Neuroradiology
https://doi.org/10.1007/s00234-025-03832-6

RESEARCH

Visual versus region-of-interest based diffusion evaluation and their diagnostic impact in adult-type diffuse gliomas

Aynur Azizova[1,2] · Yeva Prysiazhniuk[3,4] · Marcus Cakmak[1,5] · Elif Kaya[7] · Jan Petr[1,8] · Frederik Barkhof[1,6,9] · Ivar J. H. G. Wamelink[1,2] · Vera C. Keil[1,2,6]

Received: 7 January 2025 / Accepted: 26 October 2025

Abstract

Purpose To evaluate the comparability and reproducibility of standardized visual versus region-of-interest (ROI)-based diffusion assessment and their prediction capacity for isocitrate dehydrogenase (IDH) mutation status in adult gliomas.

Methods Preoperative MRI scans, including diffusion-weighted imaging (DWI), of grade 2–4 adult-type diffuse gliomas ($n = 303$) were evaluated by three raters and repeated after one month. Visual assessment used the categorization of the Visually AcceSAble Rembrandt Images-feature 17 classes (facilitated, dubious, restricted). ROI-based assessment placed circular ROI on the visually perceived lowest apparent diffusion coefficient (ADC) areas (absolute/aADC) and contralateral normal-appearing white matter (normalized/nADC). Agreement and correlation analysis between visual and ROI-based assessments were performed. Logistic regression was conducted for IDH prediction in the subgroup of 99 non-necrotic and non-hemorrhagic cases, selected from the full cohort with available IDH status.

Results ROI-based assessment demonstrated superior inter- and intra-rater agreement (intraclass correlation coefficient $\geq$ 0.56 (95%-CI: 0.48–0.63)) than visual assessment (Kendall's W/Cohen's weighted kappa $\geq$ 0.34 (95%-CI: 0.26–0.42)). Thresholds of 1,090 and 623×10^{-6} mm^2/s for aADC, and 1.38 and 0.80 for nADC, distinguishing facilitated, dubious, and restricted diffusion, significantly correlated with visual assessments ($P < .001$). IDH classification accuracy of visual assessment was comparable to that of the ROI-based method using thresholds of aADC $1{,}048 \times 10^{-6}$ mm^2/sn and nADC 1.38 (visual vs. aADC/nADC: 69% vs. 73%/70%). However, neither method achieved a balanced performance between specificity (99% vs. 81%/75%) and sensitivity (14% vs. 57%/61%).

Conclusion ROI-based diffusion assessment guided by visual input showed superior reproducibility than visual assessment alone. Although visual assessment demonstrated strong correlation with ADC thresholds and comparable overall IDH prediction accuracy, the two methods differ in clinical profile: visual assessment offered high specificity but low sensitivity, whereas ROI-based assessment improved sensitivity at the cost of reduced specificity.

Keywords Glioma · Magnetic resonance imaging · Diffusion-weighted imaging · Isocitrate dehydrogenase

Introduction

MRI is the primary modality for monitoring adult-type diffuse gliomas, the most prevalent malignant brain tumors in adults, providing essential diagnostic and prognostic information [1–3]. Conventionally, radiologists interpret MRI data visually, and several visually assessed MRI biomarkers are now part of routine clinical practice [4]. Quantitative MRI sequences and their standardized or quantitative evaluation are gaining attention, with growing evidence demonstrating their efficacy in distinguishing tumor characteristics [5–7].

Diffusion-weighted imaging (DWI) is the most established quantitative MRI technique and the only one recommended as a standard component of brain tumor imaging protocol [1]. It measures the Brownian motion of water molecules in tissues, providing insights into cellular density [8]. Commonly, DWI is analyzed by visually comparing the b-1000 isotropic map juxtaposed to an apparent diffusion coefficient (ADC) map. A European glioma imaging

Extended author information available on the last page of the article

Published online: 08 November 2025

survey reported that 78% of surveyed neuroradiologists prefer visual assessment of potential diffusion restriction [9]. While visual analysis is quick and widely adopted, it is inherently subjective. To address this, the Visually AcceSAble Rembrandt Images (VASARI) glioma imaging features set [10] introduced a standardized approach, categorizing diffusion assessment (feature 17) into three classes: restricted, dubious, and facilitated. In clinical practice, radiologists can opt against this plain visual analysis and choose a region-of-interest (ROI)-based assessment of diffusivity that produces absolute but normalizable values. Absolute ADC values directly reflect the diffusion properties within the ROI but are affected by technical factors. Normalized ADC values, the ratio of the absolute ADC to that of normal-appearing white matter, reduce variability across sequences and scanners but depend on accurate reference region selection and may obscure direct comparability of absolute values.

Diffusion assessment is critical in radiological decision-making, often alongside other MRI sequences [11, 12]. However, its independent diagnostic relevance has also been evaluated to better understand its clinical impact [13, 14]. One important application is isocitrate dehydrogenase (IDH) status prediction, a key biomarker for classifying adult-type diffuse gliomas. While visual assessment has been explored for its IDH genotyping, a recent meta-analysis [15] found limited evidence of its significance compared to other visual imaging features. In contrast, ROI-based assessments show strong potential for distinguishing IDH-wildtype tumors from IDH-mutant tumors [16–18], although the lack of consensus on ADC thresholds remains a significant limitation [19]. Almost all of these studies included tumors with necrosis and hemorrhage, defining features of IDH-wildtype tumors [4, 12, 20–23], which degrade ADC map quality and limit the reliability of diffusion assessments. Excluding such tumors could enhance the predictive accuracy and clinical applicability of DWI. Moreover, there is scant evidence in the literature regarding the correlation between visual and ROI-based approaches and the comparative diagnostic accuracy of these methods.

This study aims to evaluate if the visual assessment of diffusion - represented by VASARI feature 17 - is comparable with an ROI-based assessment and similarly reproducible. To explore a possible diagnostic impact of the diffusion evaluation method in daily practice, we assess the methods' capacity to predict IDH status in adult-type diffuse gliomas, excluding tumors with necrosis and hemorrhage.

Methods

Study cohort

The medical ethics review committee (VUmc_2021-0437) approved this retrospective single-center study and waived informed consent. Eligible patients between January 2010 and January 2021 were taken from a cohort presented in previous publications [12, 24]. The study cohort was sourced from the pseudonymized hospital glioma database (IMAGO) by I.W, a fourth-year neuro-oncology Ph.D. student. Inclusion criteria were treatment-naïve patients with grades 2–4 adult-type diffuse gliomas according to the 5th World Health Organization Central Nervous System (WHO-CNS) Classification. Patients with preoperative MRI data consisting of pre-contrast T1-weighted, T2-weighted, T2-FLAIR, post-contrast T1-weighted images, and DWI b-0 and b-1000 images with ADC maps generated automatically on the scanner were analyzed. Exclusion criteria were incomplete histomolecular diagnosis (e.g., missing IDH status), incomplete or suboptimal preoperative MRI (e.g., motion artifacts), a more than one-month interval between preoperative MRI and surgery, suprasellar, midbrain, and cerebellar tumors due to their distinct radiophenotype, and pediatric age group.

Histomolecular diagnosis

Histomolecular diagnosis followed the 2021 WHO-CNS classification and served as the reference standard. IDH mutation status was determined using immunohistochemistry, next-generation sequencing, or methylation profiling. In IDH-wildtype tumors, the diagnosis of glioblastoma was made based on characteristic molecular features (e.g., TERT promoter mutation, EGFR amplification, or combined gain of chromosome 7 and loss of chromosome 10), along with supportive histological findings such as necrosis or microvascular proliferation. A small subset of IDH-wildtype cases ($n=16$) without available (not otherwise specified) or conclusive (not elsewhere classified) molecular data were included based on a multidisciplinary consensus diagnosis of aggressive clinical behavior. IDH-wildtype diffuse gliomas were considered grade 4 tumors based on their typically aggressive clinical course, regardless of histological grade.

MRI data and analysis

MR images of all patients ($n=303$) were acquired on three 1.5T ($n=120$) and four 3 T scanners ($n=183$); see Supplementary Table 1 for details. Three raters with different levels of radiology experience (V.K., eleven years of neuroradiology experience; A.A., five years of neuroradiology experience; M.C., a fourth-year medical student with one month of specialized radiology training for this study using a different small cohort ($n=69$) from the hospital glioma database) independently conducted imaging evaluations

using RADIANT software (version 3.4.1.13367; https://www.radiantviewer.com/). The raters evaluated the visual and ROI-based methods, focusing on the solid tumor parts. Hemorrhage, necrosis, cysts, and peritumoral edema identified from the evaluation of pre- and post-contrast T1-weighted, T2-weighted, and T2-FLAIR images were excluded from the assessments. In multifocal/multicentric glioma cases, the most aggressive-looking lesion defined by the visually lowest ADC signal was considered. Visual and ROI-based evaluations were performed twice (measurements 1 and 2) by the same raters at a one-month interval for all enrolled patients. Raters were blinded to the histomolecular diagnosis during evaluation.

Visual and ROI-based DWI assessments

Visual evaluations were conducted using the VASARI feature 17: facilitated, dubious, and restricted diffusion (Supplementary Fig. 1). Facilitated diffusion is marked by a high or low b-1000 signal with a corresponding ADC signal higher than normal brain tissue. Dubious diffusion is identified by a high signal on b-1000 images with a corresponding ADC signal resembling normal brain tissue. Restricted diffusion is characterized by a high DWI signal intensity on b-1000 images with a corresponding lower signal on ADC maps than normal brain tissue. If the lesion showed a heterogeneous diffusion pattern, the lowest diffusion score was recorded, irrespective of the relative size of the area.

ROI-based assessments included placing circular ROI on areas on the ADC map that visually appeared to have the lowest ADC (absolute ADC; aADC, mm^2/s); see Supplementary Fig. 1. The slice with the largest area of this visually lowest ADC region was exclusively considered. The mean value of the measured area was recorded. The circular ROI size was 20–40 mm^2 to standardize the measurements across the raters. The raters were instructed to cover the region with the lowest ADC as completely as possible without extending into areas with a visually different ADC. A same-size circular ROI was also positioned on the contralateral normal-appearing white matter (ADC_{NAWM}) for normalization ($aADC/ADC_{NAWM}$ = normalized ADC; nADC).

While diffusion classification was based on the combined assessment of b-1000 DWI and ADC maps, T2-weighted images were also reviewed during evaluation to check the tumor's T2 signal, particularly in cases where T2 shine-through might be a concern.

Statistical analysis

Analysis was conducted using R package 4.3.0 by Y.P., a third-year Ph.D. student in neuroscience. The significance threshold was $P<.05$.

Descriptive analysis

Visual assessment was expressed as percentages per category. ROI-based measurements were summarized with the median and interquartile range (IQR).

Rater agreement

Consistency among raters was evaluated separately for the two assessment rounds through group and pairwise inter-rater agreements. In contrast, per-rater consistency between the first and second assessments was measured using intra-rater agreement analysis. Bootstrapping with 1000 iterations was used for all agreement analyses to calculate the confidence intervals. ROI-based measurements were analyzed using an intraclass correlation coefficient with two-way random-effects and mixed-effects models for inter-rater and intra-rater agreements. In visual assessment, Kendall's W and Cohen's weighted kappa were used for group inter-rater and pairwise inter-rater/intra-rater agreements. Agreement values were interpreted as follows: 0.01–0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; and 0.81–0.99, almost perfect [25].

Correlation between visual and ROI-based DWI assessments

The overall distribution of ROI-based measurements (median, IQR) within visual assessment classes was calculated using all six measurements. Logistic regression analysis was then used to determine aADC and nADC thresholds for distinguishing different visual assessment classes (facilitated vs. dubious, dubious vs. restricted). Spearman rank correlation analysis was conducted to identify the relationship between visual assessment and thresholded ROI-based measurements. The interpretation of the correlation coefficient (ρ) was as follows: 0.00–0.09, negligible; 0.10–0.39, weak; 0.40–0.69, moderate; 0.70–0.89, strong; and 0.90–1.00, very strong [26].

Inter-method IDH classification prediction performance

Considering the IDH status of cases, descriptive analyses were repeated for both visual and ROI-based measurements. Confusion matrices were used to assess the classification performance, treating IDH-wildtype gliomas as the positive class. The radiological IDH prediction was conducted only on cases without visual hemorrhage or necrosis, as these imaging features are primarily associated with IDH-wildtype tumors [4, 12, 20–23] and provide evident descriptive

characteristics, making diffusion status assessment less relevant. Hemorrhage and necrosis were evaluated using contrast-enhanced MRI as part of a previously published study [12], where all 303 cases were independently assessed by three raters. In the current analysis, we applied those prior results and excluded cases in which at least two raters identified hemorrhage or necrosis ($n = 204$), yielding a final IDH prediction cohort of 99 patients (33 IDH-wildtype and 66 IDH-mutant).

For visual assessments, cases rated as *restricted diffusion* were classified as IDH-wildtype, while a rating of *dubious plus facilitated diffusion* was classified as IDH-mutant. The rationale is the results of a previous study showing respective assumptions to be predictive [12].

Logistic regression analysis determined the IDH classification thresholds for the ROI-based values of both aADC and nADC. To ensure robustness and prevent data leakage, measurements across all raters were included and divided into training (70%) and test (30%) sets at the patient level. This approach ensured that all measurements from the same patient were assigned exclusively to either the training or test set, preserving the independence of the datasets. The Random OverSampling Examples method [27] addressed a class imbalance regarding IDH status in the training set. Subsequently, the diagnostic performance of these thresholded ROI-based measurements was calculated.

Results

Descriptive analysis

Table 1 shows the study cohort demographics.

The distribution of visual assessment classes per rater and measurement is shown in Fig. 1 and Supplementary Table 2. There was variability among evaluations by different raters, particularly in the *restricted diffusion class*. Overall, *restricted diffusion class* was the least chosen category (restricted vs. facilitated and dubious: 7–26% vs. 74–93%), with rater 2 assigning only 7% and 9% of cases to this category in assessments 1 and 2, respectively. These percentages were higher for rater 1 (25–26%) and rater 3 (16–20%). The most frequently chosen class was the *dubious diffusion class* for rater 2 (assessment 1/2: 64%/65%) and 3 (assessment 1/2: 61%/57%), while rater 1 primarily assigned cases to the *facilitated diffusion class* (assessment 1/2: 40%/43%).

Table 1 Study cohort demographics

Sample size: number	303
Age: years (standard deviation)	56.7 (14.2)
Sex: female/male	114/189
Isocitrate dehydrogenase status: mutant (codeleted/intact)/wildtype	82 (34/48)/221
Histological World Health Organization grade: grades 2/3/4	54/22/227

Caption: Table 1 describes the main characteristics of the study cohort.

The median (IQR) of aADC ranged between 842 (401) and 898 (432) x 10^{-6} mm^2/s across all measurements. The median (IQR) of nADC ranged between 1.10 (0.49) and 1.15 (0.61); see Fig. 2 and Supplementary Table 3.

Rater agreement

Group inter-rater agreements in measurements 1 and 2 were moderate for visual assessment (0.51 (95%-CI: 0.46–0.56) and 0.52 (95%-CI: 0.47–0.56)) and substantial for both aADC (0.66 (95%-CI: 0.60–0.72) and 0.64 (95%-CI: 0.56–0.70)) and nADC (0.62 (95%-CI: 0.56–0.68) and 0.62 (95%-CI: 0.55–0.68)). Pairwise inter-rater agreements were fair-moderate ($\geq$ 0.34 (95%-CI: 0.26–0.42) and moderate-substantial ($\geq$ 0.56 (95%-CI: 0.48–0.63) for visual and ROI-based assessments, respectively; see Table 2 and Supplementary Fig. 2. Intra-rater inter-measurement agreements were moderate-substantial ($\geq$ 0.56 (95%-CI: 0.49–0.64)) for visual assessment and substantial-almost perfect for ROI-based assessments ($\geq$ 0.73 (95%-CI: 0.67–0.77)); see Table 3 and Supplementary Fig. 3.

Correlation between visual and ROI-based DWI assessments

The median (IQR) of aADC within visual assessment classes was as follows: facilitated 1,235 (368) x 10^{-6} mm^2/s, dubious 825 (214) x 10^{-6} mm^2/s, restricted 574 (169) x 10^{-6} mm^2/s. Logistic regression analysis yielded optimal aADC thresholds of $1{,}090 \times 10^{-6}$ mm^2/s for facilitated vs. dubious diffusion and 623×10^{-6} mm^2/s for dubious vs. restricted diffusion. The median (IQR) of nADC within facilitated, dubious, and restricted visual assessment classes were 1.59 (0.51), 1.07 (0.27), and 0.73 (0.21), respectively. The nADC thresholds for facilitated vs. dubious and dubious vs. restricted diffusion were 1.38 and 0.80, respectively.

Subsequent analysis using the calculated thresholds revealed a strong correlation between visual and ROI-based assessments, with an overall correlation coefficient of ρ=0.79 (P<.001) for visual vs. aADC and ρ=0.81 (P<.001) for visual vs. nADC (Fig. 3). The results per rater and measurement also revealed a consistently strong correlation (all P<.001); for details, see Table 4 and Supplementary Fig. 4.

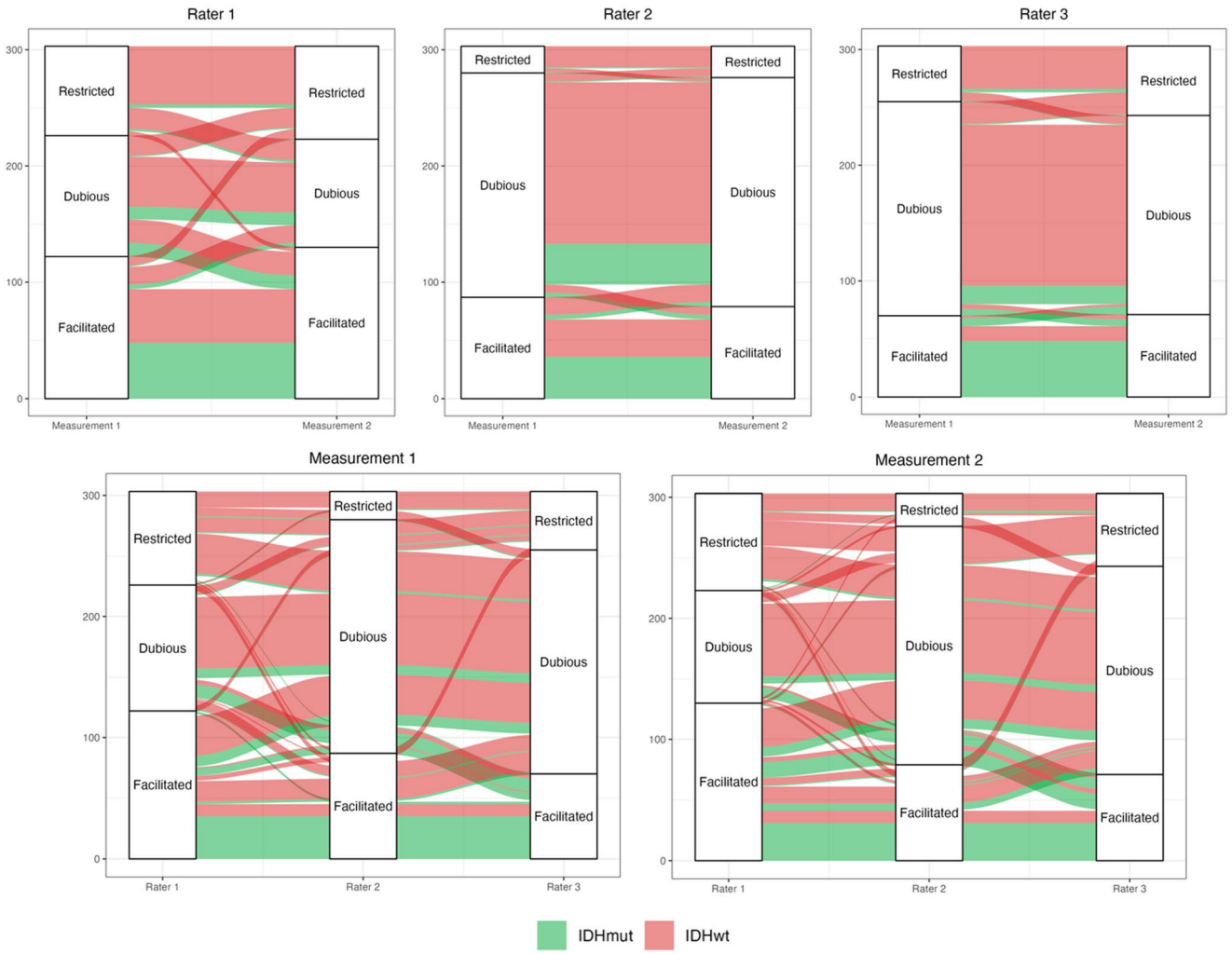

Fig. 1 Alluvial plots show the distribution of visual assessment classes per rater between measurements and per measurement between raters. Green and red colors represent isocitrate dehydrogenase-mutant (IDHmut) and -wildtype (IDHwt) gliomas, respectively

Inter-method IDH classification prediction performance

Visual assessments

In IDH-mutant gliomas, facilitated diffusion was the primary assessment class (52–80%) across all six measurements. For IDH-wildtype gliomas, facilitated (38–58%) and dubious (20–51%) classes were selected at similar rates. Restricted diffusion, the least common class overall, was more prevalent in IDH-wildtype gliomas (7–22%) compared to IDH-mutant gliomas (0–3%); see Fig. 4; Table 5.

When cases with visually restricted diffusion were accepted as IDH-wildtype and the remaining as IDH-mutant, the visual assessment achieved 69% accuracy, 14% sensitivity, 99% specificity, 89% positive predictive value, and 68% negative predictive value.

ROI-based assessments

The median (IQR) of aADC ranged from 1,006 (428) to 1,081 (425) x 10^{-6} mm^2/s for IDH-wildtype and from 1,112 (289) to 1,357 (569) x 10^{-6} mm^2/s for IDH-mutant gliomas across all six measurements. For nADC, the median (IQR) ranged from 1.22 (0.50) to 1.37 (0.45) for IDH-wildtype and from 1.48 (0.72) to 1.87 (0.80) for IDH-mutant gliomas; see Fig. 5; Table 5.

Optimal IDH classification thresholds were $1{,}048 \times 10^{-6}$ mm^2/s and 1.38 for aADC and nADC, respectively; see Fig. 5. The subsequent classification accuracy, sensitivity, specificity, and positive and negative predictive values for aADC were 73%, 57%, 81%, 58%, and 80%, respectively. The corresponding results for nADC were 70% accuracy, 61% sensitivity, 75% specificity, 53% positive predictive value, and 80% negative predictive value.

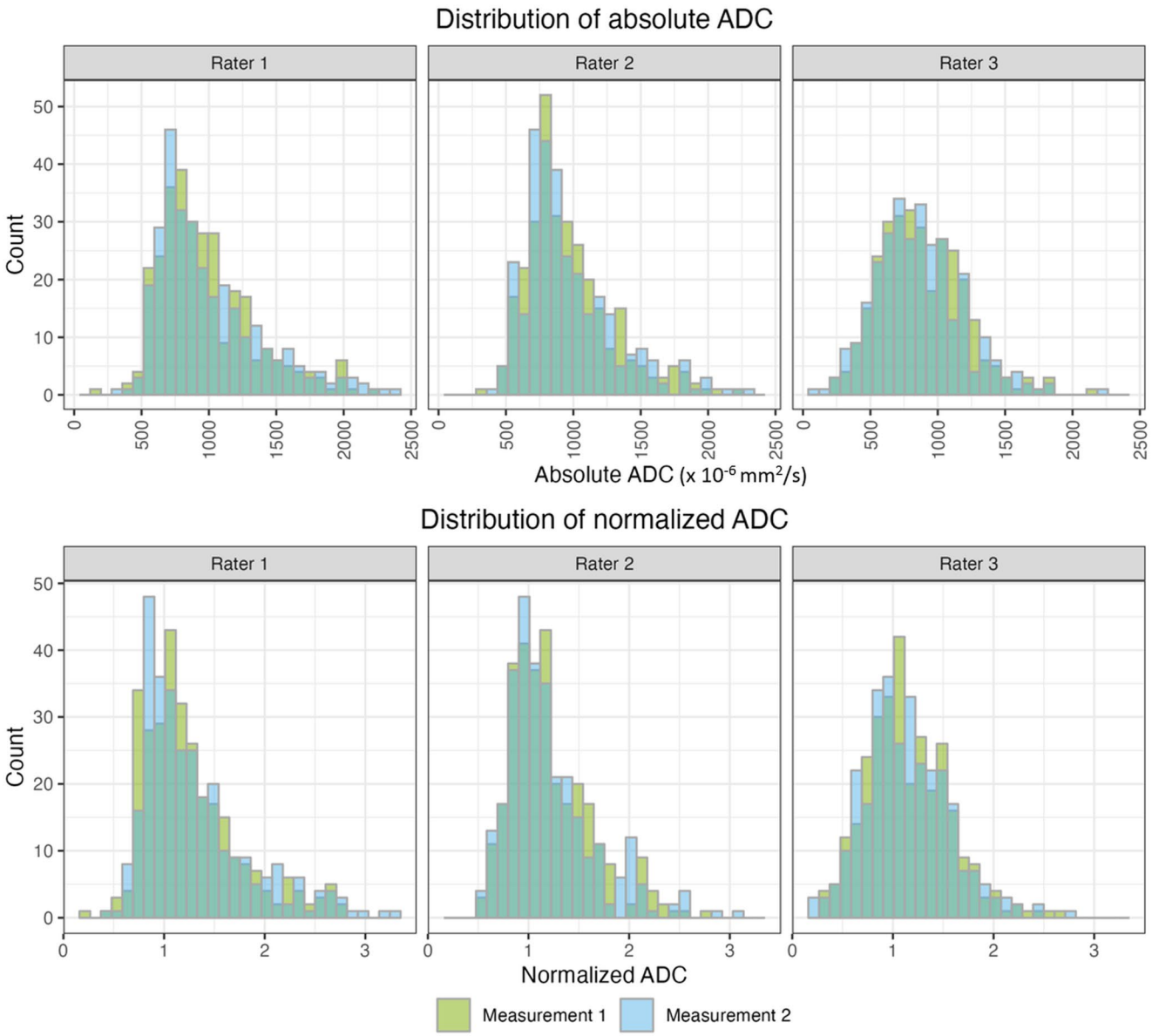

Fig. 2 Stacked histograms show the distribution of absolute and normalized ADC values per rater and measurement. Green bars represent the overlap between the first (yellow bars) and second (blue bars) measurements

Discussion

This study evaluated the reproducibility, correlation, and IDH categorization performance of visual vs. ROI-based diffusion assessment in adult-type diffuse gliomas. Our primary aim was to assess whether visual assessment of diffusion, as used in daily radiological practice, correlates well with ROI-based ADC measurements and could offer comparable reliability. We also tested whether this methodological distinction had clinical relevance by evaluating the ability of each approach to predict IDH mutation status in gliomas. Rather than developing a full diagnostic model, this analysis served to explore whether the simpler visual approach can provide similar diagnostic utility in practice.

ROI-based assessment demonstrated superior rater reproducibility, with moderate-almost perfect inter-/intra-rater agreements (≥ 0.56 (95%-CI: 0.48–0.63)) compared to fair-substantial agreements for visual assessment ($\geq$ 0.34 (95%-CI: 0.26–0.42)). ADC thresholds of 1,090 and 623×10^{-6} mm^2/s for aADC and 1.38 and 0.80 for nADC, distinguishing facilitated, dubious, and restricted diffusion, however, significantly correlated well with visual assessments. For the clinical use case of IDH classification, visual assessment, when compared to the ROI-based method at

Table 2 Pairwise inter-rater agreement in visual and ROI-based diffusion assessments

Rater pairs	Measurement 1	Measurement 2
Visual assessment*		
Rater 1 & 2	0.42 (95%-CI: 0.34–0.50)	0.42 (95%-CI: 0.34–0.49)
Rater 2 & 3	0.43 (95%-CI: 0.33–0.51)	0.37 (95%-CI: 0.28–0.46)
Rater 1 & 3	0.34 (95%-CI: 0.26–0.42)	0.39 (95%-CI: 0.31–0.47)
Absolute ADC mm^2/s**		
Rater 1 & 2	0.74 (95%-CI: 0.69–0.79)	0.69 (95%-CI: 0.63–0.74)
Rater 2 & 3	0.58 (95%-CI: 0.48–0.67)	0.58 (95%-CI: 0.47–0.67)
Rater 1 & 3	0.65 (95%-CI: 0.54–0.73)	0.63 (95%-CI: 0.47–0.74)
Normalized ADC**		
Rater 1 & 2	0.69 (95%-CI: 0.63–0.75)	0.66 (95%-CI: 0.58–0.72)
Rater 2 & 3	0.56 (95%-CI: 0.48–0.63)	0.58 (95%-CI: 0.50–0.65)
Rater 1 & 3	0.60 (95%-CI: 0.51–0.68)	0.62 (95%-CI: 0.48–0.72)

Caption: Table 2 demonstrates the pairwise inter-rater agreement results in visual and ROI-based assessments of diffusion.

*Cohen's weighted kappa

**Intraclass correlation coefficient

Table 3 Intra-rater inter-measurement agreement in visual and ROI-based diffusion assessments

Visual assessment*	
Rater 1	0.56 (95%-CI: 0.49–0.64)
Rater 2	0.75 (95%-CI: 0.68–0.82)
Rater 3	0.76 (95%-CI: 0.70–0.83)
Absolute ADC mm^2/s**	
Rater 1	0.76 (95%-CI: 0.70–0.80)
Rater 2	0.86 (95%-CI: 0.83–0.89)
Rater 3	0.79 (95%-CI: 0.75–0.83)
Normalized ADC**	
Rater 1	0.73 (95%-CI: 0.67–0.77)
Rater 2	0.85 (95%-CI: 0.81–0.87)
Rater 3	0.77 (95%-CI: 0.72–0.82)

Caption: Table 3 demonstrates intra-rater inter-measurement agreement results in visual and ROI-based assessments of diffusion.

*Cohen's weighted kappa

**Intraclass correlation coefficient

a threshold of $1{,}048 \times 10^{-6}$ mm^2/s for aADC and 1.38 for nADC, achieved the highest specificity (visual vs. aADC/nADC: 99% vs. 81%/75%), but had very low sensitivity (visual vs. aADC/nADC: 14% vs. 57%/61%). Regarding accuracy, the visual assessment showed comparable performance to the ROI-based method (69% vs. 73%/70%). This combination of results bears challenging implications for clinical practice.

Reproducibility is crucial when evaluating the consistency and reliability of imaging methods. Studies using VASARI criteria for visual assessment reported agreements ranging from fair to almost perfect (kappa 0.36–0.85), the spread highlighting the subjective nature of visual assessments [28–31]. ROI-based methods, being possibly more impartial, demonstrated superior reproducibility with intraclass correlation coefficient agreements ranging from 0.84 to 0.99 [16, 32, 33]. Our study further showed that an ROI-based evaluation offers a superior consistency (intraclass correlation coefficient 0.56–0.86) than the visual assessment (kappa 0.34–0.76). The lower agreement values for the ROI-based method in our study compared to previous studies may be attributed to methodological differences. For instance, two studies [16, 32] measured three visually defined lowest ADC areas and used the mean value of these measurements, while another study [33] included all solid tumor components with low ADC signal, potentially leading to more reliable evaluations. In contrast, our study focused on measuring only one area representing the visually lowest ADC signal to reflect the day-to-day practice and accurately correlate these two methods.

The literature lacked direct comparisons between visual and ROI-based diffusion assessments until now. A related study [34] compared both approaches but focused on glioma grading using a five-scale visual system, making direct comparison with our study challenging. Aligning with our results, they found a higher specificity (aADC 89% vs. visual 100%) and lower sensitivity (aADC 90% vs. visual 50%) for a visual evaluation. Our study established ADC thresholds for visual assessment classes in gliomas to facilitate method correlation, especially for hard-to-classify diffusion cases (aADC range: 623-1,090 $\times$ 10^{-6} mm^2/s and nADC range: 0.80–1.30), which are very common in gliomas and challenging for radiologists. These thresholds are presented to the community to streamline decision-making in clinical and research settings by integrating the reproducibility of ROI-based methods with the time efficiency of the visual method. Moreover, the results of this study could potentially be utilized to guide the application of diffusion data in advanced predictive models that incorporate artificial intelligence, which is currently hardly established.

Notably, inter-rater differences in assigning the "restricted" diffusion category were notable, ranging from 7 to 9% for rater 2 to 25–26% for rater 1 and 16–20% for rater 3. While some variation may reflect differences in experience, the trend did not follow seniority. This variability occurred despite using the standardized VASARI criteria, suggesting that individual interpretation thresholds and confidence in borderline cases continue to influence

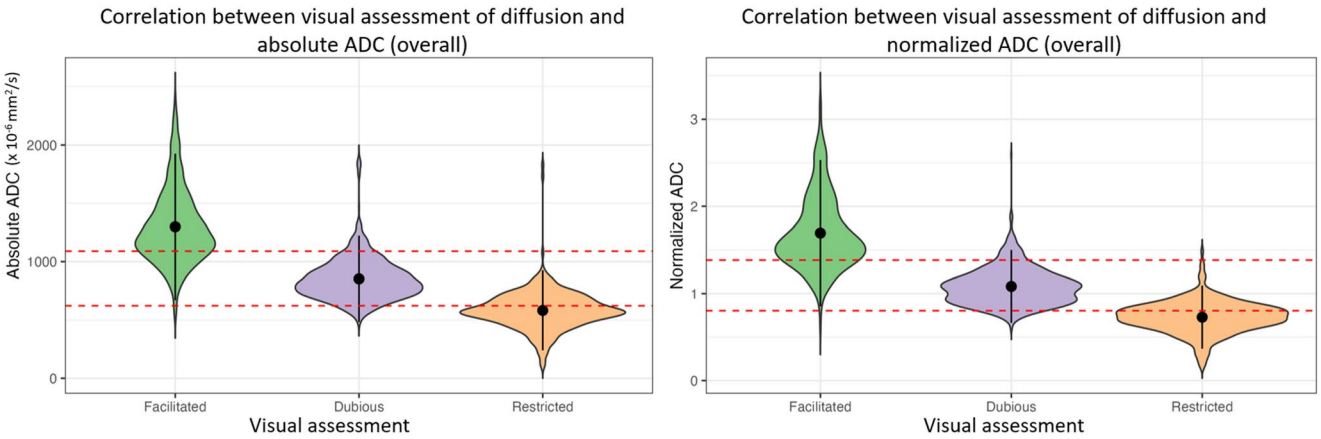

Fig. 3 Violin plots show the overall correlation between visual and ROI-based assessments of diffusion, including absolute and normalized ADC, across all measurements. Red dashed lines represent the absolute/normalized ADC distribution thresholds within visual assessment classes (facilitated, dubious, and restricted) derived from logistic regression analysis (absolute ADC: 1,090 x 10−6 mm^2/s for facilitated vs. dubious diffusion and 623 x 10−6 mm^2/s for dubious vs. restricted diffusion; normalized ADC: 1.38 for facilitated vs. dubious diffusion and 0.80 for dubious vs. restricted diffusion)

Table 4 The results of correlation analysis between visual and ROI-based diffusion assessments

Per rater visual vs. ROI-based assessments		Measurement 1	Measurement 2
		ρ* (*P*-value)	ρ* (*P*-value)
Rater 1	Visual vs. absolute ADC mm^2/s	0.78 (<0.001)	0.77 (<0.001)
	Visual vs. normalized ADC	0.78 (<0.001)	0.82 (<0.001)
Rater 2	Visual vs. absolute ADC mm^2/s	0.79 (<0.001)	0.80 (<0.001)
	Visual vs. normalized ADC	0.81 (<0.001)	0.81 (<0.001)
Rater 3	Visual vs. absolute ADC mm^2/s	0.80 (<0.001)	0.83 (<0.001)
	Visual vs. normalized ADC	0.82 (<0.001)	0.86 (<0.001)

Caption: Table 4 shows Spearman's rank correlation analysis results for each rater and measurement.

*Spearman's rank correlation coefficient

visual classification. The ADC thresholds established in this study for distinguishing between facilitated, dubious, and restricted diffusion may help guide future efforts to standardize interpretation and reduce such variability. These observations highlight the value of quantitative support in improving consistency across observers.

On the other hand, the clinical use case we applied for this study may suggest that despite lower reproducibility, the visual assessment of diffusivity in glioma and possibly other brain tumors is sufficient and diagnostically comparable to the ROI-based method, thus clinically equally powerful. Obviously, IDH status is not exclusively based on DWI in clinical practice. The IDH analysis of this study is a means to demonstrate the relevance of the choice of diffusion analysis practice. To isolate the specific diagnostic contribution of diffusion, we excluded tumors with visible necrosis or hemorrhage from the IDH classification analysis. These features are already strong indicators of IDH-wildtype gliomas and could bias the interpretation of diffusion by providing independent visual clues [4, 12, 20–23]. Moreover, they often degrade ADC map quality, limiting both visual and ROI-based assessments, even though such areas were excluded from analysis in all cases. Our aim was not to replicate routine clinical workflow, but to test how well DWI-based methods perform in diagnostically challenging cases, where diffusion signal might be one of the few available imaging clues.

Visual diffusion assessment for IDH classification was explored in several studies using the VASARI glioma imaging set [22, 23, 28–30, 35, 36]. Except for one study [28], none reported a significant predictive value for feature 17, visual assessment of diffusion, excluding it from multivariable models. Our study assessed the performance of different diffusion assessment approaches and found that the accuracy of visual assessment was comparable to that of aADC or nADC (69% vs. 73/70%). Nonetheless, both approaches revealed limitations, with the visual approach achieving high specificity but at the cost of low sensitivity, while the ROI-based method improved sensitivity but had a lower specificity. This imbalance highlights the challenge of reliably distinguishing glioma IDH characteristics using either method in isolation.

Studies using ROI-based diffusion assessment for IDH subtyping reported aADC thresholds between 900 and 1,200 × 10^{-6} mm^2/s [11, 13, 16, 37–39] and nADC thresholds between 1.28 and 1.60 [13, 16, 39]. Similarly, our study identified thresholds within this range, with an aADC

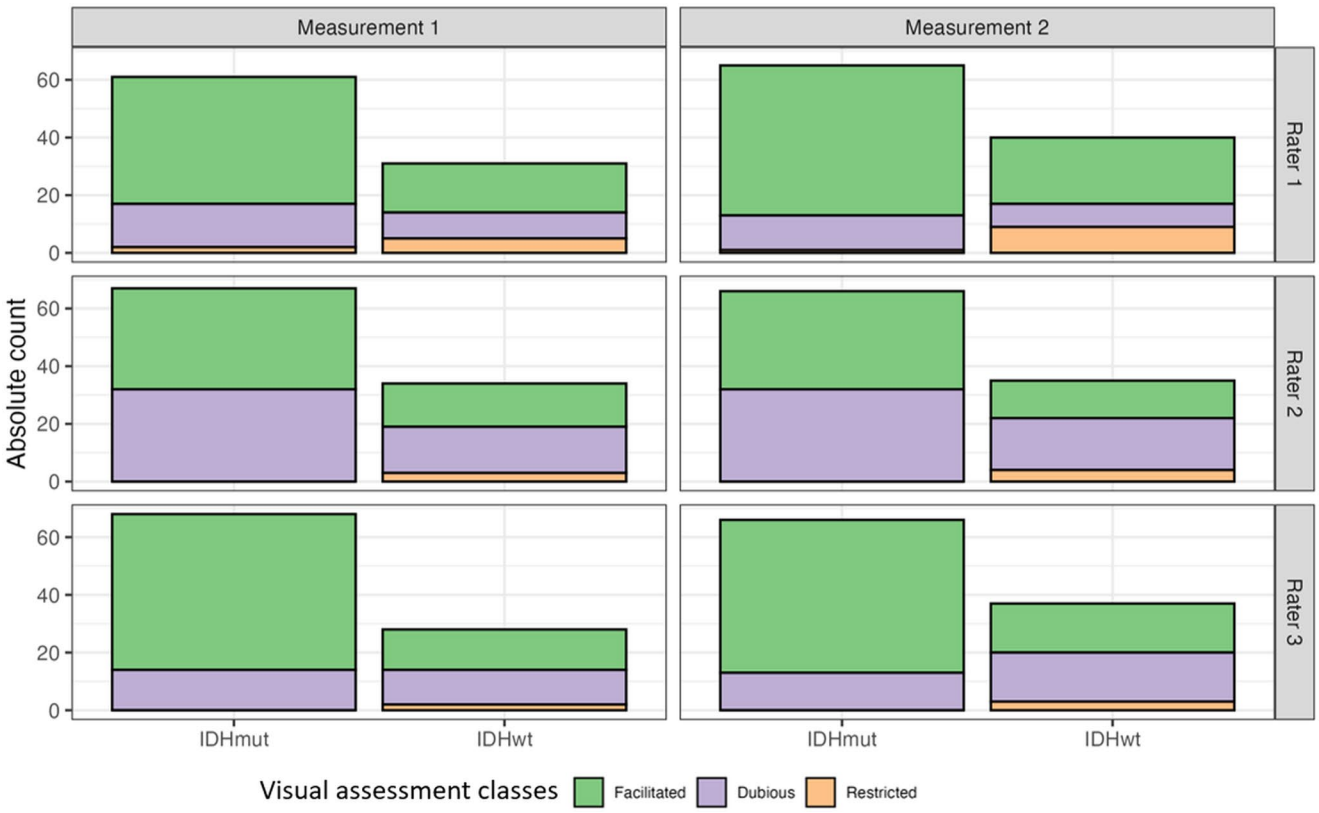

Fig. 4 Stacked bar plots show the distribution of visual diffusion assessment classes per rater and measurement in isocitrate dehydrogenase (IDH)-mutant and -wildtype gliomas using non-necrotic and non-hemorrhagic adult-type glioma cases. IDHmut = IDH-mutant, IDHwt = IDH-wildtype

Table 5 Distribution of visual and ROI-based assessment of diffusion in IDH-mutant and IDH-wildtype gliomas

		Rater 1		Rater 2		Rater 3	
		run 1	run 2	run 1	run 2	run 1	run 2
Distribution of visual assessment classes %							
IDH-mutant	Facilitated	72%	80%	52%	52%	79%	80%
	Dubious	25%	18%	48%	48%	21%	20%
	Restricted	3%	2%	0%	0%	0%	0%
IDH-wildtype	Facilitated	55%	58%	44%	38%	50%	46%
	Dubious	29%	20%	47%	51%	43%	46%
	Restricted	16%	22%	9%	11%	7%	8%
Distribution of absolute ADC median (IQR) x 10^{-6} mm^2/s							
IDH-mutant		1,288 (540)	1,357 (569)	1,165 (427)	1,166 (520)	1,162 (290)	1,112 (289)
IDH-wildtype		1,049 (566)	1,068 (544)	1,081 (425)	1,006 (428)	1,040 (333)	1,026 (404)
Distribution of normalized ADC median (IQR)							
IDH-mutant		1.62 (0.81)	1.87 (0.80)	1.49 (0.57)	1.48 (0.72)	1.56 (0.35)	1.51 (0.35)
IDH-wildtype		1.24 (0.73)	1.33 (0.68)	1.34 (0.62)	1.22 (0.50)	1.23 (0.32)	1.37 (0.45)

Caption: Table 5 describes the distribution of visual and ROI-based assessment of diffusion in IDH-mutant and IDH-wildtype gliomas using non-necrotic and non-hemorrhagic adult-type diffuse gliomas.
Abbreviation: run 1/2 = first/second assessments, IDH = isocitrate dehydrogenase, IQR = interquartile range

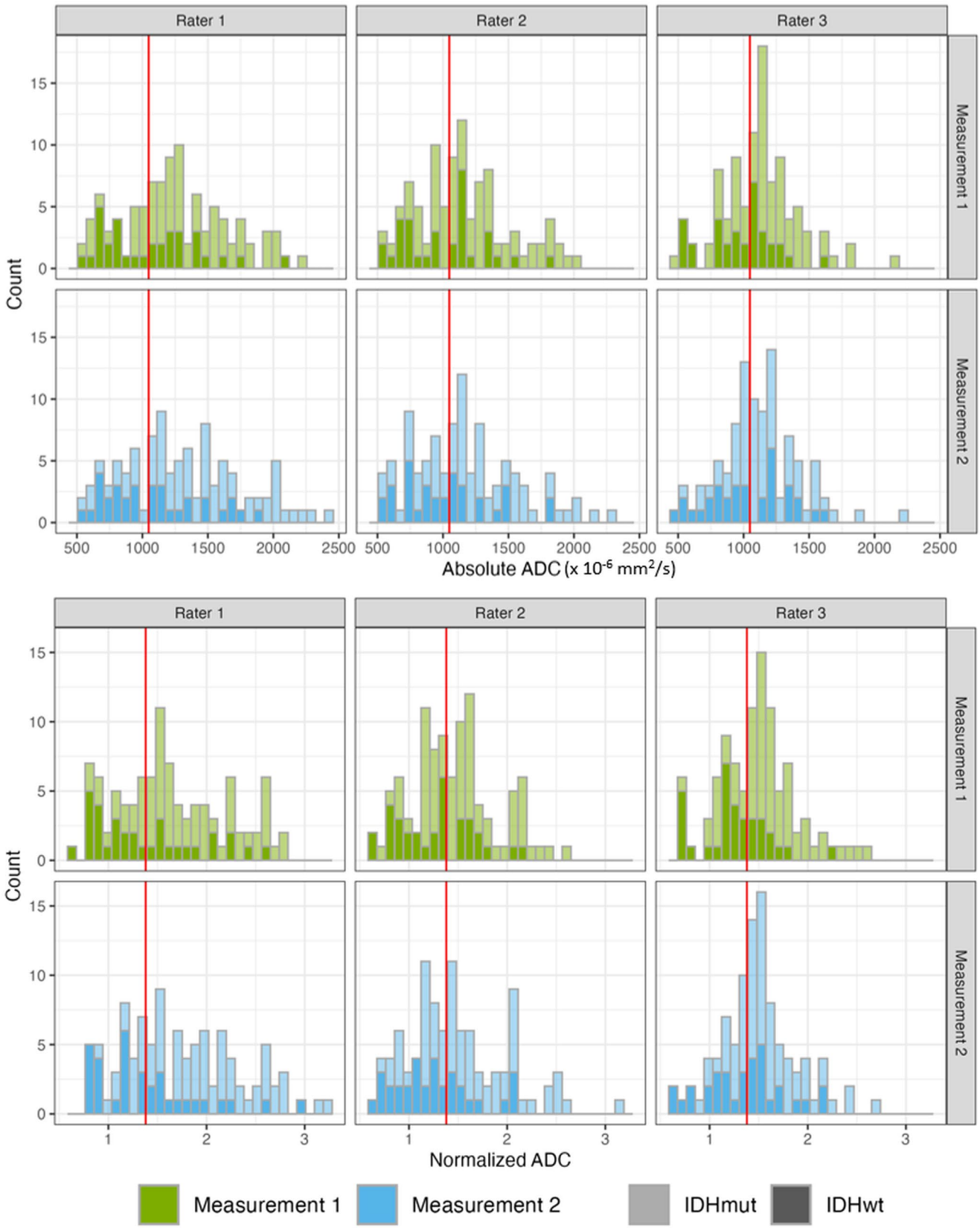

Fig. 5 Stacked histograms show the absolute and normalized ADC distribution per rater and measurement in isocitrate dehydrogenase (IDH)-mutant and -wild-type gliomas using using non-necrotic and non-hemorrhagic adult-type glioma cases. Red horizontal lines represent the optimal IDH classification thresholds of absolute/normalized ADC, derived from logistic regression analysis (absolute ADC threshold: 1,048 x 10-6 mm2/s; normalized ADC threshold: 1.38)

threshold of 1,048 × 10^{-6} mm^2/s and an nADC threshold of 1.38. The variation in reported thresholds likely reflects methodological heterogeneity across different studies; for example, Ma et al. [16] used the average of three visually identified lowest ADC areas for the tumor and a single ROI for normal-appearing white matter, reporting 65/92% sensitivity/specificity for aADC at 930 × 10^{-6} mm^2/s and 69/93% for nADC at 1.28. Another study [39] averaged four non-overlapping ROIs for the tumor and two for normal-appearing white matter, showing 84/68% sensitivity/specificity at 1,200 × 10^{-6} mm^2/s for aADC and 82/80% for nADC at 1.60. Thust et al. [13] used regional and volumetric ADC assessments with ADC_{min} measurements based on visually identified lowest ADC areas. They gathered three ROIs and took into account the mean value of the numerically lowest ADC measurement. Although cases with hemorrhage and necrosis were not excluded, the study focused on grade 2 and 3 gliomas, the majority of which were IDH-mutant (204 IDH-mutant vs. 79 IDH-wildtype), thereby reducing the likelihood of encountering these imaging features, thus closely matching our cohort. Their classification thresholds were $aADC_{min}$ at 1,070 × 10^{-6} mm^2/s, and $nADC_{min}$ at 1.40 with sensitivity/specificity values of 82/61% for $aADC_{min}$ and 86/62% for $nADC_{min}$.

This study has several limitations. First, this is a single-center study; however, it includes MRI data from multiple scanners and field strengths, reflecting normal clinical variability. Second, automated or volumetric ADC measurements were not included; instead, ROI placement was guided by visual evaluations, potentially introducing a collinearity bias. This method choice intentionally reflects real-world workflow but may limit the detection of visually subtle lowest ADC areas. However, full automation of ADC readings is unlikely to be implemented soon, particularly when excluding biasing hemorrhagic or necrotic areas is necessary. Additionally, the interrater variability in ADC measurements may have hindered the determination of a single optimal threshold for IDH classification. Future studies should aim to harmonize distributions across raters to mitigate interrater effects, incorporate external datasets with varied ADC quantification methods, and include other pathologies to increase the clinical impact of this study by validating and refining its findings.

Conclusions

ROI-based diffusion assessment with visual guidance in adult-type diffuse gliomas provided more reproducible results than visual assessment alone, and both techniques showed strong correlation and comparable overall accuracy in predicting IDH status. However, their diagnostic profiles differed: visual assessment was highly specific but lacked sensitivity, while ROI-based assessment improved sensitivity at the cost of lower specificity. Clinicians can, therefore, rely on visual DWI assessment in daily practice but should consider supplementing it with ROI measurements, particularly when a more balanced detection performance or reproducibility are required.

Author contributions All authors contributed to the study conception and design. Conceptualization, supervision and critical review were performed by J.P, F.B, and V.C.K. Material preparation, data collection, and analysis were performed by A.A., Y.P., M.C., E.K., I.J.H.G.W., and V.C.K. The first draft of the manuscript was written by A.A., and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This study has received funding from the Hanarth Foundation. F.B. is supported by the National Institute for Health and Care Biomedical Research Center at University College London Hospitals. Y.P., J.P., and V.C.K. are supported by a Czech Health Research Council grant (NU23-08-00460). A.A. is supported by the European Society of Neuroradiology Research Fellowship Grant.

Data availability No datasets were generated or analysed during the current study.

Declarations

Ethical approval and consent to participate Ethical approval was granted by the institutional medical ethics review committee. All patients had provided consent for using their data in research during their initial admission to the institution. The requirement for obtaining study-specific informed consent was waived because of this study's retrospective nature.

Competing interests The authors declare no competing interests.

References

1. Ellingson BM, Bendszus M, Boxerman J et al (2015) Consensus recommendations for a standardized brain tumor imaging protocol in clinical trials. Neuro Oncol 17:1188–1198. https://doi.org/10.1093/neuonc/nov095

2. Wen PY, van den Bent M, Youssef G et al (2023) RANO 2.0: update to the response assessment in neuro-oncology criteria for high- and low-grade gliomas in adults. J Clin Oncol 41:5187–5199. https://doi.org/10.1200/JCO.23.01059
3. Whitfield BT, Huse JT (2022) Classification of adult-type diffuse gliomas: impact of the World Health Organization 2021 update. Brain Pathol 32:e13062. https://doi.org/10.1111/bpa.13062
4. Feraco P, Franciosi R, Picori L et al (2022) Conventional MRI-derived biomarkers of adult-type diffuse glioma molecular subtypes: a comprehensive review. Biomedicines. https://doi.org/10.3390/biomedicines10102490
5. Hangel G, Schmitz-Abecassis B, Sollmann N et al (2023) Advanced MR techniques for preoperative glioma characterization: part 2. J Magn Reson Imaging 57:1676–1695. https://doi.org/10.1002/jmri.28663
6. Hirschler L, Sollmann N, Schmitz-Abecassis B et al (2023) Advanced MR techniques for preoperative glioma characterization: part 1. J Magn Reson Imaging 57:1655–1675. https://doi.org/10.1002/jmri.28662
7. Prysiazhniuk Y, Server A, Leske H et al (2024) Diffuse glioma molecular profiling with arterial spin labeling and dynamic susceptibility contrast perfusion MRI: a comparative study. Neuro Oncol Adv Vdae 113. https://doi.org/10.1093/noajnl/vdae113
8. Huisman TAGM (2010) Diffusion-weighted and diffusion tensor imaging of the brain, made easy. Cancer Imaging 10 Spec no A:S163–71. https://doi.org/10.1102/1470-7330.2010.9023
9. Thust SC, Heiland S, Falini A et al (2018) Glioma imaging in Europe: a survey of 220 centres and recommendations for best clinical practice. Eur Radiol 28:3306–3317. https://doi.org/10.1007/s00330-018-5314-5
10. VASARI research project - the cancer imaging archive (TCIA) public access - cancer imaging archive wiki. https://wiki.cancerimagingarchive.net/display/Public/VASARI+Research+Project. Accessed 19 Dec 2023
11. Xing Z, Yang X, She D et al (2017) 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 38:1138–1144. https://doi.org/10.3174/ajnr.A5171
12. Azizova A, Prysiazhniuk Y, Wamelink IJHG et al (2024) Preoperative prediction of diffuse glioma type and grade in adults: a gadolinium-free MRI-based decision tree. Eur Radiol. https://doi.org/10.1007/s00330-024-11140-5
13. Thust SC, Maynard JA, Benenati M et al (2021) Regional and volumetric parameters for diffusion-weighted WHO grade II and III glioma genotyping: a method comparison. AJNR Am J Neuroradiol 42:441–447. https://doi.org/10.3174/ajnr.A6965
14. Zhang L, Min Z, Tang M et al (2017) The utility of diffusion MRI with quantitative ADC measurements for differentiating high-grade from low-grade cerebral gliomas: evidence from a meta-analysis. J Neurol Sci 373:9–15. https://doi.org/10.1016/j.jns.2016.12.008
15. Azizova A, Prysiazhniuk Y, Wamelink IJHG et al (2024) Ten years of VASARI glioma features: systematic review and meta-analysis of their impact and performance. AJNR Am J Neuroradiol. https://doi.org/10.3174/ajnr.A8274
16. Ma X, Cheng K, Cheng G et al (2023) Apparent diffusion coefficient as imaging biomarker for identifying IDH mutation, 1p19q codeletion, and MGMT promoter methylation status in patients with glioma. J Magn Reson Imaging 58:732–738. https://doi.org/10.1002/jmri.28589
17. Wu J, Su R, Qiu D et al (2022) Analysis of DWI in the classification of glioma pathology and its therapeutic application in clinical surgery: a case-control study. Transl Cancer Res 11:805–812. https://doi.org/10.21037/tcr-22-114
18. Cindil E, Sendur HN, Cerit MN et al (2022) Prediction of IDH mutation status in high-grade gliomas using DWI and high T1-weight DSC-MRI. Acad Radiol 29 Suppl 3:S52–S62. https://doi.org/10.1016/j.acra.2021.02.002
19. Smits M (2021) MRI biomarkers in neuro-oncology. Nat Rev Neurol 17:486–500. https://doi.org/10.1038/s41582-021-00510-y
20. Guo X, Gu L, Li Y et al (2023) Histological and molecular glioblastoma, IDH-wildtype: a real-world landscape using the 2021 WHO classification of central nervous system tumors. Front Oncol 13:1200815. https://doi.org/10.3389/fonc.2023.1200815
21. Ahn SS, An C, Park YW et al (2021) Identification of magnetic resonance imaging features for the prediction of molecular profiles of newly diagnosed glioblastoma. J Neurooncol 154:83–92. https://doi.org/10.1007/s11060-021-03801-y
22. Sahu A, Patnam NG, Goda JS et al (2022) Multiparametric magnetic resonance imaging correlates of isocitrate dehydrogenase mutation in WHO high-grade astrocytomas. J Pers Med. https://doi.org/10.3390/jpm13010072
23. Gemini L, Tortora M, Giordano P et al (2023) Vasari scoring system in discerning between different degrees of glioma and IDH status prediction: a possible machine learning application? J Imaging Sci Technol. https://doi.org/10.3390/jimaging9040075
24. Azizova A, Wamelink IJHG, Prysiazhniuk Y et al (2024) Human performance in predicting enhancement quality of gliomas using gadolinium-free MRI sequences. J Neuroimaging. https://doi.org/10.1111/jon.13233
25. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174
26. Schober P, Boer C, Schwarte LA (2018) Correlation coefficients: appropriate use and interpretation. Anesth Analg 126:1763–1768. https://doi.org/10.1213/ANE.0000000000002864
27. Lunardon N, Menardi G, Torelli N (2014) ROSE: a package for binary imbalanced learning. R J 6:79
28. Park YW, Han K, Ahn SS et al (2018) Prediction of IDH1-mutation and 1p/19q-codeletion status using preoperative MR imaging phenotypes in lower grade gliomas. AJNR Am J Neuroradiol 39:37–42. https://doi.org/10.3174/ajnr.A5421
29. Zhou H, Vallières M, Bai HX et al (2017) MRI features predict survival and molecular markers in diffuse lower-grade gliomas. Neuro Oncol 19:862–870. https://doi.org/10.1093/neuonc/now256
30. Hyare H, Rice L, Thust S et al (2019) Modelling MR and clinical features in grade II/III astrocytomas to predict IDH mutation status. Eur J Radiol 114:120–127. https://doi.org/10.1016/j.ejrad.2019.03.003
31. Su C-Q, Lu S-S, Han Q-Y et al (2019) Intergrating conventional MRI, texture analysis of dynamic contrast-enhanced MRI, and susceptibility weighted imaging for glioma grading. Acta Radiol 60:777–787. https://doi.org/10.1177/0284185118801127
32. Maynard J, Okuchi S, Wastling S et al (2020) World health organization grade II/III glioma molecular status: prediction by MRI morphologic features and apparent diffusion coefficient. Radiology 296:111–121. https://doi.org/10.1148/radiol.2020191832
33. Hu Y-C, Yan L-F, Sun Q et al (2017) Comparison between ultra-high and conventional mono b-value DWI for preoperative glioma grading. Oncotarget 8:37884–37895. https://doi.org/10.18632/oncotarget.14180
34. Phuttharak W, Thammaroj J, Wara-Asawapati S, Panpeng K (2020) Grading gliomas capability: comparison between visual assessment and apparent diffusion coefficient (ADC) value measurement on diffusion-weighted imaging (DWI). Asian Pac J Cancer Prev 21:385–390. https://doi.org/10.31557/APJCP.2020.21.2.385
35. Su C-Q, Lu S-S, Zhou M-D et al (2019) Combined texture analysis of diffusion-weighted imaging with conventional MRI for non-invasive assessment of IDH1 mutation in anaplastic gliomas.

Clin Radiol 74:154–160. https://doi.org/10.1016/j.crad.2018.10.002
36. Cao M, Suo S, Zhang X et al (2021) Qualitative and quantitative MRI analysis in IDH1 genotype prediction of lower-grade gliomas: a machine learning approach. BioMed Res Int 2021(1):1235314. https://doi.org/10.1155/2021/1235314
37. Villanueva-Meyer JE, Wood MD, Choi BS et al (2018) MRI features and IDH mutational status of grade II diffuse gliomas: impact on diagnosis and prognosis. AJR Am J Roentgenol 210:621–628. https://doi.org/10.2214/AJR.17.18457
38. Wasserman JK, Nicholas G, Yaworski R et al (2015) Radiological and pathological features associated with IDH1-R132H mutation status and early mortality in newly diagnosed anaplastic astrocytic tumours. PLoS One 10:e0123890. https://doi.org/10.1371/journal.pone.0123890
39. Liu S, Zhang Y, Kong Z et al (2022) Feasibility of evaluating the histologic and genetic subtypes of WHO grade II-IV gliomas by diffusion-weighted imaging. BMC Neurosci 23:72. https://doi.org/10.1186/s12868-022-00750-8

Authors and Affiliations

Aynur Azizova[1,2] · Yeva Prysiazhniuk[3,4] · Marcus Cakmak[1,5] · Elif Kaya[7] · Jan Petr[1,8] · Frederik Barkhof[1,6,9] · Ivar J. H. G. Wamelink[1,2] · Vera C. Keil[1,2,6]

✉ Vera C. Keil
v.c.w.keil@amsterdamumc.nl

Aynur Azizova
a.azizova@amsterdamumc.nl

Yeva Prysiazhniuk
yeva.prysiazhniuk@lfmotol.cuni.cz

Marcus Cakmak
m.cakmak1@amsterdamumc.nl

Elif Kaya
elifkayamedicine@gmail.com

Jan Petr
j.petr@hzdr.de

Frederik Barkhof
f.barkhof@amsterdamumc.nl

Ivar J. H. G. Wamelink
i.j.wamelink@amsterdamumc.nl

1 Amsterdam UMC location Vrije Universiteit Amsterdam, Radiology & Nuclear Medicine Department, Amsterdam, Netherlands

2 Cancer Center Amsterdam, Imaging and Biomarkers, Amsterdam, Netherlands

3 Charles University, The Second Faculty of Medicine, Department of Pathophysiology, Prague, Czech Republic

4 Motol University Hospital, Prague, Czech Republic

5 Vrije Universiteit Amsterdam, University Medical Center, Amsterdam, Netherlands

6 Amsterdam Neuroscience, Brain Imaging, Amsterdam, Netherlands

7 Ankara Yıldırım Beyazıt University, Faculty of Medicine, Ankara, Turkey

8 Helmholtz-Zentrum Dresden-Rossendorf, Institute of Radiopharmaceutical Cancer Research, Dresden, Germany

9 Queen Square Institute of Neurology and Center for Medical Image Computing, University College London, London, United Kingdom