PIONEER
ACADEMIC PUBLISHING LIMITED

Current Research in Medical Sciences
ISSN 2958-0390
www.pioneerpublisher.com/crms
Volume 3 Number 2 June 2024

A Clinical Fusion Model Based on Radiomics Features and Deep Learning for Predicting CDKN2A/B Homozygous Deletion Status in IDH-mutant Diffuse Astrocytoma

Linling Wang[1], Yao Tang[2], Hongyu Pan[3], Zhipeng Wen[4], Xu Cao[5], Zhi Liu[6], Ming Wen[1] & Liqiang Zhang[1]

[1] Department of Radiology, The First Affiliated Hospital of Chongqing Medical University, Chongqing 400016, China
[2] Department of Oncology, People's Hospital of Chongqing Hechuan, Chongqing, China
[3] College of Computer & Information Science, Southwest University, Chongqing 400715, China
[4] Department of Radiology, Sichuan Cancer Hospital, Chengdu 610042, China
[5] School of Medical and Life Sciences Chengdu University of Traditional Chinese Medicine, Chengdu 610032, China
[6] Department of Radiology, Chongqing Hospital of Traditional Chinese Medicine, Chongqing 400021, China
Correspondence: Zhi Liu, Department of Radiology, Chongqing Hospital of Traditional Chinese Medicine, Chongqing 400021, China; Ming Wen, Department of Radiology, The First Affiliated Hospital of Chongqing Medical University, Chongqing 400016, China; Liqiang Zhang, Department of Radiology, The First Affiliated Hospital of Chongqing Medical University, Chongqing 400016, China.

doi:10.56397/CRMS.2024.06.07

Abstract

Purpose: To construct a fusion model for predicting CDKN2A/B homozygous deletion status in patients with isocitrate dehydrogenase (IDH)-mutant diffuse astrocytoma by combining the radiomics features and deep learning (DL). ***Methods:*** A total of 200 IDH-mutant astrocytoma (103 CDKN2A/B homozygous deletion (HD) and 97 CDKN2A/B non-homozygous deletion (NHD)) patients were retrospectively enrolled in the training cohort (n = 140) and the external test cohort (n = 60) for the prediction of CDKN2A/B homozygous deletion status in patients with IDH-mutant astrocytoma. DL model was constructed by SE-Net model, radiomics features of different regions (edema, tumor and overall lesion) were extracted using Pyradiomics, and radiomics model was built by selecting 4 features in the edema region and 7 features in the tumor region by the least absolute shrinkage and selection operator (LASSO). Finally, a fusion model was jointly constructed by the DL model, radiomics model, and clinical features. The predictive performance of the 3 models was evaluated using calibration curves and decision curves, and compared with the fusion model. ***Results:*** Based on the results of the different models, we finally selected a fusion model consisting of DL model, radiomics model, and clinical features. The fusion model showed the best performance with an area under the curve (AUC) of 0.958 in the training cohort and 0.914 in the test cohort. ***Conclusions:*** The

clinical fusion model based on radiomics features and DL features showed good performance in predicting CDKN2A/B homozygous deletion status in patients with IDH-mutant diffuse astrocytoma. ***Key Points:*** 1) Used DL and radiomics to non-invasively predict the CDKN2A/B homozygous deletion status. 2) The model can predict CDKN2A/B homozygous deletion status in IDH-mutant astrocytoma patients. 3) Our result improved classification accuracy and demonstrated better performance in the fusion model.

Keywords: glioma, radiomics, deep learning, Cyclin-Dependent Kinase Inhibitor 2A/B(CDKN2A/B)

1. Introduction

Glioma is the most common primary malignant brain tumor in adults (Ostrom QT, Gittleman H, Truitt G, et al., 2018). Astrocytoma, IDH-mutant, is a glioma subtype, defined as an IDH mutation, excluding the combination of 1p and 19q whole-arm deletions (Ostrom QT, Gittleman H, Truitt G, et al., 2018). In the fifth edition of the World Health Organization (WHO) classification of Central Nervous System (CNS) tumors published in 2021, this subtype was listed as a single type according to histological morphology and molecular characteristics (Louis DN, Perry A, Wesseling P, et al., 2021; Louis DN, Giannini C, Capper D, et al., 2018). CDKN2A/B (Cyclin-dependent kinase inhibitor), which is a cell-cycle regulator, is associated with poor prognosis in a number of other tumors, including ovarian (Xia L, Zhang W & Gao L., 2019), melanoma (Guo Y, Long J & Lei S., 2019), and bladder cancer. IDH-mutant astrocytoma with homozygous deletion of CDKN2A/B have the worst prognosis, and thus are classified directly into WHO grade 4, regardless of histological appearance (Louis DN, Perry A, Wesseling P, et al., 2021). Thus, the prediction of CDKN2A/B homozygous deletion is essential for the evaluation of prognosis and management of IDH-mutant astrocytoma.

CDKN2A/B homozygous deletion is usually detected using FISH, qPCR, or methylation chip techniques, but these techniques are invasive methods based on surgical excision or biopsy (Lu VM, O'Connor KP, Shah AH, et al., 2020). In addition, sampling accuracy is greatly affected by temporal and spatial heterogeneity (Fack F, Tardito S, Hochart G, et al., 2017). Therefore, it is necessary to develop a reliable, non-invasive, and reliable presurgical approach for the identification of CDKN2A/B homozygous deletion status in patients with IDH-mutant astrocytoma. MRI is a non-invasive method that has been widely used to diagnose and locate gliomas, because many of its radiomics signature are related to the genomic changes of gliomas (Ellingson BM., 2015), and can significantly improve the diagnostic efficiency of tumor typing by reflecting the pathology of the lesions (Wei J, Yang G, Hao X, et al., 2019; Ren Y, Zhang X, Rui W, et al., 2019; Liang S, Zhang R, Liang D, et al., 2018). DL and radiomics based on MRI have been extensively studied in the differential diagnosis, grading, genotyping, and prognosis of gliomas (Qian Z, Li Y, Sun Z, et al., 2018; Buda M, AlBadawy EA, Saha A & Mazurowski MA., 2020; Matsui Y, Maruyama T, Nitta M, et al., 2020). For example, DL and radiomics methods have good diagnostic performance and high accuracy in predicting gliomas IDH status, 1p/19q codeletion status and O^6-methylguanine-DNA-methyltransferase (MGMT) promoter methylation status (Choi YS, Bae S, Chang JH, et al., 2021; Chang P, Grinband J, Weinberg BD, et al., 2018; Sun C, Fan L, Wang W, et al., 2022; Cao M, Suo S, Zhang X, et al., 2021). Additionally, visually accessible rembrandt images (VASARI) is a comprehensive feature set with high reproducibility, and it has been shown that some of the features in VASARI can predict gene status in gliomas with good performance (Clark K, Vendt B, Smith K, et al., 2013; Arita H, Yamasaki K, Matsushita Y, et al., 2016; Arita H, Kinoshita M, Kawaguchi A, et al., 2018; Avants BB, Tustison NJ, Song G, et al., 2011; Shinohara RT, Sweeney EM, Goldsmith J, et al., 2014).

In the present study, we intend to combine DL and radiomics to extract deep features of glioma MR images, and then add VASARI related clinical information to the fusion model, with our ultimate goal to develop and validate a model that can noninvasively predict CDKN2A/B homozygous deletion status in patients with IDH-mutant astrocytoma.

2. Materials and Methods

2.1 Patient Enrollment and Image Acquisition

The MRI data and clinicopathological information were derived from the Cancer Imaging Archive (TCIA)/ the Cancer Genome Atlas (TCGA) and the First Affiliated Hospital of Chongqing Medical University (Clark K, Vendt B, Smith K, et al., 2013; Ceccarelli M, Barthel FP, Malta TM, et al., 2016). Inclusion and exclusion criteria were as follows: Case inclusion criteria: (1) IDH-mutant astrocytoma confirmed by pathology; (2) Conventional MRI sequences, including T2 fluid-attenuated inversion recovery (T2-FLAIR) and contrast-enhanced T1 weighted imaging (CE-T1WI); (3) available CDKN2A/B homozygous deletion status information; (4) available clinical characteristics, including age, gender, tumor location, degree of enhancement, percentage of edema, etc. Exclusion criteria: (1) images with serious artifacts; (2) previous radiotherapy, stereotactic radiosurgery, anti-vascular therapy or surgical treatment; (3) The homozygous deletion status of CDKN2A/B was unknown. A total of 200 patients was shuffled and split into a training cohort (CDKN2A/B homozygous deletion (HD): n=71) VS CDKN2A/B non-homozygous deletion (NHD): n=69), and a test cohort (CDKN2A/B HD: n= 32) VS CDKN2A/B NHD: n=28) at a ratio of 7:3.

2.2 Selection of Clinical Variables

We selected six semantic descriptors of image features from VASARI annotations based on MR imaging findings of high-grade gliomas: f4, Enhancement quality; f7, Proportional necrosis; f11, Thickness of enhancement margins; fl2, Definition of enhancement margin; f14, Proportion of edema; and f15, Crosses edema. (In the following table, it is successively expressed as Intensity, Necrosis percentage, Edge thickness, Edge definition, Edema percentage, and Edema midline). Each tumor was independently scored by a radiologist (Xu) and a radiologist (Wen). VASARI scores were used as clinical variables along with age, gender, and WHO classification.

2.3 Data Preprocessing and Segmentation

We pre-processed all images, including bias correction, registration and intensity normalization, to reduce the variation in the use of imaging parameters between hospitals. We performed a mixed white stripe method using the ANTsR and white stripe packages in R to perform intensity normalization, incorporating a statistically principled process of image normalization that preserves inter-tissue grades and matches tissue strength without disrupting the natural balance of tissue strength (Shinohara RT, Sweeney EM, Goldsmith J, et al., 2014; Wang R, Chaudhari P & Davatzikos C., 2021; Egger J, Kapur T, Fedorov A, et al., 2013). We rigidly aligned each subject's image volume to T2-FLAIR, and resampled to isotropic resolution using a linear interpolator to re-interpolate all images to 1mm × 1mm × 1mm pixels in the normalized axes.

Semi-automatic segmentation was performed on T2-FLAIR sequences using the 3D slicer (version 4.3, https://www.slicer.org) (Egger J, Kapur T, Fedorov A, et al., 2013), as T2-FLAIR sequences were widely accepted in identifying glioma boundaries. Two radiologists who had 10 years of experience separated the tumor profile by hand and ignore the final pathological outcome. The internal correlation coefficient (ICC) was used to assess the consistency of observations among observers. ICC≤0.40 indicates poor consistency, 0.40<ICC≤0.75, and high consistency of ICC is more than 0.75. The region of interest (ROI) excludes the worst and best slices that include these structures, to minimize partial volume effects (Egger J, Kapur T, Fedorov A, et al., 2013). If the split is below 5%, the final ROI is defined as the overlap zone of both ROIs. Otherwise, it is decided by the third highly experienced radiologist.

Slices with the largest tumor area were selected from CE-T1WI and T2-FLAIR respectively according to the sectioned mask files, and ROI regions were clipped out from them. The clipped slices were converted into JPG images and enlarged to 224x224 to adapt to the input of the DL model. Random flipping, scaling and contrast enhancement were used to enhance the data, and the number of images was expanded to reduce the overfitting during model training. The process of the whole experiment after image pre-processing is shown in Figure 1.

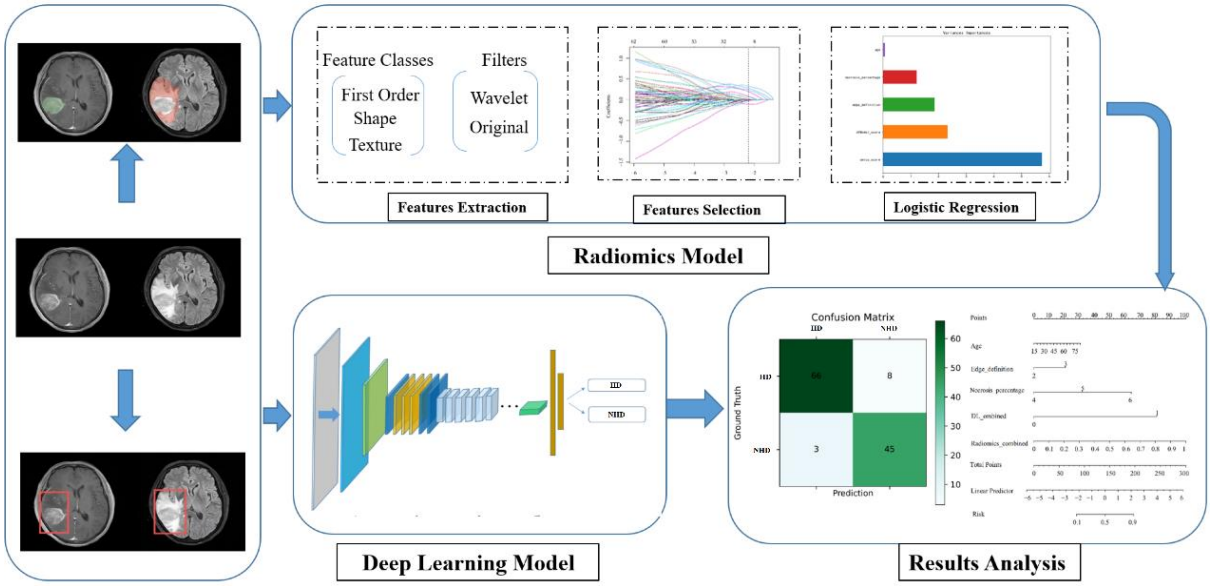


Figure 1. The experimental workflow is divided into four steps: ROI segmentation acquisition, radiomics model construction, deep learning model construction, and predictive model analysis

2.4 Clinical Model Building

Clinical variables included age, gender, WHO grading, tumor centroid site, tumor location, and 6 characteristics as measured by the VASARI score. We constructed clinical models using univariate and multivariate logistic regression. First, univariate logistic regression was used to select the clinical features, and the variables with $p<0.05$ were considered to be statistically significant. Then, a clinical prediction model was built with multivariate logistic regression. The prediction ability of the clinical model was described by the receiver operating characteristic curves (ROC).

2.5 DL-Based Model Building

ResNet has been widely used in various medical image tasks, and it has good classification performance. The SE-Net model used in this experiment is an improved version based on ResNet, which adds a lightweight attention branching structure to the original ResNet. Previous studies of convolution usually focus on the correlation of spatial information, while SE-Net can enhance the overall performance of the model by modeling the correlation between feature channels and reinforcing the important features.

The pictures in the training cohort are expanded by cropping, zooming, flipping and contrast enhancement techniques, and each picture is enhanced by 4 times, and the test cohort does not perform data expansion. Then the picture is input into the network for training. After the network passes through a series of convolutional layers, the final predicted class probability is output by the fully connected layer.

We used an Adam optimizer with an initial learning rate of 0.001, $\beta1$ of 0.9 and $\beta2$ of 0.999, using cross entropy as a loss function and batch-size of 32, and stopping the training when the accuracy of the test cohort was stable. The model was implemented on the PyTroch framework 1.10.0 and python3.7.0, and the GPU is GeForce RTX 3080.

2.6 Radiomics-Based Model Building

We used Pyradiomics to extract radiomics features. 993 features were extracted from the edema and tumor regions of CE-T1WI and T2-FLAIR sequences respectively, including first-order statistical features, shape and intensity features, high-order texture features such as gray-level cooccurrence matrix (GLCM), gray-level run-length matrix (GLRLM), gray-level size zone matrix (GLSZM), gray-level dependence matrix (GLDM) and neighborhood gray-tone difference matrix (NGTDM).

To reduce the size of the features and improve the generalization of the model, we first used the Student's t-test to filter the features, selected the significant features with $p<0.05$, and then used LASSO to further reduce the dimension, and determined the most optimal model after 10-fold

cross-validation. The hyperparameter lambda, and then select a subset of features whose coefficients are not 0 to build a radiomics model.

2.7 Fusion Model

Based on the idea of ensemble learning, we combined the prediction results of clinical features, DL and radiomics models to construct the final multiple logistic regression model for predicting CDKN2A/B homozygous deletion status. A nomogram was used to describe the relationship among the variables, and calibration and decision curves were used to evaluate the predictive performance of the models.

2.8 Statistical Analysis

All statistical analyses were performed with R software (version 4.2.0, https://www.rproject.org/). For continuous variables, the t-test or Mann-Whitney U-tests were used, and the classification of variables was chi-squared. All the statistical tests were two-sided, and the statistical significance was $p<0.05$.

3. Results

3.1 Clinical Characteristics

Details of patient characteristics are summarized in Table 1. Age (p = 0.872), gender (p = 0.427), and WHO grading (p = 0.331) was not significantly different, but there were significant differences on the side ($p<0.05$). In the training cohort, the definition of enhancement margin, thickness of enhancement margins and proportional necrosis have significant differences ($p<0.05$).

Table 1. The clinical characteristics of patients in the training and test cohorts

Clinical characteristics	**Training cohort(N=140)**			**Test cohort(N=60)**			***P*(Inter)**
	HD(n=71)	**NHD(n=69)**	***p*(Intra)**	**HD(n=32)**	**NHD(n=28)**	***p*(Intra)**	
Age			0.124			0.943	0.821
<52	41(57.75)	30(43.48)		17(56.67)	12(42.86)		
≥52	30(42.25)	39(56.52)		15(43.33)	16(57.14)		
Gender			0.927			0.321	0.427
Male	39(54.93)	34(49.28)		14(43.75)	17(60.71)		
Female	32(45.07)	35(50.72)		17(56.25)	11(39.29)		
WHO grading			0.443			0.526	0.331
2-3	38(53.52)	45(65.22)		11(34.38)	12(42.86)		
4	33(46.48)	24(34.78)		21(65.63)	16(57.14)		
Side			0. 123			0.512	<0.05
left	31(43.66)	33(47.83)		12(37.50)	16(57.14)		
right	40(56.34)	36(52.17)		20(62.50)	12(42.86)		
Location			0.098			0.332	0.313
occipital	19(26.76)	17(24.64)		7(21.88)	6(21.43)		
temporal	18(25.35)	19(27.54)		5(15.63)	7(25.00)		
parietal	21(29.58)	18(26.09)		9(28.13)	5(17.86)		
frontal	13(18.31)	15(21.74)		11(34.38)	10(35.71)		
Edge definition			<0.05			0.562	0.421
Well-defined	34(47.89)	37(53.62)		13(40.62)	14(50.00)		
Poorly-defined	37(52.11)	32(46.38)		19(59.38)	14(50.00)		

Clinical characteristics	Training cohort(N=140)			Test cohort(N=60)			*P*(Inter)
	HD(n=71)	NHD(n=69)	*p*(Intra)	HD(n=32)	NHD(n=28)	*p*(Intra)	
Edema midline			0.674			0.433	0.341
No	36(50.70)	36(52.17)		16(50.00)	16(57.14)		
Yes	35(49.30)	33(47.83)		16(50.00)	12(42.86)		
Intensity	2.13±0.3	2.39±0.4	0.6	2.18±0.31	2.2±0.31	0.43	0.822
Edge thickness	3.22±0.42	3.25±0.38	<0.05	3.53±0.3	3.33±0.59	0.144	0.431
Necrosis percentage	5.42±0.52	4.74±0.33	<0.05	5.42±0.75	4.32±0.73	0.218	0.252
Edema percentage	4.34±0.5	4.54±0.5	0.321	4.35±0.52	4.43±0.26	0.045	0.347

The clinical features were analyzed by univariate and multivariate logistic regression, and the results were presented in Table 2. We selected 3 clinical features with p<0.05: age, definition of enhancement margin and proportional necrosis to construct the clinical prediction model.

Table 2. Univariate and multivariate analysis of clinical characteristics

	Univariate		Multivariate	
	OR(95%CI)	P value	OR(95%CI)	P value
Age	0.96(0.41-2.1)	＜0.05	0.96(0.41-2.1)	＜0.05
Gender	0.93(0.42-2.12)	0.65	0.52 (0.19-1.61)	0.41
WHO grading	1.81(0.61-4.21)	0.32	1.32 (0.43-6.21)	0.25
Side	0.53(0.18-1.24)	0.25	0.41 (0.26-1.62)	0.42
Location	1.12 (0.41-3.21)	0.91	2.21 (0.42-4.2)	0.46
Edge definition	4.41 (3.2-10.1)	＜0.05	8.42 (2.19-21.1)	＜0.05
Edema midline	1.0 (0.45-2.7)	0.743	1.02 (0.44-2.82)	0.78
Intensity	1.1 (0.51-2.2)	0.723	1.1(0.29-2.72)	0.342
Edge thickness	3.2 (2.1-10.1)	0.241	2.51 (1.3-11.2)	0.421
Necrosis percentage	1.1 (1.2-4)	＜0.05	1.84 (1.42-3.43)	0.442
Edema percentage	1.42(0.61-2.91)	0.231	3.21(1.41-9.81)	0.21

3.2 Predictive Performance of the Clinical, DL, Radiomics and Fusion Models

A multivariate logistic regression model was constructed using the screened clinical characteristics. Figure 2A shows that the AUCs of the clinical model in the training cohort and test cohort are 0.826 (95%CI: 0.745-0.909) and 0.659 (95%CI: 0.514-0.801), respectively. Figure 2B shows that the AUCs of the DL model in the training cohort and test cohort are 0.873 (95%CI: 0.818-0.931) and 0.774 (95%CI: 0.634-0.871), respectively. Table 3 summarizes the predictive performance of each model.

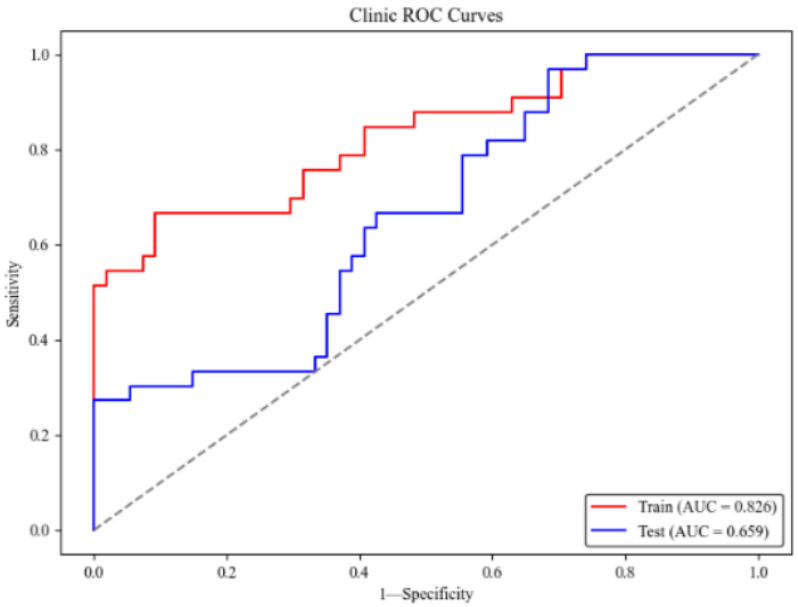


Figure 2A. The ROC curve of clinical model in the training cohort and test cohort

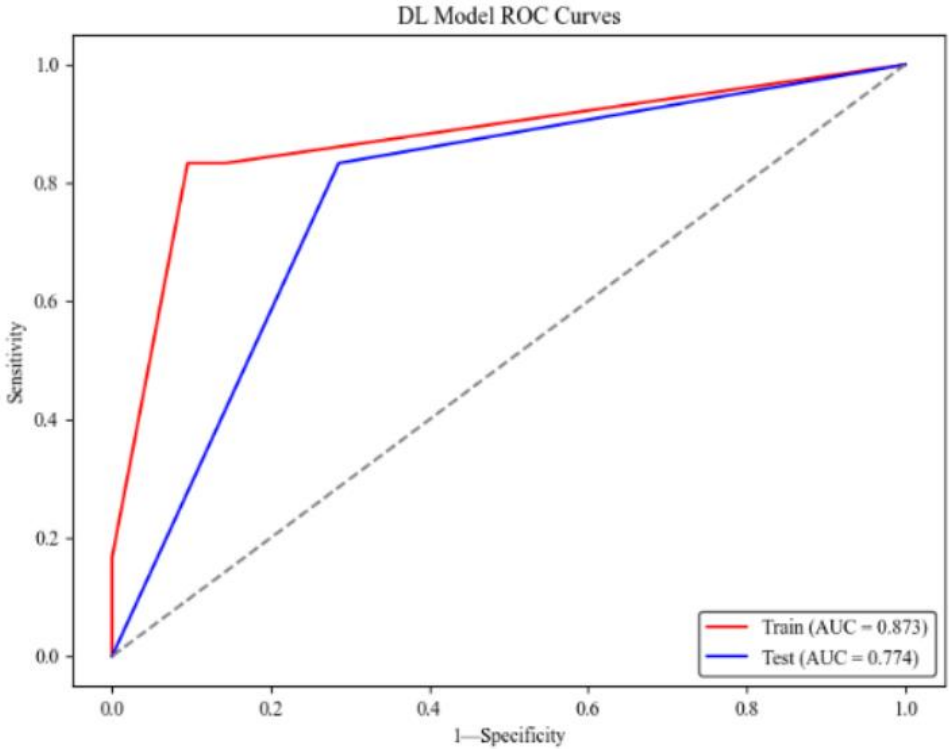


Figure 2B. The ROC curve of DL model in the training cohort and test cohort

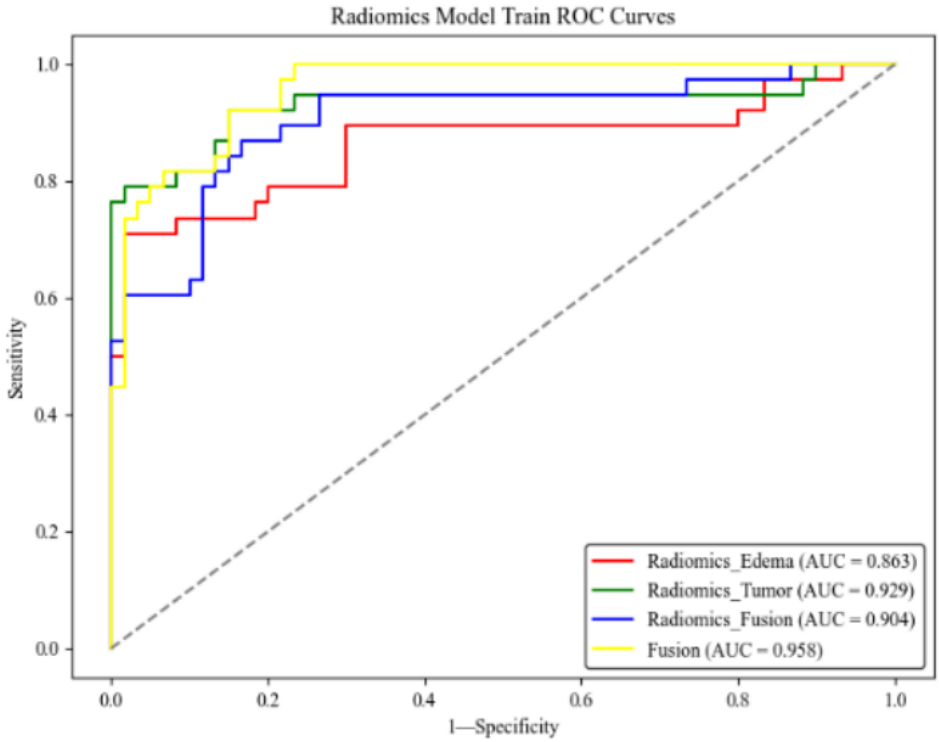


Figure 2C. The ROC curve of radiomics models of edema region and tumor region in the training cohort

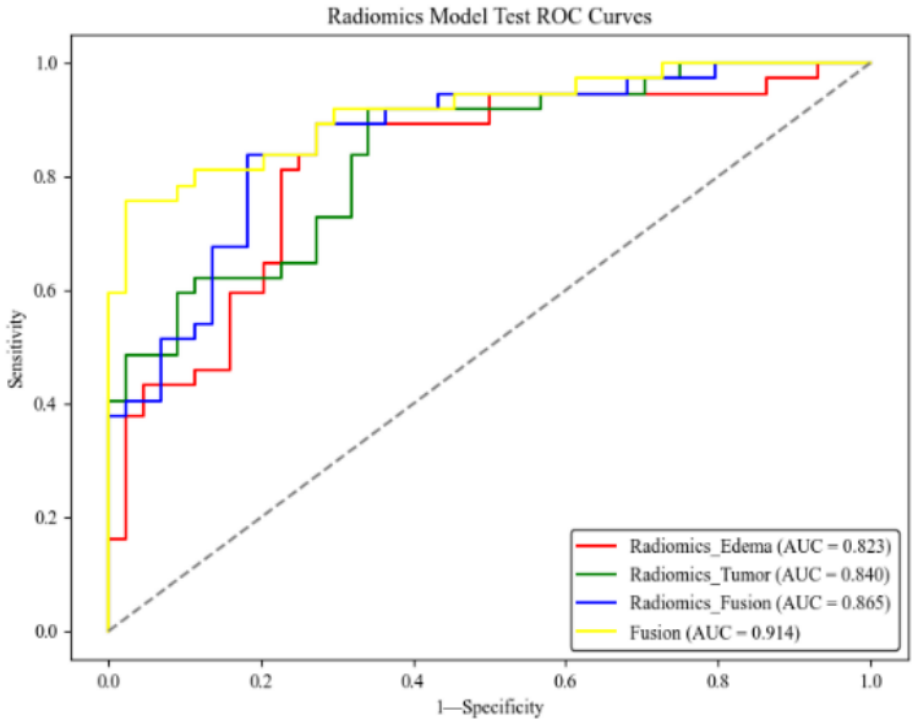


Figure 2D. The ROC curve of radiomics models of edema region and tumor region in the test cohort

Table 3. DL model, radiomics edema model, radiomics tumor model, radiomics fusion model, clinical model and fusion model performance

Models	Training cohort(N=140)	Test cohort(N=60)

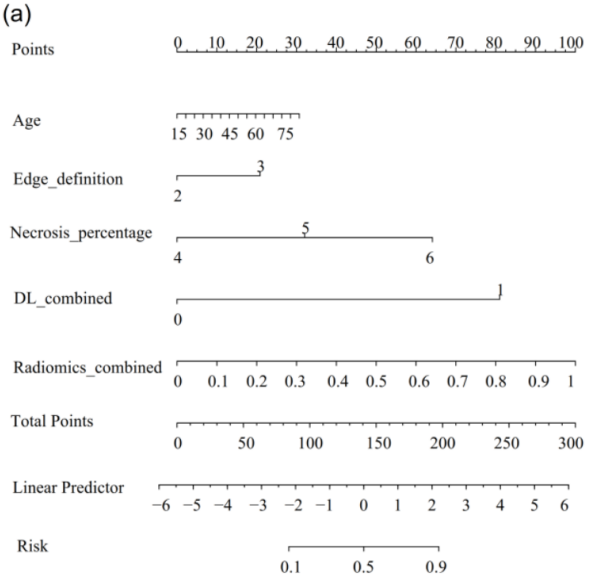
	Sensitivity	Specificity	Accuracy	AUC(95%CI)	Sensitivity	Specificity	Accuracy	AUC(95%CI)
DL Model	0.838	0.864	0.872	0.873(0.818-0.931)	0.731	0.732	0.741	0.774(0.634-0.871)
Radiomics Edema	0.689	0.765	0.713	0.863(0.801-0.901)	0.616	0.853	0.652	0.823(0.723-0.941)
Radiomics Tumor	0.865	0.864	0.863	0.929(0.886-0.965)	0.715	0.821	0.759	0.840(0.737-0.950)
Radiomics Fusion	0.827	0.808	0.816	0.904(0.851-0.950)	0.611	0.861	0.704	0.865(0.764-0.962)
Clinic Variable	0.721	0.773	0.756	0.826(0.745-0.909)	0.628	0.665	0.666	0.659(0.514-0.801)
Fusion	0.903	0.922	0.913	0.958(0.899-0.982)	0.804	0.903	0.838	0.914(0.816-0.987)

For the DL model, we input the ROI images cropped from the CE-T1WI and T2-FLAIR sequences and used the SE-Net model to train in the training cohort for 26 epochs to achieve convergence, and the accuracy of the test cohort is 0.741. The ROC curve (Figure 2B) shows that the AUC of the DL model in the training cohort and the test cohort is 0.873 (95%CI: 0.818-0.931) and 0.774 (95%CI: 0.634-0.871), the DL model shows better predictive power than the clinical model.

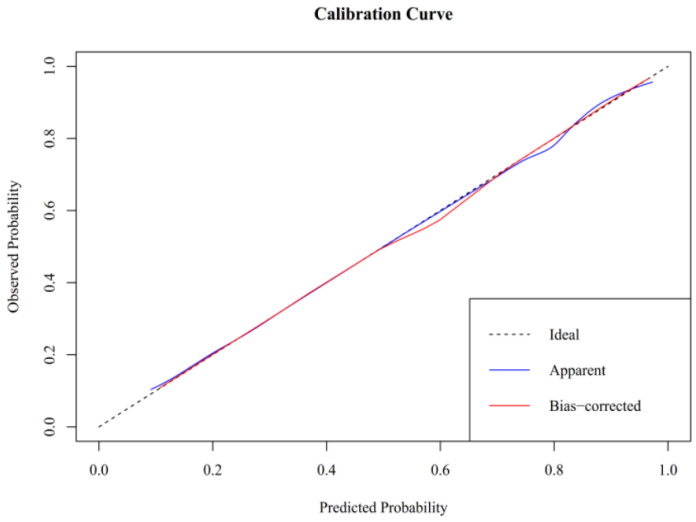
We extracted the radiomics features from the edema region and tumor region respectively, and constructed the tumor-based, edema-based and mixed region-based radiomics models. Figure 2C and 2D shows ROC curves of different types of radiomics models. The AUC of radiomics tumor model in test cohort was 0.840(0.737-0.950), and the AUC of radiomics edema tumor model in test cohort was 0.823(0.723-0.941). Then, we built a radiomics fusion model using 4 features from the edema region and 7 features from the tumor region, which further improved the predictive performance with an AUC of 0.865(0.764-0.962).

To explore the impact of different models on the prediction of CDKN2A/B homozygous deletion status, we combined clinical features, DL model and radiomics model to construct a final fusion model. The AUC of the fusion model was 0.914(0.816-0.987), and the accuracy was 0.838. After the Delong test, the clinical model and the fusion model had a significant difference ($p<0.05$). We drew the nomogram of the fusion model to show the value of predicting CDKN2A/B homozygous deletion status (Figure 3A). The calibration curve showed that the fusion model had good predictive performance (Figure 3B), and the decision curve analysis (DCA) curve showed that the overall net benefit of the fusion model was better than any single model (Figure 3C). Figure 3D represented the risk stratification of 1000 people using the fusion model, showing the comparison between the number of people classified as high risk by the fusion model and the number of people who are actually high risk under different threshold probabilities.

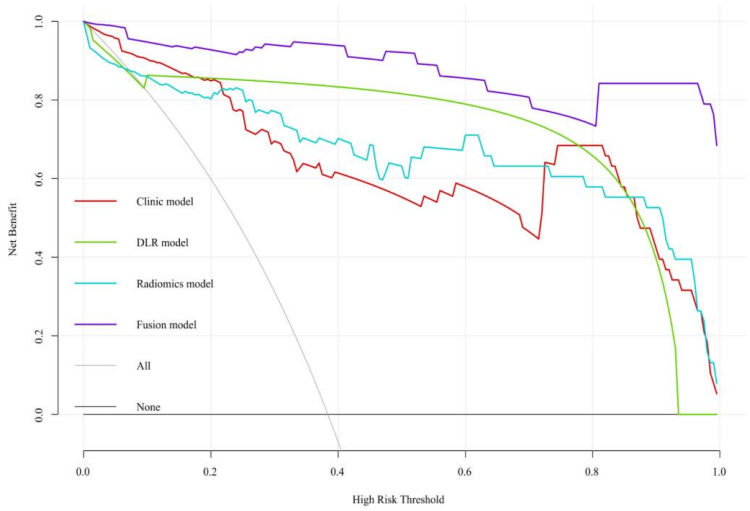
(a)


Points
0 10 20 30 40 50 60 70 80 90 100
Age
15 30 45 60 75
Edge_definition
2 3
Necrosis_percentage
4 5 6
DL_combined
0 1
Radiomics_combined
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Total Points
0 50 100 150 200 250 300
Linear Predictor
−6 −5 −4 −3 −2 −1 0 1 2 3 4 5 6
Risk
0.1 0.5 0.9

(b)


Calibration Curve
Observed Probability
Predicted Probability
0.0 0.2 0.4 0.6 0.8 1.0
Ideal
Apparent
Bias−corrected

(c)


Net Benefit
High Risk Threshold
0.0 0.2 0.4 0.6 0.8 1.0
Clinic model
DLR model
Radiomics model
Fusion model
All
None

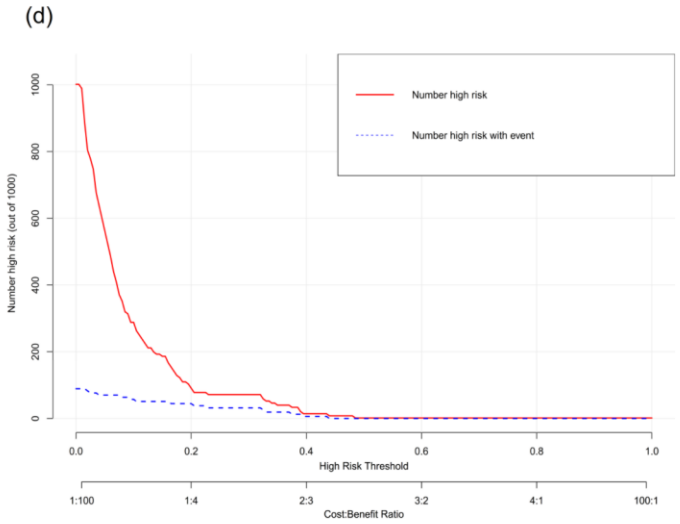


Figure 3 A. Clinical nomogram established using the fusion model; **B.** The calibration curves of the fusion model on the test set; **C.** DCA curves of the fusion model, overall hybrid and the clinic model, The X-axis means the threshold probability; the Y-axis shows the model benefit; **D.** The Clinical Impact Curve of the fusion model, the red curve (Number high risk) indicates the number of people who are classified as positive(high-risk) by the fusion model at each threshold probability; the blue curve (Number high risk with event) is the number of true positives at each threshold probability

4. Discussion

In this study, we combined DL, radiomics, and clinical information to build prediction models, and ultimately selected the optimal feature set of each fusion model for noninvasive prediction of CDKN2A/B homozygous deletion status in IDH-mutant astrocytoma. The results showed that the fusion model produced the best outcome in the forecast process, with AUC values of 0.958 and 0.914 for the training and test cohort. The fusion model can significantly improve the prediction performance of genetic biomarkers, and this noninvasive prediction is a promising approach in clinical practice.

Previous studies have shown that 30-80% of glioma patients have mistakes or defects in the CDKN2 family of tumor suppressor genes (Sauerbrei W, Royston P & Binder H., 2007; Crespo I, Vital AL, Gonzalez-Tablas M, et al., 2015). CDNK2A/B, which are located on chromosome 9, encode cyclin-dependent kinase inhibitors 2A-p16INK4a and 2B-p15INK4b, respectively, which inhibit the activity of CDK kinase and thereby control the progression of cell cycle G1 (Toyokuni S., 2010). Coin activation of CDKN2A/B is an inevitable event of cell senescence induced by Rb protein phosphorylation and oncogene, which ultimately leads to uncontrolled cell growth and proliferation and contributes to tumor invasion and metastasis. Recent studies have shown that CDKN2A/B homozygous deletion is associated with poorer overall survival in astrocytoma (Crespo I, Vital AL, Gonzalez-Tablas M, et al., 2015; Reis GF, Pekmezci M, Hansen HM, et al., 2015; Lu VM, O'Connor KP, Shah AH, et al., 2020). The presence of homozygous deletion of CDKN2A/B in IDH-mutant low-grade astrocytoma results in biological behavior consistent with that of high-grade astrocytoma (Brat DJ, Aldape K, Colman H, et al., 2020). In addition, gliomas patients with homozygous deletion of CDKN2A/B have a worse prognosis than those without homozygous deletion (Li Z, Kaiser L, Holzgreve A, et al., 2021). Therefore, the 2021 WHO classification of CNS tumor highlights the importance of the homozygous deletion status of CDKN2A/B in IDH-mutant astrocytoma.

DL and radiomics have been widely used in the prediction of glioma genes. We used CNN to predict the homozygous deletion status of CDKN2A/B, which showed superior performance to the clinical model alone, but not as good as the radiomics with manually extracted features. We presented a combination of clinical features, DL model and radiomics model, with AUC of 0.958 and 0.914 for the

training cohort and test cohort, respectively. In addition, instead of extracting the CNN features, the DL model we adopted used end-to-end input graphics and output results. Tasks that could be solved by multiple modules were modeled by a single model, which simplified the process and avoided the disadvantage of errors with the final target caused by inconsistent training objectives of multiple modules. We combined DL and radiomics, which is both convenient to obtain images, cost saving, no radiation dose, no additional damage to the patient's body, and DL can also analyze the whole images well. In general, the study method we use contributes to an efficient and reproducible workflow, thereby improving its clinical applicability.

Although DL and radiomics showed quantitative predictive performance, VASARI features were still valuable in predicting genes and are meaningful when added to the fusion model. Sun et al. (2022) and Cao et al. (2021) applied VASARI features to IDH and 1p/19q gene mutation, and the results showed that some of the VASARI features could have auxiliary predictive value for their models. The current study applied the VASIRI features to predict the homozygous deletion status of CDKN2A/B, which was not available in previous studies. The final results showed that the definition of enhancement margin and proportional necrosis had a p value <0.05 and were predictive of CDKN2A/B homozygous deletion status. Thus, it can be demonstrated that VASARI features have a positive role in predicting CDKN2A/B homozygous deletion status in IDH-mutant astrocytoma.

Several limitations of the research remain to be explored further. First, the sample size is small and the performance of the DL model cannot be fully utilized, so a large sample size and prospective validation are needed. Second, our noninvasive model was built on CE-T1WI and T2-FLAIR conventional MRI sequences and showed good performance, but still need to consider functional MRI sequences such as diffusion-weighted imaging (DWI), perfusion-weighted imaging (PWI), and diffusion kurtosis imaging (DKI), which may further improve the prediction performance. Third, our images are manually segmented, which is time-consuming and error-prone, therefore, it is necessary to obtain ROIs by semi-automatic or automatic segmentation in the future research, which may further increase the precision of prediction.

5. Conclusions

In summary, we used DL combined with radiomics to non-invasively predict the homozygous deletion status of CDKN2A/B in IDH-mutant astrocytoma, improving classification accuracy and demonstrating better performance in the fusion model.

Abbreviations

Cyclin-Dependent Kinase Inhibitor 2A/B**(**CDKN2A/B), deep learning (DL), CDKN2A/B homozygous deletion (HD), CDKN2A/B non-homozygous deletion (NHD), least absolute shrinkage and selection operator (LASSO), visually accessible Rembrandt images (VASARI), decision curve analysis (DCA), diffusion-weighted imaging (DWI), perfusion-weighted imaging (PWI), and diffusion kurtosis imaging (DKI)

References

Arita H, Kinoshita M, Kawaguchi A, et al. (2018). Lesion location implemented magnetic resonance imaging radiomics for predicting IDH and TERT promoter mutations in grade II/III gliomas. *Sci Rep*, *8*, 11773. https://doi.org/10.1038/s41598-018-30273-4

Arita H, Yamasaki K, Matsushita Y, et al. (2016). A combination of TERT promoter mutation and MGMT methylation status predicts clinically relevant subgroups of newly diagnosed glioblastomas. *Acta neuropathol commun*, *4*, 79. https://doi.org/10.1186/s40478-016-0351-2

Avants BB, Tustison NJ, Song G, et al. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage, 54*, 2033–2044. https://doi.org/10.1016/j.neuroimage.2010.09.025

Brat DJ, Aldape K, Colman H, et al. (2020). cIMPACT-NOW update 5: recommended grading criteria and terminologies for IDH-mutant astrocytomas. *Acta Neuropathol, 139*, 603–608. https://doi.org/10.1007/s00401-020-02127-9

Buda M, AlBadawy EA, Saha A, Mazurowski MA. (2020). Deep Radiogenomics of Lower-Grade Gliomas: Convolutional Neural Networks Predict Tumor Genomic

Subtypes Using MR Images. *Radiology: Artificial Intelligence, 2*, e180050. https://doi.org/10.1148/ryai.2019180050

Cao M, Suo S, Zhang X, et al. (2021). Qualitative and Quantitative MRI Analysis in IDH1 Genotype Prediction of Lower-Grade Gliomas: A Machine Learning Approach. *BioMed Research International, 2021*, 1–10. https://doi.org/10.1155/2021/1235314

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

Chang P, Grinband J, Weinberg BD, et al. (2018). Deep-Learning Convolutional Neural Networks Accurately Classify Genetic Mutations in Gliomas. *AJNR Am J Neuroradiol, 39*, 1201–1207. https://doi.org/10.3174/ajnr.A5667

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

Clark K, Vendt B, Smith K, et al. (2013). The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *J Digit Imaging, 26*, 1045–1057. https://doi.org/10.1007/s10278-013-9622-7

Crespo I, Vital AL, Gonzalez-Tablas M, et al. (2015). Molecular and Genomic Alterations in Glioblastoma Multiforme. *The American Journal of Pathology, 185*, 1820–1833. https://doi.org/10.1016/j.ajpath.2015.02.023

Egger J, Kapur T, Fedorov A, et al. (2013). GBM Volumetry using the 3D Slicer Medical Image Computing Platform. *Sci Rep, 3*, 1364. https://doi.org/10.1038/srep01364

Ellingson BM. (2015). Radiogenomics and Imaging Phenotypes in Glioblastoma: Novel Observations and Correlation with Molecular Characteristics. *Curr Neurol Neurosci Rep, 15*, 506. https://doi.org/10.1007/s11910-014-0506-0

Fack F, Tardito S, Hochart G, et al. (2017). Altered metabolic landscape in IDH -mutant gliomas affects phospholipid, energy, and oxidative stress pathways. *EMBO Mol Med, 9*, 1681–1695. https://doi.org/10.15252/emmm.201707729

Guo Y, Long J, Lei S. (2019). Promoter methylation as biomarkers for diagnosis of melanoma: A systematic review and meta-analysis. *Journal Cellular Physiology, 234*, 7356–7367. https://doi.org/10.1002/jcp.27495

Li Z, Kaiser L, Holzgreve A, et al. (2021). Prediction of TERTp-mutation status in IDH-wildtype high-grade gliomas using pre-treatment dynamic [18F]FET PET radiomics. *Eur J Nucl Med Mol Imaging, 48*, 4415–4425. https://doi.org/10.1007/s00259-021-05526-6

Liang S, Zhang R, Liang D, et al. (2018). Multimodal 3D DenseNet for IDH Genotype Prediction in Gliomas. *Genes 9*, 382. https://doi.org/10.3390/genes9080382

Louis DN, Giannini C, Capper D, et al. (2018). cIMPACT-NOW update 2: diagnostic clarifications for diffuse midline glioma, H3 K27M-mutant and diffuse astrocytoma/anaplastic astrocytoma, IDH-mutant. *Acta Neuropathol, 135*, 639–642. https://doi.org/10.1007/s00401-018-1826-y

Louis DN, Perry A, Wesseling P, et al. (2021). The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro-Oncology, 23*, 1231–1251. https://doi.org/10.1093/neuonc/noab106

Lu VM, O'Connor KP, Shah AH, et al. (2020). The prognostic significance of CDKN2A homozygous deletion in IDH-mutant lower-grade glioma and glioblastoma: a systematic review of the contemporary literature. *J Neurooncol, 148*, 221–229. https://doi.org/10.1007/s11060-020-03528-2

Matsui Y, Maruyama T, Nitta M, et al. (2020). Prediction of lower-grade glioma molecular subtypes using deep learning. *J Neurooncol, 146*, 321–327. https://doi.org/10.1007/s11060-019-03376-9

Ostrom QT, Gittleman H, Truitt G, et al. (2018). CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011–2015. *Neuro-Oncology, 20*, iv1–iv86. https://doi.org/10.1093/neuonc/noy131

Qian Z, Li Y, Sun Z, et al. (2018). Radiogenomics of lower-grade gliomas: a radiomic signature as a biological surrogate for

survival prediction. *Aging*, *10*, 2884–2899. https://doi.org/10.18632/aging.101594

Reis GF, Pekmezci M, Hansen HM, et al. (2015). *CDKN2A* Loss Is Associated With Shortened Overall Survival in Lower-Grade (World Health Organization Grades II–III) Astrocytomas. *J Neuropathol Exp Neurol*, *74*, 442–452. https://doi.org/10.1097/NEN.0000000000000188

Ren Y, Zhang X, Rui W, et al. (2019). Noninvasive Prediction of IDH1 Mutation and ATRX Expression Loss in Low-Grade Gliomas Using Multiparametric MR Radiomic Features. *Magnetic Resonance Imaging*, *49*, 808–817. https://doi.org/10.1002/jmri.26240

Sauerbrei W, Royston P, Binder H. (2007). Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Statistics in Medicine*, *26*, 5512–5528. https://doi.org/10.1002/sim.3148

Shinohara RT, Sweeney EM, Goldsmith J, et al. (2014). Statistical normalization techniques for magnetic resonance imaging. *NeuroImage Clinical*, *6*, 9–19. https://doi.org/10.1016/j.nicl.2014.08.008

Sun C, Fan L, Wang W, et al. (2022). Radiomics and Qualitative Features from Multiparametric MRI Predict Molecular Subtypes in Patients With Lower-Grade Glioma. *Front Oncol*, *11*, 756828. https://doi.org/10.3389/fonc.2021.756828

Toyokuni S. (2010). Mysterious link between iron overload and CDKN2A/2B. *J Clin Biochem Nutr*, *48*, 46–49. https://doi.org/10.3164/jcbn.11-001FR

Wang R, Chaudhari P, Davatzikos C. (2021). Harmonization with Flow-Based Causal Inference. In: De Bruijne M, Cattin PC, Cotin S, et al (eds) *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*. Springer International Publishing, Cham, pp. 181–190.

Wei J, Yang G, Hao X, et al. (2019). A multi-sequence and habitat-based MRI radiomics signature for preoperative prediction of MGMT promoter methylation in astrocytomas with prognostic implication. *Eur Radiol*, *29*, 877–888. https://doi.org/10.1007/s00330-018-5575-z

Xia L, Zhang W, Gao L. (2019). Clinical and prognostic effects of *CDKN2A*, *CDKN2B* and *CDH13* promoter methylation in ovarian cancer: a study using meta-analysis and TCGA data. *Biomarkers*, *24*, 700–711. https://doi.org/10.1080/1354750X.2019.1652685