ACCSCIENCE PUBLISHING

Cancer Plus

ORIGINAL RESEARCH ARTICLE

Pre-operative prediction of high-risk molecular subtypes of glioma based on multimodal MRI tumor habitat imaging

Yincheng Su[1,2†], Qijian Li[2†], Zihao Liu[2], Bo Peng[2], Zhiyuan Wang[3], Xiaoli Jin[4], Wenju Niu[2*], and Xiangli Yang[1*]

[1]Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Third Hospital of Shanxi Medical University, Tongji Shanxi Hospital, Taiyuan, Shanxi, China

[2]College of Medical Imaging, Shanxi Medical University, Taiyuan, Shanxi, China

[3]Department of Sports Rehabilitation, Department of Physical Education, Shanxi Medical University, Taiyuan, Shanxi, China

[4]Cancer Center, Shanxi Bethune Hospital (Shanxi Academy of Medical Sciences, Tongji Shanxi Hospital), Third Hospital of Shanxi Medical University, Taiyuan, Shanxi, China

Abstract

Gliomas are the most common malignant brain tumors driven by genetic and microenvironmental factors, with newly recognized high-risk molecular subtypes requiring aggressive treatment. This study aims to address the limitations of biopsy-based molecular typing by developing a non-invasive multimodal magnetic resonance imaging habitat imaging model to predict high-risk subtypes, thereby improving early detection and guiding treatment. Data of 204 glioma patients retrieved from The Cancer Genome Atlas public database were retrospectively analyzed. Habitat imaging based on K-means clustering was applied to three habitat regions in pre-operative T1CE and T2FLAIR sequences, extracting 10,416 radiomics features. Analysis of variance was used to assess the correlation between features and labels, screening radiomics features significantly associated with high-risk molecular subtypes. A support vector machine classifier was employed to construct a habitat radiomics model. Logistic regression (LR) was used to identify relevant clinical features, and a clinical prediction model was established, followed by performance evaluation. A combined model was developed by integrating the habitat radiomics model and the clinical model using multivariate LR. The predictive performance of the three models was evaluated and compared using metrics such as the area under the receiver operating characteristic curve (AUC), decision curve analysis (DCA), and calibration curves. The combined model achieved AUC = 0.943 and 0.912 in the training and test sets, respectively, outperforming the clinical model (training set: AUC = 0.830; test set: AUC = 0.841) and the habitat radiomics model (training set: AUC = 0.914; test set: AUC = 0.864). In DCA, the combined model demonstrated significantly higher and more stable net benefits within a reasonable clinical threshold range compared to the other two models. Calibration curves indicated that the combined model also exhibited superior calibration performance. This study shows that combining clinical and radiomics data improves glioma risk prediction, but multicenter validation of such approach for clinical use is warranted.

Keyword: Glioma; Radiomics; Magnetic resonance imaging; Habitat; High-risk molecular

†*These authors contributed equally to this work.*

***Corresponding authors:**
Xiangli Yang
(xiangli_@163.com)
Wenju Niu
(18003457517@163.com)

Citation: Su Y, Li Q, Liu Z, *et al*. Pre-operative prediction of high-risk molecular subtypes of glioma based on multimodal MRI tumor habitat imaging. *Cancer Plus*. 2025;7(4):36-46.
doi: 10.36922/CP025220038

Received: May 30, 2025

Revised: July 30, 2025

Accepted: August 19, 2025

Published online: September 19, 2025

1. Introduction

Gliomas are the most common primary malignant tumors of the central nervous system, accounting for 80% of malignant brain tumors.[1] Their evolution is influenced by genetic factors and tumor microenvironments.[1-3] Glioblastoma evolves under genetic influences, tumor microenvironmental factors, and other mechanisms.[4-6] In the 2021 updated version of the classification guidelines for central nervous system tumors, four high-risk molecular subtypes (isocitrate dehydrogenase [IDH] wild-type with telomerase reverse transcriptase [TERT] promoter hypermethylation or epidermal growth factor receptor [EGFR] amplification or +7/−10 chromosomal alteration; IDH mutant coupled with cyclin-dependent kinase inhibitor 2A/B [CDKN2A/B] homozygous deletion) were newly added. Gliomas carrying any of these subtypes are classified as Grade 4 gliomas, which exhibit similar biological behaviors to traditional glioblastomas and require more active therapeutic intervention.[7-9]

The tumor microenvironment is composed of complex cells and molecular environments surrounding the tumor, forming an ecosystem consisting of multiple sub-communities within a unique microenvironmental niche.[10-12] Genetic variations among gliomas lead to competitive interactions between them. Under environmental selection pressure, these interactions dynamically evolve in response to the environment, similar to different species occupying specific ecological niches in an ecosystem.[13-17] However, tumor microenvironment and molecular typing assays can be confirmed through core biopsies and molecular marker testing.[18-20] However, accurate detection of molecular markers requires sufficient tissue samples and expensive materials for instruments.[21-23] In recent years, based on Darwinian dynamics in the tumor microenvironment, imaging techniques have been developed to segment the tumor area into specific ecological regions and use multimodal imaging to predict molecular typing. This provides a non-invasive approach for precise prediction of high-molecular-typing gliomas without surgery.[24-28] However, previous studies were typically limited to single molecular markers,[29-31] unable to integrate them with molecular subtypes in the overall pathological diagnosis of gliomas or explore the relationships between different molecular subtypes and tumor heterogeneity in depth.

This study aims to leverage the advantages of ecological techniques to build a prediction model based on ecological principles and combine clinical information to create a comprehensive model for non-invasive prediction of high-risk molecular subtypes in gliomas. This research will deepen our understanding of the evolutionary dynamics of gliomas, providing critical evidence for early identification of malignant or potentially malignant transformations within subgroups of glioma.

2. Methods

2.1. Study subjects

The research process is shown in (Figure 1). The study was conducted in accordance with the Declaration of Helsinki. The requirement for informed consent was waived by the Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences review board (2021GLL131). A retrospective analysis was conducted on 204 glioma patients retrieved from The Cancer Genome Atlas (TCGA) (https://www.cancer.gov/ccg/research/genome-sequencing/tcga). Inclusion criteria: (1) Adult patients with diffusely infiltrating gliomas pathologically confirmed according to the World Health

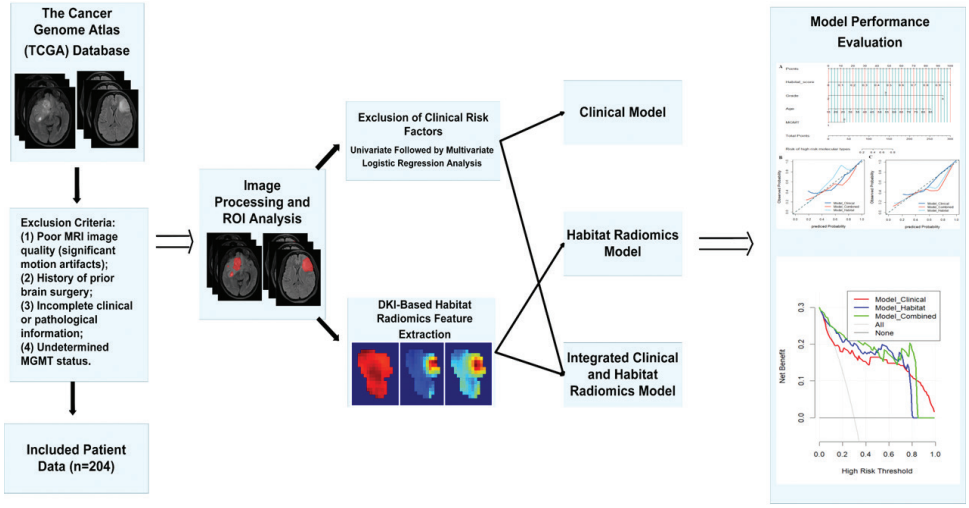

Figure 1. The research flowchart

Organization (WHO) histological classification standards; (2) No history of tumor-related treatment before surgery; (3) Clear and complete pre-operative post-contrast enhanced T1-weighted (T1CE) and T2-weighted fluid attenuation inversion recovery (T2FLAIR) magnetic resonance imaging (MRI) sequences; (4) Complete clinical and pathological information; (5) Confirmed O6-methylguanine-DNA methyltransferase (MGMT) status. Exclusion criteria: (1) Poor MRI image quality (significant motion artifacts); (2) Previous history of brain surgery; (3) Incomplete clinical or pathological information; (4) Undetermined MGMT status. Ultimately, 204 patients from the TCGA database were included in the study and randomly divided into training and testing sets in a 7:3 ratio (Figure 2). The definition of high-risk subtypes is based on the 2021 WHO Classification of Tumors of the Central Nervous System. This classification includes traditional glioblastoma and four high-risk molecular subtypes: IDH-wildtype with TERT promoter mutation, IDH-wildtype with EGFR amplification, IDH-wildtype with +7/−10 chromosomal alterations, and IDH-mutant with CDKN2A/B homozygous deletion. Gliomas carrying any of these subtypes are classified as Grade 4 tumors and constitute the high-risk subtypes (Figure 3).

2.2. MRI data acquisition and preprocessing

T1CE and T2FLAIR MRI sequences, along with clinical and pathological data for each patient, were obtained from the TCGA database. Since the images in the TCGA database were acquired using different MRI scanners and protocols, spatial resolution and grayscale intensity normalization were performed to ensure comparability. Open-source software FAE (Feature Explorer Pro, version 0.5.13, http://github.com/salan668/FAE), based on the Pyradiomics package (https://github.com/Radiomics/pyradiomics), was used to resample all MRI images to $1 \times 1 \times 1$ m^3, eliminating inconsistencies in spatial resolution due to different MRI scanner models. In addition, intensity normalization was performed using SimpleITK in Python 3.10.4 (https://www.python.org/) to mitigate heterogeneity bias.

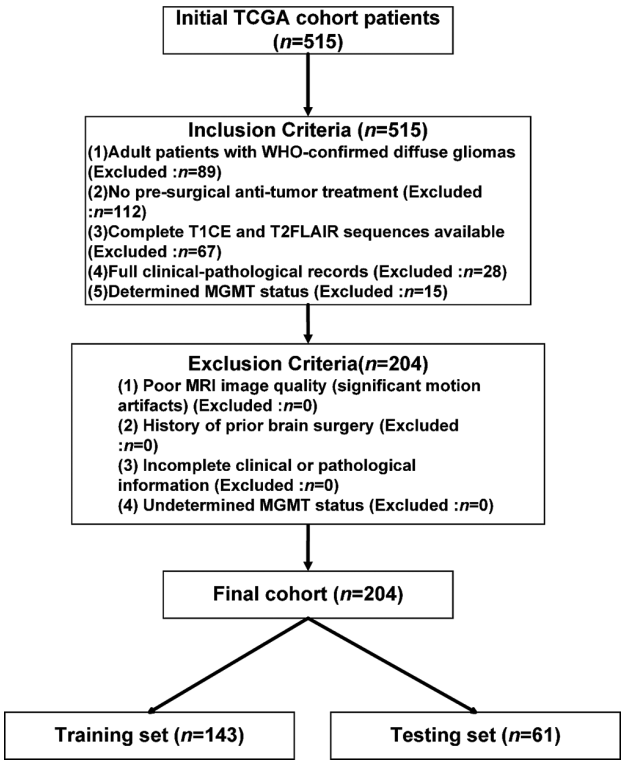

Figure 2. Flowchart for patient selection

2.3. MRI image processing and segmentation

Registration was strictly performed using the T1CE sequence as a template. Regions of interest (ROIs) were manually delineated on the registered MRI images using ITK software (version 3.6.0). Specifically, the tumor region was selected as the ROI on T1CE images, carefully excluding peritumoral edema. The delineated ROIs were then registered to FLAIR images for further analysis. ROI delineation was performed by a neuroradiologist with over 10 years of experience. To ensure accuracy and consistency, another neuroradiologist with over 15 years of experience reviewed the ROIs, cross-validated the results, and provided additional insights when necessary. Both radiologists were blinded to the patients' clinical, pathological, or molecular information during ROI delineation and review. In cases where the two radiologists encounter discrepancies, they will first attempt to reach a consensus. If disagreement persists, a more specialized and senior director-level neuroradiologist will be consulted for final adjudication (Figure 4).

2.4. K-means-based habitat imaging

A data-driven K-means clustering technique was employed to automatically segment tumor subregions characterized by consistent signal intensity patterns across multiparametric MRI. To ensure signal homogeneity within each subregion, all voxels within the segmentation mask were aggregated, combining image information from all sequences. Sixteen radiomics features were computed for each voxel to extract local information. Voxels within each cluster were grouped based on their similarity and dissimilarity, using the squared Euclidean distance between voxel intensities and voxel-level radiomics features as similarity measures. Clustering was performed at the cohort level rather than the patient level to ensure consistent assignments across patients. To determine the optimal number of clusters, cluster numbers ranging from 2 to 10 were tested. The clustering results were evaluated over 100

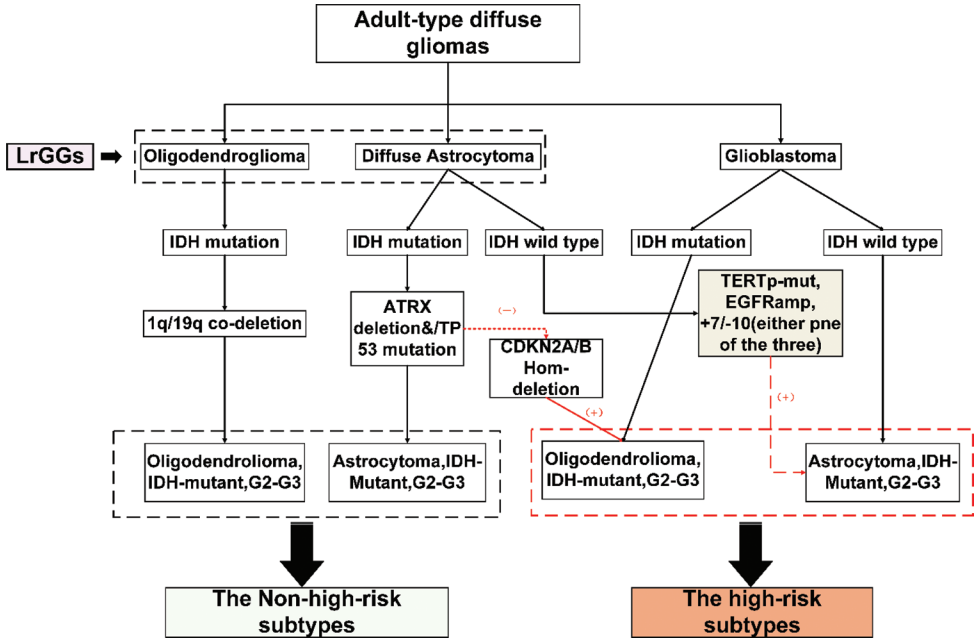

Figure 3. The schematic diagram of adult diffuse glioma classification in our study
Notes: (+): Indicates positive; (−): Indicates negative.
Abbreviations: ATRX: Glioma-associated gene; +7/10: Chromosome 7 gain and chromosome 10 loss; CDKN2A/B hom-deletion: Cyclin-dependent kinase inhibitor 2A/B homozygous deletion; EGFRamp: Epidermal growth factor receptor amplification; G2/3/4: Grade 2/3/4; IDH: Isocitrate dehydrogenase; LrGGs: Lower-grade gliomas; 1p/19q: Short arm of chromosome 1 and long arm of chromosome 19; TERTp-mut: Telomerase reverse transcriptase promoter mutation.

repeated trials using the average Calinski-Harabasz score for each k-value. Ultimately, three clusters were selected, as they exhibited the best Calinski-Harabasz score, effectively demonstrated habitat imaging differences, and avoided over-parameterized model construction (Table 1).

Table 1. Calinski-Harabasz scores

K-means clustering numbers	Calinski-Harabasz scores
2	3.921×10^6
3	5.667×10^6
4	3.282×10^6
5	2.736×10^6
6	2.666×10^6
7	2.297×10^6
8	1.894×10^6
9	1.705×10^6
10	1.642×10^6

2.5. Feature extraction

For each patient, radiomics features were extracted from three ROIs (corresponding to the three habitat ROIs defined in section 2.4) in both T1CE and T2FLAIR sequences using the open-source software FAE (version 0.5.5, http://github.com/salan668/FAE) based on the Pyradiomics package (https://github.com/Radiomics/pyradiomics). The extracted features were categorized into four groups: (1) Shape features (n = 14), quantifying ROI shape and size (e.g., maximum 3D diameter, volume, surface area); (2) First-order statistical features (n = 18), quantifying tumor signal heterogeneity based on grayscale histogram distributions (e.g., mean, standard deviation, skewness, kurtosis); (3) Texture features, describing intratumoral texture distribution to quantify microscopic heterogeneity, including Gray-Level Co-occurrence Matrix (GLCM, n = 22), Gray-Level Run-Length Matrix (GLRLM, n = 16), Gray-Level Size Zone Matrix (GLSZM, n = 16), Gray-Level Dependence Matrix (GLDM, n = 14), and Neighboring Gray-Tone Difference Matrix (NGTDM, n = 5); (4) Filtered features (n = 1638), derived by applying eight filters [LoG, Wavelet, Square, SquareRoot, Logarithm, Exponential, Gradient, LocalBinaryPattern (3D)] to the original images, followed by first-order statistical and texture feature extraction from the filtered ROIs. A total of 17,360 radiomics features were extracted per patient (two sequences × three ROIs × 1,736 features per ROI).

2.6. Habitat radiomics feature selection and model construction

First, feature matrices were normalized. Specifically, for each feature vector, the arithmetic mean and standard

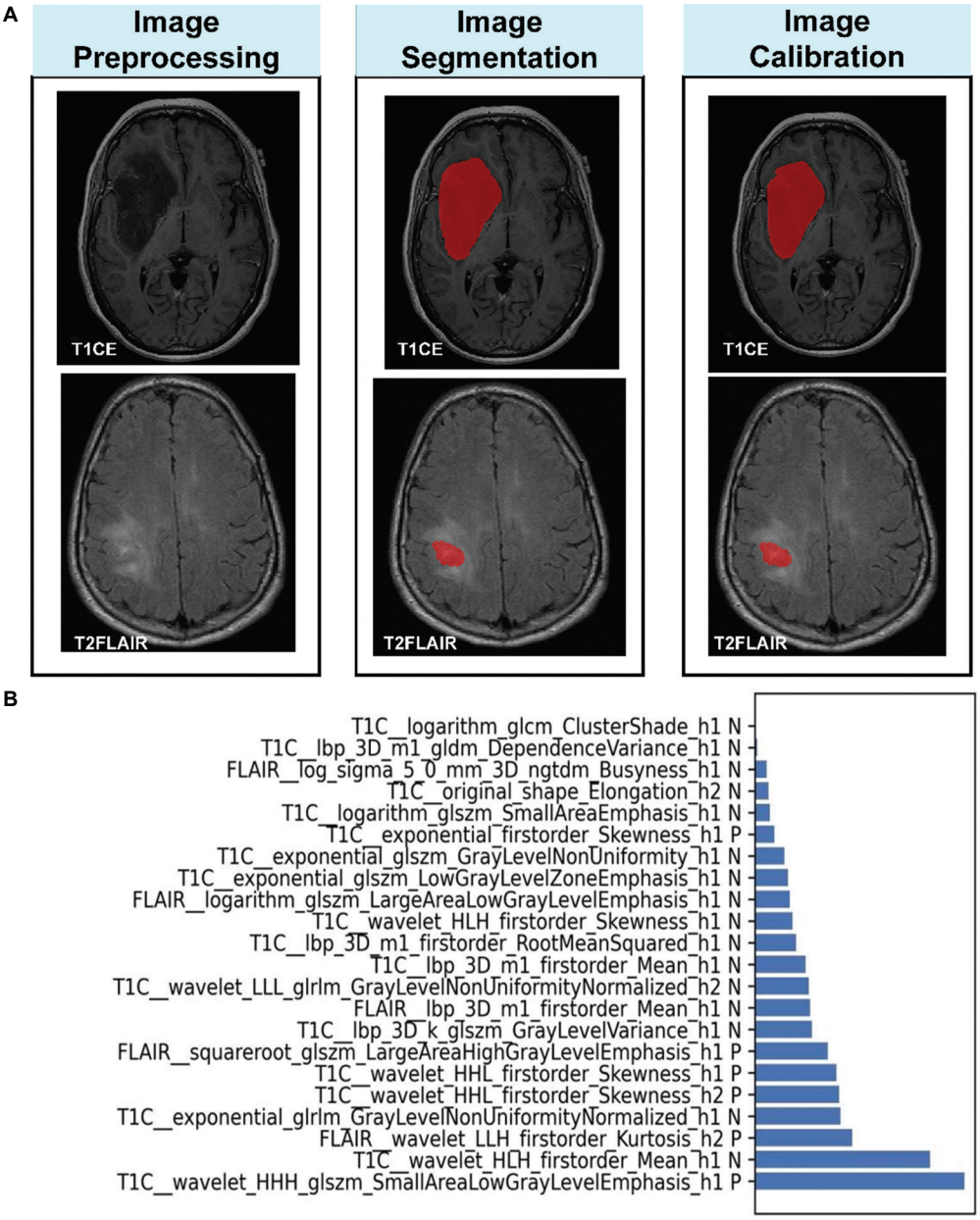

Figure 4. Magnetic resonance imaging image processing pipeline and LASSO-selected feature weights**.** (A) From left to right: T1CE and T2 FLAIR raw images, ROI-segmented images, and corrected ROI-segmented images. (B) Twenty-two radiomic features and their weights selected by LASSO regression Abbreviations: ROI: Region of interest.

deviation were computed, and each feature value was then mean-centered and scaled by the standard deviation, resulting in zero-mean and unit-variance normalized feature vectors. Given the high dimensionality of the feature space, Pearson correlation coefficient (PCC) was calculated to assess feature similarity. Feature pairs with PCC > 0.990 were considered redundant, and one was removed to reduce dimensionality while preserving feature independence. To further refine feature selection, analysis of variance was used to evaluate feature-label correlations. Features were ranked based on their F-scores, and the most discriminative subset was selected for model construction.

For classification, a support vector machine (SVM) was employed as a supervised generalized linear classifier, identifying the maximum-margin hyperplane for binary classification. Model performance was evaluated using 10-fold cross-validation, yielding the radiomics signature (Rad signature). This study employed FAE version 0.5.x (http://github.com/salan668/FAE) with the following SVM

classifier configurations: regularization parameter C values of [0.1, 0.3, 1.0, 3.0], kernel types including "linear," "rbf," and "poly," gamma parameter values of [0.1, 0.3, 1.0, 3.0], polynomial degrees of [2, 3, 4,] and a fixed random seed of 0. For additional technical specifications, please refer to the FAE repository (http://github.com/salan668/FAE). In developing the clinical prediction model, this study employed a rigorous two-stage logistic regression (LR) approach for variable selection. Initially, univariate analysis was conducted using R software to evaluate four candidate predictors (age, MGMT status, sex, and tumor grade), identifying age, grade, and MGMT status as statistically significant (p<0.05). These significant variables were subsequently entered into multivariate LR analysis, where all three maintained statistical significance (p<0.05 for each). Consequently, the final model incorporated age, tumor grade, and MGMT status as independent predictors, while sex was excluded due to non-significant association (p>0.05). This analytical strategy ensured the model exclusively included robust predictors demonstrating both statistical significance and clinical relevance. Finally, a LR model integrated the tumor habitat model and clinical model into a combined model. Predictive performance was compared among the clinical, habitat, and combined models using metrics including area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), decision curve analysis (DCA), and calibration curve analysis.

3. Results

3.1. Patient characteristics

A total of 204 patients meeting the inclusion and exclusion criteria were enrolled in this study and randomly divided into training (n = 143) and testing (n = 61) sets at a 7:3 ratio. High-risk patients accounted for 65.20% of the cohort, with 45.59% and 19.61% in the training and testing sets, respectively (Table 2). Significant differences were observed in age, pathological grade, and MGMT status between the training and testing sets (p<0.05), whereas no significant difference was found in gender distribution.

3.2. Screening of clinical factors and model construction

Using univariate followed by multivariate LR (implemented via R software), pathological grade was identified as a significant risk factor associated with the target variable. The clinical model was subsequently constructed based on the selected clinical risk factors. During the development of the clinical prediction model, we employed a rigorous two-stage variable selection approach for LR analysis. First, univariable LR analysis was performed using R software on four candidate variables (age, MGMT status, sex, and grade). The results indicated that age, grade, and MGMT status were statistically significant (p<0.05). These significant variables were subsequently included in a multivariable LR analysis, where all three retained statistical significance (p<0.05 for each). Consequently, the final model incorporated age, grade, and MGMT status as independent predictors, while the sex variable was excluded due to lack of significance (p>0.05). This analytical approach ensured that the model contained only robust predictors with both statistical and clinical relevance (Table 3).

3.3. Model performance evaluation

A combined model integrating the habitat radiomics model and clinical model was developed using LR. Performance

Table 2. Clinical patient baseline table

Characteristic	Training set			Testing set		
	Non-high-risk	High-risk	*p*-value	Non-high-risk	High-risk	*p*-value
Age	43.00±14.70	54.68±14.66	<0.001*	42.19±11.92	56.75±12.72	<0.001*
Grade			<0.001*			<0.001*
2	33 (66.00)	9 (9.68)		11 (52.38)	3 (7.50)	
3	17 (34.00)	28 (30.11)		10 (47.62)	10 (25.00)	
4	null	56 (60.22)		null	27 (67.50)	
MGMT			<0.001*			0.014
Unmethylated	6 (12.00)	40 (43.01)		1 (4.76)	15 (37.50)	
Methylated	44 (88.00)	53 (56.99)		20 (95.24)	25 (62.50)	
Sex			0.993			0.655
Female	25 (50.00)	45 (48.39)		12 (57.14)	19 (47.50)	
Male	25 (50.00)	48 (51.61)		9 (42.86)	21 (52.50)	

Note: * Indicates as a significant difference, *p*-value<0.05.
Abbreviation: MGMT: O6-methylguanine-DNA methyltransferase.

evaluation through AUC, accuracy, sensitivity, specificity, PPV, NPV, DCA, and calibration curves demonstrated the superior predictive capability of the combined clinical-habitat radiomics model (Table 4). The receiver operating characteristic (ROC) analysis revealed AUC = 0.943 (training set) and 0.912 (testing set), outperforming both the clinical model (training AUC = 0.830, testing AUC = 0.841) and the habitat radiomics model (training AUC = 0.914, testing AUC = 0.864) (Figure 5). DCA indicated that the combined model provided significantly higher and more stable net benefits across clinically reasonable threshold probabilities compared to the other two models. The calibration curve further confirmed that the combined model exhibited excellent agreement with the ideal calibration line (diagonal), demonstrating superior calibration over the standalone habitat radiomics and clinical models (Figure 6).

Table 3. Univariate and multivariate screening of clinical characteristics

Characteristic	Univariate		Multivariate	
	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value
Age	1.016 (1.01–1.022)	0.000*	0.981 (0.963–1)	0.104
MGMT	1.205 (0.862–1.685)	0.362		
Grade	1.368 (1.239–1.51)	0.000*	2.052 (1.459–2.883)	0.001*
Sex	1.92 (1.28–2.881)	0.008*	0.577 (0.306–1.087)	0.153

*$p<0.05$ is regarded as significantly different.
Abbreviations: CI: Confidence interval;
MGMT: O6-methylguanine-DNA methyltransferase; OR: Odds ratio.

4. Discussion

This study developed a habitat radiomics model based on conventional MRI sequences and prior research, then integrated clinical features to construct a fusion model with optimal predictive performance (training AUC = 0.943, testing AUC = 0.912).[1,32] This approach holds significant potential for non-invasive early warning of malignant transformation in gliomas.

The first highlight of this study was the inclusion of all WHO Grade 4 diffuse gliomas—defined by the 2021 classification guidelines—as a single cohort. This encompassed traditional glioblastomas and four high-risk molecular subtypes of astrocytoma (i.e., IDH-wildtype with TERT promoter mutation, IDH-wildtype with EGFR amplification, IDH-wildtype with +7/−10 chromosomal alterations, and IDH-mutant with CDKN2A/B homozygous deletion). For instance,[32,33] achieved an AUC of 0.966

Table 4. Performance of habitat radiomics models, clinical models, and integrated habitat radiomics-clinical models in the study cohort

Model	Accuracy	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV
Habitat	0.7377	0.8643 (0.7697–0.9688)	0.7	0.8095	0.875	0.5862
Clinical	0.8033	0.8411 (0.7433–0.9389)	0.7	1	1	0.6364
Combine	0.8197	0.9119 (0.8421–0.9817)	0.725	1	1	0.6562

Abbreviations: AUC: Area under the receiver operating characteristic curve; CI: Confidence interval; NPV: Negative predictive value; PPV: Positive predictive value.

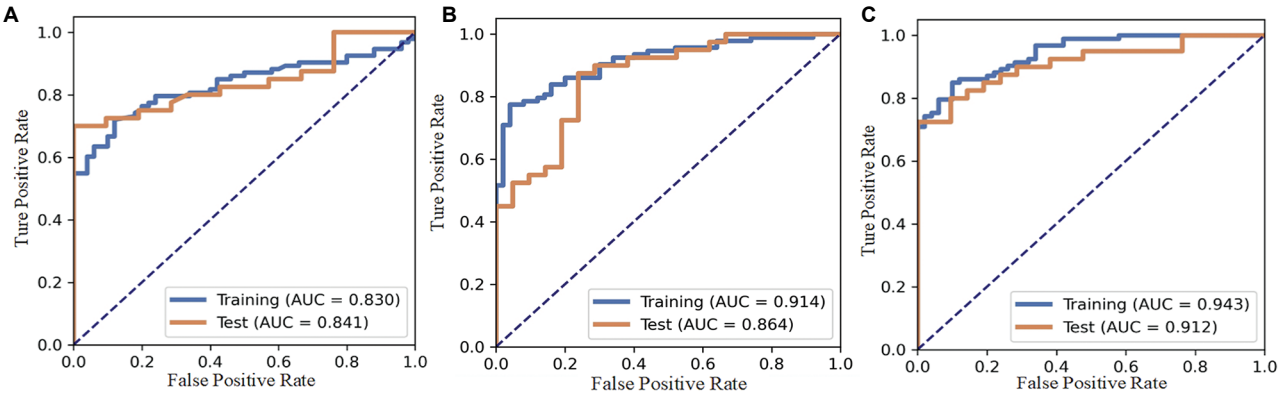

Figure 5. Model performance evaluation. (A) ROC curve of the clinical model. (B) ROC curve of the habitat radiomics model. (C) ROC curve of the integrated habitat radiomics-clinical model.
Abbreviation: AUC: Area under the curve; ROC: Receiver operating characteristic.

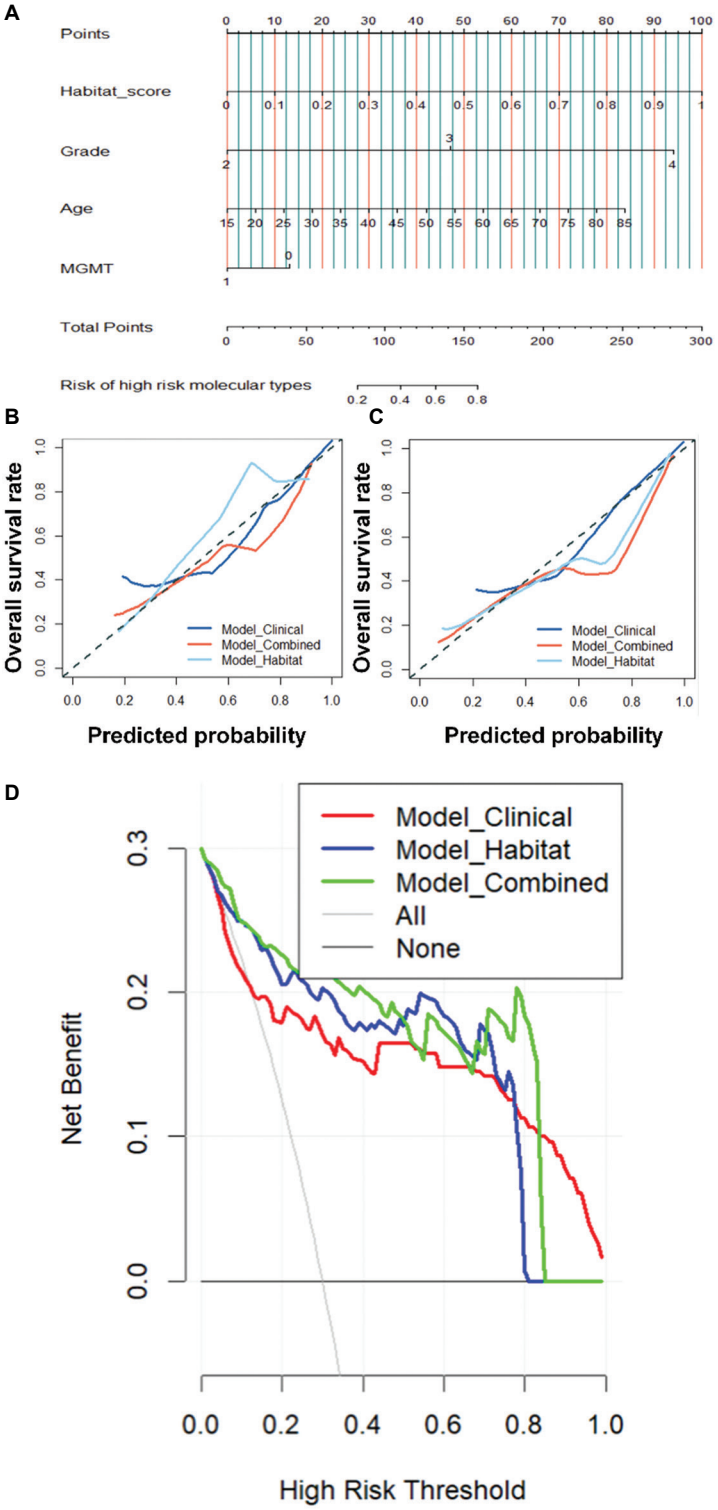

Figure 6. Predictive model validation. (A) Nomogram presents the predictive outcomes of the combined model for assessing the occurrence risk of high-risk molecular subtypes in gliomas. (B and C) Calibration curves for the training and testing cohorts, respectively, including the clinical model, habitat radiomics model, and integrated habitat radiomics-clinical model. The black diagonal dashed line represents the ideal reference standard; closer proximity to this line indicates better model performance. The Hosmer-Lemeshow test for the calibration curve was performed using R software, yielding a p=0.2468 for the test set and 0.467 for the training set. (D) Decision curve analysis compares three models, demonstrating the superior clinical net benefit of the combined model, where the red, blue, and green curves correspond to the clinical model, habitat radiomics model, and integrated habitat radiomics-clinical model, respectively. The integrated habitat radiomics-clinical model demonstrates the highest net benefit across the evaluated threshold probabilities.

in non-invasively identifying CDKN2A/B homozygous deletion in IDH-mutant astrocytomas, though their study focused solely on this subtype.[34] Similarly,[35] demonstrated that EGFR amplification significantly increased tumor aggressiveness (p<0.0001), while linked TERT promoter mutations to poorer overall survival (Hazard ratio = 1.38, 95% Confidence interval: 1.15–1.67). By consolidating all Grade 4 gliomas into one cohort, our study provides broader clinical insights.

The most notable innovation was the development of a habitat analysis model combined with clinical features. Building on prior work,[36] which explored correlations between habitat heterogeneity and high-risk molecular subtypes using diffusion kurtosis imaging-based habitat modeling (AUC: 0.977 and 0.902), our approach leveraged habitat subregion differentiation to enhance predictive power. This may explain why the habitat model substantially outperformed conventional radiomics models.[37]. Analogously,[38] integrated deep learning[12,19,21-24,37,39-41] and habitat radiomics to predict immunotherapy response in advanced NSCLC (AUC = 0.865), while[42] combined habitat features with periablational characteristics for assessing colorectal cancer lung metastases post-RFA (AUC = 0.870). The Transformer neural network architecture has recently exhibited robust predictive performance in the molecular subtyping of gliomas. Based on large multi-center retrospective study cohorts, Niu *et al.*[43] adopted a multi-dimensional, multi-channel transformer and achieved a high AUC in predicting IDH-wildtype gliomas with TERT promoter mutations. The incremental value of clinical variablesting IDH-wildtype gliomas with TERT promoter mutations. The incremental value of clinical variables—particularly pathological grade, age, and MGMT status—further optimized our combined model's performance, highlighting its translational potential.

This study also has some limitations: The study's retrospective design and relatively small sample size may constrain generalizability. Future work will expand patient recruitment, incorporate additional clinical/pathogenetic data, and validate the model prospectively to strengthen its robustness.

5. Conclusion

This study demonstrates that models integrating clinical and radiomics features show potential for improving the prediction of high-risk glioma subtypes, and their clinical utility requires validation through prospective multicenter studies.

Acknowledgments

None.

Funding

This study was supported by Fundamental Research Program of Shanxi Province (202403021221241).

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Yincheng Su, Xiangli Yang, Qijian Li, Wenju Niu
Investigation: Zihao Liu, Zhiyuan Wang, Xiaoli Jin, Bo Peng
Methodology: Yincheng Su, Wenju Niu
Formal analysis: Yincheng Su, Qijian Li, Bo Peng, Zihao Liu
Visualization: Yincheng Su, Qijian Li
Writing–original draft: Yincheng Su, Qijian Li, Zihao Liu, Bo Peng
Writing–review & editing: Zhiyuan Wang, Xiaoli Jin

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki. The requirement for informed consent was waived by the Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences review board (2021GLL131).

Consent for publication

All authors consent to the publication of this study.

Availability of data

The image data were acquired from the TCGA public database (https://www.cancer.gov/ccg/research/genome-sequencing/tcga). Clinical information of patients is publicly available (accessed via https://www.cancer.gov/ccg/research/genome-sequencing/tcga).

References

1. Le VH, Minh T, Kha QH, *et al*. A transfer learning approach on MRI-based radiomics signature for overall survival prediction of low-grade and high-grade gliomas. *Med Biol Eng Comput*. 2023,61(10):2699-2712.

   doi: 10.1007/s11517-023-02875-2

2. Choi Y, Nam Y, Jang J, *et al*. Radiomics may increase the prognostic value for survival in glioblastoma patients when combined with conventional clinical and genetic prognostic models. *Eur Radiol*. 2021;31(4):2084-2093.

   doi: 10.1007/s00330-020-07335-1

3. Ostrom QT, Price M, Neff C, *et al*. CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the united states in 2016-2020. *Neuro Oncol*. 2023,25(12 Suppl 2):iv1-iv99.

doi: 10.1093/neuonc/noad149

4. Wirsching HG, Galanis E, Weller M. Glioblastoma. *Handb Clin Neurol*. 2016;134:381-397.

doi: 10.1016/b978-0-12-802997-8.00023-2

5. Schaff LR, Mellinghoff IK. Glioblastoma and other primary brain malignancies in adults: A review. *JAMA*. 2023;329(7):574-587.

doi: 10.1001/jama.2023.0023

6. Wen PY, Weller M, Lee EQ, *et al*. Glioblastoma in adults: A society for neuro-oncology (SNO) and European society of neuro-oncology (EANO) consensus review on current management and future directions. *Neuro Oncol*. 2020,22(8):1073-1113.

doi: 10.1093/neuonc/noaa106

7. 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

8. Weller M, Van den Bent M, Preusser M, *et al*. Author correction: EANO guidelines on the diagnosis and treatment of diffuse gliomas of adulthood. *Nat Rev Clin Oncol*. 2022;19(5):357-358.

doi: 10.1038/s41571-022-00623-3

9. Perry A, Wesseling P. Histologic classification of gliomas. *Handb Clin Neurol*. 2016,134:71-95.

doi: 10.1016/B978-0-12-802997-8.00005-0.

10. Tomaszewski W, Sanchez-Perez L, Gajewski TF, *et al*. Brain tumor microenvironment and host state: Implications for Immunotherapy. *Clin Cancer Res*. 2019,25(14):4202-4210.

doi: 10.1158/1078-0432.CCR-18-1627.

11. Gisina A, Kholodenko I, Kim Y, *et al*. Glioma stem cells: Novel data obtained by single-cell sequencing. *Int J Mol Sci*. 2022;23(22):14224.

doi: 10.3390/ijms232214224

12. Tian J, Bai X, Quek C. Single-cell informatics for tumor microenvironment and immunotherapy. *Int J Mol Sci*. 2024;25(8):4485.

doi: 10.3390/ijms25084485

13. Sampson JH, Gunn MD, Fecci PE, *et al*. Brain immunology and immunotherapy in brain tumours. *Nat Rev Cancer*. 2020,20(1):12-25.

doi: 10.1038/s41568-019-0224-7

14. Desland FA, Hormigo A. The CNS and the brain tumor microenvironment: Implications for glioblastoma immunotherapy. *Int J Mol Sci*. 2020;21(19):7358.

doi: 10.3390/ijms21197358

15. Galbraith K, Snuderl M. Molecular pathology of gliomas. *Surg Pathol Clin*. 2021;14(3):379-386.

doi: 10.1016/j.path.2021.05.003

16. Nicholson JG, Fine HA. Diffuse glioma heterogeneity and its therapeutic implications. *Cancer Discov*. 2021;11(3):575-590.

doi: 10.1158/2159-8290.cd-20-1474

17. De la Fuente MI. Adult-type diffuse gliomas. *Continuum (Minneap Minn)*. 2023;29(6):1662-1679.

doi: 10.1212/con.0000000000001352

18. Xu S, Tang L, Li X, *et al*. Immunotherapy for glioma: Current management and future application. *Cancer Lett*. 2020;476:1-12.

doi: 10.1016/j.canlet.2020.02.002

19. Park YW, Vollmuth P, Foltyn-Dumitru M, *et al*. The 2021 WHO classification for gliomas and implications on imaging diagnosis: Part 1-key points of the fifth edition and summary of imaging findings on adult-type diffuse gliomas. *J Magn Reson Imaging*. 2023;58(3):677-689.

doi: 10.1002/jmri.28743

20. Fan H, Luo Y, Gu F, *et al*. Artificial intelligence-based MRI radiomics and radiogenomics in glioma. *Cancer Imaging*. 2024;24(1):36.

doi: 10.1186/s40644-024-00682-y

21. Cao L, Shi K, Liu Y, *et al*. Identification of specific genes as molecular markers for rapid and accurate detection of oil-tea *Camellia anthracnose* pathogen Colletotrichum fructicola in China. *Front Microbiol*. 2024;15:1442922.

doi: 10.3389/fmicb.2024.1442922.

22. Radaic A, Kamarajan P, Cho A, *et al*. Biological biomarkers of oral cancer. *Periodontol 2000*. 2024;96(1):250-280.

doi: 10.1111/prd.12542

23. Andryukov BG, Lyapun IN, Matosova EV, *et al*. Biosensor technologies in medicine: From detection of biochemical markers to research into molecular targets (review). *Sovrem Tekhnol Med*. 2021;12(6):70-83.

doi: 10.17691/stm2020.12.6.09

24. Zhu Y, Wang J, Xue C, *et al*. Deep learning and habitat radiomics for the prediction of glioma pathology using multiparametric MRI: A multicenter study. *Acad Radiol*. 2025;32(2):963-975.

doi: 10.1016/j.acra.2024.09.021

25. 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-e530.

doi: 10.1016/j.wneu.2023.03.136

26. Madan E, Palma AM, Vudatha V, *et al*. Cell competition in carcinogenesis. *Cancer Res*. 2022;82(24):4487-4496.

doi: 10.1158/0008-5472.can-22-2217

27. Caiado F, Silva-Santos B, Norell H. Intra-tumour heterogeneity - going beyond genetics. *FEBS J*. 2016;283(12):2245-2258.

doi: 10.1111/febs.13705

28. Brychtová V, Valík V, Vojtěšek B. Variability in the solid cancer cell population and its consequences for cancer diagnostics and treatment. *Klin Onkol*. 2018,31(Suppl 2):5-13.

doi: 10.14735/amko20182s5

29. Lin D, Shen L, Luo M, *et al*. Circulating tumor cells: Biology and clinical significance. *Signal Transduct Target Ther*. 2021;6(1):404.

doi: 10.1038/s41392-021-00817-8

30. Ludwig K, Kornblum HI. Molecular markers in glioma. *J Neurooncol*. 2017;134(3):505-512.

doi: 10.1007/s11060-017-2379-y

31. Hsieh WC, Budiarto BR, Wang YF, *et al*. Spatial multi-omics analyses of the tumor immune microenvironment. *J Biomed Sci*. 2022,29(1):96.

doi: 10.1186/s12929-022-00879-y

32. Sun C, Jiang C, Wang X, *et al*. MR-based radiomics predicts CDK6 expression and prognostic value in high-grade glioma. *Acad Radiol*. 2024;31(12):5141-5153.

doi: 10.1016/j.acra.2024.06.006

33. Gao J, Liu Z, Pan H, *et al*. Preoperative discrimination of CDKN2A/B homozygous deletion status in isocitrate dehydrogenase-mutant astrocytoma: A deep learning-based radiomics model using MRI. *J Magn Reson Imaging*. 2024;59(5):1655-1664.

doi: 10.1002/jmri.28945

34. Bale TA, Jordan JT, Rapalino O, *et al*. Financially effective test algorithm to identify an aggressive, EGFR-amplified variant of IDH-wildtype, lower-grade diffuse glioma. *Neuro Oncol*. 2019;21(5):596-605.

doi: 10.1093/neuonc/noy201

35. Vuong HG, Altibi A, Duong U, *et al*. TERT promoter mutation and its interaction with IDH mutations in glioma: Combined TERT promoter and IDH mutations stratifies lower-grade glioma into distinct survival subgroups-A meta-analysis of aggregate data. *Crit Rev Oncol Hematol*. 2017;120:1-9.

doi: 10.1016/j.critrevonc.2017.09.013

36. Yang X, Niu W, Wu K, *et al*. Diffusion kurtosis imaging-based habitat analysis identifies high-risk molecular subtypes and heterogeneity matching in diffuse gliomas. *Ann Clin Transl Neurol*. 2024;11(8):2073-2087.

doi: 10.1002/acn3.52128

37. Caii W, Wu X, Guo K, *et al*. Integration of deep learning and habitat radiomics for predicting the response to immunotherapy in NSCLC patients. *Cancer Immunol Immunother*. 2024;73(8):153.

doi: 10.1007/s00262-024-03724-3

38. Liu H, Hou CJ, Wei M, *et al*. High-risk habitat radiomics model based on ultrasound images for predicting lateral neck lymph node metastasis in differentiated thyroid cancer. *BMC Med Imaging*. 2025;25(1):16.

doi: 10.1186/s12880-025-01551-1

39. Wu J, Meng H, Zhou L, *et al*. Habitat radiomics and deep learning fusion nomogram to predict EGFR mutation status in stage I non-small cell lung cancer: A multicenter study. *Sci Rep*. 2024;14(1):15877.

doi: 10.1038/s41598-024-66751-1

40. Jiang Y, Zhou K, Sun Z, *et al*. Non-invasive tumor microenvironment evaluation and treatment response prediction in gastric cancer using deep learning radiomics. *Cell Rep Med*. 2023;4(8):101146.

doi: 10.1016/j.xcrm.2023.101146

41. Avanzo M, Wei L, Stancanello J, *et al*. Machine and deep learning methods for radiomics. *Med Phys*. 2020;47(5):e185-e202.

doi: 10.1002/mp.13678

42. Huang H, Chen H, Zheng D, *et al*. Habitat-based radiomics analysis for evaluating immediate response in colorectal cancer lung metastases treated by radiofrequency ablation. *Cancer Imaging*. 2024;24(1):44.

doi: 10.1186/s40644-024-00692-w

43. 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(1):89.

doi: 10.1038/s41698-025-00884-y