European Journal of Radiology 183 (2025) 111900

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Ensemble learning-based radiomics model for discriminating brain metastasis from glioblastoma

Qi Zeng [a,1], Fangxu Jia [a,1], Shengming Tang [a,1], Haoling He [b], Yan Fu [a,c], Xueying Wang [a], Jinfan Zhang [a], Zeming Tan [d], Haiyun Tang [a,*], Jing Wang [e,**], Xiaoping Yi [a,b,*], Bihong T. Chen [f]

[a] Department of Radiology, Xiangya Hospital, Central South University, Changsha 410008, Hunan, PR China
[b] Department of Radiology, First Affiliated Hospital of Guangxi Medical University, Nanning 530021, Guangxi, PR China
[c] National Clinical Research Center for Geriatric Disorders (Xiangya Hospital), Changsha 410008, Hunan, PR China
[d] Department of Neurosurgery, Xiangya Hospital, Central South University, Changsha 410008, Hunan, PR China
[e] Department of Neurology, Xiangya Hospital, Central South University, Changsha 410008, Hunan, PR China
[f] Department of Diagnostic Radiology, City of Hope National Medical Center, Duarte, CA, USA

ARTICLE INFO

Keywords:
Brain metastasis
Glioblastoma
Radiomics
Machine learning
Prediction model

ABSTRACT

Objective: Differentiating between brain metastasis (BM) and glioblastoma (GBM) preoperatively is challenging due to their similar imaging features on conventional brain MRI. This study aimed to enhance diagnostic accuracy through a machine learning model based on MRI radiomics data.

Methods: This retrospective study included 235 patients with confirmed solitary BM and 273 patients with GBM. Patients were randomly assigned to the training (n = 356) or the validation (n = 152) cohort. Conventional brain MRI sequences including T1-weighted imaging (T1WI), contrast-enhanced_T1WI, and T2-weighted imaging (T2WI) were acquired. Brain tumors were delineated on all three sequences and segmented. Features were selected from demographic, clinical, and radiomic data. An integrated ensemble machine learning model, i.e., the elastic regression-SVM-SVM model (ERSS) and a multivariable logistic regression (LR) model combining demographic, clinical, and radiomic data were built for predictive modeling. Model efficiency was evaluated using discrimination, calibration, and decision curve analyses. Additionally, external validation was performed using an independent cohort consisting of 47 patients with GBM and 43 patients with isolated BM to assess the ERSS model generalizability.

Results: The ERSS model demonstrated more optimal classification performance (AUC: 0.9548, 95% CI: 0.9337–0.9734 in training cohort; AUC: 0.9716, 95% CI: 0.9485–0.9895 in validation cohort) as compared to the LR model according to the receiver operating characteristic (ROC) curve and decision curve for the internal cohort. The external validation cohort had less optimal but still robust performance (AUC: 0.7174, 95% CI: 0.6172–0.8024). The ERSS model with integration of multiple classifiers, including elastic net, random forest and support vector machine, produced robust predictive performance and outperformed the LR method.

Conclusion: The results suggested that the integrated machine learning model, i.e., the ERSS model, had the potential for efficient and accurate preoperative differentiation of BM from GBM, which may improve clinical decision-making and outcomes of patients with brain tumors.

Abbreviations: ACC, accuracy; AUC, area under the curve; BM, brain metastasis; CE_T1WI, contrast-enhanced T1-weighted imaging; GBM, glioblastoma; DICOM, Digital Imaging and Communications in Medicine; DSC, dynamic susceptibility contrast; DTI, diffusion tensor imaging; DWI, diffusion-weighted imaging; ERSS, elastic regression – SVM – SVM; ICC, intra-class correlation coefficient; LASSO, least absolute shrinkage and selection operator; LR, logistic regression; MRI, magnetic resonance imaging; PACS, Picture Archiving and Communication System; ROC, receiver operating characteristic; ROI, region of interest; SEN, sensitivity; SPE, specificity; SVM, support vector machine; T1WI, T1-weighted imaging; T2WI, T2-weighted imaging.

* Corresponding author at: Department of Radiology, Xiangya Hospital, Central South University, Changsha 410008, Hunan, PR China.
** Corresponding author at: Department of Neurology, Xiangya Hospital, Central South University, Changsha 410008, Hunan, PR China.
E-mail addresses: 405016@csu.edu.cn (H. Tang), 15211132385@163.com (J. Wang), yixiaoping@csu.edu.cn (X. Yi).
[1] These authors have contributed equally to this work and share first authorship.

https://doi.org/10.1016/j.ejrad.2024.111900
Received 26 April 2024; Received in revised form 24 October 2024; Accepted 21 December 2024
Available online 24 December 2024

1. Introduction

Brain metastasis (BM) is the most common type of brain tumors in adults and represents one of the most prevalent neurological presentations in patients with cancer [1]. The incidence of BM is increasing, attributed to increased overall cancer survival, enhanced treatment, effective screening techniques, and an aging population [2,3]. The incidence and survival rates for BM are intricately linked to the primary tumor type, emphasizing the critical role of early identification of the primary tumor site in formulating tailored clinical strategies to optimize the prognosis [4]. On the other hand, glioblastoma (GBM; World Health Organization grade IV glioma) is the most common and aggressive primary brain tumor in adults [5–7]. BM and GBM have different etiologies and are treated differently. Currently, the standard treatment for BM is stereotactic radiotherapy after evaluation, with the sensitivity of the primary tumor influencing the choice of targeted agents. In contrast, the most effective treatment for GBM is surgery. Therefore, preoperative differential diagnosis of BM and GBM is important to guide therapeutic decisions [8,9]. Definitive pathology remains the gold standard for tumor diagnosis [10,11], but it is not always feasible to get tissue diagnosis, especially for brain tumors.

Brain magnetic resonance imaging (MRI) is the most commonly used imaging modality for both BM and GBM [12]. When multiple brain lesions are present, conventional brain MRI in conjunction with the patient's clinical history improves the diagnostic accuracy for both BM and GBM. However, for solitary BM or GBM, overlapping imaging features such as intra-tumoral necrosis and peritumoral edema [13,14], complicate diagnosis. Advanced quantitative brain MRI techniques, such as diffusion-weighted imaging (DWI), diffusion tensor imaging (DTI), dynamic susceptibility contrast (DSC), and other perfusion scans, have been useful in the differential diagnosis [9,14–16]. However, there are issues associated with these advanced neuroimaging techniques such as increased scanning time, additional cost, guarded diagnostic efficacy and limited resource in community hospitals, which has hindered their clinical application.

Machine learning algorithm-based strategies using radiomics have great potential for improving a non-invasive diagnosis of brain tumors [17–19]. Radiomics can extract a large number of high-level quantitative features from medically acquired images to obtain high-dimensional information about tumor heterogeneity beyond the limitations of the human eye [20]. Radiomics has shown significant potential for diagnosis, prognosis, and treatment response with significantly improved performance, and may play a key role in assessing the tumor before, during, and after treatment [21]. Machine learning models can combine a large number of variables from different data types into a single model to maximize the efficacy of predictive modeling. Machine learning techniques have been used for the identification, classification, and diagnosis of various types of brain tumors including primary brain tumors and BM [19,22–24]. Prior studies have mined conventional MRI sequences for radiomic texture analysis to differentiate between BM and GBM [9,25,26]. However, several issues in the prior studies such as the availability of selected features, machine learning algorithms, and the number of samples for predictive modeling render the models less optimal, which needs to be addressed[27,28]. In addition, there is a clinical necessity to accurately diagnosis BM and GBM given their differences in treatment and prognosis. Combining multiple conventional MRI sequences with clinical features may add clinical relevance to machine learning and improve model performance.

In this study, we combined radiomic features obtained from conventional brain MRI sequences including the T1-weighted imaging (T1WI), the contrast-enhanced T1-weighted imaging (CE_T1WI), and the T2-weighted imaging (T2WI), and used these features along with clinical laboratory data to classify BM and GBM. We used an integrated ensemble machine learning model, i.e., the elastic regression-SVM-SVM model (ERSS) and a multivariable logistic regression model (LR model) for predictive modeling. We hypothesized that the radiomic predictive models could be used to differentiate between BM and GBM and the ERSS model would outperform the LR model.

2. Methods

2.1. Patient information

All patients with brain tumors who underwent preoperative brain MRI scans in our hospital from September 2012 to September 2021 were evaluated for eligibility for this retrospective study. Inclusion criteria were as follows: (I) patients with a confirmed diagnosis of solitary BM or GBM based on surgical pathology specimens; (II) patients who had not undergone any cancer treatments such as chemotherapy, radiotherapy, or brain surgery before the brain MRI scans; (III) patients with preoperative brain MRI scans acquired within 14 days of brain surgery; (IV) patients without multifocal brain lesions (e.g., multicentric/multifocal GBM, multiple brain metastases). The study cohort and selection criteria are presented in Fig. 1. The blood samples for the laboratory data were collected within a 15-day interval of the brain MRI scans. We included only cases with solitary BM because the presence of multifocal brain lesions increased the possibility of BM, which may skew the models. On the other hand, solitary BM was more likely to share similar imaging features with GBM, presenting as a diagnostic dilemma in clinical practice. The solitary BM tends to be more challenging for diagnosis than multiple BM, which can be studied using radiomics and machine learning.

We performed external validation using an independent cohort of patients from another institution (Second Xiangya Hospital of Central South University, P.R. China), and this cohort consisted of 47 patients with GBM and 43 patients with isolated BM. Patients in the independent external cohort underwent preoperative MRI scans of the brain and met the study inclusion criteria. The external cohort allowed us to test the model on an unseen dataset, providing an assessment of the model performance outside of the training environment.

This study was performed with the approval from the institutional review board and the Medical Ethics Committee in our hospital (IRB No: 201709995). This study was conducted according to the ethical requirements of the Helsinki Declaration of 1964. The written informed consent was waived due to the retrospective nature of this study.

2.2. Reassessment of neuropathological findings

For each patient, all neuropathological slides for both BM and GBM were independently reassessed by two neuropathologists (GG and HY) who had 11 and 30 years of experience in neuropathology, respectively. Both neuropathologists were blinded to patient information, including radiological and clinical-pathological data. In cases of discrepancy, consensus was reached through discussion between the two study neuropathologists.

2.3. Brain MRI data analysis

All patients underwent brain MRI scan on one of the two scanners, a 3 T MRI scanner (Discovery MR750w, GE Healthcare) or a 1.5 T MRI scanner (MAGNETOM Avanto, Siemens Healthineers, Erlangen, Germany). In this study, we focused on conventional MRI sequences including T1WI, CE_T1WI, and T2WI. The MRI sequences were acquired in 2D. T1WI and CE_T1WI were acquired with a repetition time (TR) of ~ 1600 ms and an echo time (TE) of ~ 15 ms, with an in-plane resolution of 320 x 192 pixels and a slice thickness of 5.0 mm. T2WI used a TR of ~ 8000 ms, a TE of ~ 105 ms, a resolution of 320 x 224 pixels, and a slice thickness of 5.0 mm.

Brain MRI images were retrieved from our institution's Picture Archiving and Communication System (PACS, Carestream, Canada), and all images were stored in the Digital Imaging and Communications in Medicine (DICOM) file format. For the independent external cohort, the

brain MRI images were obtained from the PACS (Carestream, Canada) of the other institution (Second Xiangya Hospital of Central South University, P. R. China). Each patient's MRI images were independently reviewed by two neuroradiologists (Reader 1 [FZ] and Reader 2 [CC], with 15 and 25 years of experience in neuroimaging, respectively). They were blinded to patient information, including radiological and clinical-pathological data. The discrepancies in diagnosis were discussed in a group setting with an additional senior neuroradiologist (ZH, with over 35 years of experience in neuroimaging). Reader 1 and Reader 2 recorded conventional brain MRI imaging features, including tumor size measurements such as transverse dimension (mm), anteroposterior dimension (mm), and craniocaudal dimension (mm).

For each brain tumor, the two neuroradiologists (Reader 1 and 2) used ITK-SNAP 3.8.0 software (https://www.itksnap.org/) to delineate the tumor regions of interest (ROIs) layer by layer on CE_T1WI axial, sagittal, and coronal slices, as well as T2WI and T1WI sequences. The entire tumor was segmented using the ROI approach which included both enhancing and non-enhancing components, such as hemorrhage, necrosis, and cystic degeneration within the tumor. The peritumoral edema was not segmented. The study used the z-score normalization method with the AK software (GE Healthcare, AnalysisKit, Version 3.2.0) for normalization of images and radiomic feature extraction[29]. Multiple quantitative texture features were captured from grayscale histogram, gradient, run-length matrix, co-occurrence matrix, autoregressive model, and wavelet transform analysis. To ensure the repeatability of radiomic features, we calculated the inter-class correlation coefficient (ICC) between the two neuroradiologists and retained features with an ICC greater than 0.75.

2.4. Predictive modeling and model performance

Two types of predictive models were developed for this study. One model was established using the ERSS ensemble machine learning approach similar to the models developed by our team in our previous studies [30–34], while the other model used the traditional multivariable LR method.

Briefly, for the ERSS algorithm, we performed model construction and evaluation with the pandas package (version: 2.1.1), scikit-learn package (version: 1.3.1), and numpy package (version: 1.26.0) in Python 3.11.1. Multiple features from radiomic analysis, demographic information and laboratory data were included in the machine learning analysis. The model was developed in the training cohort and was tested in the validation cohort for the internal data set. The efficiency of each model was assessed for both the training and validation cohorts by calculating receiver operating characteristic (ROC) curves. A calibration curve (Hosmer-Lemeshow H test) was used to evaluate the model performance. The clinical utility of the model in the validation cohort was assessed using decision curve analysis. Fig. 2 presents the workflow for tumor segmentation, radiomic feature extraction, feature selection and model performance.

During data analysis, each patient was randomly assigned to either the training cohort or the validation cohort in a 7:3 ratio. During the model development process, we employed the Elastic Net (EN) method for feature selection, which was conducted solely on the training cohort to prevent any data leakage and to ensure an unbiased assessment of model performance on the validation cohort. Each patient's features were combined linearly and weighted by their respective coefficients to produce an EN-score. Based on these selected features, we developed a classification model using the Random Forest (RF) method, which generated an RF-score. Following this, we constructed a new combined classification model using a Support Vector Machine (SVM) approach. This model integrated the EN-score and RF-score, alongside two additional SVM scores derived from demographic and laboratory test features. The final ERSS model was then built using the SVM algorithm,

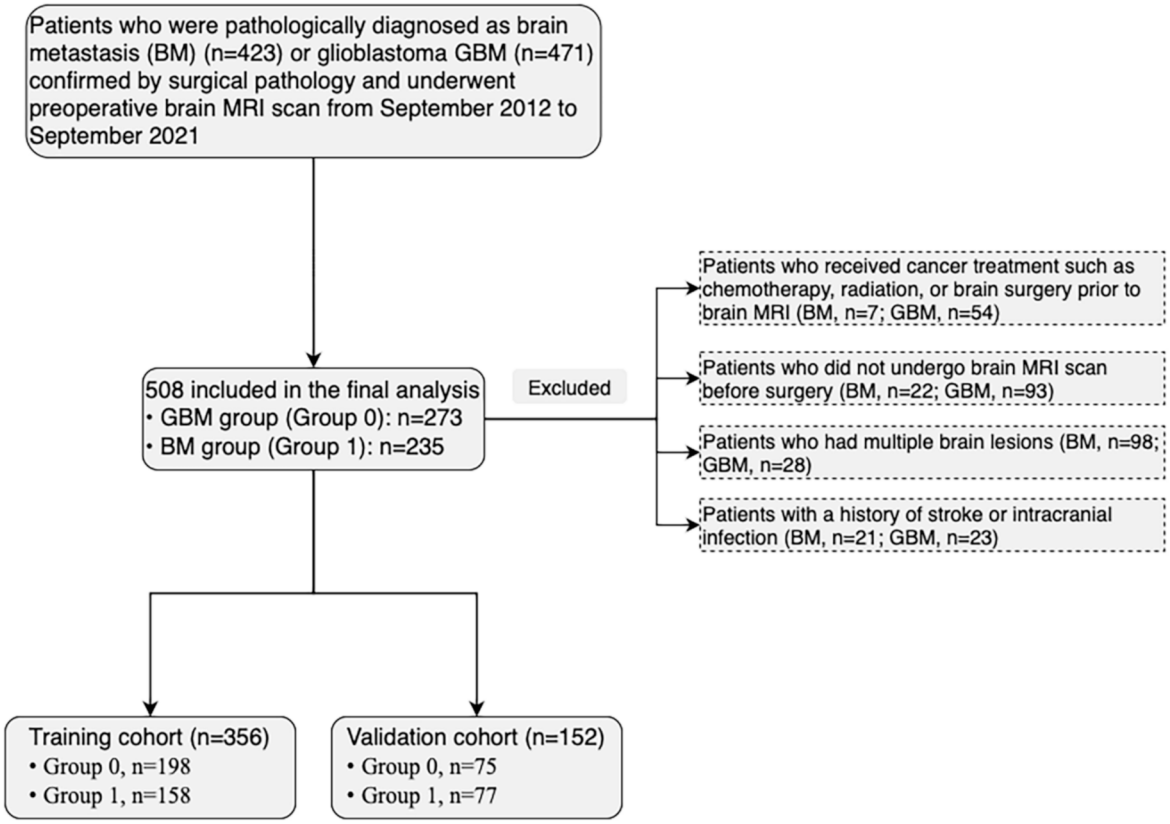

Fig. 1. Flow chart presenting the recruitment process for the study cohort. (Abbreviations: MRI, magnetic resonance imaging; GBM, glioblastoma; BM, brain metastasis).

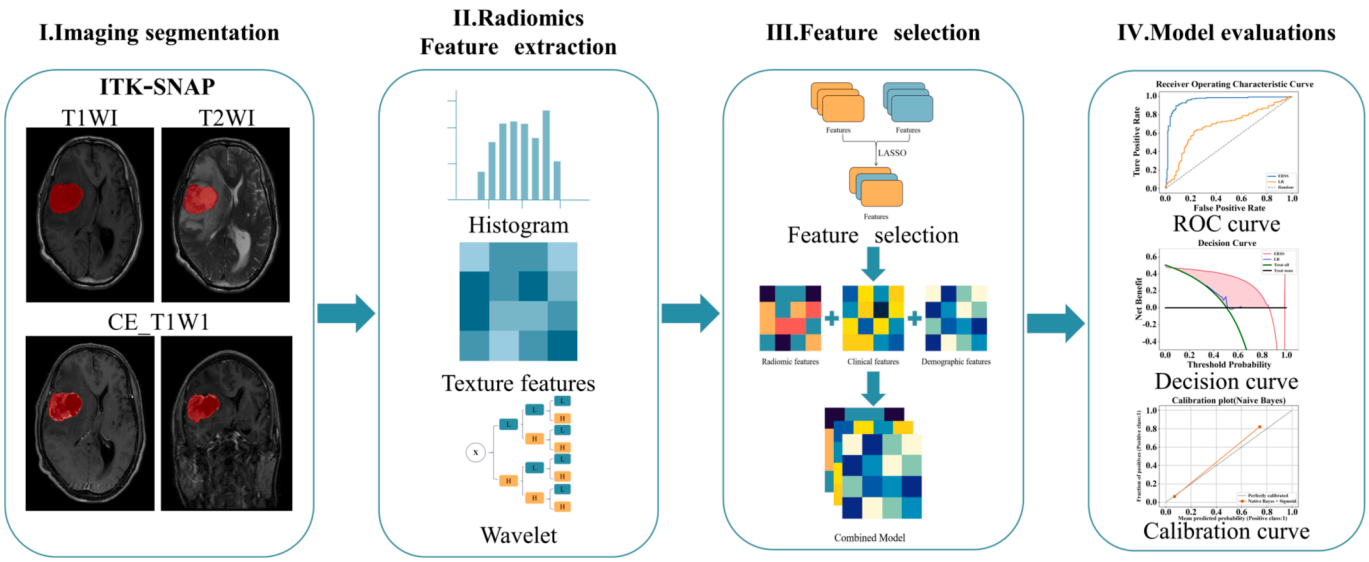

Fig. 2. Workflow for radiomic data extraction as well as feature reduction and predictive modeling. (Abbreviations: ROC, receiver operating characteristic).

producing three types of SVM scores from the steps. These scores, representing different data types, served as weak classifiers. Since each SVM score was derived from independently regressed variables, regressing them together on a two-dimensional plane using SVM yielded better results compared to regressing each score individually. Ultimately, the regressed scores were binarized for further prediction.

Table 1
Demographic information and laboratory characteristics of the study cohort.

Characteristic	Total (n = 508)	BM (n = 235)	GBM (n = 273)	p-value	Training Cohort (n = 356)	Validation Cohort (n = 152)	p-value
Demographics							
Gender, n (%)				0.131			0.697
Male	314(61.8 %)	137(58.3 %)	177(64.8 %)		222(62.4 %)	92(60.5 %)	
Female	194(38.2 %)	98(41.7 %)	96(35.2 %)		134(37.6 %)	60(39.5 %)	
Age (y)	52(44–61)	55(48–62)	49(38–58)	*<0.001	52(44–61)	53(43–61)	0.841
KPS	80(70–100)	70(60–80)	100(80–100)	*<0.001	80(70–100)	80(60–100)	0.177
Routine blood parameters							
RBC (10^{12}/L)	$4.41(3.82–5.00)_{mean}$	$4.32(3.72–4.93)_{mean}$	$4.48(3.91–5.04)_{mean}$	*0.003	$4.42(3.82–5.02)_{mean}$	$4.38(3.82–4.94)_{mean}$	0.546
HGB (g/L)	133(121–145)	131(121–142)	134(122–146)	*0.027	134(121–144)	132(123–145)	0.729
WBC (10^9/L)	7.5(5.7–9.7)	7.0(5.4–9.0)	7.8(6.0–10.6)	*<0.001	7.6(5.7–9.5)	7.5(5.6–10.2)	0.890
Neu(10^9/L)	5.10(3.33–7.10)	4.60(3.10–6.80)	5.30(3.70–8.10)	*0.001	5.10(3.40–6.98)	4.90(3.20–8.10)	0.818
Lym(10^9/L)	1.5(1.1–2.0)	1.5(1.0–2.0)	1.6(1.1–2.0)	0.164	1.5(1.0–2.0)	1.5(1.1–1.9)	0.935
Mono(10^9/L)	0.5(0.3–0.7)	0.5(0.3–0.7)	0.5(0.4–0.7)	0.129	0.5(0.4–0.7)	0.5(0.3–0.7)	0.411
Eos(10^9/L)	0.1(0.0–0.2)	0.1(0.0–0.2)	0.1(0.0–0.1)	*0.023	0.1(0.0–0.2)	0.1(0.0–0.2)	0.346
PLT (10^9/L)	203.0(162.2–252.0)	196.0(159.0–252.0)	207.0(163.5–253.0)	0.263	200.5(159.0–248.5)	207.5(168.3–255.80	0.214
RDW (%)	13.1(12.5–13.9)	13.3(12.6–14.2)	13.0(12.4–13.7)	*0.002	13.2(12.5–13.9)	13.1(12.4–13.9)	0.306
PCT (ng/mL)	0.20(0.17–0.26)	0.20(0.14–0.21)	0.20(0.20–0.29)	*<0.001	0.20(0.16–0.25)	0.20(0.19–0.27)	0.144
MPV (fL)	9.71(8.60–11.00)	9.35(8.40–10.70)	9.9(8.7–11.1)	*0.004	9.7(8.6–11.0)	9.8(8.6–10.9)	0.836
Blood biochemical indexes							
ALT (U/L)	18.55(13.00–18.55)	19.10(13.30–28.00)	17.80(12.70–29.05)	0.48	18.20(13.0–28.1)	18.90(12.90–28.25)	0.696
AST (U/L)	19.40(16.23–24.20)	20.70(16.90–24.90)	18.70(15.30–23.40)	*0.003	19.40(16.40–24.20)	19.40(16.10–24.28)	0.979
Albumin (g/L)	41.0(38.3–44.0)	40.3(38.0–42.7)	41.9(38.9–44.6)	*<0.001	41.0(38.3–43.8)	41.0(38.1–44.1)	0.965
Globulin (g/L)	26.5(23.9–29.0)	27.1(24.4–29.8)	25.7(23.5–28.5)	*0.001	26.5(23.9–28.9)	26.5(24.0–29.2)	0.780
BUN (mmol/L)	5.090(4.033–6.235)	5.050(3.750–6.370)	5.100(4.170–6.075)	0.499	5.075(4.040–6.300)	5.095(3.985–6.085)	0.936
Scr (μmol/L)	79.0(69.9–89.0)	77.6(69.0–87.0)	81.0(70.5–91.0)	*0.023	79.0(70.0–90.0)	80.0(69.0–88.0)	0.904
Blood coagulation parameters							
PT (s)	12.70(12.20–13.30)	12.60(12.20–13.30)	12.80(12.20–13.40)	0.078	12.70(12.20–13.30)	12.80(12.20–13.30)	0.749
APTT (s)	31.60(28.60–34.28)	31.90(28.70–34.50)	31.10(28.50–34.00)	0.155	31.50(28.60–34.10)	31.80(28.83–34.48)	0.436
INR	0.99(0.95–1.04)	0.98(0.94–1.04)	0.99(0.95–1.04)	0.086	0.99(0.95–1.04)	0.99(0.95–1.04)	0.869

Note: Data are presented as n (%) or # mean (mean-SD to mean + SD)$_{mean}$, or presented as median (IQR). * indicates a p-value < 0.05.
Abbreviations: BM, brain metastasis; GBM, glioblastoma; KPS, Karnofsky performance status; RBC, red blood cell; HGB, hemoglobin; WBC, white blood cells; Neu, Neutrophils; Lym, lymphocytes; Mono, monocytes; Eos, Eosinophils; PLT, platelets; RDW, red cell distribution width; PCT, Platelet specific volume; MPV, mean platelet volume; ALT, alanine transaminase; AST, aspartate transaminase; BUN, blood urea nitrogen; Scr, serum creatinine; PT, prothrombin time; APTT, activated partial thromboplastin time; INR, international normalized ratio.

Models were subsequently built using the training cohort and tested on the validation cohort. The ERSS model was externally validated on the independent external cohort from the Second Xiangya Hospital of Central South University, P. R. China. The discriminative performance of the model on the independent dataset was assessed by calculating the ROC curve.

2.5. Statistical analysis

The data analysis was performed using SPSS Statistics for Windows, version 22.0 (IBM Corp., Armonk, NY, USA). The Shapiro-Wilk test was used to assess the distribution of data. Continuous, normally distributed data were presented as mean (standard deviation [SD]) and were assessed using Student's *t*-test, while continuous, non-normally distributed data were expressed as median (inter-quartile range [IQR]) and were analyzed using the Mann-Whitney *U* test. Categorical variables were compared using Pearson's chi-square test or Fisher's exact test. A p-value < 0.05 was considered statistically significant. To ensure high quality radiomic research, a dedicated checklist using the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) criteria for radiomics research was added to the Supplementary Material [35–37].

3. Results

3.1. Patient information

This study retrospectively enrolled 235 patients with solitary BM and 273 patients with GBM. Detailed information on the patient recruitment process is presented in Fig. 1, and a summary of the demographic and clinical characteristics of the study subjects is presented in Table 1. For radiological analysis, each patient was randomly assigned to either the training cohort (n = 356) or the validation cohort (n = 152) in a 7:3 ratio (Fig. 1).

There were no significant differences between the training cohort and the validation cohort regarding the demographic data and the laboratory data (all $p > 0.05$). However, there were significant differences between the patients with BM and the patients with GBM in age and Karnofsky performance status (KPS) ($p < 0.001$), and in blood parameters such as red blood cells (RBC) ($p = 0.003$) and hemoglobin (HGB) ($p = 0.027$).

3.2. Model construction, performance and application

We used the elastic net (EN) method to select a set of features for building a predictive model leveraging demographic information, clinical values, and radiomic data. The selected features from the EN method included one demographic feature (gender), 24 radiomic features from analyzing the conventional radiological images such as wavelet-HLH_glrlm_ShortRunLowGrayLevelEmphasis, and 16 clinical laboratory features such as alanine transaminase (ALT). For model building, we used demographic data and clinical laboratory data together as clinical data to construct the machine learning model. Scores (weak classifiers) obtained by fitting the clinical and imaging data were utilized to construct the ERSS model (strong classifier) using the support vector machine (SVM) method. Table 2 shows the integrated machine learning model combining clinical and imaging data having superior overall performance compared to the other three models that used a single data source (clinical data, T1 imaging data, or T2 imaging data).

The ROC curves of the models in the training and validation cohorts (Fig. 3 A&B) showed the classification performance and discriminative ability of the models. The AUC for the ERSS model was 0.9548 (95 % CI: 0.9337–0.9734) for the training cohort, and was 0.9716 (95 % CI: 0.9485–0.9895) for the validation cohort.

The ROC curves (Fig. 3 A&B) and the decision curves (Fig. 4) showed the accuracy of the ERSS model being higher than that of the LR model for both the training and validation cohorts, enabling improved differentiation between the BM and GBM groups. The calibration curves demonstrated good consistency between predictions and observations (Fig. 3 C&D). The decision curves for the predictive models (Fig. 4) indicated that the ERSS model had greater clinical benefit than the LR model.

The ROC curves of the ERSS model in the independent external cohort showed reasonable classification performance and discriminative ability of the model (Fig. 5). For the independent external cohort, the ERSS model had an AUC of 0.7174 (95 % CI: 0.6172–0.8024), an accuracy of 0.6932 (95 % CI: 0.6136–0.7727), sensitivity of 0.6957 (95 % CI: 0.5789–0.8085), and specificity of 0.6905 (95 % CI: 0.5714–0.8049).

4. Discussion

In this study, we established an integrated ERSS machine learning model combining the radiomic features, clinical laboratory data and demographic information to distinguish patients with BM from patients with GBM. This study showed that the ERSS model may potentially assist in the preoperative diagnosis of BM and GBM.

The study results indicated that incorporating radiomic features from routine MRI sequences such as T1WI, CE_T1WI and T2WI into the model construction enhanced the robustness of the predictive model in the validation cohort with an AUC reaching 0.9548. For instance, a prior

Table 2
Performance data for the machine learning model using different data source and external validation.

	Training cohort		Validation cohort	
Metrics	Clinical	T2	Clinical	T2
AUC	0.93, 95 %CI (0.90–0.95)	0.85, 95 %CI (0.81–0.88)	0.92, 95 %CI (0.88–0.96)	0.82, 95 %CI (0.76–0.88)
Accuracy	0.88, 95 %CI (0.85–0.90)	0.77, 95 %CI (0.73–0.81)	0.89, 95 %CI (0.85–0.93)	0.77, 95 %CI (0.71–0.82)
Sensitivity	0.84, 95 %CI (0.79–0.88)	0.86, 95 %CI (0.82–0.90)	0.83, 95 %CI (0.75–0.89)	0.85, 95 %CI (0.79–0.92)
Specificity	0.93, 95 %CI (0.90–0.96)	0.66, 95 %CI (0.60–0.72)	0.95, 95 %CI (0.90–0.99)	0.69, 95 %CI (0.60–0.78)
	Training cohort		Validation cohort	
Metrics	T1	Combine	T1	Combine
AUC	0.80, 95 %CI (0.75–0.83)	0.95, 95 %CI (0.93–0.97)	0.76, 95 %CI (0.69–0.83)	0.97, 95 %CI (0.95–0.99)
Accuracy	0.76, 95 %CI (0.72–0.80)	0.88, 95 %CI (0.85–0.90)	0.69, 95 %CI (0.63–0.76)	0.89, 95 %CI (0.85–0.93)
Sensitivity	0.87, 95 %CI (0.83–0.91)	0.84, 95 %CI (0.79–0.88)	0.83, 95 %CI (0.75–0.90)	0.83, 95 %CI (0.75–0.89)
Specificity	0.62, 95 %CI (0.56–0.68)	0.93, 95 %CI (0.90–0.96)	0.56, 95 %CI (0.47–0.65)	0.95, 95 %CI (0.90–0.99)
External Validation				
Metric	Value			
AUC	0.72, 95 %CI (0.62–0.80)			
Accuracy	0.69, 95 %CI (0.61–0.77)			
Sensitivity	0.70, 95 %CI (0.58–0.81)			
Specificity	0.69, 95 %CI (0.57–0.80)			

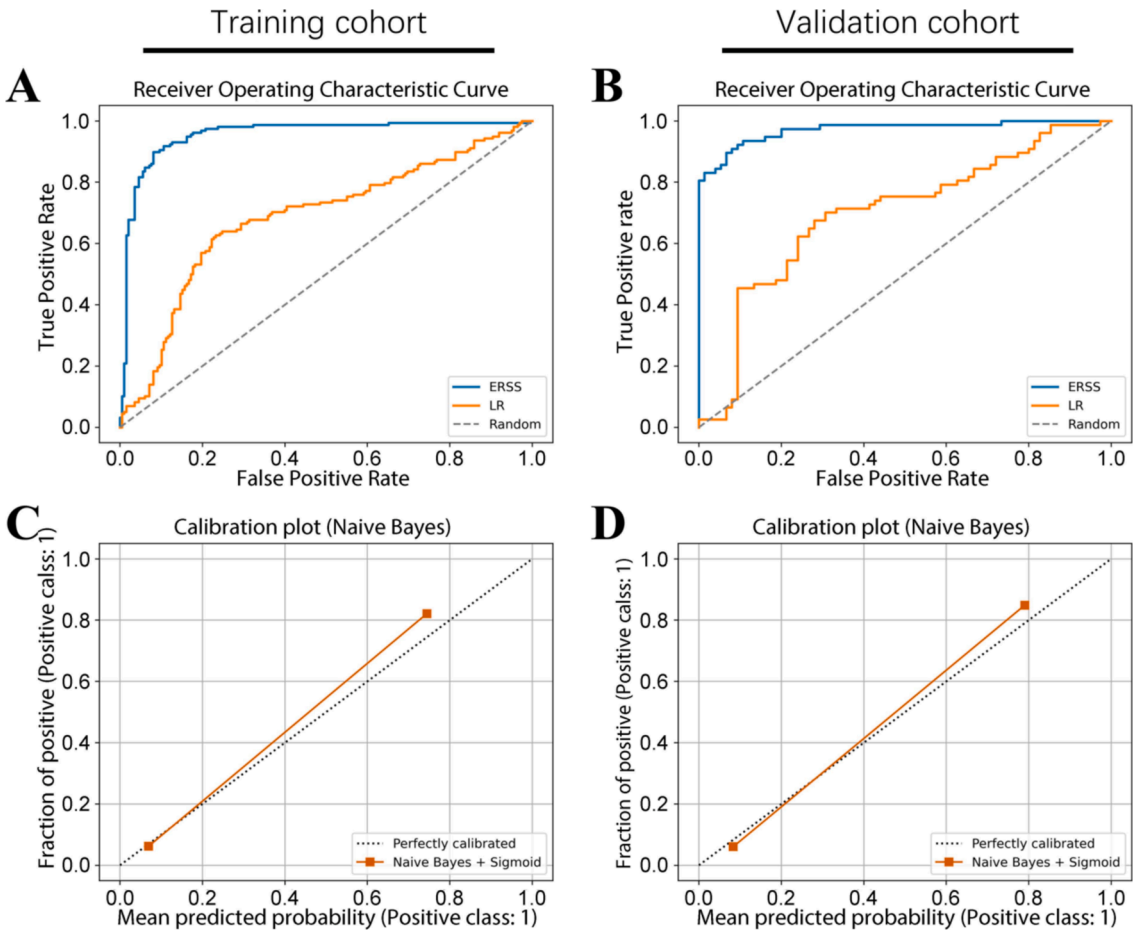

Fig. 3. Receiver operating characteristic (ROC) and calibration curves for the ERSS (Elastic Regression-SVM-SVM) model and the multivariable logistic regression (LR) model in the training cohort (A and C, respectively) and the validation cohort (B and D, respectively).

study constructed classifiers based on the contrast-enhanced sequence using an LR algorithm, resulting in a lower AUC of 0.80 in the validation cohort [38]. Various MRI sequences contribute to the multimodal imaging of brain tumors. Combining information from multiple MRI sequences effectively supplements information across different dimensions such as structural information and for different types of lesions, thereby enhancing the classification performance of the radiomic model [9,39]. Taking T2WI images as an example, T2WI provides relevant features related to heterogeneous vascularization and may reflect information about edema [40]. Research indicates that observing the peritumoral region improves the accuracy of discriminating brain tumors, encompassing vasogenic edema and cellular infiltration [41]. BM shows almost no histological infiltration and is primarily surrounded by edema beyond the enhanced edge on brain MRI. In GBM, the enhanced edge may also have corresponding abnormal FLAIR signal due to infiltrating tumor cells with vasogenic edema [42–44]. In addition, features extracted from T2WI images provide detailed information about peritumoral edema, which is critical for diagnosis of brain tumor. Nevertheless, extracting features from multiple MRI sequences is a time-consuming process and demands greater computation in the feature selection process. In clinical practice, clinicians often rely on the information directly from multiple MRI sequences to assist in clinical decision making. Therefore, it is reasonable to assess the conventional MRI sequences in our study for their usefulness in model development.

Our study integrated the radiomic features with clinical variables, which should have enabled the model to more accurately encompass the overall condition of patients, thus with more optimal classification performance. Artzi and colleagues extracted radiomic features from CE_T1WI sequences and incorporated them with clinical and imaging features. They constructed an SVM classifier with an AUC of 0.96, supporting the notion of combination of features improving model performance [45]. The ERSS model in this study incorporated multiple clinical features, including white blood cell count, platelet count, international normalized ratio (INR), activated partial thromboplastin time (APTT), and other indicators commonly used to assess inflammation processes and coagulation function, contributing to early disease diagnosis [46]. The methods for obtaining these parameters were straightforward, cost-effective, and commonly acquired in preoperative clinical settings. Studies indicate that a diagnostic model constructed using eight preoperative blood test indicators, such as Platelet specific volume (PCT), INR, and Thrombin Time (TT), could differentiate between BM and GBM with accuracies of 88.2 % and 76.1 %, respectively [47].

Our study showed the model combining clinical and imaging data having more robust performance compared to the models using a single data source being either imaging data or clinical data. However, the model performance did not significantly improve when comparing the combined model to the clinical model, which we believe may reflect the inherent complexities involved in integrating radiomic features with clinical data. Radiomic features were generated with extensive computational analysis, encompassing texture features, wavelet and histograms of the tumors, which were not directly related to the easily understood clinical features. The predictive capabilities of radiomic features may be limited given the heterogeneity of these features. Nevertheless, radiomics and clinical variables should be complimentary and their combined use in prediction model should increase the model performance. More research is needed to enhance the generalizability and robustness of the prediction models across different patient populations and imaging protocols.

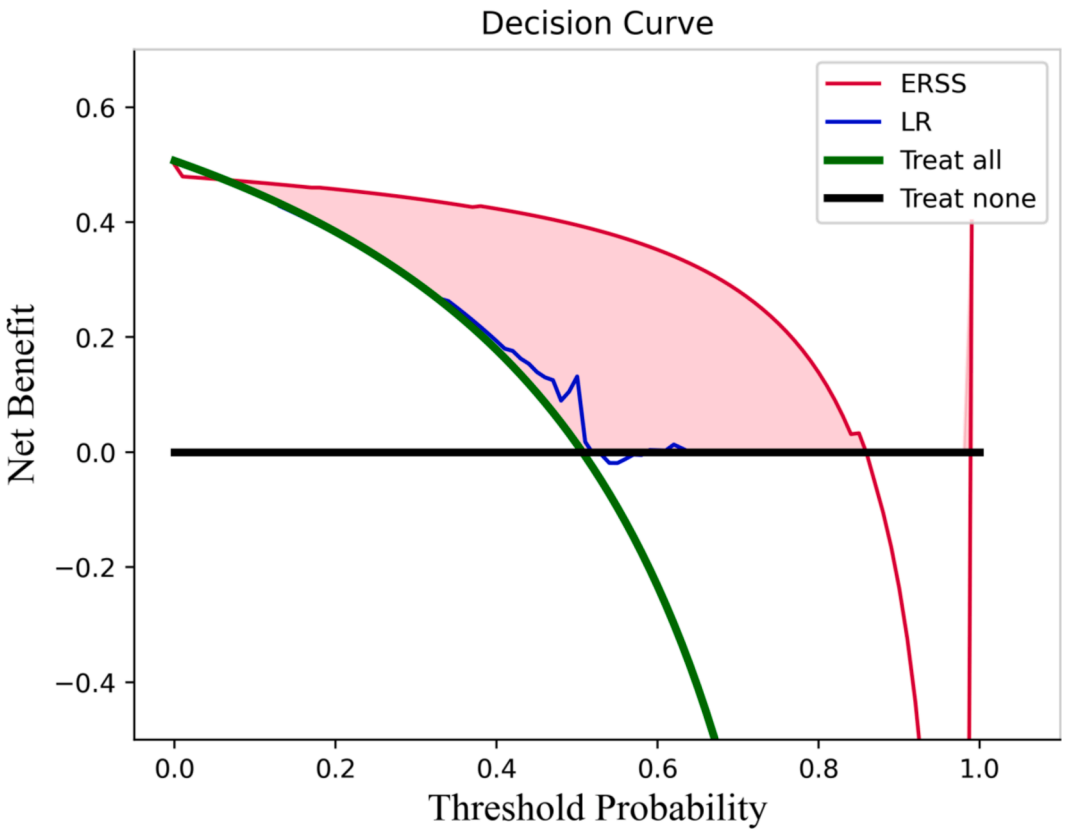

Fig. 4. Decision curve analysis for the ERSS (Elastic Regression-SVM-SVM) model and the multivariable logistic regression (LR) model. The y-axis measures the net benefit. The horizontal grey line represents the assumption that no patients would receive treatment and the oblique green line represents the assumption that all patients would receive treatment.

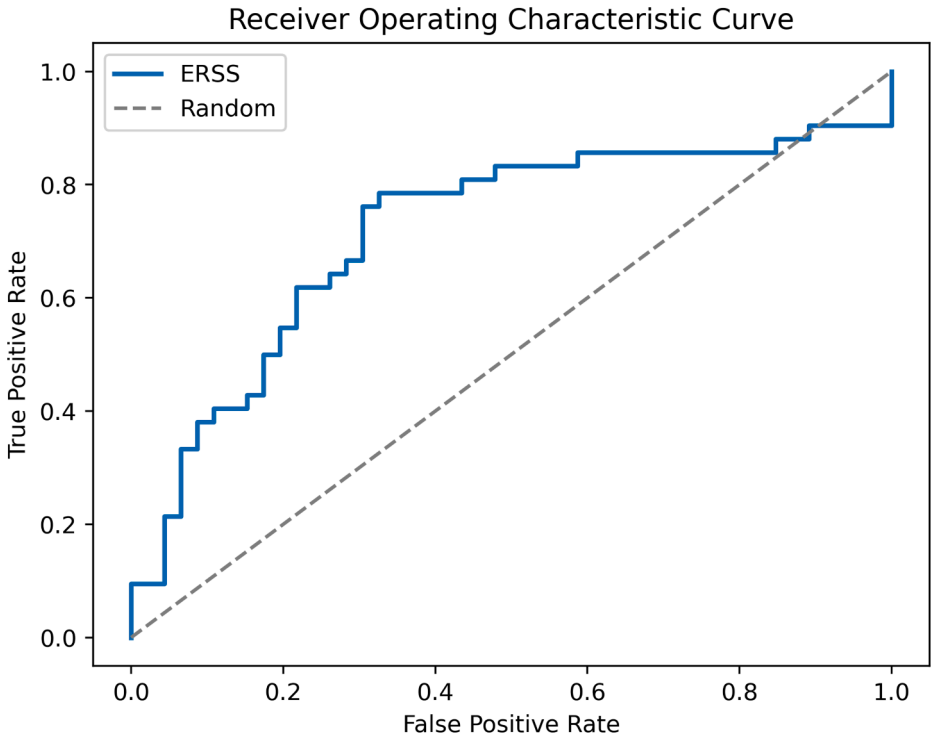

Fig. 5. Receiver operating characteristic (ROC) curves of the ERSS (Elastic Regression-SVM-SVM) model for the independent external validation cohort.

Our study indicated that the predictive performance of the ensemble machine learning model, i.e., the ERSS model, surpassed that of the LR model in the preoperative differentiation of BM and GBM. The ERSS model constructed from multiple weakly supervised models by integrating classifiers, and forming a strong supervised model using SVM classifiers. Throughout the model construction process, each classifier had distinct advantages and limitations [48–50]. Early studies have suggested that SVM exhibited optimal diagnostic efficiency for BM and GBM when constructing a single classifier model [51–53]. For instance, Qian and colleagues achieved a predictive performance using SVM algorithms for feature extraction from several imaging sequences from brain MRI scan, applying the SVM algorithm with least absolute shrinkage and selection operator (LASSO) feature selection and reaching an AUC of 0.90 [54]. The ERSS model developed for this study with an AUC of 0.95 had a similar high performance as the previously published study [54].

The strength of the ensemble machine learning model such as the ERSS model was in its ability to effectively integrate multiple weak classifiers, partially overcoming overfitting issues in machine learning [55]. Additionally, ensemble machine learning enhanced model robustness [56] and mitigated the impact of data noise. It addressed challenges with machine learning like class imbalance and dimensionality [57]. It is noteworthy that our study with an innovative classification algorithm and a reasonable sample size should lead to better generalizability as compared to previous research [58,59].

The external validation results showed reasonable performance of the ERSS model on unknown data, generally in line with literature [60]. However, the performance of our external validation cohort was lower as compared to the internal validation cohort, which may be related to factors such as heterogeneity of data sources [61], differences in imaging devices, and insufficient sample size of the validation cohort. Our study showed the necessity to validate prediction models in external cohorts and to overcome the limitations of radiomic prediction models. It is prudent to make these models more robust and more generalizable to various tumors in different institutions. Future studies should aim to further validate and enhance the clinical applicability of these models through multicenter collaborations and the use of larger, more diverse external datasets.

This study, however, had a few limitations. First, the clinic-demographic information and brain tumor data might be relatively homogeneous in this single center study. Also, the retrospective nature of this study may not fully represent the overall population, leading to case selection bias. To further improve the reliability and applicability of our findings, prospective multi-center studies will be necessary. The generalizability of our machine learning model was limited by the lack of cross-validation. This study was constrained by the inherent complexity of machine learning models and for their inherent lack of interpretability, which made it difficult to apply traditional computer methods to fully understand the model performance. In addition, the study results could not be easily interpreted nor explained relative to their clinical significance. Further research is needed to explore the interpretability of machine learning based on a large cohort data in the future.

5. Conclusion

In summary, we developed a robust ensemble machine learning model for preoperative differentiation of BM from GBM. Leveraging clinical brain MRI images alongside clinicodemographic data and radiomic features, this model showed the potential in preoperative diagnosis of BM and GBM, which should assist in clinical decision-making for personalized care of patients with a primary or a metastatic brain tumor.

Data Availability Statements

As the algorithms are proprietary to us, the datasets generated and/or analyzed in this study, along with the source code used, are not publicly available. However, they can be obtained from the corresponding author, Xiaoping Yi, upon reasonable request.

CRediT authorship contribution statement

Qi Zeng: Writing – review & editing, Writing – original draft. **Fangxu Jia:** Writing – review & editing, Writing – original draft. **Shengming Tang:** Writing – review & editing, Writing – original draft. **Haoling He:** Resources, Data curation. **Yan Fu:** Methodology, Conceptualization. **Xueying Wang:** Methodology, Investigation. **Jinfan Zhang:** Visualization, Data curation. **Zeming Tan:** Resources, Data curation. **Haiyun Tang:** Writing – review & editing. **Jing Wang:** Writing – review & editing. **Xiaoping Yi:** Writing – review & editing, Visualization, Methodology, Conceptualization. **Bihong T. Chen:** Writing – review & editing.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the staff of the Department of Radiology, Neurology, and Pathology at Xiangya Hospital for their efforts in collecting information for this study.

Scientific contributor Nancy Linford, PhD, provided editorial assistance for this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejrad.2024.111900.

References

[1] J. Gállego Pérez-Larraya, J. Hildebrand, Brain metastases. In: Handbook of Clinical Neurology. Vol 121. Elsevier; 2014:1143-1157. doi: 10.1016/B978-0-7020-4088-7.00077-8.
[2] S. Habbous, K. Forster, G. Darling, et al., Incidence and real-world burden of brain metastases from solid tumors and hematologic malignancies in Ontario: a population-based study. Neuro-Oncology Advances. 2021;3(1):vdaa178. doi: 10.1093/noajnl/vdaa178.
[3] J.K. Tabor, A. Onoichenco, V. Narayan, A.G. Wernicke, R.S. D'Amico, M. Vojnic, Brain metastasis screening in the molecular age. Neuro-Oncology Advances. 2023; 5(1):vdad080. doi: 10.1093/noajnl/vdad080.
[4] M. Parker, K. Jiang, J. Rincon-Torroella, et al., Epidemiological trends, prognostic factors, and survival outcomes of synchronous brain metastases from 2015 to 2019: a population-based study. Neuro-Oncology Advances. 2023;5(1):vdad015. doi: 10.1093/noajnl/vdad015.
[5] H.G. Wirsching, E. Galanis, M. Weller, Glioblastoma. In: Handbook of Clinical Neurology. Vol 134. Elsevier; 2016:381-397. doi: 10.1016/B978-0-12-802997-8.00023-2.
[6] D.N. Louis, A. Perry, G. Reifenberger, et al., The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary, Acta Neuropathol. 131 (6) (2016) 803–820, https://doi.org/10.1007/s00401-016-1545-1.
[7] S. Grochans, A.M. Cybulska, D. Simińska, et al., Epidemiology of glioblastoma multiforme–literature review, Cancers 14 (10) (2022) 2412, https://doi.org/10.3390/cancers14102412.

[8] W. Hall, H. Djalilian, E. Nussbaum, K. Cho, Long-term survival with metastatic cancer to the brain, Med. Oncol. 17 (4) (2000) 279–286, https://doi.org/10.1007/BF02782192.
[9] J. Wu, F. Liang, R. Wei, et al., A multiparametric MR-based radiofusionomics model with robust capabilities of differentiating glioblastoma multiforme from solitary brain metastasis, Cancers 13 (22) (2021) 5793, https://doi.org/10.3390/cancers13225793.
[10] N. Shastri-Hurst, M. Tsegaye, D.K. Robson, J.S. Lowe, D.C. Macarthur, Stereotactic brain biopsy: an audit of sampling reliabilityin a clinical case series, Br. J. Neurosurg. 20 (4) (2006) 222–226, https://doi.org/10.1080/02688690600875507.
[11] R. Di Bonaventura, N. Montano, M. Giordano, et al., Reassessing the role of brain tumor biopsy in the era of advanced surgical, molecular, and imaging techniques—a single-center experience with long-term follow-up, Journal of Personalized Medicine. 11 (9) (2021) 909, https://doi.org/10.3390/jpm11090909.
[12] P.O.P.E. Wb, Brain metastases: neuroimaging, Handb. Clin. Neurol. 149 (2018) 89–112, https://doi.org/10.1016/B978-0-12-811161-1.00007-4.
[13] R.R. Lonser, A.O. Vortmeyer, J.A. Butman, et al., Edema is a precursor to central nervous system peritumoral cyst formation, Ann. Neurol. 58 (3) (2005) 392–399, https://doi.org/10.1002/ana.20584.
[14] G. Yang, T.L. Jones, T.R. Barrick, F.A. Howe, Discrimination between glioblastoma multiforme and solitary metastasis using morphological features derived from the p:q tensor decomposition of diffusion tensor imaging, NMR Biomed. 27 (9) (2014) 1103–1111, https://doi.org/10.1002/nbm.3163.
[15] E. Cindil, H.N. Sendur, M.N. Cerit, et al., Validation of combined use of DWI and percentage signal recovery-optimized protocol of DSC-MRI in differentiation of high-grade glioma, metastasis, and lymphoma, Neuroradiology 63 (3) (2021) 331–342, https://doi.org/10.1007/s00234-020-02522-9.
[16] C.H. Suh, H.S. Kim, S.C. Jung, C.G. Choi, S.J. Kim, Perfusion MRI as a diagnostic biomarker for differentiating glioma from brain metastasis: a systematic review and meta-analysis, Eur. Radiol. 28 (9) (2018) 3819–3831, https://doi.org/10.1007/s00330-018-5335-0.
[17] J. Valdebenito, F. Medina, Machine learning approaches to study glioblastoma: A review of the last decade of applications, Cancer Reports. 2 (6) (2019) e1226.
[18] M. Kocher, M.I. Ruge, N. Galldiks, P. Lohmann, Applications of radiomics and machine learning for radiotherapy of malignant brain tumors, Strahlenther Onkol. 196 (10) (2020) 856–867, https://doi.org/10.1007/s00066-020-01626-8.
[19] C. Cui, X. Yao, L. Xu, et al., Improving the classification of PCNSL and brain metastases by developing a machine learning model based on 18F-FDG PET, Journal of Personalized Medicine. 13 (3) (2023) 539, https://doi.org/10.3390/jpm13030539.
[20] G. Lee, H. Park, S.H. Bak, H.Y. Lee, Radiomics in lung cancer from basic to advanced: current status and future directions, Korean J. Radiol. 21 (2) (2020) 159–171, https://doi.org/10.3348/kjr.2019.0630.
[21] H. Abdollahi, E. Chin, H. Clark, et al., Radiomics-guided radiation therapy: opportunities and challenges. Phys Med Biol. 2022;67(12):12TR02. doi: 10.1088/1361-6560/ac6fab.
[22] D. Wong, S. Yip, Machine learning classifies cancer, Nature 555 (7697) (2018) 446–447, https://doi.org/10.1038/d41586-018-02881-7.
[23] D. Painuli, S. Bhardwaj, köse U., Recent advancement in cancer diagnosis using machine learning and deep learning techniques: a comprehensive review, Comput. Biol. Med. 146 (2022) 105580, https://doi.org/10.1016/j.compbiomed.2022.105580.
[24] J. Kang, Z. Ullah, J. Gwak, MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers, Sensors 21 (6) (2021) 2222, https://doi.org/10.3390/s21062222.
[25] W.F. Wu, C.W. Shen, K.M. Lai, Y.J. Chen, E.C. Lin, C.C. Chen, The application of DTCWT on MRI-derived radiomics for differentiation of glioblastoma and solitary brain metastases, Journal of Personalized Medicine. 12 (8) (2022) 1276, https://doi.org/10.3390/jpm12081276.
[26] S. Bijari, A. Jahanbakhshi, P. Hajishafiezahramini, P. Abdolmaleki, Differentiating glioblastoma multiforme from brain metastases using multidimensional radiomics features derived from MRI and multiple machine learning models, Biomed Res. Int. 2022 (2022) e2016006, https://doi.org/10.1155/2022/2016006.
[27] K. Skogen, A. Schulz, E. Helseth, B. Ganeshan, J.B. Dormagen, A. Server, Texture analysis on diffusion tensor imaging: discriminating glioblastoma from single brain metastasis, Acta Radiol. 60 (3) (2019) 356–366, https://doi.org/10.1177/0284185118780889.
[28] C.Q. Su, X.T. Chen, S.F. Duan, et al., A radiomics-based model to differentiate glioblastoma from solitary brain metastases, Clin. Radiol. 76 (8) (2021) 629.e11–629.e18, https://doi.org/10.1016/j.crad.2021.04.012.
[29] T. Tang, X. Li, Q. Zhang, et al., Development of a novel multiparametric MRI radiomic nomogram for preoperative evaluation of early recurrence in resectable pancreatic cancer, Magn. Reson. Imaging. 52 (1) (2020) 231–245, https://doi.org/10.1002/jmri.27024.
[30] X. Yi, Q. Xiao, F. Zeng, et al., Computed tomography radiomics for predicting pathological grade of renal cell carcinoma, Front Oncol. 10 (2021) 570396, https://doi.org/10.3389/fonc.2020.570396.
[31] X. Yi, Y. Liu, B. Zhou, et al., Incorporating SULF1 polymorphisms in a pretreatment CT-based radiomic model for predicting platinum resistance in ovarian cancer treatment, Biomed. Pharmacother. 133 (2021) 111013, https://doi.org/10.1016/j.biopha.2020.111013.
[32] X. Yi, Q. Pei, Y. Zhang, et al., MRI-Based radiomics predicts tumor response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer, Front. Oncol. 9 (2019) 552, https://doi.org/10.3389/fonc.2019.00552.
[33] Y. Fu, X. Wang, X. Yi, et al., Ensemble machine learning model incorporating radiomics and body composition for predicting intraoperative HDI in PPGL. The Journal of Clinical Endocrinology & Metabolism. Published online September 14, 2023:dgad543. doi: 10.1210/clinem/dgad543.
[34] Z. Zhang, X. Yi, Q. Pei, et al., CT radiomics identifying non-responders to neoadjuvant chemoradiotherapy among patients with locally advanced rectal cancer, Cancer Med. 12 (3) (2023) 2463–2473, https://doi.org/10.1002/cam4.5086.
[35] B. Kocak, B. Baessler, S. Bakas, et al., CheckList for EvaluAtion of Radiomics research (CLEAR): a step-by-step reporting guideline for authors and reviewers endorsed by ESR and EuSoMII, Insights Imaging. 14 (1) (2023) 75, https://doi.org/10.1186/s13244-023-01415-8.
[36] P. Lambin, R.T.H. Leijenaar, T.M. Deist, et al., Radiomics: the bridge between medical imaging and personalized medicine, Nat Rev Clin Oncol. 14 (12) (2017) 749–762, https://doi.org/10.1038/nrclinonc.2017.141.
[37] K.G.M. Moons, D.G. Altman, J.B. Reitsma, et al., Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration, Ann Intern Med. 162 (1) (2015) W1–W73, https://doi.org/10.7326/M14-0698.
[38] C. Chen, X. Ou, J. Wang, W. Guo, X. Ma, Radiomics-based machine learning in differentiation between glioblastoma and metastatic brain tumors, Front Oncol. 9 (2019) 806, https://doi.org/10.3389/fonc.2019.00806.
[39] G. Bathla, S. Priya, Y. Liu, et al., Radiomics-based differentiation between glioblastoma and primary central nervous system lymphoma: a comparison of diagnostic performance across different MRI sequences and machine learning techniques, Eur Radiol. 31 (11) (2021) 8703–8713, https://doi.org/10.1007/s00330-021-07845-6.
[40] J.W. Hopewell, W. Calvo, R. Jaenke, H.S. Reinhold, M.E.C. Robbins, E.M. Whitehouse, Microvasculature and Radiation Damage. In: Hinkelbein W, Bruggmoser G, Frommhold H, Wannenmacher M, eds. Acute and Long-Term Side-Effects of Radiotherapy. Vol 130. Recent Results in Cancer Research. Springer Berlin Heidelberg; 1993:1-16. doi: 10.1007/978-3-642-84892-6_1.
[41] J.E. Villanueva-Meyer, R.F. Barajas, M.C. Mabray, et al., Differentiation of brain tumor-related edema based on 3D T1rho imaging, Eur. J. Radiol. 91 (2017) 88–92, https://doi.org/10.1016/j.ejrad.2017.03.022.
[42] S.C.L. Deoni, Quantitative relaxometry of the brain, Top. Magn. Reson. Imaging. 21 (2) (2010) 101–113, https://doi.org/10.1097/RMR.0b013e31821e56d8.
[43] K.S. Opstad, M.M. Murphy, P.R. Wilkins, B.A. Bell, J.R. Griffiths, F.A. Howe, Differentiation of metastases from high-grade gliomas using short echo time ^{1}H spectroscopy, Magn. Reson. Imaging. 20 (2) (2004) 187–192, https://doi.org/10.1002/jmri.20093.
[44] D.N. Louis, H. Ohgaki, O.D. Wiestler, et al., The 2007 WHO classification of tumours of the central nervous system, Acta Neuropathol. 114 (2) (2007) 97–109, https://doi.org/10.1007/s00401-007-0243-4.
[45] M. Artzi, I. Bressler, B.D. Ben, Differentiation between glioblastoma, brain metastasis and subtypes using radiomics analysis, Magn. Reson. Imaging. 50 (2) (2019) 519–528, https://doi.org/10.1002/jmri.26643.
[46] P. Jarmuzek, K. Kozlowska, P. Defort, M. Kot, A. Zembron-Lacny, Prognostic values of systemic inflammatory immunological markers in glioblastoma: a systematic review and meta-analysis, Cancers (Basel) 15 (13) (2023) 3339, https://doi.org/10.3390/cancers15133339.
[47] S.Q. Wang, Q. Yuan, G.T. Zhang, et al., Preoperative blood testing for glioblastoma, brain metastases, and primary central nervous system lymphoma differentiation, Transl. Cancer Res. 11 (1) (2022), https://doi.org/10.21037/tcr-21-1957.
[48] J.E. Bibault, P. Giraud, A. Burgun, Big Data and machine learning in radiation oncology: State of the art and future prospects, Cancer Lett. 382 (1) (2016) 110–117, https://doi.org/10.1016/j.canlet.2016.05.033.
[49] Y. Xiao, J. Wu, Z. Lin, X. Zhao, A deep learning-based multi-model ensemble method for cancer prediction, Comput. Methods Programs Biomed. 153 (2018) 1–9, https://doi.org/10.1016/j.cmpb.2017.09.005.
[50] B.J. Erickson, P. Korfiatis, Z. Akkus, T.L. Kline, Machine learning for medical imaging, Radiographics 37 (2) (2017) 505–515, https://doi.org/10.1148/rg.2017160130.
[51] A. Westcott, D.P.I. Capaldi, D.G. McCormack, A.D. Ward, A. Fenster, G. Parraga, Chronic obstructive pulmonary disease: thoracic CT texture analysis and machine learning to predict pulmonary ventilation, Radiology 293 (3) (2019) 676–684, https://doi.org/10.1148/radiol.2019190450.
[52] W.H. Lo-Ciganic, J.L. Huang, H.H. Zhang, et al., Evaluation of machine-learning algorithms for predicting opioid overdose risk among medicare beneficiaries with opioid prescriptions, JAMA Netw Open. 2 (3) (2019) e190968, https://doi.org/10.1001/jamanetworkopen.2019.0968.
[53] S.A.S. Souza, R.A.C. Guassu, A.F.F. Alves, et al., Texture analysis: a potential tool to differentiate primary brain tumors and solitary brain metastasis. Multimed Tools Appl. Published online October 3, 2023. doi: 10.1007/s11042-023-17139-2.
[54] Z. Qian, Y. Li, Y. Wang, et al., Differentiation of glioblastoma from solitary brain metastases using radiomic machine-learning classifiers, Cancer Lett. 451 (2019) 128–135, https://doi.org/10.1016/j.canlet.2019.02.054.
[55] Y. Wu, Y. Guo, J. Ma, Y. Sa, Q. Li, N. Zhang, Research progress of gliomas in machine learning, Cells. 10 (11) (2021) 3169, https://doi.org/10.3390/cells10113169.
[56] P. Alves, S. Liu, D. Wang, M. Gerstein, Multiple-swarm ensembles: improving the predictive power and robustness of predictive models and its use in computational biology, IEEE/ACM Trans. Comput. Biol. Bioinf. 15 (3) (2018) 926–933, https://doi.org/10.1109/TCBB.2017.2691329.

[57] L. Jekel, W.R. Brim, M. Von Reppert, et al., Machine learning applications for differentiation of glioma from brain metastasis—a systematic review, Cancers 14 (6) (2022) 1369, https://doi.org/10.3390/cancers14061369.
[58] M. Tateishi, T. Nakaura, M. Kitajima, et al., An initial experience of machine learning based on multi-sequence texture parameters in magnetic resonance imaging to differentiate glioblastoma from brain metastases, J. Neurol. Sci. 410 (2020) 116514, https://doi.org/10.1016/j.jns.2019.116514.
[59] N.C. Swinburne, J. Schefflein, Y. Sakai, et al., Machine learning for semi-automated classification of glioblastoma, brain metastasis and central nervous system lymphoma using magnetic resonance advanced imaging, Ann Transl Med. 7 (11) (2019) 232, https://doi.org/10.21037/atm.2018.08.05.
[60] Q.D. Strotzer, T. Wagner, P. Angstwurm, et al., Limited capability of MRI radiomics to predict primary tumor histology of brain metastases in external validation, Neurooncol Adv. 6 (1) (2024) vdae060, https://doi.org/10.1093/noajnl/vdae060.
[61] C.L. Ramspek, K.J. Jager, F.W. Dekker, C. Zoccali, M. Van Diepen, External validation of prognostic models: what, why, how, when and where? Clin. Kidney J. 14 (1) (2021) 49–58, https://doi.org/10.1093/ckj/sfaa188.