Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Interpretable Deep Learning Enables MRI-Based Virtual Biopsy for Molecular Stratification of Diffuse Gliomas

Barbara Iadarola
Siena Imaging Srl
Francesco Sforazzini
Siena Imaging Srl
Ludovico Luchetti
Siena Imaging Srl
Matteo Leoncini
Siena Imaging Srl
Giordano Gentile
Siena Imaging Srl
Gerben Borst
University of Manchester
Pietro Ivo D'Urso
Manchester Centre of Clinical Neuroscience
Federico Roncaroli
University of Manchester
Marco Battaglini
battaglini@sienaimaging.it

Siena Imaging Srl
Romina D'Aurizio
romina.daurizio@iit.cnr.it

CNR

Research Article

Keywords: Non-invasive biopsy, Whole-brain 3D structural MRI, Explainable deep learning, Glioma molecular stratification, Personalized medicine

Posted Date: October 29th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7965791/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Interpretable Deep Learning Enables MRI-Based Virtual Biopsy for Molecular Stratification of Diffuse Gliomas

Barbara Iadarola[1,2], Francesco Sforazzini[1], Ludovico Luchetti[1], Matteo Leoncini[1], Giordano Gentile[1], Gerben Borst[3], Pietro Ivo D'Urso[4], Federico Roncaroli[5], Marco Battaglini[1,6,#], Romina D'Aurizio[2,#]

[1]Siena Imaging Srl, Siena, Italy

[2]Institute of Informatics and Telematics, National Research Council, Pisa, Italy

[3]Division of Cancer Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, United Kingdom

[4]Department of Neurosurgery, Manchester Centre of Clinical Neuroscience, Northern Care Alliance NHS Foundation Trust, Salford Royal, Salford, United Kingdom

[5]Geoffrey Jefferson Brain Research Centre, Division of Neuroscience, School of Biology, Faculty of Biology, Medicine and Health, University of Manchester, United Kingdom.

[6]Department of Medicine, Surgery and Neuroscience, University of Siena, Siena, Italy

[#]Corresponding authors

Background: The integration of molecular profiling in the classification of gliomas is essential to inform on patients' prognosis and guide treatment. Radiomics and Deep Learning (DL) algorithms proved cost-effective non-invasive alternatives to predict genetic alterations. However, these methodologies have limitations due to tumor heterogeneity, variability in imaging acquisition protocols, the use of tumor segmentation masks and the lack of external validation cohorts. In this study, we developed a DL-based methodology to achieve a robust molecular stratification of adult gliomas based on whole-brain 3D diagnostic, structural MRI scans.

Methods: Pre-operative T1-weighted before and after gadolinium administration, T2-weighted and T2-weighted fluid-attenuated inversion recovery sequences and molecular information (IDH and 1p19q mutation status) of 870 gliomas across 14 centers were collected. Two independent DL models were implemented, optimized, and subsequently integrated into a tiered framework for classifying tumors as IDH-WT, IDH-mutant 1p/19q intact and IDH-mutant / 1p/19q co-deleted. The final classifier was tested on 227 patients across 4 centers, which were not included in the training phase.

Results: In the external test set, the tiered model achieved an AUC of 0.93 [IDH-WT], 0.91 [IDH-mutant 1p/19q intact] and 0.89 [IDH-mutant 1p/19q co-deleted], improving on previous approaches. Generated attention maps of enhancing tumor regions identified IDH status; maps of non-enhancing regions distinguished 1p/19q intact from co-deleted tumors.

Conclusions: We demonstrated that the tired DL model developed in this study is effective to stratify gliomas according to their IDH and 1p/19q status, relying on standard 3D MRI scans from multicenter validation cohort. This strategy has the potential to be implemented in clinical practice to predict a broader range of mutations and epigenetic alterations and avoid an invasive procedure in patients with comorbidities or tumors in eloquent areas. The application of the tired DL model will be instrumental to implement clinical trials testing the efficacy and safety of neoadjuvant therapy in patients with glioma.

Keywords: Non-invasive biopsy, Whole-brain 3D structural MRI, Explainable deep learning, Glioma molecular stratification, Personalized medicine.

Key points:

- MRI-based explainable DL predicted IDH and 1p/19q status with AUC ≥ 0.89 on an external multicenter test set.
- Patient-level attention maps highlighted informative regions both within and outside the FLAIR signal.
- Our work supports MRI-based virtual biopsy as a non-invasive effective tool for WHO 2021 glioma molecular classification.

1. Introduction

The classification of glial tumors continuously evolves to incorporate new genetic and epigenetic data and redefine known tumor types or characterize new entities[1]. Molecular profiling is essential to reach an integrated diagnosis and an accurate stratification of patients to inform on outcomes and guide treatment (WHO CNS5; cIMPACT-NOW). However, molecular profiling requires tissue from surgical resections or biopsies, it depends on tissue quality, can be affected by tumor heterogeneity, and requires referral to genomics labs. In addition, patients may not always be eligible for surgery due to comorbidities, lesions can involve eloquent areas and patients can be enrolled in novel clinical trials administering neoadjuvant treatment[2,3,4]. Magnetic resonance imaging (MRI) provides essential information on tumor location, and extent of brain involvement[5,6]. More recently, advances in radiomics[7] and Deep Learning (DL) strategies[8] have converted MRI into a valuable source to predict diagnostic, prognostic and predictive genetic alterations. The value of radiomics[9,10,11], and DL[12,13,14,15,16,17,18] in isolation or combined[19,20,21,22] for diagnostic or prognostic purposes has already been demonstrated in gliomas. However, these modalities have limitations. Radiomics relies on a predefined set of quantitative features extracted from regions of interest (ROI)[23], overlooking tumor microenvironment in disease progression and treatment response. DL models have limited interpretability of their decision-making processes as the consequence of abstraction of features in deeper layers. Deep learning and radiomics models can achieve relatively high accuracy in predicting IDH mutation whilst the identification of the 1p/19q co-deleted status remains challenging without prior knowledge of IDH mutation or tumor grade[11,24,25,26]. Imaging features correlated to the co-deleted are generally more heterogeneous patients, resulting in a lower predictive power and AUC values compared to the classification of the IDH mutation status[27]. Tumor heterogeneity, MRI acquisition protocols and small sample size, all impact on the performance of radiomics and DL modalities. This study has been designed to exploit DL to develop non-invasive prediction of IDH mutation and 1p/19q codeletion using routine pre-operative, structural whole-brain 3D MRI scans. The approach predicted these mutations with high accuracy indicating its applicability to inform on prognosis and treatment strategies in patients who are at high-risk of complications for an invasive procedure and patients who are recruited in trials designed to administer neoadjuvant treatment.

2. Materials and methods

2.1 Data collection

MRI scans were obtained from The Cancer Imaging Archive (TCIA)[28], encompassing data from multiple cohorts: 257 subjects from TCGA-GBM (The Cancer Genome Atlas - Glioblastoma), 199 from TCGA-LGG (The Cancer Genome Atlas - Lower Grade Glioma), 495 from UCSF-PDGM (University of California San Francisco - Preoperative Diffuse Glioma MRI), and 630 from UPENN-GBM (Multi-parametric MRI scans of de novo glioblastoma patients from the University of Pennsylvania Health System). MRI acquisition sites and scanner details for the multi-center TCGA cohort have been previously described[29] while the details for the UPENN and UCSF cohorts are available in TCIA[28]. Following the exclusion of non-preoperative scans, we further excluded 337 subjects to ensure that only those with T1-weighted (T1-w) before and after gadolinium administration (T1-Gd), T2-weighted (T2-w) and T2-weighted fluid-attenuated inversion recovery (FLAIR) were included in our study. Clinical and molecular annotations were retrieved from the study by Ceccarelli et al.[30] (TCGA cohort) and from the TCIA clinical tables (UCSF-PDGM and UPENN-GBM datasets). IDH mutations and 1p/19q co-deletion were used to reclassify tumors according to the 2021 WHO classification into three molecular types: IDH-wildtype (IDH-WT), "IDH-mutant 1p/19q intact", and "IDH-mutant 1p/19q co-deleted". Additional 98 patients were excluded due to incomplete molecular profiling and six due to missing clinical annotations. The resulting dataset comprised 870 patients (Figure 1).

2.2 Study design

The cohort was split into two training sets ("IDH status" and "1p19q co-deleted status") and an external test set. To improve generalizability and ensure robust multicenter validation, a subset of cases from the four institutions (TCGA-CS, TCGA-HT, TCGA-FG, and UPENN-GBM) and with a balanced representation of the three molecular subtypes was set aside for the external testing. It was ensured that enough 1p/19q-co-deleted tumors were included in the external testing set to avoid underrepresentation of this class. Since the participating centers employed different MRI scanners and acquisition protocols[29], this strategy preserved data heterogeneity and better reflected the real-world clinical scenario. The IDH status group included 643 patients, subdivided into 497 IDH-WT and 146 IDH-mutant; 146 IDH-mutant tumors were stratified into 30 1p/19q-co-deleted and 116 1p/19q-intact. The external test set comprised 227 patients, subdivided into 191 IDH-WT, 25 IDH-mutant 1p/19q-intact and 11 IDH-mutant and 1p/19q-co-deleted.

2.3 MRI Processing

MRI scans from TCGA and UPENN-GBM were pre-processed converting the T1-w, T2-w, T1-Gd, and T2-FLAIR DICOM sequences into NIfTI format. Images were re-oriented to standard orientation using fslreorient2std FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki) function, and intensities were re-scaled to a range from 0 to 10,000 with the ImageMath ANTs (https://stnava.github.io/ANTs/) function for 3-dimensional operations. For each patient, T1-w, T2-w and T2-FLAIR were registered to the T1-Gd space, previously resampled to a resolution of 1x1x1 mm^3, using NiftyReg (https://github.com/KCL-BMEIS/niftyreg). Finally, brain extraction of all sequences was performed using a combination of an automated in-house DL tool and FSL. MRI scans from UCSF were already pre-processed as described on the TCIA website[28]. To harmonize this dataset with the other cohorts, re-orienting to standard and intensity re-scaling from 0 to 10,000 have been performed.

2.4 Deep learning model for tiered molecular classification

We adopted a DL-based tiered classification strategy, starting with the prediction of IDH mutations and followed by the assessment of 1p/19q co-deleted. The training phase of the tiered model consisted of two independent modules ("Model A" and "Model B" in Figure 1), automatically optimized through the exploration of different architecture configurations, including the ResNet network structure, the learning rate, the optimizer weight decay, the dropout size and the use of combined MRI modalities (Supplementary Table S1). All networks were trained with 80 epochs and a batch size of 2, using a standard cross-entropy loss function and exploiting the stochastic gradient descent (SGD) to automatically optimize the trainable parameters. Class imbalance was mitigated through oversampling of the minority classes within the training set, applying class weights (0.4 for the majority class, 0.6 for the minority class). For each module, 324 different configurations were inspected and evaluated on 20% of the entire training set (validation set), stratified to account for class imbalance and center distribution. The best-performing configurations for both models (i.e. Model A for IDH status and Model B for 1p/19q co-deleted status prediction), as identified on the corresponding validation set with an AUC > 0.8, were used to perform inference on the external test set. All possible tiered combinations of the two models were then evaluated. The final cascade model ("Tiered Model" in Figure 1) was selected as the model achieving the best balanced accuracy to account for class imbalance in the classification of the three molecular subtypes: "IDH-WT", "IDH-mutant 1p/19q intact" and "IDH-mutant 1p/19q co-deleted".

2.5 Model explainability

Guided backpropagation was applied to enhance the explainability of the tiered model. Starting from the last layer of the residual network, single-patient attention maps (AM) were generated during the inference to highlight the most relevant features that influenced the model's prediction. Patients' maps were z-scored for standardization by subtracting the mean and dividing by the standard deviation of

only voxels with positive values. For visualization purposes, maps were smoothed using a Gaussian kernel (σ = 1 voxel) with FSL's ImageMaths and masked with the subject's brain mask to include only within-brain voxels. To filter smaller clusters in attention maps as they may represent noise, DLH and volume smoothness parameters were estimated using FSL's SmoothEstimate and applied as inputs for a cluster-based thresholding with FSL's Cluster. Only voxels above the 95th percentile of the z-score distribution contributed to the clustering. Among these, only clusters surviving a cluster-wise significance threshold of $p < 0.01$ and with a minimum volume of 300 mm^3 were retained in the attention map following thresholding.

2.6 Tumor segmentation

To identify the extent of overlap of each thresholded map with the glioma, tumor segmentation masks were generated using a nnU-Net framework[31]. The dataset was converted into the specific format required by the framework and then used for segmentation. Pretrained weights from the KAIST MRI Lab's BraTS 2021 submission[32] were adopted using the nnU-Net trainer "V2BraTSRegions_DA4_BN_BD_largeUnet_Groupnorm". The obtained masks' labels were the following: "edema" (ED), "enhancing tumor" (ET), "non-enhancing tumor" (NET), and "outside-of-the-tumor" (OT). The label 'outside-of-the-tumor' (OT) referred to brain regions located beyond the FLAIR hyperintense peritumoral edema, corresponding to radiologically normal-appearing tissue.

2.7 Statistical Analysis

The percentage of voxels from each patient's attention map falling within the four tumor labels was calculated. Then, Kruskal-Wallis (KW) test was performed between the three molecular groups, considering only correctly predicted patients from the external test set. Normality of the distributions was previously assessed using the Shapiro-Wilk test. Following KW, Dunn's post-hoc test was performed for pairwise comparisons between glioma classes, with Benjamini-Hochberg false discovery rate (FDR) correction applied to control for multiple comparisons.

3. Results

3.1 Patients characteristics

The dataset consisted of 688 IDH-WT and 182 IDH-mutant patients including 141 1p/19q-intact and 41 1p/19q-co-deleted subjects). Median age for IDH-WT patients was 62 years (ranging 18-94 years), 37 years for IDH-mutant 1p/19q-intact (range 17-71 years) and 53 years for IDH-mutant 1p/19q-co-deleted ones (ranging 19-75 years) (Table 1). Male and female patients were equally distributed across tumor types, with a slight predominance of male patients. Median survival was higher for IDH-mutant

1p/19q-intact and co-deleted classes (26.5 and 25 months, respectively) and lower for the IDH-WT group (12.8 months).

3.2 The tiered model performance

The optimal combination of the two independent DL modules for the prediction of IDH mutations and 1p/19q co-deleted status achieved an overall 84% accuracy, 75% balanced accuracy and 85% F1-weighted score on the external test set. When evaluated separately in the validation set, the IDH prediction module reached an accuracy and balanced accuracy of 86%, a F1-weighted score of 87% and an AUC of 0.93. This model leveraged the combination of the T1-Gd and FLAIR sequences, a ResNet-10 network, an optimizer weight decay of 3e-04, a learning rate of 1e-04 and a dropout of 0.3. The 1p/19q prediction module achieved 83% accuracy, 78% balanced accuracy, a F1-weighted score of 83% and an AUC of 0.81, exploiting a combination of the T1-Gd, T1-w, T2-w and FLAIR sequences, a ResNet-18 network, an optimizer weight decay of 3e-04, a learning rate of 1e-04 and a dropout of 0.3. Per-class metrics were also calculated. The tiered model reached an accuracy of 86% for the IDH-WT, 86% for the IDH-mutant 1p/19q-intact and 96% for the IDH-mutant 1p/19q-co-deleted classes, with an AUC of 0.93 [0.88-0.97], 0.91 [0.84-0.96] and 0.89 [0.79-0.98], respectively (Figure 2).

3.3 Model explainability

The generated patient-specific attention maps showed that the model focused on recurrent MRI patterns, primarily derived from the T1-Gd and FLAIR sequences. The maps showed higher values within contrast uptake regions (on the T1-Gd sequence), inhomogeneous intensity regions on the FLAIR sequence, or in putative tumor-induced deformations of the image (i.e. midline shift and compression of the ventricles). Figure 3 shows a representative example of three patients from the external test set, correctly classified by the model as IDH-WT, IDH-mutant 1p/19q-intact and IDH-mutant 1p/19q-co-deleted. In the IDH-WT group, the AM was most significantly focused on the peritumoral region (z-scores=12-30), which is known to be infiltrated and highly vascular in most of these tumors and it described a jagged pattern. In the IDH-mutant group, AMs mainly highlighted regions within the tumor epicenter: the 1p/19q-intact map showed higher significance in both enhancing and non-enhancing areas, as well as in a tumor-induced deformation near the mass (z-scores=10-26), while the 1p/19q-co-deleted map focused more on signal heterogeneity in the non-enhancing region (z-scores=8-22). AMs across the axial, coronal and sagittal planes and slices as well as sequences are shown in Supplementary Materials (Figure S1, S2 and S3).

AMs of misclassified patients still focused on the tumor mass and its surroundings. However, cases exhibiting atypical attention map patterns within their molecular subtype, particularly regarding peritumoral contour shapes or hyperintense regions, tended to be incorrectly classified by the model. Misclassified cases included a large tumors IDH-WT tumor extending to a ventricle and with smooth

tumor margins and homogeneous texture of peritumoral brain on T1-Gd that was incorrectly predicted by the model as IDH-mutant 1p/19q-intact (Supplementary Figure S4), an IDH-mutant 1p/19q-intact patient with a large, non-enhancing tumor misclassified as IDH-mutant 1p/19q-co-deleted (Supplementary Figure S5), and an IDH-mutant 1p/19q-co-deleted case with an intense and jagged peritumoral region, misclassified as IDH-WT (Supplementary Figure S6). Lastly, the attention maps identified discriminative patterns within brain regions appearing normal on conventional imaging—extending from areas adjacent to the tumor toward more distant regions—suggesting that subtle cues beyond the visible tumor margins contributed to molecular classification (Figure 4).

We also assessed the distribution of AMs in each molecular group using four tumor segmentation masks (i.e. "edema" (ED), "enhancing tumor" (ET), "non-enhancing tumor" (NET), and "outside-of-the-tumor" (OT)) to check for the presence of per class-significant differences. Only model correctly classified patients from the external test set (n=163 IDH-WT, n=21 IDH-Mutant 1p/19q-intact, n=6 IDH-Mutant 1p/19q-co-deleted) were considered. For the IDH-WT class, most of the attention maps' voxels were located inside the enhancing tumor (ET: 53%) and outside of the tumor regions (OT: 22%) (ED: 18%; NET: 7%) (Table 2). For both the IDH-mutant classes, attention maps revealed a principal focus on the non-enhancing tumor region (1p/19q intact: 45%; 1p/19q co-deleted: 26%) and on regions outside of the tumor (1p/19q intact: 31%; 1p/19q co-deleted: 62%). Edema and enhancing tumor regions showed lower importance (ED 1p/19q intact: 22%; ED 1p/19q co-deleted: 10%; ET 1p/19q intact: 3%; ET 1p/19q co-deleted: 2%). The distribution of attention across tumor regions differed significantly between molecular subgroups in the non-enhancing and enhancing tumor regions (p-adj: 3.14e-05 and 9.52e-14, respectively). A significant difference was present in the NET region for the "IDH-WT vs IDH-mutant 1p/19q intact" group (p-adj: 0.000017) and for the "IDH-mutant 1p/19q intact vs IDH-mutant 1p/19q co-deleted" group (p-adj: 0.026840) (Table 2). In the ET region, the IDH-WT class significantly differentiated from the IDH-mutant 1p/19q-intact (p-adj: 1.234712e-11) and from the IDH-mutant 1p/19q-co-deleted class (p-adj: 0.000154).

4. Discussion

This study explored the feasibility of performing MRI-based virtual biopsies using a tiered DL framework to predict IDH mutation and 1p/19q co-deletion and allow pre-operative accurate molecular stratification of patients with glioma. The present methodology integrated two distinct predictive modules, each optimized independently to maximize their performance on a specific molecular target and then combined the prediction of IDH and 1p19q co-deleted. The final tiered model used routine structural MRI scans (T1-Gd, FLAIR, T2-w and T1-w) and two different ResNet architectures. The model achieved 84% accuracy in the external test set. Individual-class AUC confirmed the high performance reached for the three groups (AUC of 0.93 [95% CI: (0.88, 0.97)] for IDH-WT, 0.91 [95% CI: (0.84, 0.96)] for 1p/19q intact and 0.89 [95% CI: (0.79, 0.98)] for 1p/19q co-deleted). A similar study on IDH and 1p/19q using interpretable DL[16] reported a prediction performance of 23% in the estimation of the IDH-mutant 1p/19q-co-deleted class in the TCGA external independent test cohort using a baseline logistic regression model. Accuracy dropped to 0% when employing a tiered or 3-class anatomical model whilst our model achieved 96% accuracy for the same tumor subtype, proving the effectiveness of the approach. Tiered approaches for the classification of IDH and 1p/19q co-deleted status are not widely adopted. Most studies primarily addressed the detection of each individual biomarker[17,33,13,18]. Previous studies had limitations including lack of independent datasets, smaller cohorts, previous knowledge of IDH mutation status or grade, lack of model's interpretability and use of 2D-imaging datasets. In contrast, the strategy proposed in this study offered several advantages compared to a simple multitarget classification; it reduces the complexity of classification analyses by enabling task-specific specialization, it targets interpretability, mitigating the risk of overfitting and allowing generalization. Furthermore, the modular structure of the tiered model has the potential to perform on other genetic alterations still remaining agnostic to tumor grade. Another significant strength of this approach lies in its use of conventional 3D anatomical MRI sequences, which are routinely acquired in clinical settings. Unlike methods requiring advanced MRI modalities, our model achieved high performance using standard imaging sequences, especially for the prediction of 1p/19q co-deleted status, enhancing its potential use in a clinical environment. Imaging analysis in similar studies used T1 post-contrast (T1-Gd) and FLAIR sequences only, due to their capability to highlight different tumor features. However, in our optimization process, we found that the native T1-weighted and T2-weighted sequences also carry significant discriminative information for biomarker classification, and more so for IDH-mutant tumors. This finding suggests that conventional MRI sequences, when properly optimized, can provide valuable molecular insights. Radiomics or DL segmentation methods directly focus on the tumor ROI and therefore potentially miss relevant features from the microenvironment of the normal appearing brain. Our method proved that distant-from-the-tumor regions are also informative to identify IDH and 1p/19 status. The generation of explainable single patient AMs, paired with the original sequences, allowed the identification of the regions that proved relevant to molecular

stratification. The maps highlighted tumor and peritumoral regions, especially in IDH-WT cases, and non-enhancing tumor areas and regions supposed to be impacted by tumor-induced changes, especially in IDH-mutant cases, a finding consistent with other studies[34,16,17]. Statistical analysis showed that regions distant from the tumor bulk significantly discriminated IDH-WT from IDH-mutant cases, whereas NET areas allowed to discern 1p/19q intact from co-deleted tumors. Conversely, misclassified patients frequently showed unusual imaging features for molecular subtype. These cases may reflect both biological variability or tumor heterogeneity, which could effectively reduce the accuracy of the model. Interestingly, some maps highlighted infiltration pathways starting from the tumor area and proceeding towards more distant areas. This result should be further evaluated to assess an association between distant infiltrations and progression or relapse.

The option to exploit diagnostic, structural imaging to perform a virtual biopsy and identify genetic mutations has the potential to implement clinical practice in the field of neuro-oncology. The reasons are multifold: i) although surgery remains the cornerstone for gliomas, it carries risks of neurological deficits particularly for tumors located in eloquent areas; ii) patients with co-morbidities are at high risk of developing surgery-related complications and iii) deep-seated locations are difficult to access. In these instances, patients are often elected for a diagnostic biopsy, but tissue may not always be representative of the whole tumor or is too small to perform genetic and epigenetic tests. Finally, recent evidence proved that mechanical injury and hypoxia caused by surgical manipulation can impact on tumor cells and TME, increasing invasion and proliferation and pro-tumoral inflammatory response[35]. Neoadjuvant modalities aimed to administer chemotherapy or radiotherapy prior to surgery are well-established in other types of solid tumor and commonly delivered to patients with brain metastasis but only now emerging as promising modalities to improve surgical outcomes in patients with glioma[36,37]. Their rationale is to target neoplastic cells in their native TME that are unaffected by surgical injury. Non-invasive treatments rely on accurate molecular profiling to ensure appropriate tumor classification and to maximize therapeutic efficacy[36,37]. The neoadjuvant trials **P**re**O**perative **B**rain **I**rradiation in **G**lioblastoma (POBIG) (NCT03582514)[38] and the **P**re**O**pe**R**a**T**ive pre**RA**diotherapy Ttfields (PORTRAIT) (NTC06136611) are currently recruiting in the United Kingdom.

The present study has limitations. Despite data collection included multiple cohorts, a limitation of this study relates to the small number of IDH-mutant 1p/19q-co-deleted cases. A strategy to mitigate such limitation was to account for class imbalance and to perform oversampling of low frequent classes during the training phase. However, future works should explore the use of generative artificial intelligence for synthetic data generation, especially for rare cases.

In conclusion, this study represents a step toward introducing MRI-based virtual biopsy as a complementary tool to overcome the challenges of tissue sampling and to support neoadjuvant treatment trials where patients enrollment is based on pre-operative neuroimaging. Our proof-of-concept study represents the foundation to expand virtual testing to other genes that are essential to the diagnosis, and prognostic and predictive stratification.

5. Funding

This work was financed by RI&SS ARTES 4.0 N.5 - 2023 – Financed by European Union – Next Generation EU.

Conflict of interest statement. The authors have no conflicts of interest to report with respect to the content of this manuscript.

6. References

1. Louis DN, Perry A, Wesseling P, et al. The 2021 WHO classification of tumors of the central nervous system: A summary. *Neuro Oncol*. 2021;23(8):1231-1251. doi:10.1093/neuonc/noab106
2. Pasternak KA, Schwake M, Warneke N, et al. Evaluation of 311 contemporary cases of stereotactic biopsies in patients with neoplastic and non-neoplastic lesions-diagnostic yield and management of non-diagnostic cases. doi:10.1007/s10143-020-01394-0/Published
3. Katzendobler S, Do A, Weller J, et al. Diagnostic Yield and Complication Rate of Stereotactic Biopsies in Precision Medicine of Gliomas. *Front Neurol*. 2022;13. doi:10.3389/fneur.2022.822362
4. Alhalabi OT, Sahm F, Unterberg AW, Jakobs M. The molecular diagnostic yield of frame-based stereotactic biopsies in the age of precision neuro-oncology: a cross-sectional study. *Acta Neurochir (Wien)*. 2023;165(9):2479-2487. doi:10.1007/s00701-023-05742-z
5. Weller M, van den Bent M, Preusser M, et al. EANO guidelines on the diagnosis and treatment of diffuse gliomas of adulthood. *Nat Rev Clin Oncol*. 2021;18(3):170-186. doi:10.1038/s41571-020-00447-z
6. 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;28(8):3306-3317. doi:10.1007/s00330-018-5314-5
7. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.Nature Publishing Group*. 2017;14(12):749-762. doi:10.1038/nrclinonc.2017.141
8. Najjar R. Redefining Radiology: A Review of Artificial Intelligence Integration in Medical Imaging. *Diagnostics.Multidisciplinary Digital Publishing Institute (MDPI)*. 2023;13(17). doi:10.3390/diagnostics13172760
9. Fan H, Luo Y, Gu F, et al. Artificial intelligence-based MRI radiomics and radiogenomics in glioma. *Cancer Imaging.BioMed Central Ltd*. 2024;24(1). doi:10.1186/s40644-024-00682-y
10. Lu CF, Hsu FT, Hsieh KLC, et al. Machine learning–based radiomics for molecular subtyping of gliomas. *Clinical Cancer Research*. 2018;24(18):4429-4436. doi:10.1158/1078-0432.CCR-17-3445
11. Sun C, Fan L, Wang W, et al. Radiomics and Qualitative Features From Multiparametric MRI Predict Molecular Subtypes in Patients With Lower-Grade Glioma. *Front Oncol*. 2022;11. doi:10.3389/fonc.2021.756828
12. Chang K, Bai HX, Zhou H, et al. Residual convolutional neural network for the determination of IDH status in low- and high-grade gliomas from mr imaging. *Clinical Cancer Research*. 2018;24(5):1073-1081. doi:10.1158/1078-0432.CCR-17-2236
13. Chang P, Grinband J, Weinberg BD, et al. Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. *American Journal of Neuroradiology*. 2018;39(7):1201-1207. doi:10.3174/ajnr.A5667

14. Holtkamp M, Parmar V, Hosch R, et al. AI-guided virtual biopsy: Automated differentiation of cerebral gliomas from other benign and malignant MRI findings using deep learning. *Neurooncol Adv*. 2025;7(1). doi:10.1093/noajnl/vdae225

15. Yogananda CGB, Shah BR, Yu FF, et al. A novel fully automated MRI-based deep-learning method for classification of 1p/19q co-deletion status in brain gliomas. *Neurooncol Adv*. 2020;2:IV42-IV48. doi:10.1093/noajnl/vdaa066

16. Cluceru J, Interian Y, Phillips JJ, et al. Improving the noninvasive classification of glioma genetic subtype with deep learning and diffusion-weighted imaging. *Neuro Oncol*. 2022;24(4):639-652. doi:10.1093/neuonc/noab238

17. Van Der Voort SR, Incekara F, Wijnenga MMJ, et al. Combined molecular subtyping, grading, and segmentation of glioma using multi-task deep learning. *Neuro Oncol*. 2023;25(2):279-289. doi:10.1093/neuonc/noac166

18. Chakrabarty S, Lamontagne P, Shimony J, Marcus DS, Sotiras A. MRI-based classification of IDH mutation and 1p/19q codeletion status of gliomas using a 2.5D hybrid multi-task convolutional neural network. *Neurooncol Adv*. 2023;5(1). doi:10.1093/noajnl/vdad023

19. Li Z, Wang Y, Yu J, Guo Y, Cao W. Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma. *Sci Rep*. 2017;7(1). doi:10.1038/s41598-017-05848-2

20. Choi YS, Bae S, Chang JH, et al. Fully automated hybrid approach to predict the IDH mutation status of gliomas via deep learning and radiomics. *Neuro Oncol*. 2021;23(2):304-313. doi:10.1093/neuonc/noaa177

21. Haubold J, Hosch R, Parmar V, et al. Fully automated mr based virtual biopsy of cerebral gliomas. *Cancers (Basel)*. 2021;13(24). doi:10.3390/cancers13246186

22. Zhang H, Fan X, Zhang J, et al. Deep-learning and conventional radiomics to predict IDH genotyping status based on magnetic resonance imaging data in adult diffuse glioma. *Front Oncol*. 2023;13. doi:10.3389/fonc.2023.1143688

23. Wang Z, Wang L, Wang Y. Radiomics in glioma: emerging trends and challenges. *Ann Clin Transl Neurol.John Wiley and Sons Inc*. 2025;12(3):460-477. doi:10.1002/acn3.52306

24. van der Voort SR, Incekara F, Wijnenga MMJ, et al. Predicting the 1p/19q codeletion status of presumed low-grade glioma with an externally validated machine learning algorithm. *Clinical Cancer Research*. 2019;25(24):7455-7462. doi:10.1158/1078-0432.CCR-19-1127

25. Yan J, Zhang S, Sun Q, et al. Predicting 1p/19q co-deletion status from magnetic resonance imaging using deep learning in adult-type diffuse lower-grade gliomas: a discovery and validation study. *Laboratory Investigation*. 2022;102(2):154-159. doi:10.1038/s41374-021-00692-5

26. Li F, Li Z, Xu H, et al. Prediction of 1p/19q state in glioma by integrated deep learning method based on MRI radiomics. *BMC Cancer*. 2025;25(1). doi:10.1186/s12885-025-14454-9

27. Richter V, Ernemann U, Bender B. Novel Imaging Approaches for Glioma Classification in the Era of the World Health Organization 2021 Update: A Scoping Review. *Cancers (Basel).Multidisciplinary Digital Publishing Institute (MDPI)*. 2024;16(10). doi:10.3390/cancers16101792

28. The Cancer Imaging Archive (TCIA). Accessed August 11, 2025. https://www.cancerimagingarchive.net/browse-collections/

29. Bakas S, Akbari H, Sotiras A, et al. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci Data*. 2017;4. doi:10.1038/sdata.2017.117

30. Ceccarelli M, Barthel FP, Malta TM, et al. Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. *Cell*. 2016;164(3):550-563. doi:10.1016/j.cell.2015.12.028

31. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods*. 2021;18(2):203-211. doi:10.1038/s41592-020-01008-z

32. Brats21_KAIST_MRI_Lab. Accessed August 11, 2025. http://github.com/rixez/Brats21_KAIST_MRI_Lab

33. Byeon Y, Park YW, Lee S, et al. Interpretable multimodal transformer for prediction of molecular subtypes and grades in adult-type diffuse gliomas. *NPJ Digit Med*. 2025;8(1). doi:10.1038/s41746-025-01530-4

34. Villanueva-Meyer JE, Bakas S, Tiwari P, et al. Artificial Intelligence for Response Assessment in Neuro Oncology (AI-RANO), part 1: review of current advancements. *Lancet Oncol.Elsevier Ltd*. 2024;25(11):e581-e588. doi:10.1016/S1470-2045(24)00316-4

35. Pöhlmann J, Weller M, Marcellusi A, et al. High costs, low quality of life, reduced survival, and room for improving treatment: an analysis of burden and unmet needs in glioma. *Front Oncol.Frontiers Media SA*. 2024;14. doi:10.3389/fonc.2024.1368606

36. Nabian N, Ghalehtaki R, Zeinalizadeh M, Balaña C, Jablonska PA. State of the neoadjuvant therapy for glioblastoma multiforme—Where do we stand? *Neurooncol Adv.Oxford University Press*. 2024;6(1). doi:10.1093/noajnl/vdae028

37. Waqar M, Trifiletti DM, McBain C, et al. Early Therapeutic Interventions for Newly Diagnosed Glioblastoma: Rationale and Review of the Literature. *Curr Oncol Rep.Springer*. 2022;24(3):311-324. doi:10.1007/s11912-021-01157-0

38. Waqar M, Roncaroli F, Djoukhadar I, et al. Study protocol: PreOperative Brain Irradiation in Glioblastoma (POBIG) – A phase I trial. *Clin Transl Radiat Oncol*. 2023;39. doi:10.1016/j.ctro.2023.100585

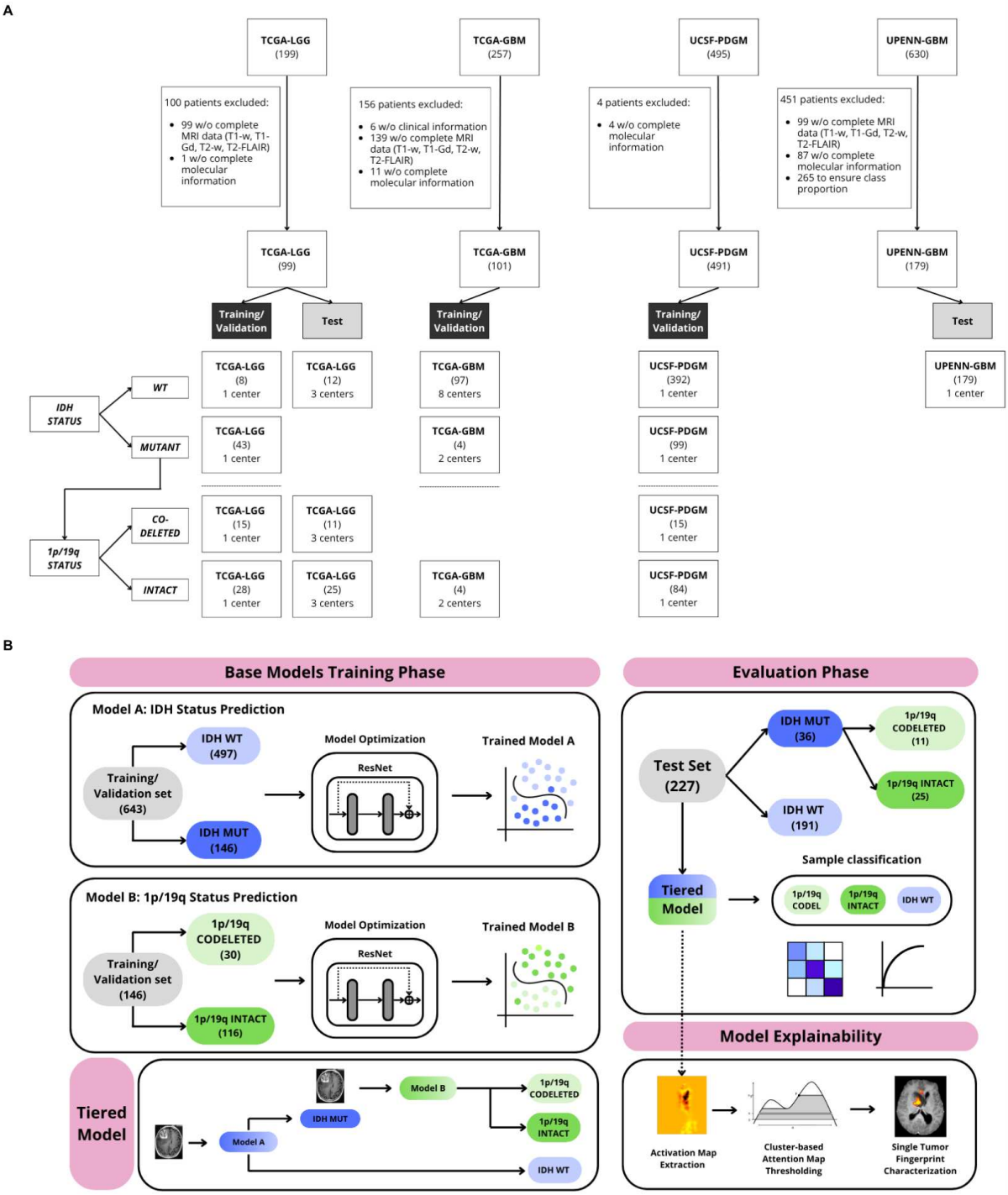


Figure 1. (A) Overview of exclusion criteria and final patient distribution across training/validation and external test sets for the IDH and 1p/19q-codel status prediction. (B) Computational workflow for the classification of glioma patients into "IDH-WT", "IDH-mutant 1p/19q co-deleted" and "IDH-mutant 1p/19q intact" molecular subtypes. The training phase included two independent optimized modules (A and B) to discriminate between IDH and 1p/19q co-deletion status, respectively. Finally, the evaluation phase classified patients in the external test set using sequentially the two optimized modules.

A

Class	AUC (C.I.)	Precision	Sensitivity (Recall)	Specificity	FPR	FNR	Accuracy	Balanced Accuracy	F1-score
IDH-WT	0.93 [0.88-0.97]	0.97	0.85	0.86	0.14	0.15	0.86	0.86	0.91
IDH-Mut 1p/19q intact	0.91 [0.84-0.96]	0.44	0.84	0.87	0.13	0.16	0.86	0.85	0.58
IDH-Mut 1p/19q codel	0.89 [0.79-0.98]	0.55	0.55	0.98	0.02	0.45	0.96	0.76	0.55

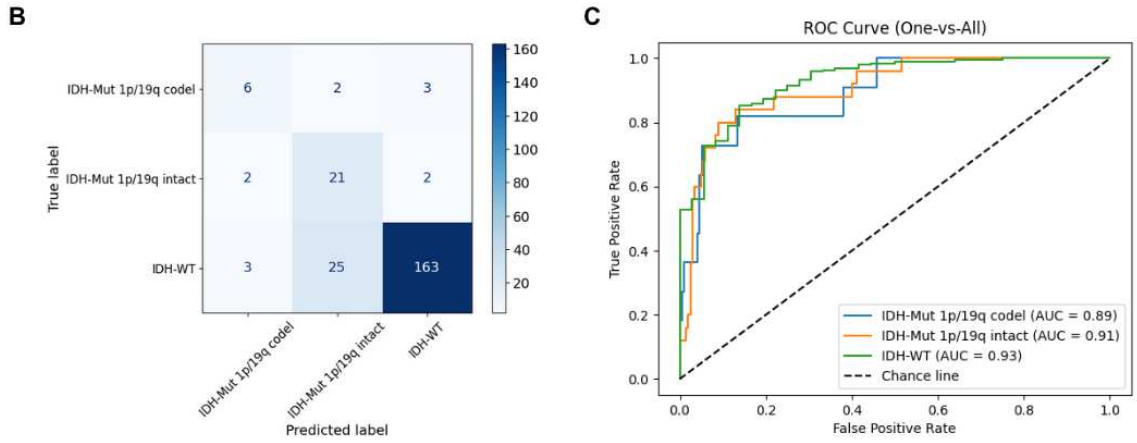


Figure 2. (A) Performance metrics of the tiered classification model (AUC with 95% confidence interval, precision, sensitivity, specificity, False Positive Rate (FPR), False Negative Rate (FNR), accuracy, balanced accuracy and F1-score), calculated for each class. (B) Confusion matrix evaluated on the external test set. (C) Receiver Operating Characteristic (ROC) curves and corresponding AUC values evaluated on the external test set.

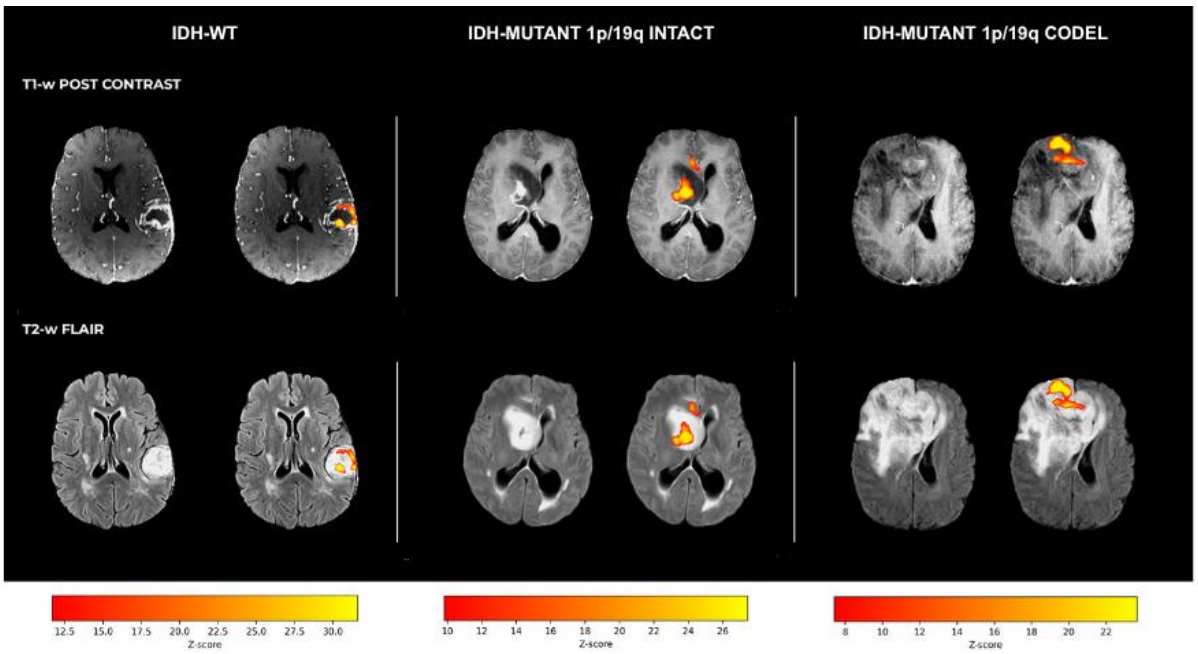


Figure 3. Patient-level attention maps for an IDH-WT, IDH-Mutant 1p/19q intact and IDH-Mutant 1p/19q-codel patient, correctly classified.

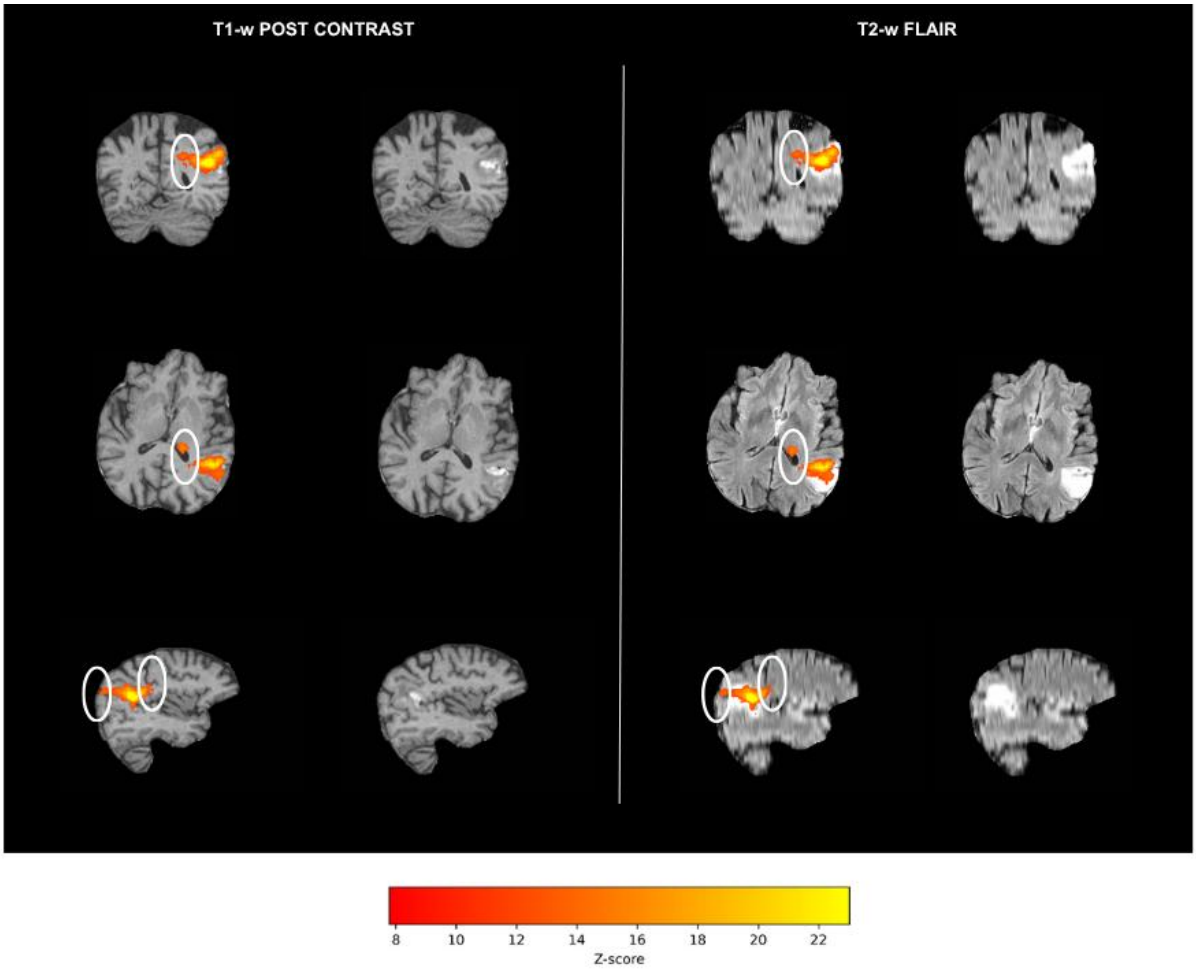


Figure 4. Infiltration pathways (highlighted in white circles) proceeding towards distant-from-the-tumor areas, for a correctly classified IDH-Mutant 1p/19q-codel case.

Molecular subtype	Cases(%)	Median Age (years)	Gender	Grade	Median Survival (months)
IDH-WT	688 (79%)	62 (18-94)	M: 409 (59%) F: 279 (41%)	G2: 14 (2%) G3: 30 (4%) G4: 644 (94%)	12.8 (0.2-135.6)
IDH-Mutant 1p/19q intact	141 (16%)	37 (17-71)	M: 77 (55%) F: 64 (45%)	G2: 62 (44%) G3: 52 (37%) G4: 26 (19%)	26.5 (0.1-158.4)
IDH-Mutant 1p/19q codeleted	41 (5%)	53 (19-75)	M: 22 (54%) F: 19 (46%)	G2: 24 (59%) G3: 15 (37%) G4: 2 (4%)	25.0 (0.2-96.4)

Table 1. Clinical characteristics of the sample set arranged by IDH and 1p/19q co-deletion status.

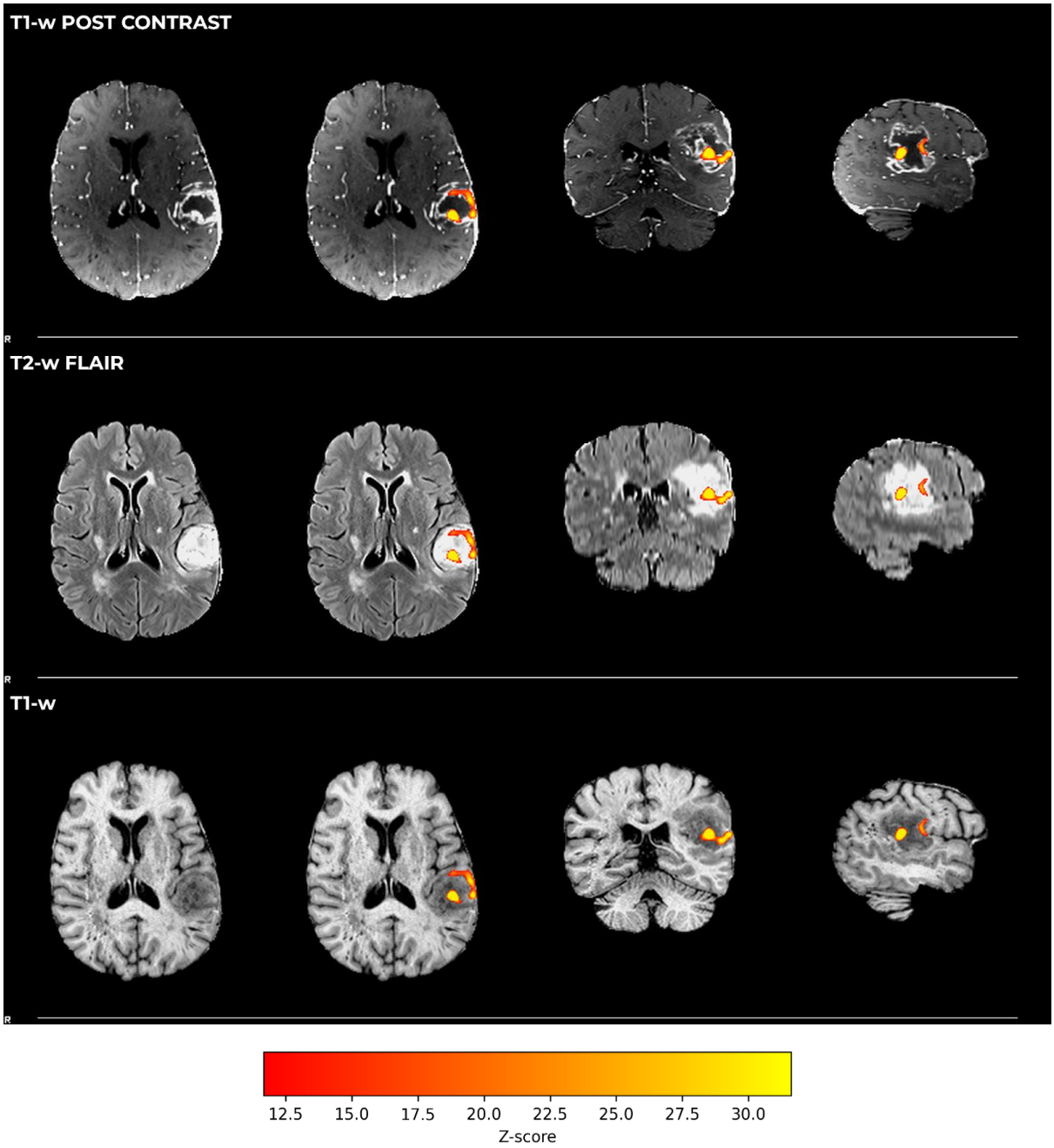
A

Class	Edema [ED] (%)	Non-Enhancing Tumor [NET] (%)	Enhancing Tumor [ET] (%)	Outside-of-the-tumor [OT] (%)
IDH-WT	17.54 ± 15.21	6.86 ± 6.51	53.27 ± 26.54	22.33 ± 24.04
IDH-Mutant 1p/19q intact	21.66 ± 13.79	44.82 ± 29.23	3.02 ± 5.36	30.5 ± 37.08
IDH-Mutant 1p/19q codel	10.14 ± 11.06	25.75 ± 38.0	2.11 ± 4.53	62.0 ± 41.19

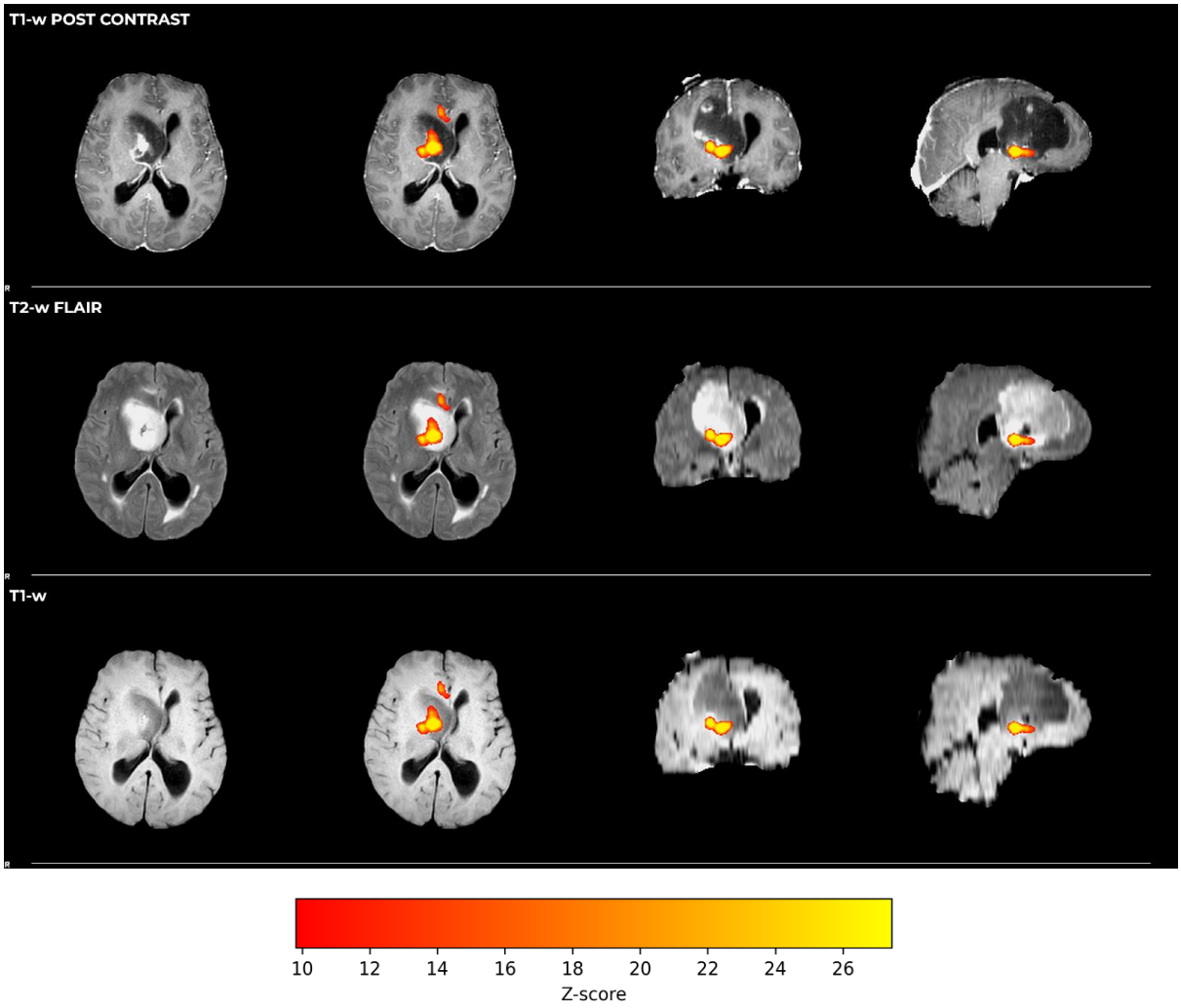
B

Edema [ED]			
	IDH-WT	**IDH-Mutant 1p/19q**	**IDH-Mutant 1p/19q**
IDH-WT	1	0,17367	0,17367
IDH-Mutant 1p/19q	-	1	0,13359
IDH-Mutant 1p/19q	-	-	1
Non-Enhancing Tumor [NET]			
	IDH-WT	**IDH-Mutant 1p/19q**	**IDH-Mutant 1p/19q**
IDH-WT	1	**0,00017**	0,91520
IDH-Mutant 1p/19q	-	1	**0,02684**
IDH-Mutant 1p/19q	-	-	1
Enhancing Tumor [ET]			
	IDH-WT	**IDH-Mutant 1p/19q**	**IDH-Mutant 1p/19q**
IDH-WT	1	**0,00001**	**0,00015**
IDH-Mutant 1p/19q	-	1	0,98805
IDH-Mutant 1p/19q	-	-	1
Outside-of-the-tumor [OT]			
	IDH-WT	**IDH-Mutant 1p/19q**	**IDH-Mutant 1p/19q**
IDH-WT	1	0,88898	0,05515
IDH-Mutant 1p/19q	-	1	0,06092
IDH-Mutant 1p/19q	-	-	1

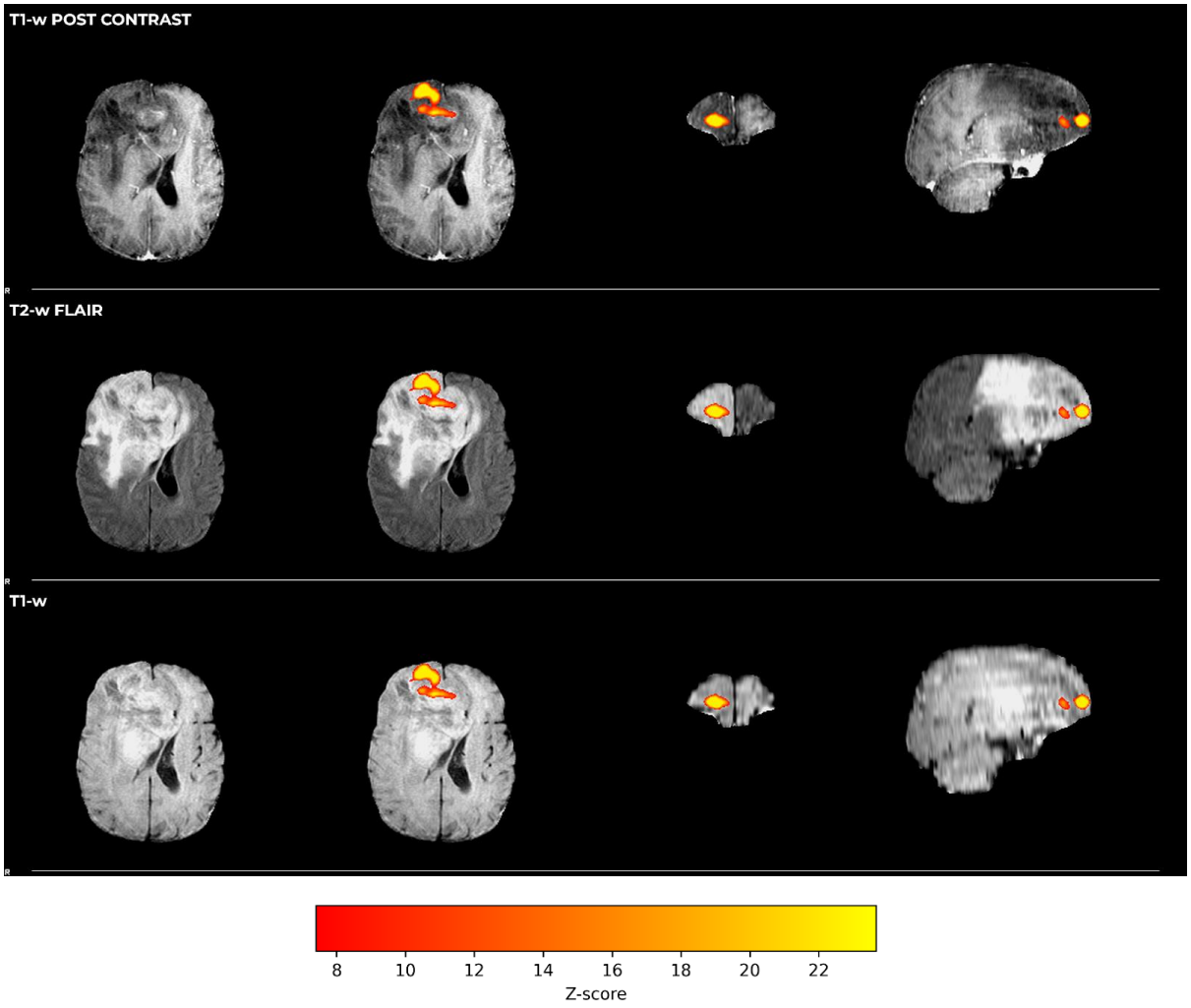
Table 2. (A) Mean percentage and standard deviation of attention map voxels intersecting with outside-of-the-tumor (OT), edema (ED), non-enhancing tumor (NET), and enhancing tumor (ET) regions. (B) Post-hoc Dunn's tests pairwise comparisons between glioma molecular subgroups for overlap percentages in ED, NET, ET and OT regions. Significant differences are indicated in bold (p-adj < 0.05).



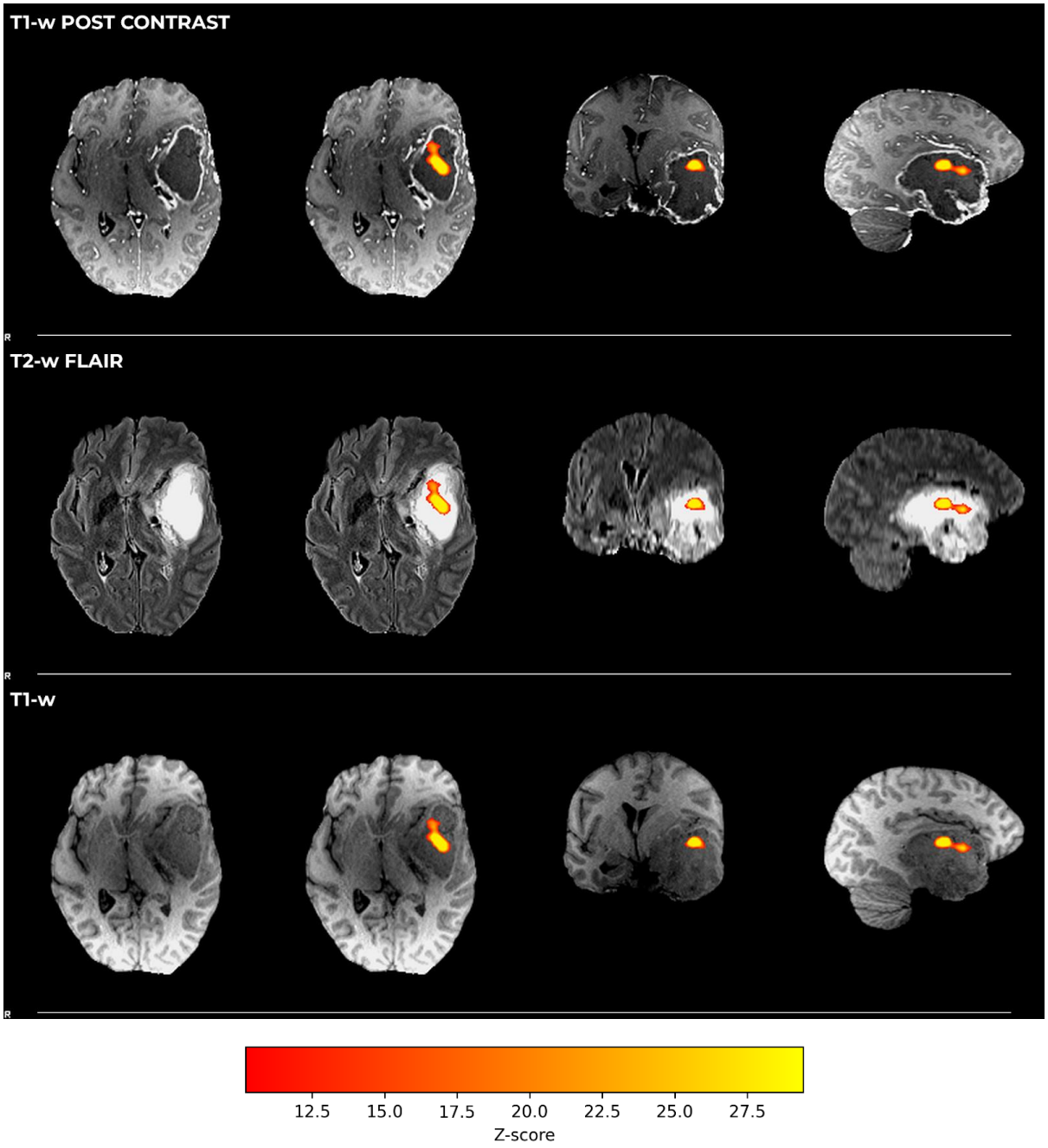
Supplementary Figure S1. Patient-level attention map for an IDH-WT patient correctly classified.



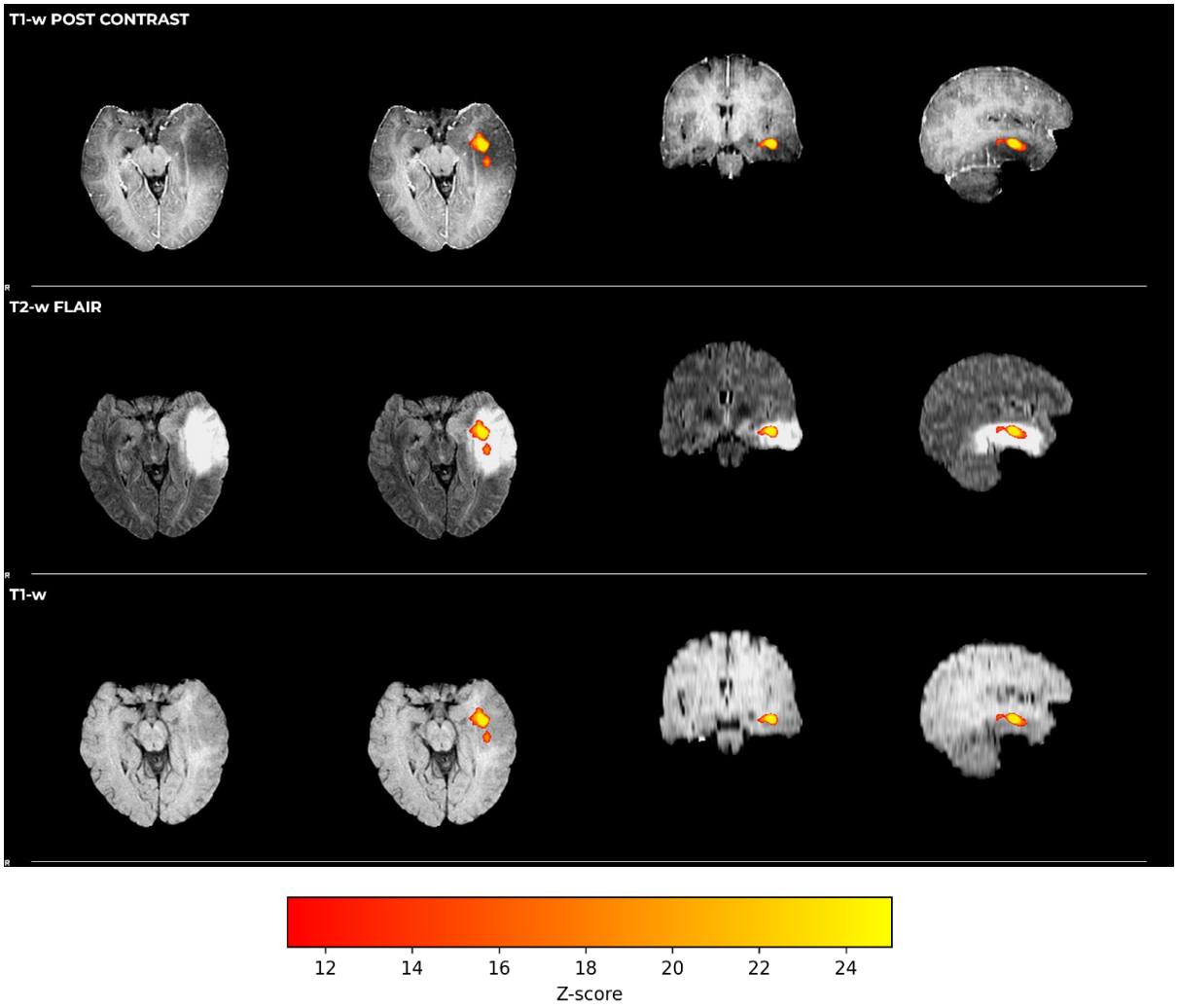
Supplementary Figure S2. Patient-level attention map for an IDH-Mutant 1p/19q-intact patient correctly classified.



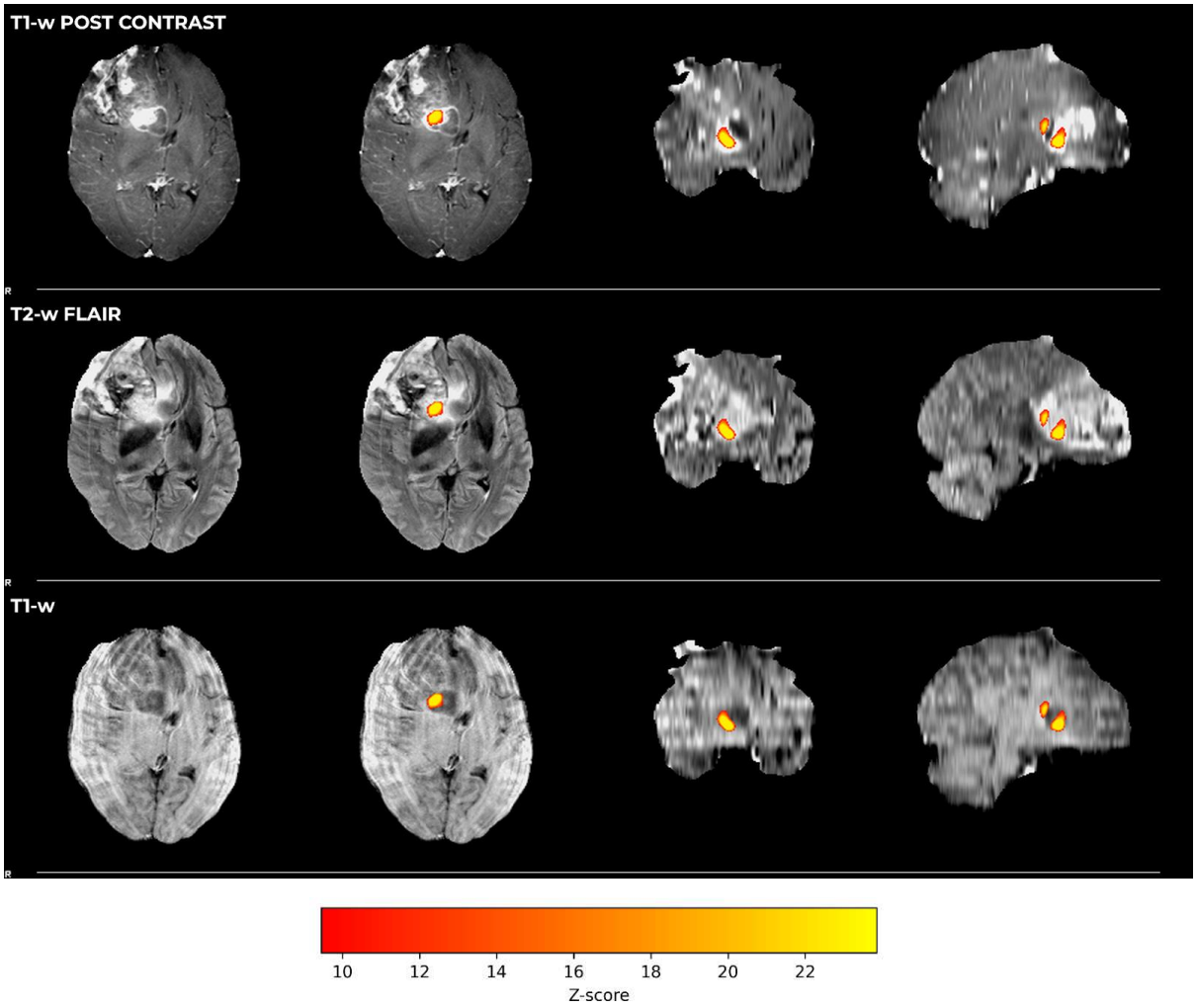
Supplementary Figure S3. Patient-level attention map for an IDH-Mutant 1p/19q-codel patient correctly classified.



Supplementary Figure S4. Patient-level attention map for an IDH-WT patient misclassified as IDH-Mutant 1p/19q-intact.



Supplementary Figure S5. Patient-level attention map for an IDH-Mutant 1p/19q-intact patient misclassified as IDH-Mutant 1p/19q-codel.



Supplementary Figure S6. Patient-level attention map for an IDH-Mutant 1p/19q-codel patient misclassified as IDH-WT.

Parameters	Values
Batch size	2
Epoch number	80
Optimizer	SGD
Scheduler	Polynomial
Input modalities	[T1-Gd, FLAIR]; [T1-Gd, FLAIR, T2-w]; [T1-Gd, FLAIR, T1-w]; [T1-Gd, FLAIR, T1-w, T2-w]
Network	ResNet-10; ResNet-18; ResNet-34
Dropout size	0.3; 0.5; 0.7
Optimizer weight decay	3e-04; 3e-05; 3e-06
Learning rate	1e-03; 1e-04; 1e-05

Supplementary Table S1. Hyperparameters selected during the optimization phase of both modules.