



# The Nomogram of MRI-based Radiomics with Complementary Visual Features by Machine Learning Improves Stratification of Glioblastoma Patients: A Multicenter Study

Yuyun Xu, MD,<sup>1</sup>  Xiaodong He, MD,<sup>1</sup> Yumei Li, MD,<sup>1</sup> Peipei Pang, MD,<sup>2</sup> Zhenyu Shu, MD,<sup>1\*</sup> and Xiangyang Gong, PhD<sup>1</sup> 

**Background:** Glioblastomas (GBMs) represent both the most common and the most highly malignant primary brain tumors. The subjective visual imaging features from MRI make it challenging to predict the overall survival (OS) of GBM. Radiomics can quantify image features objectively as an emerging technique. A pragmatic and objective method in the clinic to assess OS is strongly in need.

**Purpose:** To construct a radiomics nomogram to stratify GBM patients into long- vs. short-term survival.

**Study Type:** Retrospective.

**Population:** One-hundred and fifty-eight GBM patients from Brain Tumor Segmentation Challenge 2018 (BRATS2018) were for model construction and 32 GBM patients from the local hospital for external validation.

**Field Strength/Sequence:** 1.5 T and 3.0 T MRI Scanners, T<sub>1</sub>WI, T<sub>2</sub>WI, T<sub>2</sub>FLAIR, and contrast-enhanced T<sub>1</sub>WI sequences

**Assessment:** All patients were divided into long-term or short-term based on a survival of greater or fewer than 12 months. All BRATS2018 subjects were divided into training and test sets, and images were assessed for ependymal and pia mater involvement (EPI) and multifocality by three experienced neuroradiologists. All tumor tissues from multiparametric MRI were fully automatically segmented into three subregions to calculate the radiomic features. Based on the training set, the most powerful radiomic features were selected to constitute radiomic signature.

**Statistical Tests:** Receiver operating characteristic (ROC) curve, sensitivity, specificity, and the Hosmer–Lemeshow test.

**Results:** The nomogram had a survival prediction accuracy of 0.878 and 0.875, a specificity of 0.875 and 0.583, and a sensitivity of 0.704 and 0.833, respectively, in the training and test set. The ROC curve showed the accuracy of the nomogram, radiomic signature, age, and EPI for external validation set were 0.858, 0.826, 0.664, and 0.66 in the validate set, respectively.

**Data Conclusion:** Radiomics nomogram integrated with radiomic signature, EPI, and age was found to be robust for the stratification of GBM patients into long- vs. short-term survival.

**Level of Evidence:** 3

**Technical Efficacy Stage:** 2

J. MAGN. RESON. IMAGING 2021.

Glioblastomas (GBMs) represent the most common and highly malignant primary brain tumors associated with a particularly poor median survival duration of 14 months, and a 5-year survival rate of 5.5%.<sup>1,2</sup> Approximately 13,000 GBM cases in the United States have been confirmed each

year, with an incidence rate as low as 3.2 per 100,000 members of the population.<sup>2</sup>

The standard treatment for GBM is surgical resection, followed by radiation therapy and/or chemotherapy.<sup>3</sup> However, the overall survival (OS) time widely varies across

View this article online at [wileyonlinelibrary.com](http://wileyonlinelibrary.com). DOI: 10.1002/jmri.27536

Received Sep 22, 2020, Accepted for publication Jan 16, 2021.

\*Address reprint requests to: Z.S., 310013 Hangzhou, China. E-mail: [coojuty@hotmail.com](mailto:coojuty@hotmail.com)

Zhenyu Shu and Xiangyang Gong contributed equally to this work.

Contract grant sponsor: Fund of Health Commission of Zhejiang Province; Contract grant number: 2017KY230.

From the <sup>1</sup>Department of Radiology, Zhejiang Provincial People's Hospital, Affiliated People's Hospital of Hangzhou Medical College, Hangzhou, China; and <sup>2</sup>GE Healthcare (China), Shanghai, China

Additional supporting information may be found in the online version of this article

individuals due to the inherent heterogeneity of GBMs.<sup>4</sup> The radiological appearance of GBM reflects its aggressive nature, which contributes to its heterogeneity; thus, physicians commonly rely on multiparametric magnetic resonance imaging (MRI) to aid in prognostication.<sup>5</sup> The visual imaging features derived from MRI have previously been used to predict of survival.<sup>6</sup> However, this assessment has been difficult, as image interpretation is highly subjective, challenging, and time-consuming. Therefore, objective quantification of the image characteristics may help stratify patients based on predictions for long- vs. short-term survival, yielding valuable information for clinicians who can then offer precise management strategies.

Radiomics, which is a promising and rapidly growing discipline, is a method that can be used to extract a large number of quantitative features from medical images and converts the information into mineable data.<sup>7</sup> These data can be subsequently analyzed to construct biomarkers for disease diagnosis, prediction, stratification, and prognosis through feature selection. Radiomics has the benefit of using routine, clinical MRI sequences to evaluate GBM images, which could be highly cost-effective and efficient as it is a further investigation on the basis of existing medical image and requires no additional examination.<sup>8</sup> Previous studies have indicated that imaging features extracted from conventional MRI in routine clinic can predict survival in patients with GBM before the treatment.<sup>9, 10</sup> Furthermore, previous studies showed that intratumoral radiomic features outperformed conventional prognostic factors and provided complementary value to genomic information for prognosis.<sup>11,12</sup>

Radiomic features can be employed to quantify radiological appearance, including heterogeneity, size, and intensity; however, it is less effective at determining ependymal and pia mater involvement (EPI). EPI has been shown to be a risk factor for OS. In one recent study, it was noted that ependymal involvement manifested as extensive contact with the ventricle in GBM patients; as such, the cerebrospinal fluid (CSF) may allow for widespread tumor cell dissemination, potentially affecting survival.<sup>13</sup> In one study, pia mater involvement was reported to occur in up to 23.4% of GBM patients (75 of 321), and this was associated with poor survival<sup>14</sup>. Therefore, the assessment of EPI in GBM patients is crucial because it would influence the treatment decision. However, radiomics from automatic or manually segmented tumors cannot reflect the involvement of both.

Machine-learning techniques enable the analysis of large quantities of imaging features to extract highly predictive imaging characteristics, which is increasingly used to generate genetic and prognostic biomarkers for GBMs.<sup>15</sup> Multiple models of survival analysis that apply MR imaging features have also been generated to predict patient outcomes in GBM.<sup>16</sup> Several studies have applied support vector machines (SVMs) on MRI data to predict survival, whereas Zacharaki

et al used classification trees to imaging features from preoperative and postoperative MRI scans to predict short- vs. long-term survival in patients with high-grade gliomas.<sup>17–19</sup> The feasibility of these predictive models is limited by their inability to directly calculate and visualize these outcomes, which could be further improved by using nomograms in the clinic. Nomogram is actually a pictorial representation of a complex mathematical formula, using biologic and clinical variables to graphically depict a statistical prognostic model that generates a probability of a clinical event, such as OS for a given individual. Nomograms are commonly used tools with rapid computation through user friendly digital interfaces, together with high accuracy and easily understood prognose, allowing for better prognosis estimation and clinical decision-making in oncology.<sup>20</sup>

Thus this study sought to investigate whether combining radiomic features and specific visual features could better stratify patients into long- vs. short-term survival sets.

## Materials and Methods

### Patients

This retrospective study was approved by our institutional review board, and the requirements for informed consent were waived. The preoperative clinic routine brain MRI scans of 158 GBM patients from Brain Tumor Segmentation Challenge 2018 (BRATS2018) and external validation dataset of 32 GBM patients from the local hospital were retrospectively analyzed.<sup>21–23</sup> Inclusion criteria were: (1) pathologically confirmed WHO grade IV GBM; (2) no prior history of treatment, including surgery, chemotherapy, or radiation therapy; (3) pre-treatment MRI with conventional routine sequences available. Patients with insufficient quality image were excluded. In addition, the patients of BRATS2018 were divided into two sets according to alphabetical order of the labeled name in the dataset, the training set and testing set, with a ratio of 7:3.

All patients were pathologically confirmed to have a diagnosis of GBM. Clinicodemographic data, including age and treatment were obtained. MRI scans were performed at multiple institutions ( $N = 20$ , 19 institutions for the BRATS2018 data and 1 local institutions for the external validation) with different clinical protocols and various scanners. The routine sequences included  $T_1$ -weighted imaging ( $T_1W$ ),  $T_2$ -weighted imaging ( $T_2W$ ),  $T_2FLAIR$  (FLAIR), and contrast-enhanced  $T_1W$  ( $T_1CE$ ). The detailed parameters of all the MR images can be found in Table S1 in the Supplementary Material. All patients were divided into long-term (12 months or longer) or short-term (12 months or less) survival sets. Their mean age was 64.6 years old. Of all patients, 97 had a long-term OS and 93 had a short-term OS. In this study, data from BRATS2018 were used to build the model, and data from the local medical center were used to validate the model.

### Visual Features Assessment

All MR images were reviewed and assessed for multifocality and EPI status by three neuroradiologists (Yuyun Xu, Xiaodong He, and Zhenyu Shu of 14, 16, 20 years of experience, respectively) who were blinded to the clinical data using ITK-SNAP (<http://www.>

itksnap.org/) and assessed the images independently. Discordant interobserver interpretations were resolved by discussion until a consensus was achieved. Ependymal involvement was defined if the contrast-enhanced tumor lesion contacted the lining of the ventricle.<sup>24</sup> Pia mater involvement was defined as diffuse or nodular enhancement of the meninges, subarachnoid space, intracranial nerves, or spinal cord surface in T<sub>1</sub> and/or T<sub>2</sub> FLAIR gadolinium-enhanced images.<sup>25</sup> EPI positive was defined as the involvement of the ependyma and/or pia mater. Multifocality was defined as having more than one area of tumor enhancement, which was located separately from the other enhanced areas on a postcontrast T<sub>1</sub>W image.<sup>26</sup>

### Image Preprocessing

All BRATS multimodal scans are available as NIfTI files (.nii.gz). In light of the different imaging parameters, the MR images were preprocessed to standardize data analysis across patients and were co-registered to the same anatomical template, interpolated to the same resolution ( $1 \times 1 \times 1 \text{ mm}^3$ ), and skull stripped.<sup>22, 27</sup> Gaussian filtering was used to reduce noise in the images; magnetic field migration correction was performed to reduce external interference factors; and histogram mapping and intensity standardization were performed to limit the gray level of all images to 0–255, and the image grayscale intensity level was then discretized and normalized by downsampling each image into 32 bins to reduce image noise.

With such fixed values and numbers of bins, the image gray range was divided into equally spaced intervals. Therefore, the bin size and intensity resolution of the discretized volumes depended on the grayscale value (i.e., four bin sizes for each grayscale).

### Extraction of Radiomic Features

All MRI images were imported into the AGK software (Artificial-Intelligent Radio-Genomics Kits; GE Healthcare, Chicago, IL, USA) to extract radiomics features. Fully automatic GBM segmentation was performed for both datasets, and three subregions were obtained, including edema, enhancing, necrotic and nonenhancing area, the details of the algorithm and process can be found in Figs. S2 and S3 in the Supplemental Material). The corresponding radiomic features were also extracted from the dataset from the manually segmented subregions by experts provided by BRATS2018 dataset. The segmented regions of interest of the subregions were further analyzed by a function tool of AGK software for feature extraction. Radiomic features were calculated, including histogram, Haralick, form factor, gray-level co-occurrence matrix (GLCM), and run-length matrix, and gray level size zone matrix. See Supplemental Material for details.

### Construction of the Radiomic Signature

Analysis of variance of the extracted features was performed based on the training set. The variance of each feature was calculated, and then the features greater than the threshold 1 were retained. Then, the maximum relevance minimum redundancy (mRMR) algorithm was used to extract the most robust features in the training set. The maximum-relevance selection aimed to select features that had a maximal correlation to the actual survival time. At the same time, the minimum-redundancy selection ensured that the selected features had minimal redundancy among each other. Subsequently, an

emerging gradient boosting decision tree (GBDT) algorithm was used to further reduce the dimensionality of the remaining features. The detailed dimensionality reduction method of GBDT can be found in the Supplemental Material. Finally, logistic regression (LR) analysis was used to obtain robust features that participate in the construction of the radiomic signature. For the details of feature dimensionality reduction, please refer to the Supplemental Material for this article. The radiomics workflow is presented in Fig. 1.

### Assessment of Radiomics Signature

Based on the training dataset, three subregions were automatically segmented from each sequence of T<sub>2</sub>WI, T<sub>1</sub>WI, T<sub>2</sub>FLAIR, and T<sub>1</sub>CE, and the features of each subregion were extracted separately, which constituted a total feature set of 12 subregional features from these four sequences, and then dimensionality reduction was performed on this feature set. After feature dimensionality reduction, a joint feature set containing three automatically segmented subregions of each sequence was used to build radiomic signature. In order to demonstrate the correlation between the radiomic features and OS, scores for each patient in the training set were calculated using the radiomics signature formula to reflect the actual survival probability, which was defined as the Rad-score.

The radiomic signature of the test set was calculated using the same formula used in the training set.

Furthermore, a feature set containing the features of four sequences based on each subregion was extracted as well as the established radiomic signatures of the three subregions.

Finally, based on the radiomics features extracted by manual segmentation, the Rad-score of each patient's survival was calculated, using the radiomic signature formula constructed by radiomic features based on automatic segmentation.

### Nomogram Construction and Evaluation

Univariate LR analysis was performed on each potential predictor variable, including multifocality, EPI, age, and radiomics signature in the training set, selecting the independent predictors to construct the models. In this study, five machine-learning methods were applied including LR, SVM, as well as the Bayes and K nearest neighbor algorithms, and random forests to build the prediction model based on the filtered prediction factors, and then used the test set data to calculate the predictive performance of the prediction model based on the training set. Receiver operating characteristic (ROC) was used to evaluate the model's diagnostic accuracy and finally utilized the DeLong test to select the optimal machine-learning model. As a result, the LR machine-learning method was used to construct the prediction model and nomogram was built. A radiomics nomogram was then built on the basis of the optimal machine-learning model.

Nomogram efficiency was verified by the training set, the test set and the external verification set, including calibration efficiency, diagnostic accuracy, and net value, which were evaluated by the calibration curve, ROC curve, and decision curve analysis (DCA), respectively. According to the nomogram for each patient in the external verification data set, the predicted survival for each patient was calculated, then all the subjects were divided into long-term survival set and short-term survival set by the cutoff value of the ROC curve. Finally, the clinical effect was determined by the actual

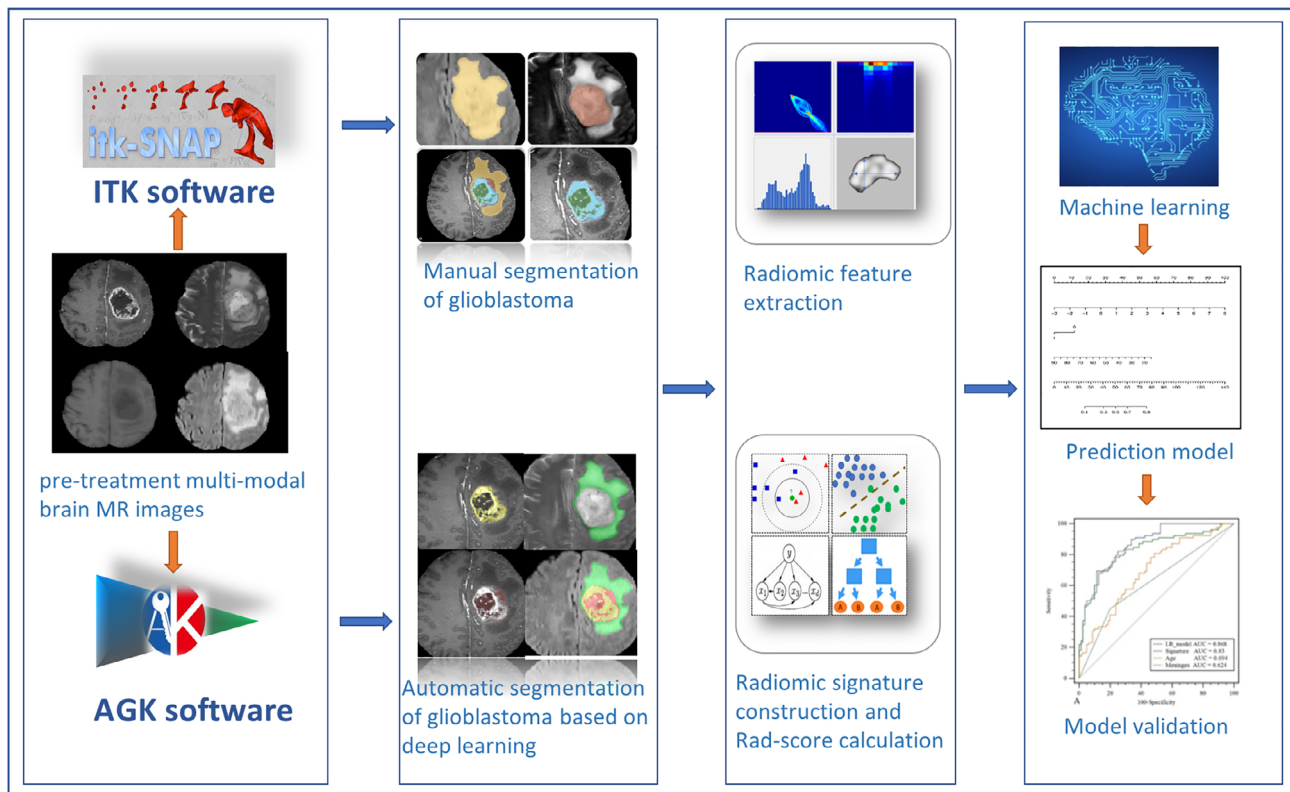


FIGURE 1: Flow chart of this study.

survival state in the different sets. In this study, the best cutoff value is determined by using the Yoden index, which is equal to the sensitivity - (1 - specificity). The maximum value of this index value is the best cutoff value. Finally, a DCA was conducted for the different predictors in the BRATS2018 set and the external validation set.

**Statistical Analysis**

SPSS 17.0 software (IBM Corporation, Armonk, NY, USA) was used to perform the Kolmogorov–Smirnov test to evaluate the normality of the distribution of the data, and the chi-squared test was used for categorical data. Parametric data were assessed using an independent-samples *t*-test, whereas nonparametric

data were assessed using the Mann–Whitney *U* test between two sets.

One-way analysis of variance was used to compare the continuous variables among the three sets. The interobserver agreement of visual features by three neuroradiologists was evaluated by intraclass correlation coefficient (ICC). The area under the curve (AUC) of the ROC curve was used to evaluate the accuracy of the radiomic signatures in the training and test sets. The diagnostic efficacy of Rad-score in all patients was evaluated by ROC curve, and the DeLong test was used to compare the difference in the performance of radiomic signature based on automatic segmentation and manual segmentation in all patients. The MedCalc15.8 software (MedCalc, Ostend, Belgium) was used to

TABLE 1. Patient Characteristics in the Primary and Internal Validation Cohorts

Variable		Training set (N = 110)	Test set (N = 48)	Validation set (N = 32)	P-value
Age (years)		60 ± 14	61 ± 11	58 ± 14	0.663
Radiomic signature		0.139 ± 2.076	-0.039 ± 2.114	0.208 ± 2.383	0.852
Multifocality (N [%])	No	84 (76.4)	34 (70.8)	25 (78.1)	0.702
	Yes	26 (23.6)	14 (29.2)	7 (21.9)	
EPI (N [%])	No	34 (30.9)	17 (35.4)	11 (34.4)	0.837
	Yes	76 (69.1)	31 (64.6)	21 (65.6)	

EPI = ependymal and pia mater involvement.

assess the ROC curves and calculate the cutoff value, and the differences between AUCs were compared with the DeLong test. The R statistical software package, version 3.4.1, was used for all other statistical analyses. The “mRMRe” packages were used for mRMR analyses. Calibration plots and the radiomics nomogram were established with the “rms” package, and DCA with the “dca.R” package. Two-sided *P*-values <0.05 were considered statistically significant.

## Results

### Patients’ Clinical Characteristics

The MR images showed that 128 patients were EPI-positive and 62 were EPI-negative. The ICCs between the three neuroradiologists of multifocality and EPIs were 0.885 (95% CI, 0.854–0.911) and 0.839 (95% CI, 0.797–0.875), respectively. The differences in age, radiomic signature, multifocality, and EPIs of the patients in the training, testing, and validation sets were not statistically significant with *P* value of 0.663, 0.852, 0.702, 0.837, respectively, all *P* > 0.05 (Table 1). The differences in age, radiomic signature, and meninges between patients with short- and long-term OS were statistically significant in training, testing, and validation sets (*P* < 0.05 in all cases), whereas the differences in multifocality were not statistically significant *P* > 0.05 (Table 2).

### Construction and Assessment of the Radiomics Signature

A single region extracted 378 features, and so 4536 features were extracted from each patient’s four MRI sequences based on three subregions. Finally, seven GLCM features (Haralick correlations [*N* = 2], inverse difference moments [*N* = 3]; inertia [*N* = 1]; and cluster prominence [*N* = 1]) were selected by dimension reduction to construct the radiomic signature. The radiomic signature model constructed based on three separate subregions of the tumor showed favorable diagnostic performance. In addition, compared with the signature model constructed on a single subregion of the tumor, the signature model constructed based on the combined area had a higher diagnostic performance in the training set and the test set (see Fig. 2 and Table 3 for details). The AUC values of the radiomic signature constructed based on automatic segmentation and manual segmentation in all datasets were 0.826 and 0.717, the sensitivities were 0.692 and 0.564, and the specificities were 0.8 and 0.836, respectively. The DeLong test showed that the AUC values of the two were significantly different (*P* < 0.05).

TABLE 2. Patient Characteristics of the Training and Validation Sets for Survival

Variable	Training set (N = 110)			Test set (N = 48)			Validation set (N = 32)		
	Short survival (N = 56)	Long survival (N = 54)	<i>P</i> -value	Short survival state (N = 24)	Long survival (N = 24)	<i>P</i> -value	Short survival (N = 13)	Long survival (N = 19)	<i>P</i> -value
Age(year)	65 ± 10	57 ± 11	<0.01*	65 ± 10	55 ± 15	0.014*	67 ± 12	53 ± 13	0.006
Radiomic signature	-0.965 ± 1.267	1.285 ± 2.137	<0.01*	-1.064 ± 1.012	0.986 ± 2.431	0.001*	-1.479 ± 0.962	1.362 ± 2.389	<0.01*
Multifocality (N [%])	No Yes	44 (78.6) 12 (21.4)	0.579	16 (31) 8 (69)	18 (56.6) 6 (43.4)	0.525	12 (92.3) 1 (7.7)	13 (68.4) 6 (31.6)	0.108
EPI (N [%])	No Yes	11 (19.6) 45 (80.4)	0.009*	5 (65.5) 19 (34.5)	12 (70) 12 (30)	0.035*	2 (15.4) 11 (84.6)	10 (52.6) 9 (47.4)	0.033*

EPI = ependymal and pia mater involvement.

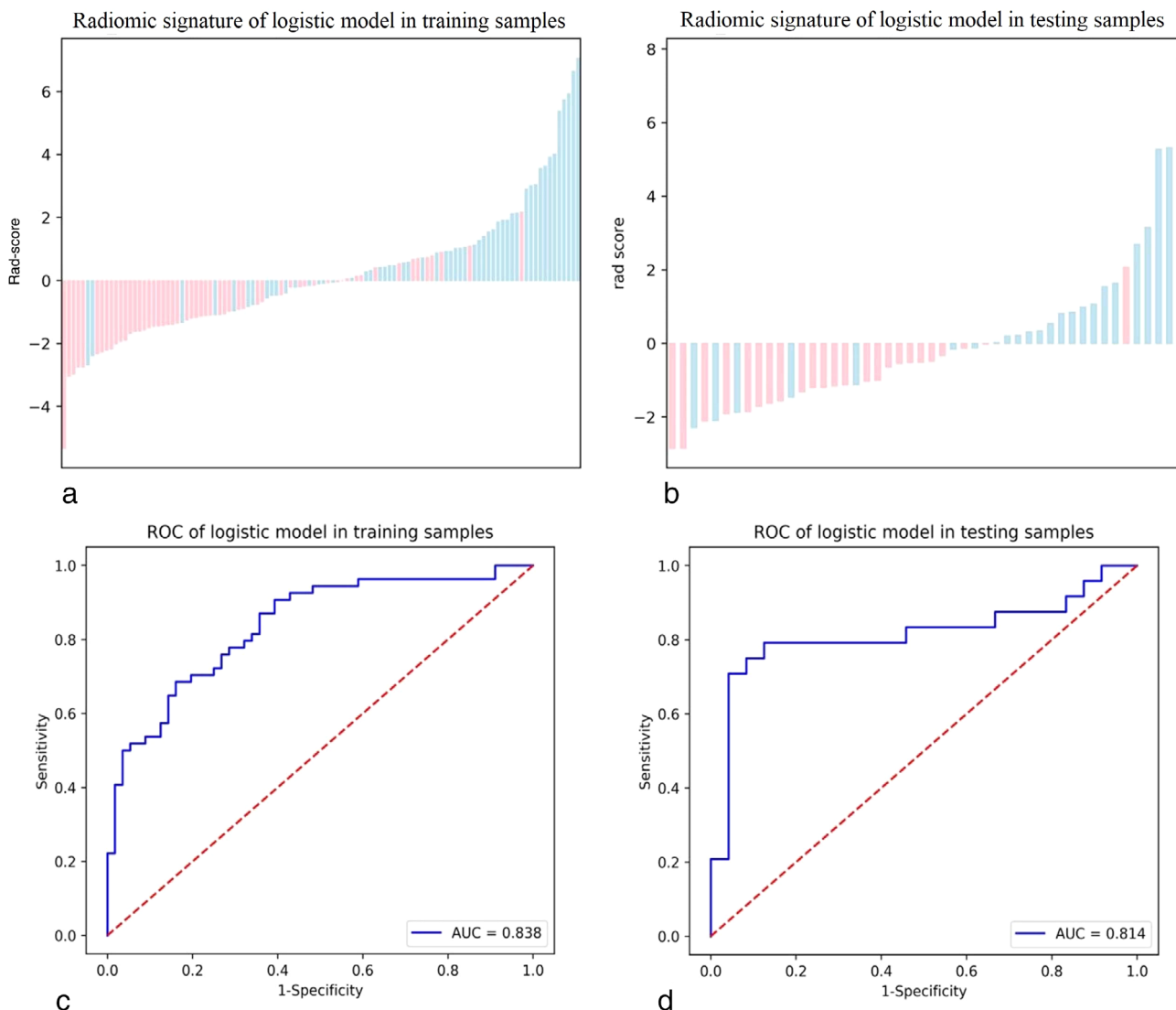


FIGURE 2: Score diagrams of the radiomics signature in (a) the training set and (b) the test set. (c, d) The diagnostic accuracy of the prediction scores of the radiomic signature in the training and test sets

**Radiomics Nomogram Construction and Validation**

Univariate LR analysis results showed that age, EPI, and radiomic signature were independent predictors of OS, as

shown in Table 4. The accuracy, specificity, and sensitivity of the independent predictors in the training, test, and validation set are shown in Table 5. The ROC curve showed

**TABLE 3. Diagnostic Efficacy of Radiomic Signature based on Different Tumor Subregions**

Set	Performance features	Edema	Enhancing	Necrotic +nonenhancing	Merge area
Training set	AUC	0.815	0.81	0.823	0.838
	Sensitivity	0.685	0.685	0.648	0.704
	Specificity	0.786	0.661	0.786	0.786
Test set	AUC	0.786	0.715	0.795	0.814
	Sensitivity	0.708	0.458	0.75	0.708
	Specificity	0.792	0.75	0.751	0.958

AUC = area under curve.

**TABLE 4. Logistic Regression Analyses to Predict the Survival State**

Variable	Univariate logistic regression	
	OR (95% CI)	P-value
Age (per 1 year)	0.93 (0.893–0.968)	<0.001*
Multifocality (no vs. yes)	1.283 (0.531–3.1)	0.579
EPI (no vs. yes)	0.329 (0.141–0.772)	0.011*
Radiomic signature (per 0.1 increase)	2.718 (1.796–4.114)	<0.001*

EPI = ependymal and pia mater involvement.

that the prediction model constructed by the LR machine-learning method had the highest diagnostic efficiency in the training and test sets (AUC = 0.878), AUCs for SVM, KNN, Forest, Bayes were 0.872, 0.822, 0.860, 0.877, respectively. The DeLong test showed no statistically significant difference between the AUC value of the LR and other machine-learning methods with  $P > 0.05$  (see Fig. 3 for details). The prediction model based on LR machine-learning method was constructed and the corresponding nomogram was built, as shown in Fig. 4. The nomogram had a survival prediction accuracy of 0.878 and 0.875, a specificity of 0.875 and 0.583, and a sensitivity of 0.704 and 0.833, respectively, in the training and test set. The calibration curve showed good calibration performance in the training and test sets, indicating good consistency between the predicted OS using the nomogram and the actual OS in the training and test sets. The DCA also showed good net benefits in the training and test sets (Fig. 5).

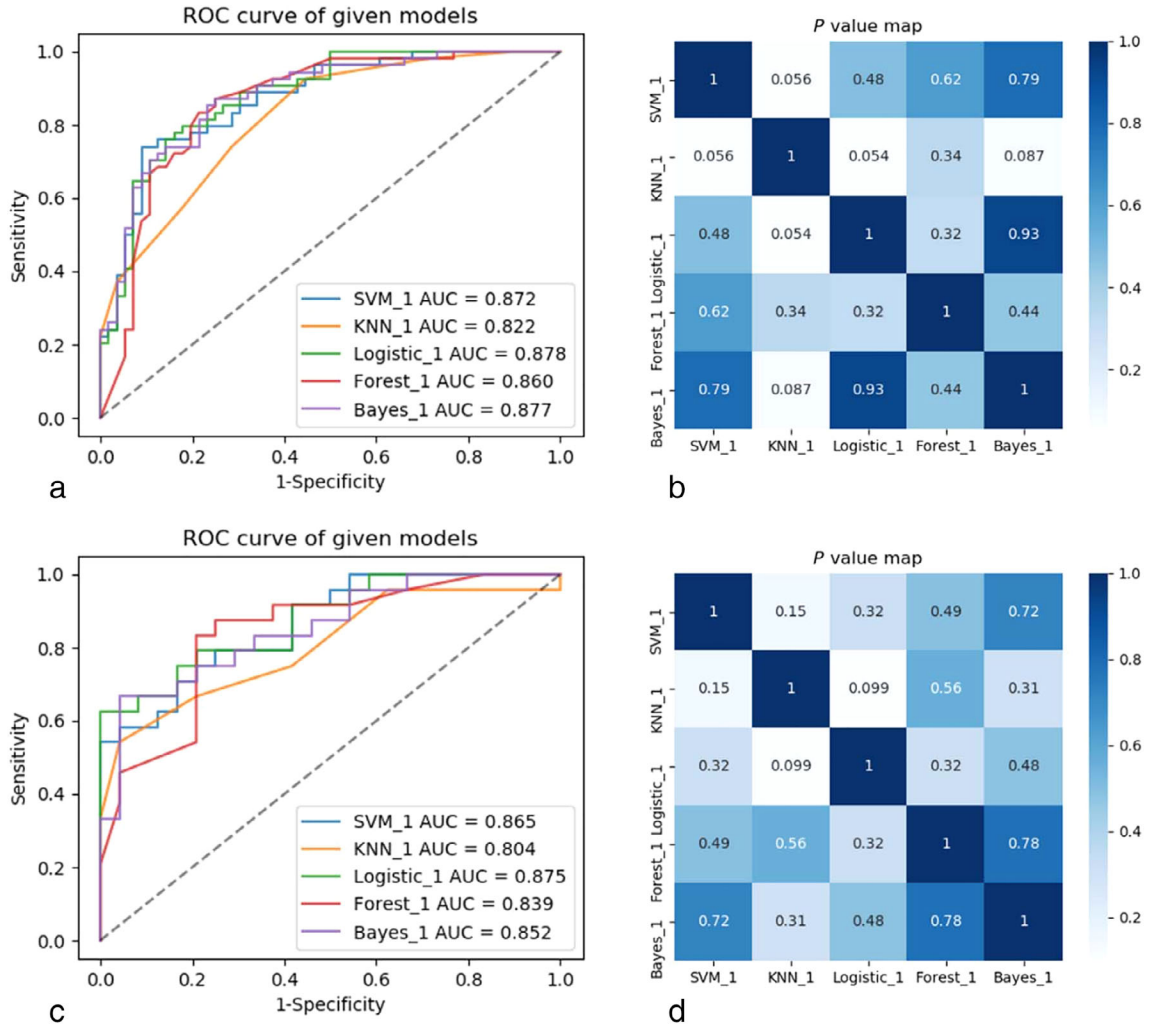
**Overall External Validation of the Radiomics Nomogram**

The AUC value of the ROC curve showed the accuracy of the nomogram, radiomic signature, age, EPI, signature + age and signature + EPI for external validation set were 0.858 (95% CI 0.69–0.956), 0.826 (95% CI 0.651–0.936), 0.664 (95% CI 0.476–0.82), 0.660 (95% CI 0.472–0.817), 0.838 (95% CI 0.665–0.944), and 0.846 (95% CI 0.675–0.949) respectively. According to the best diagnostic cutoff values for the nomogram (cutoff value: 0.264), the patients were divided into long-term and short-term OS sets. There was a significant difference in the number of GBM cases between the long- and short-term OS sets ( $P < 0.05$ ; Fig. 6). The results of DCA show that the net benefit of nomogram is similar to the radiomic

**TABLE 5. The Diagnostic Efficiency of Independent Predictors in Different Sets**

Item	Training set			Test set			Validation set		
	Age	EPI	Signature	Age	EPI	Signature	Age	EPI	Signature
AUC (95% CI)	0.692 (0.597–0.777)	0.615 (0.517–0.706)	0.838 (0.756–0.901)	0.694 (0.545–0.819)	0.646 (0.495–0.778)	0.813 (0.675–0.911)	0.689 (0.509–0.813)	0.66 (0.472–0.817)	0.827 (0.725–0.971)
Specificity	0.518	0.804	0.839	0.625	0.792	0.958	0.539	0.846	0.937
Sensitivity	0.815	0.426	0.685	0.75	0.51	0.708	0.895	0.474	0.737

EPI = ependymal and pia mater involvement.



**FIGURE 3:** (a, b) The ROC curves of different machine-learning methods in the training set and the heat maps of their *P*-values of compared AUCs. (c, d) The ROC curves of different machine-learning methods in the test set and the heat maps of their *P*-values of compared AUCs.

signature in both datasets. In addition, the calibration curve of the variables showed good consistency in both datasets, as shown in Fig. 7.

### Discussion

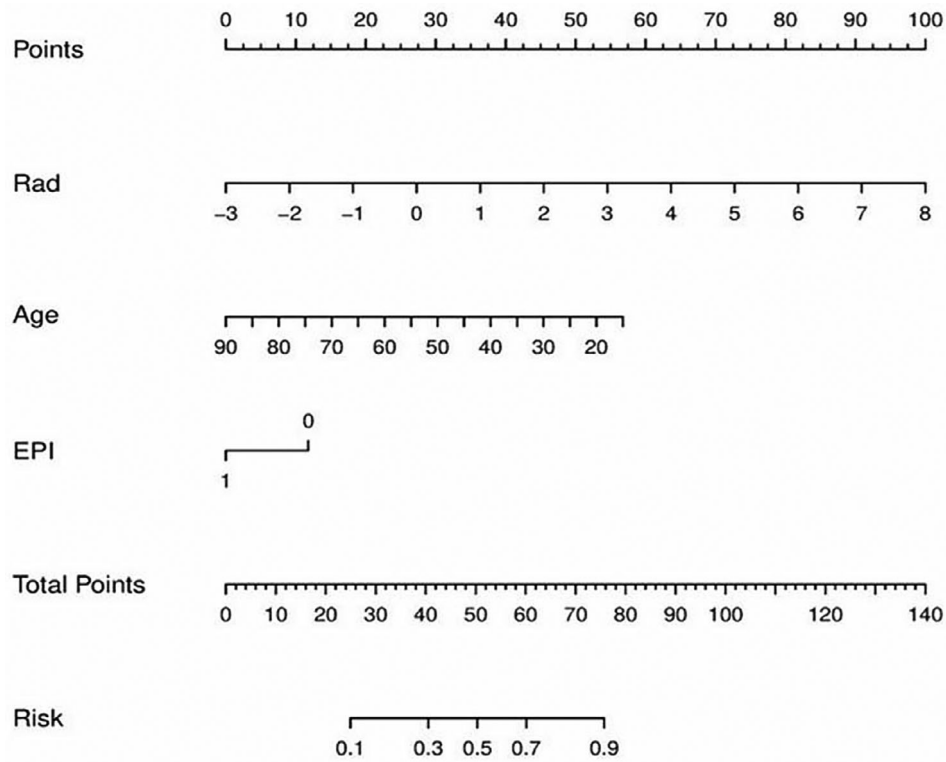
OS prediction is important when tailoring treatment options to reduce burden and improve quality of life, as well as to guide the development of novel therapies. In this study, we investigated the role of integrated nomograms on the OS stratification of GBM patients. EPI, age, and radiomic signature are found to be independent predictors of OS for GBM patients. The nomogram was created by integrating the three independent predictors, had the best performance when stratifying GBM patients into long- vs. short-term survival, which could help clinicians develop optimal treatment plans.

Our results showed that the nomogram based on radiomic features could predict the survival state of GBM patients very well. Interestingly, seven GLCM features survived as robust radiomic signatures and participated in the

construction of the nomogram, whereas the other feature types were not selected. GLCM features were defined to measure the relationship between two neighboring pixels and to reflect the local heterogeneity of the tumor. Tumor heterogeneity, which is often characterized by high cellularity, diffuse infiltration, and necrosis, is one of the main causes of treatment resistance in GBM, directly affecting patient prognosis (i.e., OS time).<sup>28</sup> Our results are consistent with those of a previous study, which indicated that local heterogeneity played a crucial role in survival stratification in GBM patients.<sup>29</sup>

EPI was also found to be an independent predictor. EPI positivity indicates the probability with which tumor seeds through CSF dissemination, which is a potential contributing factor of a poor clinical outcome, and can be reflected by MRI images as involvement of the ependyma and/or pia mater.<sup>30</sup> A recent report confirmed in a multivariate analysis of 647 GBM patients, which adjusted for age, gender, extent of resection, postoperative treatment, and tumor volume, that





**FIGURE 4: Nomogram of the prediction model based on the LR machine-learning method.**

ependymal involvement was an independent predictor of survival.<sup>31</sup> Previous studies showed meningeal involvement (as detected on a postoperative MRI) with a diagnostic specificity of 77% is associated with poor survival outcomes.<sup>32</sup> However, the correlation between pia mater involvement and survival is a little controversial. Park et al showed that meningeal involvement did not affect the clinical results in GBM patients; however, this may be due to the small number of patients.<sup>24</sup> In addition, age is a predictor of poor prognosis in GBM patients, in line with other GBM studies.<sup>33</sup> Multifocality did not appear to be an independent predictor, consistent with the findings of one previous study.<sup>34</sup>

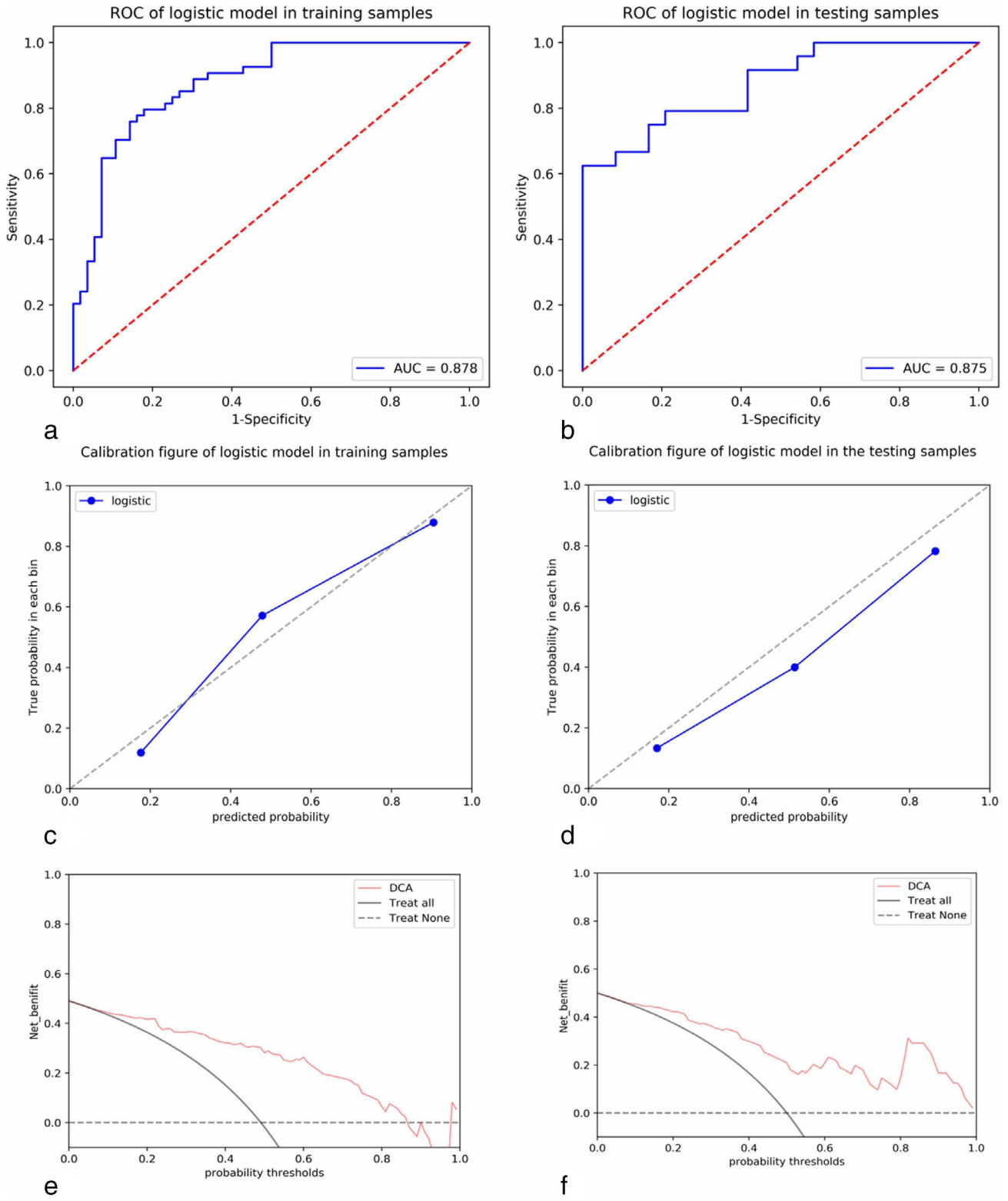
After taking the radiomic signature, age, and EPI together, the nomogram yielded the best performance on the OS stratification of GBM patients, which might mean they are complementary factors, thus improving classification performance. Although the DCA shows that the net benefit of the nomogram is much greater than the age or EPI, the net benefit of nomogram and radiomic signature in the BRATS2018 dataset was very close, while in the external validation dataset the net benefit of the nomogram is only slightly higher than that of radiomics, which may further reveal that the nomogram’s prediction mechanism is largely dependent on radiomic signature.

A prediction model is an important component of radiomics analysis. Highly accurate and reliable models are needed for clinicians to provide decision support. Machine

learning can provide highly accurate and reliable models to improve clinical decision-making in clinical oncology, which has played an increasingly important role in prediction and oncology.<sup>35</sup> In this study, multiple machine-learning methods were used to select an optimal method, which greatly improved the accuracy of nomogram. The joint multi-sequence and multi-region features were selected to construct the radiomic signature, which may comprehensively reflect the heterogeneity of the tumor, and had a good predictive performance. The result supports previous report that a joint model has improved performance when compared with clinical and radiologic models.<sup>36</sup> Furthermore, our nomogram’s performance outperformed some of those noted in previous studies, perhaps due to the joint features we used. In addition, multicenter data were utilized, which further shows that our nomogram can be widely generalized.<sup>37</sup>

### Limitations

First, as a retrospective study, the MR imaging data used in this investigation were acquired with various imaging parameters and contrast agents, and EPI assessment precision may be of a little influence. A prospective clinical study using a larger population is needed to better understand the role of EPIs in the prognosis of GBM patients. Second, all patients were divided into short- and long-term survival, using a survival time of binary endpoints instead



**FIGURE 5:** (a, b) The ROC curves of the nomograms in the training and test sets; (c, d) the calibration curves of the nomograms in the training and test sets; and (e, f) the DCA of the nomograms in the training and test sets.

of actual survival analysis, which may result in loss of threshold-related results and information. However, machine learning prediction based on binary survival time was also used in previous literature.<sup>38, 39</sup> In further

research, we will take this into consideration and a more precise survival time prediction would need to be attempted. Last, a detailed medical history of and the treatment course for the GBM patients in this study were not

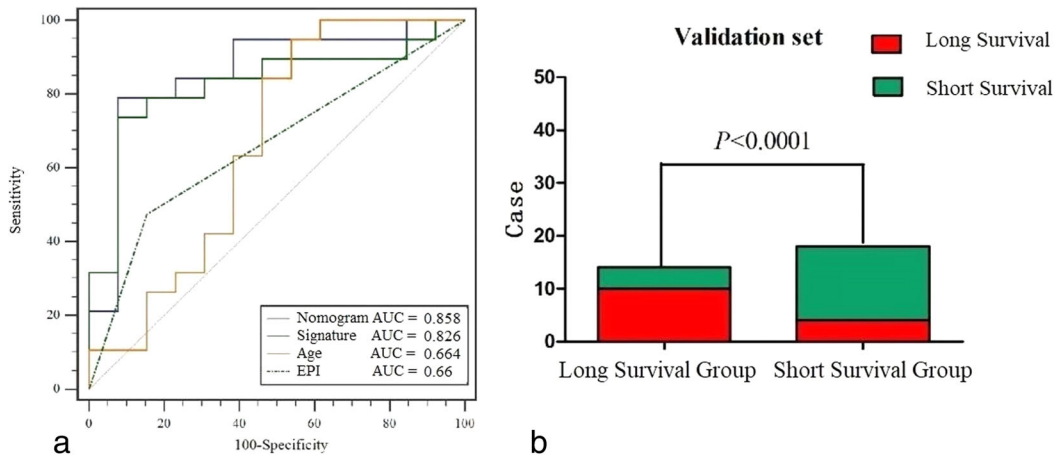


FIGURE 6: (a) ROC curves for nomogram, radiomic signature, age and meninges predicting OS in validation set. (b) The stratification performance of the nomogram in validation set.

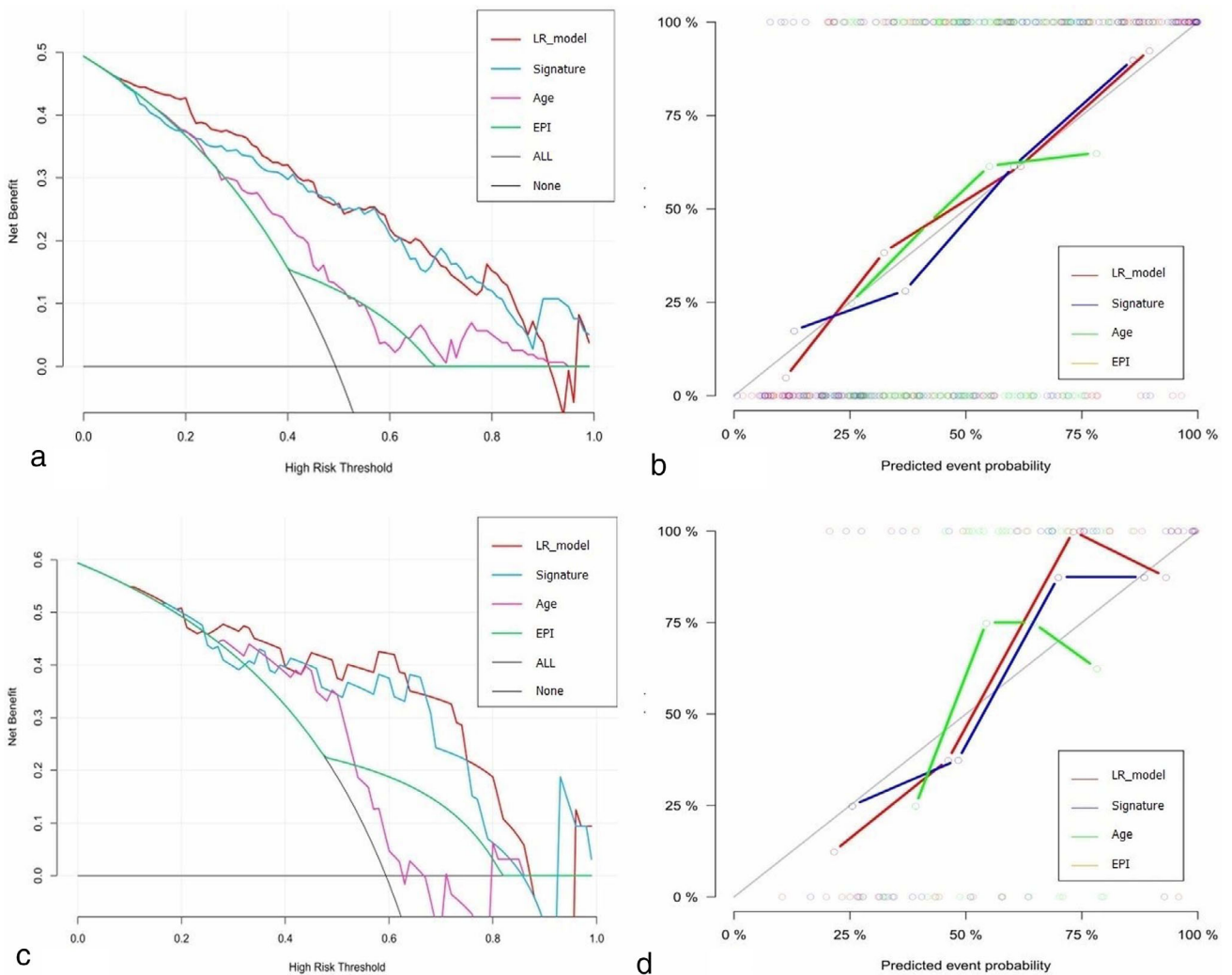


FIGURE 7: (a, b) DCA and calibration curves for LR model, radiomic signature, age, and EPIs predicting OS in BRATS2018 dataset, the graphs show that the LR model has the greatest net benefit in both datasets. (c, d) DCA and calibration curves for LR model, radiomic signature, age, and EPIs predicting OS in in validation set.

available, such as Karnofsky Performance Scale score which is well established as one of the most important predictors, which might also influence their survival prediction result and should be considered in future studies.<sup>40</sup> As this is a preliminary research study attempting to integrate radiomics with complementary visual features to predict survival, future studies may benefit from the combined use of additional clinical data, variable treatments, anatomic information, and functional images.

## Conclusion

This study demonstrates that in a radiomics nomogram integrated with a radiomic signature, complementary visual features such as EPI and age were found to be robust for the stratification of GBM patients into long- vs. short-term survival and could be useful in clinical practice.

## ACKNOWLEDGMENT

The authors are grateful to Ms. Pang Peipei for providing assistance with using the software.

## REFERENCES

- Hanif F, Muzaffar K, Perveen K, Malhi SM, Simjee Sh U. Glioblastoma multiforme: A review of its epidemiology and pathogenesis through clinical presentation and treatment. *Asian Pac J Cancer Prev* 2017;18(1):3-9.
- Ostrom QT, Gittleman H, Farah P, et al. CBTRUS statistical report: Primary brain and central nervous system tumors diagnosed in the United States in 2006–2010. *Neuro Oncol* 2013;15(Suppl 2):ii1-ii56.
- Messaoudi K, Clavreul A, Lagarce F. Toward an effective strategy in glioblastoma treatment. Part I: Resistance mechanisms and strategies to overcome resistance of glioblastoma to temozolomide. *Drug Discov Today* 2015;20(7):899-905.
- 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(10):4009-4014.
- Gutman DA, Cooper LA, Hwang SN, et al. MR imaging predictors of molecular profile and survival: Multi-institutional study of the TCGA glioblastoma data set. *Radiology* 2013;267(2):560-569.
- Mazurowski MA, Desjardins A, Malof JM. Imaging descriptors improve the predictive power of survival models for glioblastoma patients. *Neuro Oncol* 2013;15(10):1389-1394.
- Yip SS, Aerts HJ. Applications and limitations of radiomics. *Phys Med Biol* 2016;61(13):R150-R166.
- Limkin EJ, Sun R, Derclé L, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol* 2017;28(6):1191-1206.
- Zhang Z, Jiang H, Chen X, et al. Identifying the survival subtypes of glioblastoma by quantitative volumetric analysis of MRI. *J Neurooncol* 2014;119(1):207-214.
- Pérez-Beteta J, Molina-García D, Martínez-González A, et al. Morphological MRI-based features provide pretreatment survival prediction in glioblastoma. *Eur Radiol* 2019;29(4):1968-1977.
- Cui Y, Tha KK, Terasaka S, et al. Prognostic imaging biomarkers in glioblastoma: Development and independent validation on the basis of multiregion and quantitative analysis of MR images. *Radiology* 2016;278(2):546-553.
- Cui Y, Ren S, Tha KK, Wu J, Shirato H, Li R. Volume of high-risk intratumoral subregions at multi-parametric MR imaging predicts overall survival and complements molecular analysis of glioblastoma. *Eur Radiol* 2017;27(9):3583-3592.
- Chaichana KL, McGirt MJ, Frazier J, Attenello F, Guerrero-Cazares H, Quinones-Hinojosa A. Relationship of glioblastoma multiforme to the lateral ventricles predicts survival following tumor resection. *J Neurooncol* 2008;89(2):219-224.
- Noh JH, Lee MH, Kim WS, et al. Optimal treatment of leptomeningeal spread in glioblastoma: Analysis of risk factors and outcome. *Acta Neurochir* 2015;157(4):569-576.
- Molina-García D, Vera-Ramírez L, Pérez-Beteta J, Arana E, Pérez-García VM. Prognostic models based on imaging findings in glioblastoma: Human versus machine. *Sci Rep* 2019;9(1):5982.
- Kickingeder P, Bonekamp D, Nowosielski M, et al. Radiogenomics of Glioblastoma: Machine learning-based classification of molecular characteristics by using multiparametric and multiregional MR imaging features. *Radiology* 2016;281(3):907-918.
- Emblem KE, Pinho MC, Zöllner FG, et al. A generic support vector machine model for preoperative glioma survival associations. *Radiology* 2015;275(1):228-234.
- Macyszyn L, Akbari H, Pisapia JM, et al. Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques. *Neuro Oncol* 2016;18(3):417-425.
- Zacharakis EI, Morita N, Bhatt P, O'Rourke DM, Melhem ER, Davatzikos C. Survival analysis of patients with high-grade gliomas based on data mining of imaging variables. *AJNR Am J Neuroradiol* 2012;33(6):1065-1071.
- Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: More than meets the eye. *Lancet Oncol* 2015;16(4):e173-e180.
- Menze BH, Jakab A, Bauer S, et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans Med Imaging* 2015;34(10):1993-2024.
- Bakas S, Akbari H, Sotiras A, et al. Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci Data* 2017;4:170117.
- Bakas S, Reyes M, Jakab A, et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge 2018.
- Park M, Lee SK, Chang JH, et al. Elderly patients with newly diagnosed glioblastoma: Can preoperative imaging descriptors improve the predictive power of a survival model? *J Neurooncol* 2017;134(2):423-431.
- Kim H, Lim DH, Kim TG, et al. Leptomeningeal enhancement on preoperative brain MRI in patients with glioblastoma and its clinical impact. *Asia Pac J Clin Oncol* 2018;14(5):e366-e373.
- Carrillo JA, Lai A, Nghiempu PL, et al. Relationship between tumor enhancement, edema, IDH1 mutational status, MGMT promoter methylation, and survival in glioblastoma. *AJNR Am J Neuroradiol* 2012;33(7):1349-1355.
- Rohlfing T, Zahr NM, Sullivan EV, Pfefferbaum A. The SRI24 multi-channel atlas of normal adult human brain structure. *Hum Brain Mapp* 2010;31(5):798-819.
- Kalpathy-Cramer J, Gerstner ER, Emblem KE, Andronesi O, Rosen B. Advanced magnetic resonance imaging of the physical processes in human glioblastoma. *Cancer Res* 2014;74(17):4622-4637.
- Liu Y, Xu X, Yin L, Zhang X, Li L, Lu H. Relationship between glioblastoma heterogeneity and survival time: An MR imaging texture analysis. *AJNR Am J Neuroradiol* 2017;38(9):1695-1701.
- van Dijken BRJ, Jan van Laar P, Li C, et al. Ventricle contact is associated with lower survival and increased peritumoral perfusion in glioblastoma. *J Neurosurg* 2018;131(3):717-723.
- Mistry AM, Hale AT, Chambless LB, Weaver KD, Thompson RC, Ithrie RA. Influence of glioblastoma contact with the lateral ventricle on survival: A meta-analysis. *J Neurooncol* 2017;131(1):125-133.

32. Mandel JJ, Yust-Katz S, Cachia D, et al. Leptomeningeal dissemination in glioblastoma: An inspection of risk factors, treatment, and outcomes at a single institution. *J Neurooncol* 2014;120(3):597-605.
33. Wrensch M, Minn Y, Chew T, Bondy M, Berger MS. Epidemiology of primary brain tumors: Current concepts and review of the literature. *Neuro Oncol* 2002;4(4):278-299.
34. Chiang GC, Galla N, Ferraro R, Kovanlikaya I. The added prognostic value of metabolic tumor size on FDG-PET at first suspected recurrence of glioblastoma multiforme. *J Neuroimaging* 2017;27(2):243-247.
35. Liu Z, Wang S, Dong D, et al. The applications of radiomics in precision diagnosis and treatment of oncology: Opportunities and challenges. *Theranostics* 2019;9(5):1303-1322.
36. Kickingereder P, Burth S, Wick A, et al. Radiomic profiling of glioblastoma: Identifying an imaging predictor of patient survival with improved performance over established clinical and radiologic risk models. *Radiology* 2016;280(3):880-889.
37. Yang D, Rao G, Martinez J, Veeraraghavan A, Rao A. Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma. *Med Phys* 2015;42(11):6725-6735.
38. Nie D, Lu J, Zhang H, et al. Multi-channel 3D deep feature learning for survival time prediction of brain tumor patients using multi-modal neuroimages. *Sci Rep* 2019;9(1):1103.
39. Liao X, Cai B, Tian B, Luo Y, Song W, Li Y. Machine-learning based radiogenomics analysis of MRI features and metagenes in glioblastoma multiforme patients with different survival time. *J Cell Mol Med* 2019; 23(6):4375-4385.
40. Lacroix M, Abi-Said D, Fourney DR, et al. A multivariate analysis of 416 patients with glioblastoma multiforme: Prognosis, extent of resection, and survival. *J Neurosurg* 2001;95(2):190-198.