ozaki M, et al. Apparent Diffusion Coefficient Value Is Not Dependent on Magnetic Resonance Systems and Field Strength Under Fixed Imaging Parameters in Brain. J Comput Assist Tomogr. 2015;39(5):760–765.

24. Tan WL, Huang WY, Yin B, Xiong J, Wu JS, Geng DY. Can Diffusion Tensor Imaging Noninvasively Detect IDH1 Gene Mutations in Astrogliomas? A Retrospective Study of 112 Cases. Am J Neuroradiol. 2014;35(5):920–927.

25. Figini M, Riva M, Graham M, et al. Prediction of Isocitrate Dehydrogenase Genotype in Brain Gliomas with MRI: Single-Shell versus Multishell Diffusion Models. Radiology. 2018;289(3):788–796.

26. Zhao J, Wang Y-L, Li X-B, et al. Comparative analysis of the diffusion kurtosis imaging and diffusion tensor imaging in grading gliomas, predicting tumour cell proliferation and IDH-1 gene mutation status. J Neurooncol. 2019;141(1):195–203.

27. Tan Y, Zhang H, Wang X, et al. Comparing the value of DKI and DTI in detecting isocitrate dehydrogenase genotype of astrocytomas. Clin Radiol. 2019;

28. Delfanti RL, Piccioni DE, Handwerker J, et al. Imaging correlates for the 2016 update on WHO classification of grade II/III gliomas: implications for IDH, 1p/19q and ATRX status. J Neurooncol. 2017;135(3):601–609.

29. Sonoda Y, Shibahara I, Kawaguchi T, et al. Association between molecular alterations and tumor location and MRI characteristics in anaplastic gliomas. Brain Tumor Pathol. 2015;32(2):99–104.

30. Gozé C, Mansour L, Rigau V, Duffau H. Distinct IDH1/IDH2 mutation profiles in purely insular versus paralimbic WHO Grade II gliomas. J Neurosurg. 2013;118(4):866–872.

31. Xiong J, Tan W, Wen J, et al. Combination of diffusion tensor imaging and conventional MRI correlates with isocitrate dehydrogenase 1/2 mutations but not 1p/19q genotyping in oligodendroglial tumours. Eur Radiol. 2016;26(6):1705–1715.

32. Reyes-Botero G, Giry M, Mokhtari K, et al. Molecular analysis of diffuse intrinsic brainstem gliomas in adults. J Neurooncol. 2014;116(2):405–411.

33. Wu C-C, Jain R, Radmanesh A, et al. Predicting Genotype and Survival in Glioma Using Standard Clinical MR Imaging Apparent Diffusion Coefficient

Images: A Pilot Study from The Cancer Genome Atlas. AJNR Am J Neuroradiol. 2018;39(10):1814–1820.

34. Villanueva-Meyer JE, Wood MD, Choi BS, et al. MRI Features and IDH Mutational Status of Grade II Diffuse Gliomas: Impact on Diagnosis and Prognosis. AJR Am J Roentgenol. 2018;210(3):621–628.

35. Park YW, Han K, Ahn SS, et al. Prediction of IDH1-Mutation and 1p/19q-Codeletion Status Using Preoperative MR Imaging Phenotypes in Lower Grade Gliomas. AJNR Am J Neuroradiol. 2018;39(1):37–42.

36. Johnson DR, Diehn FE, Giannini C, et al. Genetically Defined Oligodendroglioma Is Characterized by Indistinct Tumor Borders at MRI. AJNR Am J Neuroradiol. 2017;38(4):678–684.

Tables

Commented [BE12]: -Throughout the tables, please make sure that all abbreviations are explained in the footnotes.

Table 1. Patient demographics, IDH and 1p19 genotypes of the glioma population

Number of patients (male/female)	All glioma subtypes	IDH^{wt}	IDH^{mut}/ 1p19q^{int}	IDH^{mut}/ 1p19q^{del}
290 (169/121)	290	82	107†	101
Age: Median (range, IQR)	40 (17–77, 19.25)	58.50 (20-77, 24.25)	35 (17-66, 13)	40 (19-76, 13.50)
Enhancement category				
Non-Enhancing	174	34	77	63
Patchy-Enhancing	89	28	28	33
Rim-Enhancing	25	20	0	5
Tumor location category				
Front or insula*	163	24	69	70
Other**	113	45	37	31
Thalamus or brainstem***	14	13	1	0
Absence of calcification: Non-calcified (calcified)	225 (31) ††	70 (4)	94 (4)	61 (23)
Absence of cyst(s): Non-cystic (cystic)	189 (101)	73 (9)	58 (49)	58 (43)
Haemorrhage: None (Petechial/ Macroscopic)	238 (7/11) †††	63 (5/5)	96 (2/2)	79 (0/4)
T2-FLAIR mismatch: Present (Absent)	51 (239)	0 (82)	46 (61)	5 (96)
Diameter ≥6 cm (Diameter < 6cm)	121 (162) ††††	32 (43)	47 (60)	42 (59)

* Indicates that the lesion was located in the frontal lobe or the insula.
** Indicates that the lesion was in a location other than the frontal lobe, insula, thalamus or brainstem.
*** Indicates that the lesion was located in the thalamus or the brainstem.
† Two cases within the IDH^{mut} 1p19q^{int} group had no post contrast imaging available for assessment.
†† Calcification status was evaluated as "uncertain" in a total of 34 cases
††† Hemorrhage status was evaluated as "uncertain" in a total of 34 cases
†††† Single largest tumour diameter could not be clearly measured in a total of 7 cases

Deleted: Haemorrhage

Table 2. Univariable and multivariable binary logistic regression model results for prediction of IDH^{wt} status

Feature	Univariable analyses	Multivariable analyses

				Model A		Model B	
	β	**p value**	**Pseudo R^2 ††**	**β**	**p value**	**β**	**p value**
ADC_{min} (s/mm²)	-0.05	<0.001	0.27	N/A		N/A	
ADC_{min}:ADC_{NAWM} ratio	-4.46	<0.001	0.31	N/A		N/A	
ADC_{mean} (s/mm²)	-0.05	<0.001	0.34	N/A		N/A	
ADC_{mean}:ADC_{NAWM} ratio	-4.39	<0.001	0.38	-5.712	<0.001	-3.225	<0.001
Age (years)	0.09	<0.001	0.34	-0.05	0.71†	-0.095	0.37†
Age² (years²)	0.01	<0.001	0.36	0.002	0.21†	0.002	0.09†
Age + age² (years²) †		<0.001	N/A	N/A	<0.001	N/A	<0.001
Enhancement (categorical)			0.17				
Non-Enhancing	Reference category			Reference		Reference	
Patchy-Enhancing	0.64	0.03		-0.315	0.58	-0.405	0.41
Rim-Enhancing	2.80	<0.001		2.956	0.02	1.66	0.02
Tumor location category			0.23				
Front or insula*	Reference category			Reference		Reference	
Other**	1.34	<0.001		0.776	0.12	0.862	0.04
Thalamus or brainstem***	4.32	<0.001		3.583	0.01	3.636	0.002
Absence of calcification	1.12	0.045	0.03	4.335	<0.001	N/A	
Absence of cyst(s)	1.86	<0.001	0.15	N/A		1.169	0.02
Constant	N/A			2.241	0.54	3.072	0.31
Pseudo R^2 ††	N/A			0.75		0.65	

† Age and Age squared are considered joint terms, hence a joint significance test was applicable. This was significant at p<0.001, which combined with the likelihood ratio test confirmed a significant contribution of age to the prediction model †† Nagelkerke Pseudo R^2 provides a summary statistic expressing the degree to which the overall model predicts the variation in the outcome (IDH^{wt} status). Please note, in the case of the univariable analyses, each individual predictor variable is analysed as a single predictive model, hence being given a pseudo R^2 value at the categorical (but not sub-categorical) level, while the multivariable models A and B each have their own overall pseudo R^2 value. * Indicates that the lesion was in the frontal lobe or the insula. ** Indicates that the lesion was in a location other than the frontal, insula, thalamus or brainstem. *** Indicates that the lesion was located in the thalamus or the brainstem.

Table 2 Footnote. Using the multivariable regression results, a formula was designed to calculate the likelihood of IDH_{wt} status for individual glioma patients as follows:

Log Odds Ratio for Model A

$L_A = (-5.712 \times ADC_{mean}{:}ADC_{NAWM}) + (-0.05 \times age) + (0.002 \times age^2) + (-0.315 \times$ solid contrast enhancement*) + (2.956 x rim contrast enhancement*) + (0.776 × tumor location = other)** + (3.583 × tumor location in thalamus or brainstem)** + (4.335 x absent calcification)*** + 2.241

Log Odds Ratio for Model B

L_B = (-3.225 × ADC_{mean}:ADC_{NAWM}) + (-0.095 x age) + (0.002 × age^2) + (- 0.405 x solid contrast enhancement*) + (1.66 x rim contrast enhancement*) + (0.862 × tumor location = other)** + (3.636 × tumor location in thalamus or brainstem)** + (1.169 x absent cyst(s))*** + 3.072

* Contrast enhancement pattern: 1 if present, 0 if absent (with each tumor assigned to one contrast enhancement category only)

** Tumor Location: 1 if in this category, 0 if not in this category

***Calcification (model A)/Cyst(s) (model B): 1 if present, 0 if absent (note the reversal is on purpose)

The probability of being IDH^{wt} was calculated, for models A and B using:

$$\text{Probability of being } IDH^{wt} = \frac{1}{(1 + e^{-L})}$$

where L is the relevant log odds-ratio.

Supplementary material

Supplementary Table S1. Anatomical and diffusion MRI parameters.

Sequence	TE* (ms)	TR* (ms)	In plane resolution (mm^2)	Slice thickness (mm)	Interslice spacing (mm)
T2	98.8 [9.4, 510.0]	4520 [800, 15000]	0.47 x 0.47 [0.26 x 0.26, 1.25 x 1.25]	5.0 [1.0, 7.0]	6.5 [0.0, 7.7]
T1 +contrast	9.0 [1.7, 293.0]	500 [5.3, 3200]	0.47 x 0.47 [0.43 x 0.43, 1.15 x 1.15]	5.0 [0.9, 7.0]	6.5 [0.0, 7.7]
DWI	89.0 [42.0, 379.0]	4000 [2538, 11500]	1.22 x 1.22 [0.64 x 0.64, 2.50 x 2.50]	5.0 [1.0, 6.0]	6.5 [0.0, 7.5]

- All values provided as median [max, min]. ** DWI performed using three diffusion gradients with b values 0 and b = 1000 s/mm^2.

Supplementary Table S2. Intraclass correlation coefficient (ICC), all gliomas versus non-enhancing gliomas

Method	ICC	All glioma cases	Non-enhancing subgroup

	measured by	Average measures ICC (95% CI)	Individual measures ICC (95% CI)	N (valid)	Average measures 95% CI)	Individual Measures (95% CI)	N (valid)
ADC_{min} (s/mm²) Observer: 1 vs 2	Consistency	0.89 (0.82 - 0.93)	0.80 (0.70 - 0.87)	75	0.89 (0.80 - 0.94)	0.80 (0.67 - 0.88)	51
	Absolute Agreement	0.89 (0.82 - 0.93)	0.80 (0.70 - 0.87)	75	0.89 (0.80 - 0.94)	0.80 (0.68 - 0.88)	51
ADC_{min} (s/mm²) Observer: 1 vs 3	Consistency	0.90 (0.86 - 0.92)	0.81 (0.76 - 0.85)	215	0.86 (0.80 - 0.90)	0.76 (0.67 - 0.82)	123
	Absolute Agreement	0.89 (0.83 - 0.92)	0.79 (0.71 - 0.85)	215	0.85 (0.75 - 0.90)	0.73 (0.60 - 0.82)	123
ADC_{mean} (s/mm²) Observer: 1 vs 2	Consistency	0.83 (0.73 - 0.89)	0.71 (0.58 - 0.81)	75	0.76 (0.57 - 0.86)	0.61 (0.40 - 0.76)	51
	Absolute Agreement	0.83 (0.73 - 0.89)	0.71 (0.58 - 0.81)	75	0.76 (0.58 - 0.86)	0.61 (0.40 - 0.76)	51
ADC_{mean} (s/mm²) Observer: 1 vs 3	Consistency	0.96 (0.94 - 0.97)	0.92 (0.89 - 0.94)	215	0.95 (0.92 - 0.96)	0.90 (0.86 - 0.93)	123
	Absolute Agreement	0.96 (0.94 - 0.97)	0.92 (0.89 - 0.94)	215	0.95 (0.92 - 0.96)	0.90 (0.86 - 0.93)	123
ADC_{min}:ADC_{NAWM} ratio Observer: 1 vs 2	Consistency	0.89 (0.83 - 0.93)	0.81 (0.71 - 0.87)	75	0.90 (0.82 - 0.94)	0.81 (0.69 - 0.89)	51
	Absolute Agreement	0.89 (0.83 - 0.93)	0.81 (0.71 - 0.87)	75	0.90 (0.82 - 0.94)	0.81 (0.69 - 0.89)	51
ADC_{min}:ADC_{NAWM} ratio Observer: 1 vs 3	Consistency	0.87 (0.83 - 0.90)	0.77 (0.71 - 0.82)	212	0.86 (0.80 - 0.90)	0.75 (0.66 - 0.82)	122
	Absolute Agreement	0.85 (0.76 - 0.90)	0.74 (0.61 - 0.82)	212	0.83 (0.69 - 0.90)	0.71 (0.53 - 0.81)	122
ADC_{mean}:ADC_{NAWM} ratio Observer: 1 vs 2	Consistency	0.86 (0.77 - 0.91)	0.75 (0.63 - 0.83)	75	0.81 (0.66 - 0.89)	0.68 (0.49 - 0.8)	51
	Absolute Agreement	0.85 (0.75 - 0.91)	0.74 (0.60 - 0.83)	75	0.80 (0.65 - 0.89)	0.67 (0.49 - 0.80)	51
ADC_{mean}:ADC_{NAWM} ratio Observer: 1 vs 3	Consistency	0.93 (0.90 - 0.94)	0.86 (0.82 - 0.89)	212	0.92 (0.89 - 0.94)	0.85 (0.80 - 0.89)	122
	Absolute Agreement	0.92 (0.90 - 0.94)	0.86 (0.81 - 0.89)	212	0.92 (0.88 - 0.94)	0.85 (0.78 - 0.89)	122
ADC_{NAWM} (s/mm²) Observer: 1 vs 2	Consistency	0.86 (0.77 - 0.91)	0.75 (0.63 - 0.83)	75	0.88 (0.79 - 0.93)	0.78 (0.65 - 0.87)	51
	Absolute Agreement	0.83 (0.65 - 0.90)	0.70 (0.48 - 0.82)	75	0.86 (0.70 - 0.92)	0.75 (0.54 - 0.86)	51
ADC_{NAWM} (s/mm²)	Consistency	0.83	0.71	212	0.85	0.74	122

Observer:		(0.78 - 0.87)	(0.64 - 0.77)		(0.79 - 0.90)	(0.65 - 0.81)	
1 vs 3	Absolute	0.82	0.70	212	0.84	0.73	122
	Agreement	(0.75 - 0.87)	(0.60 - 0.77)		(0.77 - 0.89)	(0.63 - 0.81)	

Supplementary Table S3. Kruskal-Wallis ANOVA results for the WHO 2016 glioma molecular groups, with non-enhancing gliomas (NE) additionally shown in a subgroup analysis

Diffusion parameter	Kruskal Wallis Omnibus tests		Pairwise comparisons between molecular subtypes*			Number of cases	
	p value			p value			
	All	NE		All	NE	All	NE
	Glioma	glioma		glioma	Glioma	glioma	glioma
ADC_{min} (s/mm^2)	<0.001	<0.001	0 vs 1	<0.001	<0.001	290	174
			0 vs 2	<0.001	0.01		
			1 vs 2	<0.001	<0.001		
ADC_{min}:ADC_{NAWM}	<0.001	<0.001	0 vs 1	<0.001	<0.001	287	173
			0 vs 2	<0.001	0.01		
			1 vs 2	<0.001	<0.001		
ADC_{mean} (s/mm^2)	<0.001	<0.001	0 vs 1	<0.001	<0.001	290	174
			0 vs 2	<0.001	0.02		
			1 vs 2	<0.001	<0.001		
ADC_{mean}:ADC_{NAWM}	<0.001	<0.001	0 vs 1	<0.001	<0.001	287	173
			0 vs 2	<0.001	0.01		
			1 vs 2	<0.001	<0.001		

*Molecular subtypes with nominal categories of 0=IDH wild-type (IDH^{wt}), 1=IDH mutant/1p19q intact (IDH^{mut}/$1p19q^{int}$) and 2=IDH mutant/1p19q co-deleted (IDH^{mut}/$1p19q^{del}$).

Supplementary Table S4. Eta² (η^2) results for the WHO 2016 glioma molecular groups, with non-enhancing gliomas (NE) additionally shown in a subgroup analysis.

	All gliomas	Non-enhancing gliomas	Rim-enhancing gliomas

Diffusion parameter	η^2	Effect size	N	η^2	Effect size	N	η^2	Effect size	N
ADC_{min}	0.28	Large	290	0.25	Large	174	0.01	Small	25
ADC_{min}:ADC_{NAWM} ratio	0.29	Large	287	0.25	Large	173	0.03	Small	25
ADC_{mean}	0.38	Large	290	0.42	Large	174	0.00	None	25
ADC_{mean}:ADC_{NAWM} ratio	0.38	Large	287	0.40	Large	173	0.03	Small	25

Supplementary Table S5. Cohen's kappa results for morphology categories

Observer	Kappa* (95% CI, weighted Kappa only)	Standard error	p value	Interpretation of Kappa (Agreement)	Two reader Agreement** (%)	N
Tumor Location						
1 vs 2	0.84	0.03	<0.001	Almost perfect		290

1 vs 3	0.81	0.03	<0.001	Almost perfect	97	290
2 vs 3	0.89	0.02	<0.001	Almost perfect		290
Multifocality						
1 vs 2	0.46	0.09	<0.001	Moderate		290
1 vs 3	0.20	0.05	<0.001	Slight	100	290
2 vs 3	0.37	0.06	<0.001	Fair		290
Definition of the non-enhancing margin						
1 vs 2	0.61 (0.56 - 0.67)	0.03	<0.001	Substantial		290
1 vs 3	0.58 (0.51 - 0.64)	0.03	<0.001	Moderate	94	290
2 vs 3	0.45 (0.38 - 0.52)	0.03	<0.001	Moderate		290
Haemorrhage (uncertain cases excluded***)						
1 vs 2	0.29	0.09	<0.001	Fair		145
1 vs 3	0.51	0.09	<0.001	Moderate	99	145
2 vs 3	0.50	0.13	<0.001	Moderate		145
Calcification (uncertain cases excluded***)						
1 vs 2	0.73	0.09	<0.001	Substantial		106
1 vs 3	0.67	0.09	<0.001	Substantial	N/A	106
2 vs 3	0.74	0.08	<0.001	Substantial		106
Cystic change						
1 vs 2	0.70	0.043	<0.001	Substantial		290
1 vs 3	0.66	0.045	<0.001	Substantial	N/A	290
2 vs 3	0.66	0.046	<0.001	Substantial		290
Enhancement (3 nominal categories of enhancement)						
1 vs 2	0.69 (0.62 - 0.76)	0.04	<0.001	Substantial		288[a]
1 vs 3	0.67 (0.60 - 0.75)	0.04	<0.001	Substantial	98	288[a]
2 vs 3	0.77 (0.71 - 0.83)	0.03	<0.001	Substantial		288[a]
Diameter (single longest diameter, binary, above or below 6 cm in diameter)						
1 vs 2	0.82	0.04	<0.001	Almost perfect		268[b]
1 vs 3	0.80	0.04	<0.001	Almost perfect	N/A	282[b]
2 vs 3	0.81	0.04	<0.001	Almost perfect		268[b]
FLAIR mismatch sign						
1 vs 2	0.59	0.05	<0.001	Moderate		289[c]
1 vs 3	0.44	0.06	<0.001	Moderate	N/A	289[c]
2 vs 3	0.62	0.07	<0.001	Substantial		290

* Where non-standard (weighted) Kappa has been used, confidence intervals for Kappa have been provided in parentheses. **Percentage agreement between at least 2 readers on categorization, prior to consensus read. ***Cases were considered uncertain if one or more observers rated the finding as uncertain.

Percentage agreement = N/A; where 2 observer agreement will always be 100% (for example, binary categories with 3 observers), this has been entered as N/A.

288a Two cases of enhancement were agreed as uncertain by observers, and were hence excluded

268/282b = 22/8 cases excluded as at least one observer indicated accurate measurement not possible due to diffuse margins

289c = One observer rated one case as status uncertain

Figure legends

Figure 1. An example of apparent diffusion coefficient (ADC) measurements. Axial T_2-weighted images of a right temporal IDH wild-type (IDH^{wt}) glioma (A) and ADC maps (B-D) showing the regions of interest (ROIs) used to determine minimum ADC (ADC_{min})

(perceived lowest ADC regions (3 per each patient), blue), ADC_{mean} (largest tumor cross-section measurement, red) and ADC in normal appearing white matter ($ADC_{NAWM,}$ contralateral centrum semiovale, yellow). Note that round ROIs were chosen, as this method can be replicated on most PACS systems.

Figure 2. Glioma morphology. Location: T_2-weighted images showing a temporal IDH wild-type (IDH^{wt}) glioma (A) versus a different patient with a frontal IDH mutant/1p19q co-deleted ($IDH^{mut}/1p19q^{del}$) glioma (B). Non-enhancing **tumor margins**: T_2-weighted and FLAIR images demonstrating distinct borders (also a T_2/FLAIR mismatch sign) in an IDH mutant/1p19q intact ($IDH^{mut}/1p19q^{int}$) glioma (C, D) versus the indistinct margin of a bithalamic IDH^{wt} glioma (E, F). **Cyst formation** and **enhancement patterns**: $IDH^{mut}/1p19q^{int}$ astrocytoma containing a small cyst (arrow) nearly isointense to cerebrospinal fluid (CSF) on FLAIR (G) without contrast uptake (H); T_2-weighted, FLAIR and post Gadolinium T_1-weighted images reveal small cysts (arrows) and patchy contrast uptake in a $IDH^{mut}1p19q^{del}$ oligodendroglioma (I-K); post Gadolinium contrast T_1-weighted image demonstrating rim enhancement surrounding central necrosis in an IDH^{wt} glioma (L).

Figure 3. STARD diagram of patient selection for the study.

Figure 4. Boxplots showing differences in the apparent diffusion coefficient (ADC) values between World Health Organization (WHO) grade II/III glioma molecular subtypes for ADC_{min} (A), ADC_{min}:ADC_{NAWM} ratio (B), ADC_{mean} (C) and ADC_{mean}:ADC_{NAWM} ratio (D). Abbreviations: ADC_{min} = minimum ADC, ADC_{mean} = mean ADC, ADC_{NAWM}= ADC in normal appearing white matter.

Figure 5. Univariable and multivariable logistic regression to predict IDH status. Univariable receiver operating characteristic (ROC) curves for ADC metrics (A), age and selected imaging features (B) displayed in comparison. ROC curves of the multivariable probabilities for model A and model B (C). Abbreviations: ADC_{min} = minimum ADC, ADC_{mean} = mean ADC, ADC_{min} ratio= minimum ADC:ADC in normal appearing white matter, ADC_{mean} ratio= mean ADC: ADC in normal appearing white matter (the inverse is shown for univariable AUC comparison).

Figure 6. An example of two patients in whom the contribution of age and glioma morphology resulted in **correct IDH status classification** over ADC alone: T2, FLAIR, ADC and T1+Gad images in a male **patient** aged 75 years (A-D) demonstrating an IDH wild-type (**IDH^{wt}) glioma** tumor with high solid component diffusivity (ADC_{mean}:ADC_{NAWM} ratio 2.19) and a rim-enhancement pattern. Non-contrast CT, T2, ADC and T1+Gad in a male **patient** aged 45 years (E-H), showing a calcified IDH mutant/1p19q co-deleted (**IDH^{mut}/1p19q^{del}**) oligodendroglioma (ADC_{mean}:ADC_{NAWM} ratio of 1.07). Abbreviations: ADC_{mean} = mean ADC, ADC_{NAWM}= ADC in normal appearing white matter.