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.

# Voxel-wise Radiomics Habitat Modeling of Multiparametric MRI for IDH Prediction and Survival Stratification in Glioma: a multicenter retrospective study

**Tao Zheng**
The First Hospital of Qinhuangdao
**Linsha Yang**
The First Hospital of Qinhuangdao
**Duo Zhang**
Baoding No. 1 Central Hospital
**Juan Du**
The First Hospital of Qinhuangdao
**Qinglei Shi**
Chinese University of Hong Kong (Shenzhen) & Shenzhen Research Institute of Big Data
**Defeng Liu**
defengliu2023@163.com

The First Hospital of Qinhuangdao

**Research Article**

**Keywords:** Gliomas, Multiparametric MRI, Voxel-wise Analysis, Radiomics, IDH Status

**Posted Date:** June 22nd, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-9921588/v1

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

**Additional Declarations:** No competing interests reported.

## Abstract

# Background

Noninvasive preoperative assessment of isocitrate dehydrogenase (IDH) status and survival risk is clinically important in glioma. This study aimed to develop and externally validate a voxel-wise probabilistic MRI habitat framework for IDH status assessment and survival risk stratification.

# Methods

This retrospective multicenter study included 269 adults with pathologically confirmed diffuse glioma: 174 patients from Center A for training and 95 from Center B for external validation. Preoperative multiparametric MRI included contrast-enhanced T1-weighted imaging, T2-weighted imaging, T2-FLAIR, apparent diffusion coefficient maps, and cerebral blood flow maps. Tumor and peritumoral edema were segmented using nnU-Net v2. Three-dimensional sliding-window radiomics generated voxel-wise feature maps, and Gaussian mixture models produced probabilistic habitat maps. Label-independent habitat metrics and tumor-edema differential features were combined with clinical variables for IDH classification. Survival value was assessed using Kaplan-Meier analysis and Cox regression.

# Results

The tumor-edema differential habitat model achieved an external validation AUC of 0.907 for IDH prediction. The combined clinical-imaging model achieved the highest discrimination (AUC, 0.938; 95% CI, 0.884–0.979), outperforming the clinical-only model (AUC, 0.695). Calibration was favorable (Brier score, 0.103), and decision curve analysis showed higher net benefit than clinical or habitat-only models. SHAP analysis identified perfusion- and diffusion-related entropy metrics as the dominant contributors. The habitat signature was independently associated with overall survival in multivariable analysis (hazard ratio, 4.829; 95% CI, 2.382–9.790; $P < 0.001$) and showed good prognostic discrimination (Harrell C-index, 0.847).

# Conclusions

Voxel-wise probabilistic habitat modeling of multiparametric MRI characterized spatial tumor heterogeneity and showed value for noninvasive IDH assessment and survival risk stratification in glioma.

## Background

Isocitrate dehydrogenase (IDH) status is central to the molecular classification, risk stratification, and treatment planning of diffuse glioma[1–4]. In the current integrated diagnostic framework, IDH mutation

status is closely associated with tumor biology, therapeutic response, and patient prognosis, and it directly influences clinical decision-making before and after surgery. Although IDH status is ultimately determined by histopathology and genetic testing, preoperative noninvasive assessment may support surgical planning, biopsy targeting, treatment selection, and patient counseling, especially because tissue sampling can be affected by spatial tumor heterogeneity[5].

Conventional MRI-based radiomics and deep learning have shown promise for IDH prediction by extracting quantitative imaging features from structural, diffusion, and perfusion MRI[5–10]. These methods provide a potential bridge between macroscopic imaging phenotypes and underlying molecular alterations. However, many existing approaches summarize the tumor as a single region of interest and use global intensity, shape, or texture descriptors for model construction. Such region-level representations may overlook spatially distinct subregions within the enhancing tumor, nonenhancing tumor, and peritumoral edema, thereby underrepresenting intratumoral and peritumoral ecological variation.

Habitat imaging has emerged as a promising strategy for characterizing spatial tumor heterogeneity by partitioning tumors into imaging phenotypes that may reflect differences in cellularity, perfusion, edema, angiogenesis, and microenvironmental gradients [11–13]. In glioma, habitat-based analysis may be particularly relevant because molecular alterations such as IDH mutation are associated not only with tumor cellular composition but also with vascular remodeling, infiltrative growth patterns, and changes in the surrounding microenvironment. Nevertheless, conventional habitat approaches often rely on hard clustering or region-level summaries, which can compress continuous spatial information and assign variable label meanings across patients [14]. These limitations may reduce biological interpretability and weaken cross-patient comparability.

Voxel-wise radiomics habitat modeling addresses these limitations by extracting local texture features around each voxel and modeling habitat membership probabilistically [15]. Compared with hard partitioning, probabilistic habitat modeling allows each voxel to retain graded memberships across multiple latent phenotypes, thereby better reflecting transitional tissue states and spatially continuous microenvironmental variation. We hypothesized that probabilistic voxel-wise habitat metrics derived from multiparametric MRI would improve preoperative IDH classification and provide prognostic information beyond routine clinical variables.

The aim of this study was to develop and externally validate a multicenter voxel-wise probabilistic radiomics habitat framework using structural, diffusion, and perfusion MRI for IDH status assessment and overall survival risk stratification in glioma.

## Methods

# Study design and patients

This retrospective observational study included consecutive adults with diffuse glioma treated at two tertiary hospitals between January 2016 and December 2025. Inclusion criteria were age 18 years or older, pathologic confirmation after resection or stereotactic biopsy, complete preoperative multiparametric MRI within 2 weeks before surgery, and available clinical data. Exclusion criteria were previous treatment before MRI, severe artifacts or major protocol inconsistency, and missing or indeterminate IDH testing. Of 345 initially identified patients, 269 were included: 174 from Center A for training and 95 from Center B for external validation. The study was approved by the Ethics Committee of the First Hospital of Qinhuangdao (approval number: 2023KZ065; approval date: 17 May 2023), and informed consent was waived because of the retrospective design.

IDH status was assessed using immunohistochemistry for IDH1 R132H. Immunohistochemistry-negative cases underwent Sanger sequencing or next-generation sequencing for IDH1/IDH2 mutations. Clinical variables included age, sex, WHO tumor grade, and tumor location. Tumor location was assigned by registering tumor masks to MNI152 space (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MNI) and determining overlap with atlas-defined brain regions(https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases); tumors involving two or more lobes were categorized as multilobar. All procedures complied with the Declaration of Helsinki. The overall data processing workflow is illustrated in Fig. 1.

## MRI preprocessing and segmentation

All patients underwent preoperative MRI on 1.5-T or 3.0-T scanners from Siemens, GE, or Philips. Sequences included contrast-enhanced T1-weighted imaging (T1CE), T2-weighted imaging (T2WI), T2-FLAIR, diffusion-weighted imaging with derived apparent diffusion coefficient (ADC) maps, and arterial spin labeling-derived cerebral blood flow (CBF) maps (Supplementary Table S1).

Preprocessing was performed in Python using SimpleITK and ANTsPy. Structural sequences underwent N4 bias correction. All modalities were rigidly registered to T1CE space, visually inspected, resampled to 1 x 1 x 1 $mm^3$ resolution, and z-score normalized within brain parenchyma. Masks were resampled using nearest-neighbor interpolation. Skull stripping was performed using HD-BET [16]. Tumor core and peritumoral edema were automatically segmented using nnU-Net v2 (https://github.com/MIC-DKFZ/nnUNet/tree/master/nnunetv2) trained with the 3D full-resolution configuration on BraTS 2021 data, followed by visual quality inspection.

[17].

## Voxel-wise Feature Mapping

Voxel-wise radiomics feature maps were generated using a three-dimensional sliding-window strategy. For each voxel within the tumor core and edema regions, local radiomics features were calculated from a 7 × 7 × 7 voxel neighborhood. When the window extended beyond the segmentation mask, only voxels inside the mask were used. Features were extracted independently for each MRI modality using a

customized PyRadiomics-based pipeline and mapped back to the original MRI space to generate three-dimensional feature maps. A total of 33 voxel-wise radiomics features were extracted for each modality (Supplementary Methods S1). The implementation of the radiomics habitat pipeline is publicly available on GitHub (https://github.com/zt3226073/radiomics-habitat-pipeline).

## Probabilistic Habitat Construction

Voxel-wise habitat modeling was performed separately for each MRI modality using Gaussian Mixture Models. For each modality, voxel-wise radiomics feature vectors from the tumor and edema regions were modeled using a diagonal-covariance Gaussian Mixture Model, generating posterior probability maps that represented probabilistic habitat membership rather than hard cluster labels.

The number of habitats was determined in the training set using the Bayesian Information Criterion across K = 2–8. To avoid information leakage, model parameters were estimated only from the training data and then fixed when applied to the external validation set. The posterior probabilities were reconstructed in MRI space to generate K-channel three-dimensional habitat probability maps (Supplementary Methods S2).

Because habitat labels were not assumed to be directly comparable across patients, label-independent patient-level metrics were extracted from the probability maps. These included habitat entropy, dominance ratio, effective habitat number, Moran's I, fragmentation index, and Gini coefficient. Tumor–edema gradient features, including Δentropy, ΔMoran's I, and Δfragmentation index, were further calculated to quantify spatial differences between the tumor core and peritumoral edema (Supplementary Methods S3).

## IDH Prediction Modeling

After Z-score standardization, ComBat harmonization was applied to reduce scanner-related batch effects, with scanner model defined as the batch variable. The scanner with the largest sample size in the training set was used as the reference batch. To avoid information leakage, IDH status and outcome variables were not included during harmonization, and external validation features were mapped to the training-defined reference distribution (Supplementary Methods S4)[18].

Feature selection was performed in the training set using Elastic Net logistic regression. Habitat metrics were combined with pretreatment clinical variables, including age, sex, and tumor location, for IDH classification. The regularization and mixing parameters were optimized using nested cross-validation, and features with nonzero coefficients were retained for subsequent modeling.

The model combining clinical variables and habitat metrics, termed Clin+TEΔ-ALL, was defined as the primary model. Comparative models were constructed to assess the contributions of tumor-region features, edema-region features, tumor–edema differential features, and modality-specific feature

groups. Candidate classifiers included logistic regression, random forest, support vector machine, XGBoost, multilayer perceptron, and decision tree. Hyperparameters were optimized using Optuna with the Tree-structured Parzen Estimator algorithm, using cross-validated AUC as the objective function[19]. Model development used nested stratified 5-fold cross-validation in the training set, and the external validation set was reserved for final evaluation.

A habitat signature was constructed from the selected TEΔ-ALL features by linearly combining retained habitat metrics weighted by their Elastic Net coefficients. This signature was used to quantify patient-level spatial heterogeneity and for subsequent survival analysis. Model performance was evaluated using AUC, accuracy, sensitivity, specificity, and $F1_{max}$. AUC 95% confidence intervals were estimated using 1,000 bootstrap resamples, and DeLong tests were used for model comparisons. Calibration was assessed using calibration curves and the Brier score, clinical utility using decision curve analysis, and interpretability using Shapley Additive Explanations (SHAP).

# Survival Analysis

Overall survival was defined as the interval from diagnosis to death or last follow-up. Patients without death were censored at the last follow-up, with a follow-up cutoff date of February 2026. Survival analysis was performed in patients from the external validation cohort with available follow-up data (n = 61). Patients were stratified into high- and low-risk groups according to the optimal cutoff value of the habitat signature determined using X-tile software, which identifies the cutoff by maximizing the log-rank $\chi^2$ statistic. Kaplan–Meier curves were used to compare survival between the two risk groups, and differences were assessed using the log-rank test. Associations between the habitat signature and overall survival were further evaluated using univariable and multivariable Cox proportional hazards models, with hazard ratios and 95% confidence intervals reported. Variables with $P < 0.10$ in univariable analysis were included in the multivariable model. The proportional hazards assumption was assessed using Schoenfeld residuals.

# Statistical Analysis

All statistical analyses were performed in a Python environment. Continuous variables were tested for normality using the Shapiro–Wilk test. Normally distributed variables were expressed as mean ± standard deviation and compared using the independent-samples t test; nonnormally distributed variables were expressed as median and interquartile range and compared using the Mann–Whitney U test. Categorical variables were expressed as frequency and percentage and compared using the chi-square test or Fisher’s exact test. All tests were two-sided, and $P < 0.05$ was considered statistically significant.

## Results

# Patients

The training and external validation sets included 174 and 95 patients, respectively (Table 1). Age, sex, tumor location, WHO grade and IDH status were similar between sets. IDH-mutant tumors were present in 72 of 174 training patients (41.4%) and 42 of 95 validation patients (44.2%).

Table 1
Clinical characteristics of patients in different sets

| **Characteristic** | **Training set** | **External validation set** | ***P*** |
|---|---|---|---|
| Sample size | 174 | 95 | |
| Age (year, mean ± SD) | 49.7 ± 18.5 | 49.0 ± 20.8 | 0.780 |
| Gender (%) | | | > 0.999 |
| Female | 72 (41.4%) | 39 (41.1%) | |
| Male | 102 (58.6%) | 56 (58.9%) | |
| Location (%) | | | 0.896 |
| Frontal lobe | 47 (27.0%) | 27 (28.4%) | |
| Temporal | 21 (12.1%) | 14 (14.7%) | |
| Parietal/Occipital | 16 (9.2%) | 6 (6.3%) | |
| Other | 28 (16.1%) | 16 (16.8%) | |
| Multilobar (≥ 2 lobes) | 62 (35.6%) | 32 (33.7%) | |
| Grade (%) | | | 0.126 |
| II | 38 (21.8%) | 16 (16.8%) | |
| III | 56 (32.2%) | 23 (24.2%) | |
| IV | 80 (46.0%) | 56 (58.9%) | |
| IDH (%) | | | 0.749 |
| Mutant | 72 (41.4%) | 42 (44.2%) | |
| Wild-type | 102 (58.6%) | 53 (55.8%) | |

Note: Independent samples t-test was applied in continuous variables. Chi-Squared test or Fisher's exact test was applied in categorical variables. IDH=isocitrate dehydrogenase; SD: standard deviation. The *Parietal/Occipital* group included tumors confined to either the parietal or occipital lobe. The *Other* category comprised deep or midline structures (e.g., thalamus, basal ganglia, corpus callosum), infratentorial regions (e.g., cerebellum, brainstem), and peri-rolandic areas. Tumors involving two or more lobes were categorized as *multilobar*.

# Feature Selection

Patient-level BIC analysis in the training set supported K = 3 as the optimal habitat number for most MRI sequences and ROIs; although several sequences favored K = 4, K = 3 was selected for overall model

stability (Supplementary Table S2). After ComBat harmonization, the proportion of habitat metrics with significant scanner-related differences decreased from 62.7% to 9.3% (Supplementary Figures S2 and S3).

A total of 75 habitat metrics were entered into Elastic Net feature selection. At the optimal penalty parameter, 15 features were retained, including 13 habitat imaging features and 2 clinical variables, age and tumor location (Supplementary Figure S4). The selected imaging features were mainly derived from ADC and CBF and included tumor-region, edema-region, and tumor–edema differential metrics.

Among the evaluated classifiers, logistic regression achieved the best performance, with the highest cross-validated AUC in the training cohort and the highest AUC in the external validation cohort, outperforming SVC, random forest, XGBoost, MLP, and decision tree models (Supplementary Table S3).

# Model Performance and Comparison

Multiple models were constructed to evaluate the contributions of spatial regions, imaging modalities, and tumor–edema differential features. In the external validation set, the Clin+TEΔ-ALL model achieved the best performance, with an AUC of 0.938 (95% CI: 0.884–0.979), outperforming the clinical-only model (AUC = 0.695) and demonstrating the incremental value of habitat-based imaging features (Fig. 2, Supplementary Table S4, Supplementary Figure S5).

Among habitat-based models, TEΔ-ALL significantly outperformed TE-ALL (P = 0.024), indicating that tumor–edema differential features improved discrimination. In contrast, no significant differences were observed between TE-ALL and T-ALL (P = 0.100) or between E-ALL and T-ALL (P = 0.070). In modality-based comparisons, functional models showed numerically higher AUCs than structural models, but the differences were not significant. However, TEΔ-ALL significantly outperformed both TEΔ-FUNC (P = 0.030) and TEΔ-STRUCT (P < 0.001), suggesting complementary contributions of functional and structural habitat phenotypes for IDH prediction.

# Model Calibration and Clinical Utility

In the external validation set, the Clin+TEΔ-ALL model showed good calibration, with predicted probabilities closely aligned with observed outcomes and a Brier score of 0.103 (Fig. 3). Decision curve analysis demonstrated that Clin+TEΔ-ALL provided the highest net benefit across a broad range of threshold probabilities, outperforming the clinical-only model and other imaging-based models (Fig. 3).

# SHAP Interpretability Analysis

SHAP analysis showed that functional habitat features were the main contributors to the Clin+TEΔ-ALL model, with CBF_Tumor_habitat_entropy_mean and ADC_Tumor_habitat_entropy_mean ranking as the most influential predictors. Edema-region features and tumor–edema differential metrics, including ΔMoran’s I and Δentropy, also contributed to model prediction (Fig. 4).

Representative cases showed distinct voxel-wise habitat architectures between IDH-mutant and IDH-wildtype gliomas (Fig. 5). IDH-wildtype tumors exhibited more complex perfusion- and diffusion-related heterogeneity and fragmentation, whereas IDH-mutant tumors showed relatively more homogeneous habitat distributions, consistent with the SHAP findings.

# Survival analysis

Patients were stratified into high- and low-risk groups using the optimal habitat signature cutoff identified by X-tile software in the survival cohort (cutoff, 1.09). Kaplan–Meier analysis showed significantly poorer OS in the high-risk group than in the low-risk group (log-rank $P < 0.0001$; Fig. 6).

In multivariable Cox analysis, the habitat signature was independently associated with OS (HR = 4.829, 95% CI: 2.382–9.790, $P < 0.001$). Age was also independently associated with OS (HR = 1.109, 95% CI: 1.057–1.164, $P < 0.001$). Compared with frontal lobe tumors, parietal/occipital, other-region, and multilobar tumors were associated with higher mortality risk, whereas temporal lobe tumors showed no significant difference (Supplementary Table S5). The model achieved a Harrell's C-index of 0.847, suggesting favorable prognostic discrimination.

## Discussion

This study developed and externally validated a voxel-wise probabilistic radiomics habitat framework for noninvasive molecular and prognostic assessment of diffuse glioma. By capturing spatially resolved tumor–edema habitat organization from multiparametric MRI, the proposed framework improved preoperative IDH classification beyond clinical variables and showed favorable calibration and decision curve benefit. The independent prognostic value of the habitat signature further suggests that spatial habitat architecture encodes biologically meaningful heterogeneity, providing an interpretable imaging bridge between glioma molecular status, microenvironmental organization, and survival risk.

Previous imaging-based studies have predicted IDH status using whole-tumor radiomics, deep learning models, or ROI-level habitat analysis, but these approaches may insufficiently characterize continuous spatial heterogeneity and often have limited interpretability[20–24]. Although MRI habitat studies based on perfusion, ADC, or diffusion kurtosis metrics have demonstrated differences between IDH-mutant and IDH-wildtype gliomas[9, 25–29], most rely on discrete region-level clustering and global feature extraction, which may obscure voxel-scale microenvironmental variation and reduce cross-patient comparability[25, 27, 29, 30]. The present voxel-wise probabilistic habitat framework addressed these limitations by combining sliding-window radiomics feature mapping with Gaussian Mixture Model–based soft clustering to generate continuous habitat probability maps[15]. This strategy achieved strong external validation performance without deep learning and provided interpretable spatial heterogeneity and tumor–edema gradient metrics for characterizing glioma microenvironment organization.

Model interpretability analysis showed that IDH prediction and survival stratification were mainly driven by perfusion–diffusion habitat heterogeneity and its spatial organization. The most influential features

included CBF- and ADC-derived entropy, fragmentation, Moran's I, and Gini metrics, together with tumor location. Biologically, CBF reflects tumor perfusion, angiogenesis, and microvascular remodeling, whereas ADC is associated with cellular density and microstructural alterations[31]. Previous studies have reported higher perfusion parameters in IDH-wildtype gliomas and supported ADC as an important biomarker for IDH classification[32, 33]. Our findings extend these conventional parameter-based observations by showing that IDH-related imaging phenotypes are reflected not only by mean perfusion or diffusion values, but also by their voxel-wise spatial heterogeneity. Habitat entropy and Gini coefficient quantify phenotype diversity and imbalance, whereas Moran's I and fragmentation index describe spatial autocorrelation and subregional disorganization. These metrics may capture IDH-associated differences in vascular heterogeneity, cellular architecture, infiltrative growth, and tumor–edema microenvironmental gradients[26, 27, 34].

The peritumoral edema region may also contain important molecular and prognostic information. Previous habitat analyses based on dynamic susceptibility contrast (DSC) perfusion imaging have partitioned the tumor microenvironment into multiple subregions, including infiltrative edema and vasogenic edema. Differences in rCBV across these habitats have been shown to be associated not only with IDH status but also with patient prognosis[26]. Furthermore, multimodal studies have demonstrated that the spatial distribution of subregions characterized by varying vascular permeability and cellular density can reflect the heterogeneity of the tumor microenvironment and is associated with disease progression and survival outcomes[27]. Similarly, vascular heterogeneity–based habitat analyses have indicated that the hemodynamic characteristics of the peritumoral region can predict recurrence patterns and clinical outcomes[35].

The habitat signature remained independently associated with OS after adjustment for age and tumor location, and the model achieved a Harrell's C-index of 0.847, indicating good survival discrimination. Age and non-frontal or multilobar tumor locations were also associated with poorer outcomes. Previous studies have shown that perfusion- and diffusion-defined habitats, including low-perfusion or diffusion-restricted subregions, are related to treatment resistance and survival outcomes [36]. Multimodal MRI habitat studies further suggest that the volumetric distribution of subregions with different vascular permeability and cellular density is prognostically relevant, possibly reflecting hypoxia, metabolic stress, and aggressive tumor microenvironments [27]. These findings support the concept that the spatial organization of the glioma microenvironment carries prognostic information and may complement conventional clinical risk stratification.

This study has several limitations. First, the maturity of follow-up was moderate: the median follow-up duration was 26 months, and the overall median OS was also 26 months, indicating that the set is statistically at an intermediate level of maturity. The median OS of the low-risk group (42 months) may still be influenced by censoring, and longer follow-up is required. Second, the retrospective study design may introduce potential selection bias, as the inclusion of patients with complete imaging data may favor individuals who were more likely to undergo surgery or had higher imaging quality. Future studies should employ prospective sets with consecutive enrollment and explicitly report the mechanisms of

missing data. Third, external validation for survival analysis remains limited. Although external validation was performed for IDH classification, the OS model still lacks multicenter external validation for survival outcomes, which warrants further investigation in larger independent sets.

## Conclusions

Multiparametric MRI–based voxel-wise probabilistic habitat modeling enables noninvasive characterization of tumor and peritumoral heterogeneity, supports preoperative IDH prediction, and provides prognostic information for survival risk stratification in glioma.

## List of abbreviations

ADC, apparent diffusion coefficient

ASL, arterial spin labeling

AUC, area under the receiver operating characteristic curve

BIC, Bayesian information criterion

CBF, cerebral blood flow

CI, confidence interval

DCA, decision curve analysis

FLAIR, fluid-attenuated inversion recovery

GMM, Gaussian mixture model

IDH, isocitrate dehydrogenase

MRI, magnetic resonance imaging

OS, overall survival

ROI, region of interest

SHAP, Shapley Additive Explanations

T1CE, contrast-enhanced T1-weighted imaging

T2WI, T2-weighted imaging

TEΔ-ALL, tumor-edema differential habitat features from all MRI sequences

# Declarations

Ethics approval and consent to participate

This retrospective multicenter study was approved by the Ethics Committee of the First Hospital of Qinhuangdao (approval number: 2023KZ065; approval date: 17 May 2023). The requirement for written informed consent was waived because of the retrospective study design. All procedures complied with the Declaration of Helsinki.

Consent for publication

Not applicable. The manuscript does not contain identifiable individual participant data.

Availability of data and materials

The datasets analyzed during the current study are not publicly available because they contain potentially sensitive clinical and imaging information, but may be available from the corresponding author on reasonable request and with appropriate institutional and ethics approvals. The implementation code for the radiomics habitat pipeline is available at https://github.com/zt3226073/radiomics-habitat-pipeline.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by Qinhuangdao S&T Plan Program (grant number: 202301A212). The funder had no role in the study design, data collection, analysis, interpretation of data, decision to publish, or preparation of the manuscript.

Authors' contributions

TZ and LY contributed to the study conception and design, curated the data, and drafted the original manuscript. DZ contributed to the methodology. QS developed and implemented the software. JD acquired funding for the study. DL reviewed and edited the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

Authors' information

Not applicable.

Additional files

Additional file 1. Supplementary materials (.docx). Supplementary methods, MRI acquisition parameters, habitat model selection results, feature definitions, classifier screening, model performance analyses, and supplementary figures/tables.

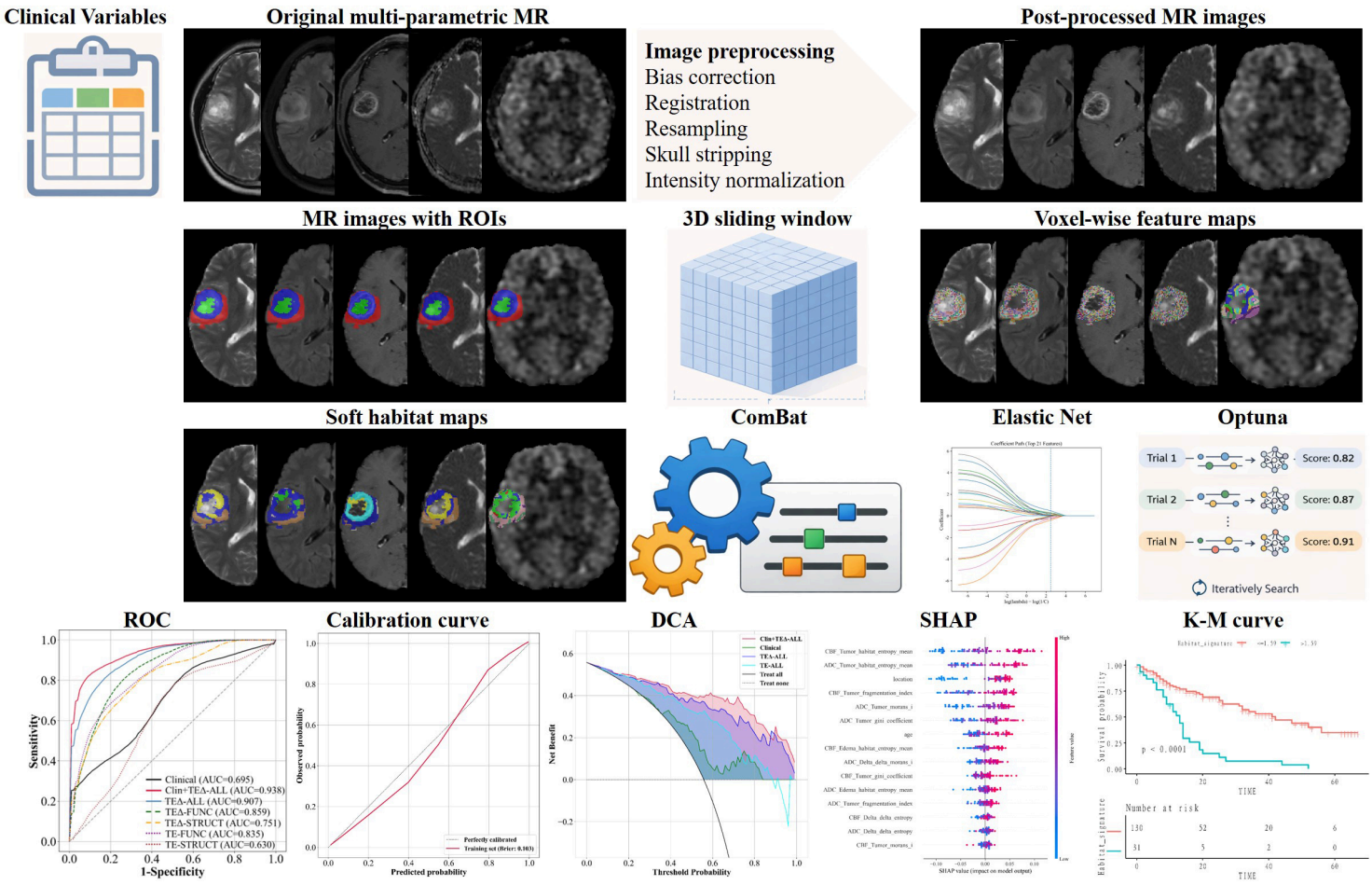
# References


1. Greenwald AC, Darnell NG, Hoefflin R, et al. Integrative spatial analysis reveals a multi-layered organization of glioblastoma. Cell. 2024;187:2485–e250126. 10.1016/j.cell.2024.03.029.
2. Weller M, Wen PY, Chang SM, et al. Glioma. Nat Rev Dis Primers. 2024;10:33. 10.1038/s41572-024-00516-y.
3. Gupta P, Dang M, Oberai S, et al. Immune landscape of isocitrate dehydrogenase-stratified primary and recurrent human gliomas. Neuro Oncol. 2024;26:2239–55. 10.1093/neuonc/noae139.
4. Zhang L, Fritah S, Nazarov PV, Kaoma T, Van Dyck E. Impact of IDH Mutations, the 1p/19q Co-Deletion and the G-CIMP Status on Alternative Splicing in Diffuse Gliomas. Int J Mol Sci. 2023;24. 10.3390/ijms24129825.
5. Chen C, Du X, Yang L, et al. Research on application of radiomics in glioma: a bibliometric and visual analysis. Front Oncol. 2023;13:1083080. 10.3389/fonc.2023.1083080.
6. Hajri R, Nicod-Lalonde M, Hottinger AF, Prior JO, Dunet V. Prediction of Glioma Grade and IDH Status Using (18)F-FET PET/CT Dynamic and Multiparametric Texture Analysis. Diagnostics (Basel). 2023;13. 10.3390/diagnostics13152604.
7. 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;25:279–89. 10.1093/neuonc/noac166.
8. Niu W, Yan J, Hao M, et al. MRI transformer deep learning and radiomics for predicting IDH wild type TERT promoter mutant gliomas. NPJ Precis Oncol. 2025;9:89. 10.1038/s41698-025-00884-y.
9. Liu Y, Wang P, Wang S, et al. Heterogeneity matching and IDH prediction in adult-type diffuse gliomas: a DKI-based habitat analysis. Front Oncol. 2023;13:1202170. 10.3389/fonc.2023.1202170.
10. Qiao J, Wu H, Liu J, et al. Spectral Analysis Based on Hemodynamic Habitat Imaging Predicts Isocitrate Dehydrogenase Status and Prognosis in High-Grade Glioma. World Neurosurg. 2023;175:e520–30. 10.1016/j.wneu.2023.03.136.
11. Beig N, Patel J, Prasanna P, et al. Radiogenomic analysis of hypoxia pathway is predictive of overall survival in Glioblastoma. Sci Rep. 2018;8:7. 10.1038/s41598-017-18310-0.
12. Sottoriva A, Spiteri I, Piccirillo SG, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proc Natl Acad Sci U S A. 2013;110:4009–14. 10.1073/pnas.1219747110.
13. Hu LS, Ning S, Eschbacher JM, et al. Radiogenomics to characterize regional genetic heterogeneity in glioblastoma. Neuro Oncol. 2017;19:128–37. 10.1093/neuonc/now135.

14. Zhang HM, Huo XB, Wang HL, Wang C. Diagnostic Performance of Dynamic Susceptibility Contrast-Enhanced Perfusion-Weighted Imaging in Differentiating Recurrence From Radiation Injury in Postoperative Glioma: A Meta-analysis. J Comput Assist Tomogr. 2022;46:938–44. 10.1097/rct.0000000000001356.
15. Bernatowicz K, Grussu F, Ligero M, Garcia A, Delgado E, Perez-Lopez R. Robust imaging habitat computation using voxel-wise radiomics features. Sci Rep. 2021;11:20133. 10.1038/s41598-021-99701-2.
16. Isensee F, Schell M, Pflueger I, et al. Automated brain extraction of multisequence MRI using artificial neural networks. Hum Brain Mapp. 2019;40:4952–64. 10.1002/hbm.24750.
17. 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;18:203–11. 10.1038/s41592-020-01008-z.
18. Orlhac F, Frouin F, Nioche C, Ayache N, Buvat I. Validation of A Method to Compensate Multicenter Effects Affecting CT Radiomics. Radiology. 2019;291:53–9. 10.1148/radiol.2019182023.
19. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: A Next-generation Hyperparameter Optimization Framework. ACM; 2019.
20. Peng H, Huo J, Li B, et al. Predicting Isocitrate Dehydrogenase (IDH) Mutation Status in Gliomas Using Multiparameter MRI Radiomics Features. J Magn Reson Imaging. 2021;53:1399–407. 10.1002/jmri.27434.
21. 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;13:1143688. 10.3389/fonc.2023.1143688.
22. Wu J, Xu Q, Shen Y, Chen W, Xu K, Qi XR. Swin Transformer Improves the IDH Mutation Status Prediction of Gliomas Free of MRI-Based Tumor Segmentation. J Clin Med. 2022;11. 10.3390/jcm11154625.
23. Yuan Y, Yu Y, Chang J, et al. Convolutional neural network to predict IDH mutation status in glioma from chemical exchange saturation transfer imaging at 7 Tesla. Front Oncol. 2023;13:1134626. 10.3389/fonc.2023.1134626.
24. Stefano A. Challenges and limitations in applying radiomics to PET imaging: Possible opportunities and avenues for research. Comput Biol Med. 2024;179:108827. 10.1016/j.compbiomed.2024.108827.
25. Kersch CN, Kim M, Stoller J, Barajas RF Jr., Park JE. Imaging Genomics of Glioma Revisited: Analytic Methods to Understand Spatial and Temporal Heterogeneity. AJNR Am J Neuroradiol. 2024;45:537–48. 10.3174/ajnr.A8148.
26. Wu H, Tong H, Du X, et al. Vascular habitat analysis based on dynamic susceptibility contrast perfusion MRI predicts IDH mutation status and prognosis in high-grade gliomas. Eur Radiol. 2020;30:3254–65. 10.1007/s00330-020-06702-2.

27. Wang X, Xie Z, Wang X, et al. Preoperative prediction of IDH genotypes and prognosis in adult-type diffuse gliomas: intratumor heterogeneity habitat analysis using dynamic contrast-enhanced MRI and diffusion-weighted imaging. Cancer Imaging. 2025;25:11. 10.1186/s40644-025-00829-5.
28. Yang Z, Lin R, Wu Z, et al. Habitat subregions analysis based on neurite orientation dispersion and density imaging enhances isocitrate dehydrogenase genotyping in glioma. BMC Med Imaging. 2025;25:384. 10.1186/s12880-025-01934-4.
29. Zheng WY, Yang RH, Wan ZF, et al. MRI-based habitat radiomics for preoperatively predicting IDH status in gliomas. Neurosurg Focus. 2025;59:E4. 10.3171/2025.5.focus25135.
30. Han Y, Wang Y, Cui W, Xiu S, Yang Y, Zhang J. IDH Mutation Classification in Nonenhancing Gliomas: A Comparison of Habitat and Whole-Tumor Transfer Learning Strategies. J Magn Reson Imaging. 2026;63:1079–89. 10.1002/jmri.70187.
31. Kim M, Kim HS. Emerging Techniques in Brain Tumor Imaging: What Radiologists Need to Know. Korean J Radiol. 2016;17:598–619. 10.3348/kjr.2016.17.5.598.
32. van Santwijk L, Kouwenberg V, Meijer F, Smits M, Henssen D. A systematic review and meta-analysis on the differentiation of glioma grade and mutational status by use of perfusion-based magnetic resonance imaging. Insights Imaging. 2022;13:102. 10.1186/s13244-022-01230-7.
33. Bhatti F, Strobel J, Tench C, et al. Apparent diffusion coefficient for genetic characterization of untreated adult gliomas: A meta-analysis stratified by methods. Neurooncol Adv. 2025;7:vdaf103. 10.1093/noajnl/vdaf103.
34. Singh G, Manjila S, Sakla N, et al. Radiomics and radiogenomics in gliomas: a contemporary update. Br J Cancer. 2021;125:641–57. 10.1038/s41416-021-01387-w.
35. Wang H, Zeng L, Wu H, et al. Preoperative vascular heterogeneity based on dynamic susceptibility contrast MRI in predicting spatial pattern of locally recurrent high-grade gliomas. Eur Radiol. 2024;34:1982–93. 10.1007/s00330-023-10149-6.
36. Li C, Yan JL, Torheim T, et al. Low perfusion compartments in glioblastoma quantified by advanced magnetic resonance imaging and correlated with patient survival. Radiother Oncol. 2019;134:17–24. 10.1016/j.radonc.2019.01.008.

# Figures




## Figure 1

**Overview of the proposed voxel-wise habitat modeling framework for IDH mutation prediction and survival analysis in diffuse glioma.**

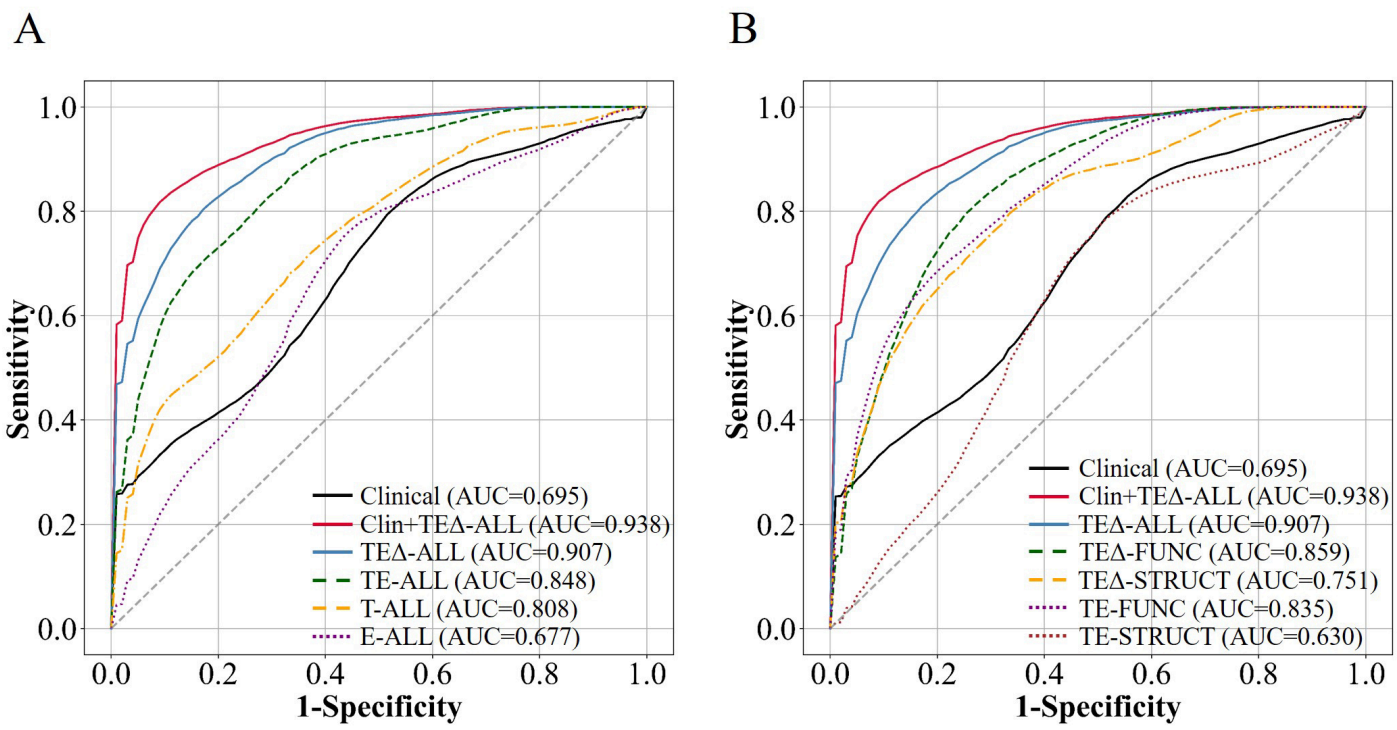
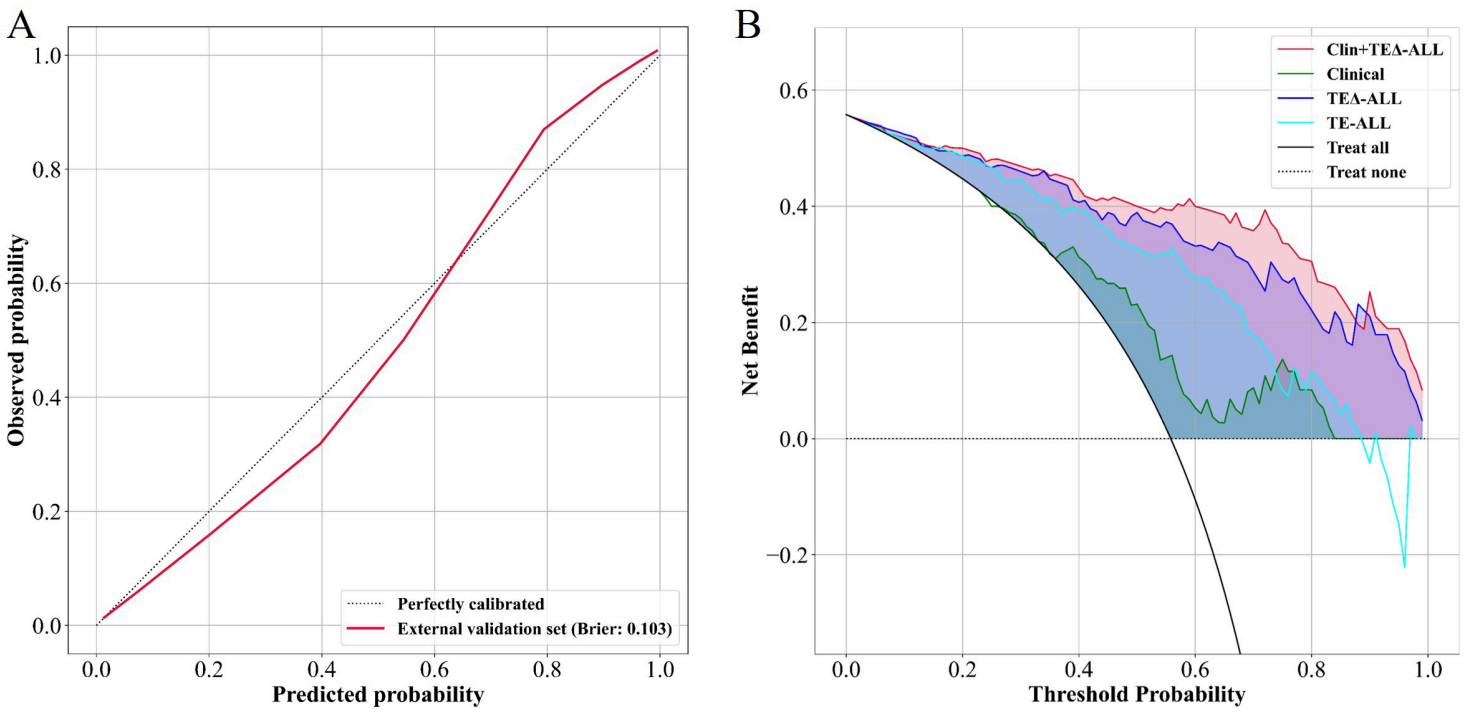


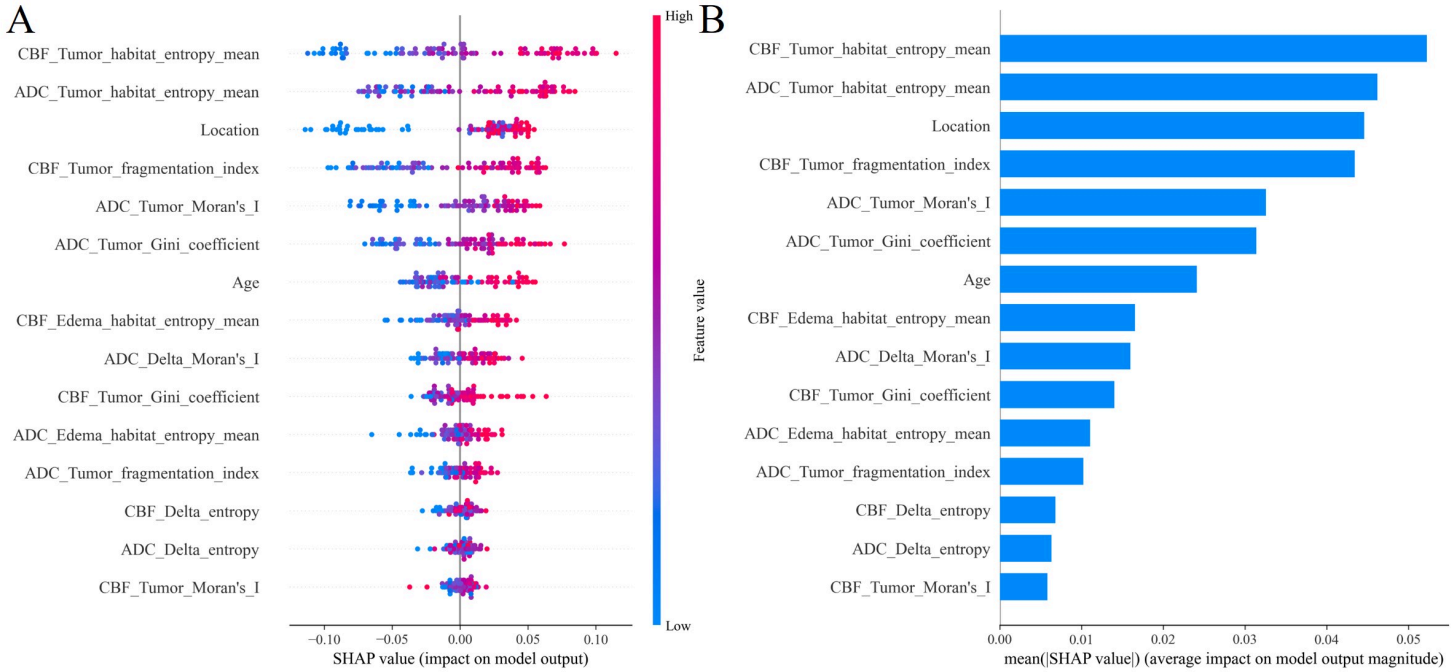

**Figure 2**

**Receiver operating characteristic (ROC) curves of different predictive models for IDH status in the external validation cohort.** (A) Comparison of habitat-based models constructed using different spatial regions. (B) Comparison of modality-based models, including functional imaging models and structural imaging models. Abbreviations: AUC, area under the receiver operating characteristic curve; T, tumor region; E, edema region; TE, combined tumor and edema regions; Δ, tumor–edema difference features; ALL, all radiomics habitat features; FUNC, functional imaging features (ADC and CBF); STRUCT, structural imaging features (T1CE, T2, and FLAIR).




## Figure 3

**Calibration and clinical utility of the predictive models in the external validation cohort. (A)** Calibration curve of the Clin+TEΔ-ALL model. **(B)** Decision curve analysis (DCA) comparing the net clinical benefit of different models across a range of threshold probabilities. Abbreviations:DCA, decision curve analysis; TE, combined tumor and edema regions; **Δ**, tumor–edema difference features; ALL, all radiomics habitat features.




## Figure 4

**SHAP-based interpretability analysis of the Clin+TEΔ-ALL model.** (A) SHAP summary (beeswarm) plot illustrating the impact of individual features on model predictions. Each point represents a patient-level observation, with the horizontal axis indicating the SHAP value (i.e., the contribution of the feature to the model output). Feature values are color-coded from low (blue) to high (red). Features are ranked according to their overall importance. (B) Bar plot showing the mean absolute SHAP values of the top features, representing their average contribution to the model output magnitude.

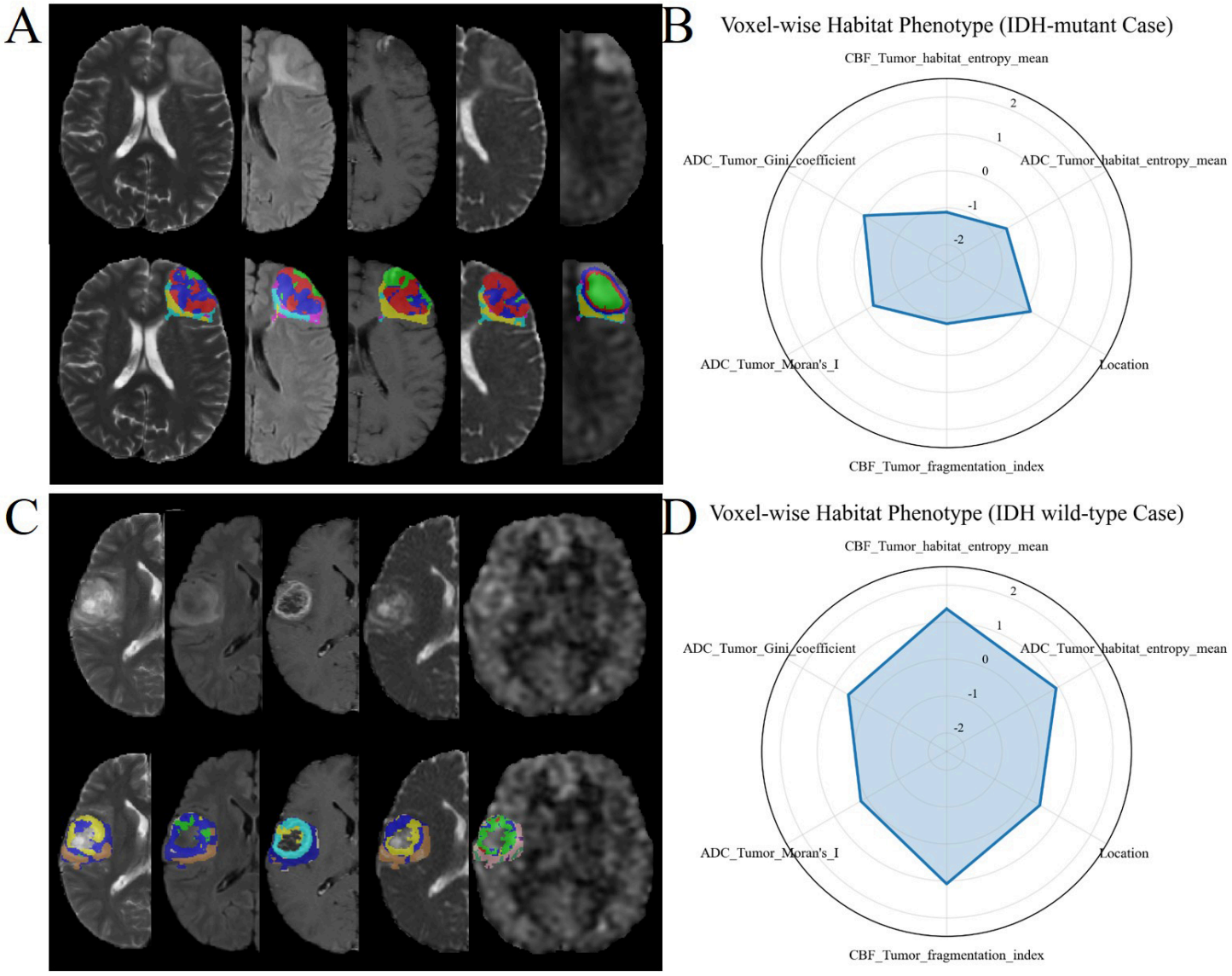



**Figure 5**

**Representative voxel-wise habitat phenotypes for IDH- mutant and wildtype gliomas.** (A&C) Representative multimodal MRI and corresponding voxel-wise habitat maps for two glioma cases with different IDH status. (A) A 45-year-old male patient with an IDH-mutant glioma located in the left frontal lobe, showing axial T2, T2-FLAIR, T1CE, ADC, and CBF images together with the corresponding voxel-wise habitat probability maps reconstructed within the tumor region. (C) A 62-year-old female patient with an IDH-wildtype glioma located in the right temporal lobe, demonstrating the same multimodal MR

sequences and the corresponding habitat maps. (B&D) Radar plots illustrating the voxel-wise habitat phenotypes for the representative cases. Six key habitat metrics were selected based on SHAP-derived global feature importance and normalized using training set z-scores. (B) The IDH-mutant case demonstrates relatively lower habitat entropy and spatial fragmentation, indicating a comparatively more homogeneous microenvironment. (D) In contrast, the IDH-wild-type case exhibits higher entropy and fragmentation across multiple metrics, reflecting increased spatial heterogeneity of tumor habitats.

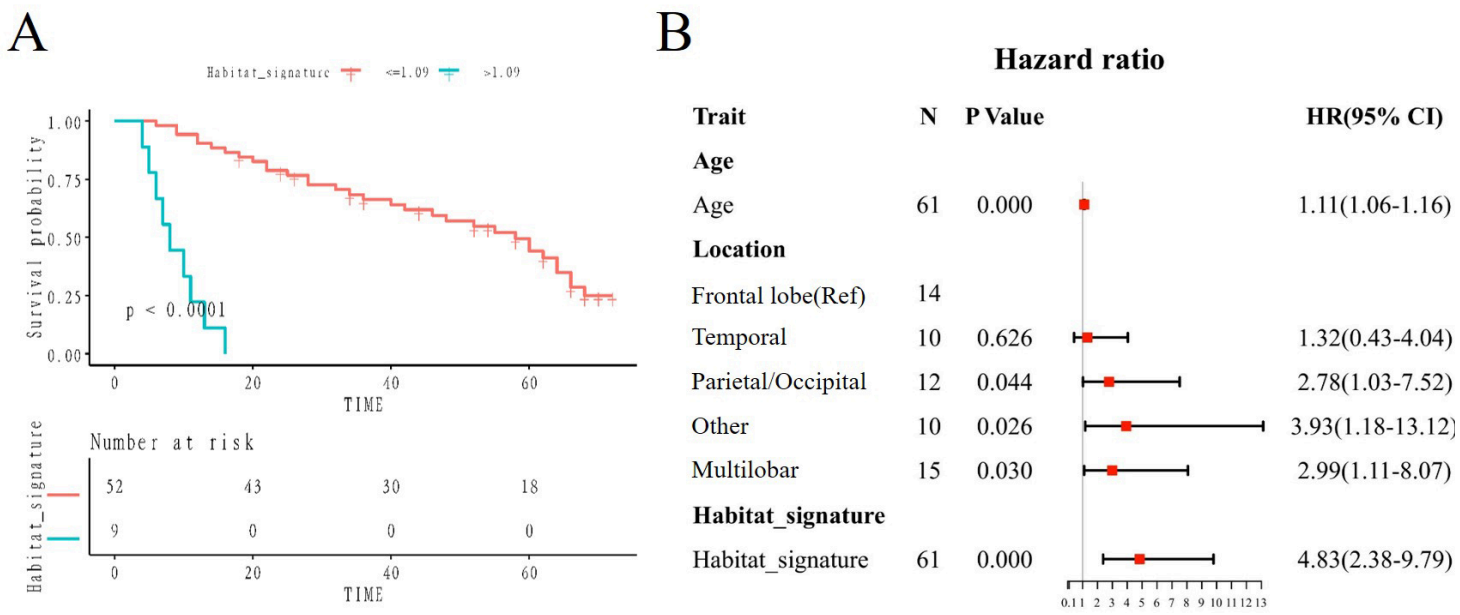



**Figure 6**

**Survival analysis based on the habitat-derived risk signature.** (A) Kaplan–Meier survival curves stratified by the habitat signature. (B)Forest plot of the multivariable Cox proportional hazards analysis for overall survival.

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementary.docx