Contrast Media & Molecular Imaging

Advancements in Radiomics in Clinical Imaging

Lead Guest Editor: Xuelei Ma
Guest Editors: Jianguo Zhou, Jie Shen, Chunxiao Guo, and Wei Li

Advancements in Radiomics in Clinical Imaging

Contrast Media & Molecular Imaging

Advancements in Radiomics in Clinical Imaging

Lead Guest Editor: Xuelei Ma
Guest Editors: Jianguo Zhou, Jie Shen, Chunxiao Guo, and Wei Li

Contents

Hindawi
Contrast Media & Molecular Imaging
Volume 2020, Article ID 2127062, 12 pages
https://doi.org/10.1155/2020/2127062

WILEY | Hindawi

Research Article

Diagnostic Accuracy of Machine Learning-Based Radiomics in Grading Gliomas: Systematic Review and Meta-Analysis

Curtis K. Sohn[1] **and Sotirios Bisdas**[1,2]

[1]*Queen Square Institute of Neurology, University College London, Queen Square 7, London WC1N 3BG, UK*
[2]*Lysholm Department of Neuroradiology, The National Hospital for Neurology and Neurosurgery, Queen Square 8-11, London WC1N 3BG, UK*

Correspondence should be addressed to Sotirios Bisdas; s.bisdas@ucl.ac.uk

Received 31 July 2020; Revised 20 November 2020; Accepted 8 December 2020; Published 18 December 2020

Academic Editor: Wei Li

Purpose. This study aimed to estimate the diagnostic accuracy of machine learning- (ML-) based radiomics in differentiating high-grade gliomas (HGG) from low-grade gliomas (LGG) and to identify potential covariates that could affect the diagnostic accuracy of ML-based radiomic analysis in classifying gliomas. *Method*. A primary literature search of the PubMed database was conducted to find all related literatures in English between January 1, 2009, and May 1, 2020, with combining synonyms for "machine learning," "glioma," and "radiomics." Five retrospective designed original articles including LGG and HGG subjects were chosen. Pooled sensitivity, specificity, their 95% confidence interval, area under curve (AUC), and hierarchical summary receiver-operating characteristic (HSROC) models were obtained. *Result*. The pooled sensitivity when diagnosing HGG was higher (96% (95% CI: 0.93, 0.98)) than the specificity when diagnosing LGG (90% (95% CI 0.85, 0.93)). Heterogeneity was observed in both sensitivity and specificity. Metaregression confirmed the heterogeneity in sample sizes ($p = 0.05$), imaging sequence types ($p = 0.02$), and data sources ($p = 0.01$), but not for the inclusion of the testing set ($p = 0.19$), feature extraction number ($p = 0.36$), and selection of feature number ($p = 0.18$). The results of subgroup analysis indicate that sample sizes of more than 100 and feature selection numbers less than the total sample size positively affected the diagnostic performance in differentiating HGG from LGG. *Conclusion*. This study demonstrates the excellent diagnostic performance of ML-based radiomics in differentiating HGG from LGG.

1. Introduction

Glioma is the most common primary malignant brain tumor that accounts for 80% of malignancies [1], and 2% of all cancers in US adults [2]. According to the World Health Organization (WHO) classification [3], gliomas are subdivided into two groups based on their malignant status low-grade glioma (LGG) for grades I to II with focal symptoms and high-grade glioma (HGG) for III to IV with generalized symptoms. Grade IV tumors called glioblastoma (GBM) account for 54% of all gliomas [4], with a median survival rate of 15 months [5]. Treatment of gliomas is essential since there is an eventual progression from LGG to HGG due to gliomas' distinctive molecular and clinical features [6]. For targeted treatment that is individualized to specific changes in individual tumors, different treatments including a near-total resection, postsurgical radiation, or temozolomide combined with radiation must be considered depending on the glioma's grade [2]. Therefore, the classification of tumor levels is crucial for intraoperative decision-making.

Magnetic resonance imaging (MRI) has been utilized to classify gliomas noninvasively for histopathological purposes. Recent studies have demonstrated the feasibility of conventional MRI sequences, especially gadolinium-based contrast-enhanced T1-weighted imaging (T1-CE) [7] when grading gliomas. With technological developments, advanced MRI sequences also contribute to physiological and metabolic assessments when classifying gliomas, such as perfusion-weighted imaging (PWI) [8] and diffusion-weighted imaging (DWI) [9]. However, previous studies on

grading gliomas were limited due to utilizing only a small number of parameters extracted from a single MRI sequence.

A capability of radiomics analysis, which maximizes the number of quantitative image features from digital images, has great potential for the assessment of tumor biology [10]. The vast quantities of radiomics data enable information to be extracted from the entire tumor. As a result, radiomics can overcome intratumoral heterogeneities in both the molecular and histopathological assessment of tumors using quantitative values [11] that contribute to evidence-based decision-making in oncology. The assessment of both mutation status and gene expression, such as O6-methylguanine-DNA-methyltransferase (MGMT) gene expression or isocitrate dehydrogenase (IDH) mutation, is essential for predicting therapeutic responses when treating gliomas. Radiomics has proven the potential for the genotype classification of prognostic factors to predict IDH status, 1/19q codeletion status, or MGMT methylation status [12–14] in glioma-related studies. However, the treatment of gliomas cannot be processed by genotype alone since IDH mutation or 1/19q codeletion status tends to be used for classifying grades II and III and MGMT promoter methylation for grade IV. Therefore, the histological grade should also be incorporated into the genotype classification of gliomas. The vast amount of quantitative image features, including first-, second-, and higher-order statistical features, can represent histological values that include intensity differences and spatial interrelationships. As a result, radiomic features can provide distinctive information about tumor phenotypes and their microenvironments. Considering the heterogeneous histopathology biomarkers of angiogenesis, apoptosis, proliferation, and cellular invasion in gliomas [15], extracting a large amount of hidden data using radiomics could be a potential tool in classifying gliomas from single- or multiparameter MRI sequences.

As far as we know, no previous research has performed a systematic evaluation of the accuracy of machine learning-(ML-) based radiomics analysis in differentiating HGG from LGG. Therefore, the purpose of the study was twofold: first, to estimate the diagnostic accuracy of ML-based radiomics analysis in classifying HGG and second, to identify the potential covariates that could affect the diagnostic accuracy of ML-based radiomics.

2. Method

A meta-analysis was performed using Meta-DiSc version 1.4 (Unit of Clinical Biostatistics Team, Hospital Universitario Ramón y Cajal, Madrid, Spain). However, the Meta-DiSc version 1.4 uses outdated statistical methods since the Moses–Littenberg method does not account for between-study variances [16]. Therefore, RStudio (version 4.0.2) using the MADA package was implemented to utilize hierarchical summary receiver-operating characteristic (HSROC) models and bivariate models.

2.1. Literature Search. This meta-analysis was performed following the Preferred Reporting Items for Systematic Reviews and Meta-Analysis guidelines [17]. A primary literature search of the PubMed database was conducted to find all related literature in English between January 1, 2009, and May 1, 2020, including Medical Subject Headings (MeSH) and non-MeSH terms (see Supplementary Materials for key terms (available (here))).

2.2. Inclusion and Exclusion Criteria. All studies were selected by the following criteria: (a) original research articles; (b) patients with histopathologically confirmed WHO grade gliomas including both lower-grade glioma and high-grade glioma; (c) ML-based with radiomics features that were applied to classify gliomas using radiomics features; and (d) information for reconstructing 2×2 tables to estimate the diagnostic sensitivity and specificity for grading gliomas was included.

Studies were excluded if (a) they did not use ML to classify the grade of gliomas; (b) did not focus on differentiating between LGG and HGG; (c) no replies were received from the authors after requesting the data related to reconstructing the 2×2 table or subgroup analysis; and (d) they had a small sample size for performing the machine learning classifier.

2.3. Quality Assessment. Two independent reviewers conducted the quality assessment (S.C. and B.S.). Four main domains including patient selection, index test, reference standard, and flow and timing were evaluated based on the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) [18].

2.4. Statistical Analysis. A meta-analysis of the performance of radiomics-based ML in differentiating HGG from LGG was conducted. Therefore, the definition of true positive (TP) was set for HGG and true negative (TN) for LGG. For consistency, the TP definition in the study of Zhao et al. had to be redefined by switching sensitivity and specificity in the calculation. If the data for reconstructing the 2×2 table and analyzing subgroup analysis were insufficient, we contacted the authors.

The performance of the studies that implemented multiple ML classifiers for grading gliomas was averaged. For studies that reported the results of both the training and testing sets, the best result was selected for the meta-analysis. The performance of ML using radiomics in differentiating HGG from LGG was performed using the bivariate random effects model.

The following methods and criteria were used to estimate heterogeneity: (a) considering the small number of studies and lower statistical strength of Cochrane's Q test, p value <0.10 (not 0.05) indicated the presence of heterogeneity (14); (b) Higgins inconsistency index (I^2) test values of 25%, 50%, and 75% defined the heterogeneity as low, moderate, or high; (c) a forest plot for assessing the heterogeneity in sensitivity and specificity and the visual assessment of the forest plots to assess the presence of threshold effect (increasing sensitivity with decreasing specificity and a positive correlation

between sensitivity and the false-positive rate); (d) a large difference between the 95% confidence region and 95% prediction region in the HSROC curve; and (e) a Spearman correlation coefficient value >0.6, that indicated a threshold effect across all studies.

Metaregression and subgroup analysis were performed to explain the possible factors that contribute to heterogeneity and the factors contributing to the diagnostic performance of ML-based radiomics when grading gliomas with the following covariates: (a) subject number, (b) sequence types, (c) selected feature numbers, and (d) inclusion of the testing set.

3. Result

3.1. Literature Search. One hundred and ninety-four studies of interest were found, and five studies were selected for this meta-analysis after considering the inclusion and exclusion criteria (Figure 1).

3.2. Risk of Bias Assessment. Overall, a high risk of bias was estimated in the studies summarized in Table 1, with detailed descriptions given in Table 2. The high risk of bias could be attributed to the nature of the retrospective study, in which the patient's outcomes are known. Therefore, a case-control design for selecting patients already represents a high risk of bias. Furthermore, using public data could also lead to a high risk of bias in the QUADAS-2 assessment because not all acquisition factors for LGG and HGG in public data can be controlled. An assessor already knows the patients' diagnosis or reference standard results because WHO-classified patients were considered for research. ML-based radiomics analysis studies tend to find a methodological justification from a previous study method with an advanced ML classifier to improve the diagnostic accuracy. Therefore, prior knowledge before implementing the index test may introduce a high risk to the "index test" in the second domain. However, the reference standard for the histological diagnosis of HGG and LGG has already taken into account the accurate grading of gliomas that leads to reducing the risk of bias in the "reference standard" domain. Finally, though it was assumed that most were related to preoperative studies, it was unclear whether there was an appropriate interval between the index test and the reference standard or whether patients received a specific therapy. Therefore, the fourth domain of "flow and timing" in the reviewed studies had an unclear general bias content. Significant heterogeneity was present in data sources regarding image acquisition, feature engineering, and ad hoc analysis. Consequently, the quality assessment was limited regarding the applicability of ML-based radiomics analysis for grading gliomas.

3.3. Data Extraction. A summary of the results is presented in Table 3, while the method-related information is summarized in Table 4. Three of five studies utilized a single MRI sequence acquired by either conventional or advanced imaging [14, 19, 20], while the remaining studies implemented both conventional and advanced ASL and DWI sequences in [21, 22]. An imbalanced ratio was observed between the HGG and LGG datasets in the studies that used a large number of samples [19, 21], while the remaining three studies had a ratio equal to the sample [14, 20, 22]. Two of the five studies selected feature numbers equal to or greater than the total sample size [14, 20], while the remaining three studies selected fewer than the total sample size [19, 21, 22]. Only two studies included a testing set [20] without reporting the sample size of the testing set [19].

3.4. Heterogeneity Assessment. A forest plot was drawn to estimate the heterogeneity in sensitivity and specificity in Figure 2. Heterogeneity was found in both sensitivity ($I^2 = 69.70\%$, $p = 0.01$) and specificity ($I^2 = 80.20\%$, $p \leq 0.01$).

A large difference between the confidence region and 95% prediction regions in the HSROC curve represents the possibility of heterogeneity across the studies in Figure 3.

3.5. Threshold Effect Assessment. The Spearman correlation coefficient between the sensitivity and false-positive rate was −0.4 ($p = 0.51$), indicating the absence of a threshold effect. A threshold effect indicates a positive correlation between sensitivities and the false-positive rate that leads to a "shoulder arm" plot in the summary receiver-operating characteristic curve space. However, the visual assessment of the HSROC indicates the absence of a threshold effect as shoulder is absent in the HSROC space.

3.6. Data Analysis. Significant heterogeneity was observed in both pooled sensitivity ($I^2 = 69.70\%$, $p = 0.0104$) and pooled specificity ($I^2 = 80.20\%$, $p \leq 0.01$) as is shown in Figure 2. Therefore, the HSROC model based on a random effect model was applied to account for both intra- and interstudy variances in analyzing the diagnostic accuracy of the ML method with radiomics based on differentiating HGG from LGG. The area under the curve (AUC) value of 0.96 indicates high diagnostic performance.

3.7. Metaregression. A bivariate metaregression with a p value-based chi-squared statistic recommended by the Cochrane *diagnostic test accuracy (DTA) handbook* [23] was performed for the metaregression. As a result, the metaregression confirmed the heterogeneity in sample size ($p = 0.05$), imaging sequence types ($p = 0.02$), and data sources ($p = 0.01$), but not in the testing set ($p = 0.19$), feature extraction number ($p = 0.36$), or selecting feature number ($p = 0.18$). Bivariate metaregression based on the random effect model was also performed to analyze the regression coefficients between two groups related to four covariates: (a) subject number sample size of fewer than 100 vs. sample size of more than 100; (b) implementing only conventional or advanced MRI sequences vs. using both advanced MRI sequences and conventional MRI sequences; (c) selecting a feature number greater than or close to the total sample size vs. smaller than the total sample size; and (d) including testing set (validation set) or not.

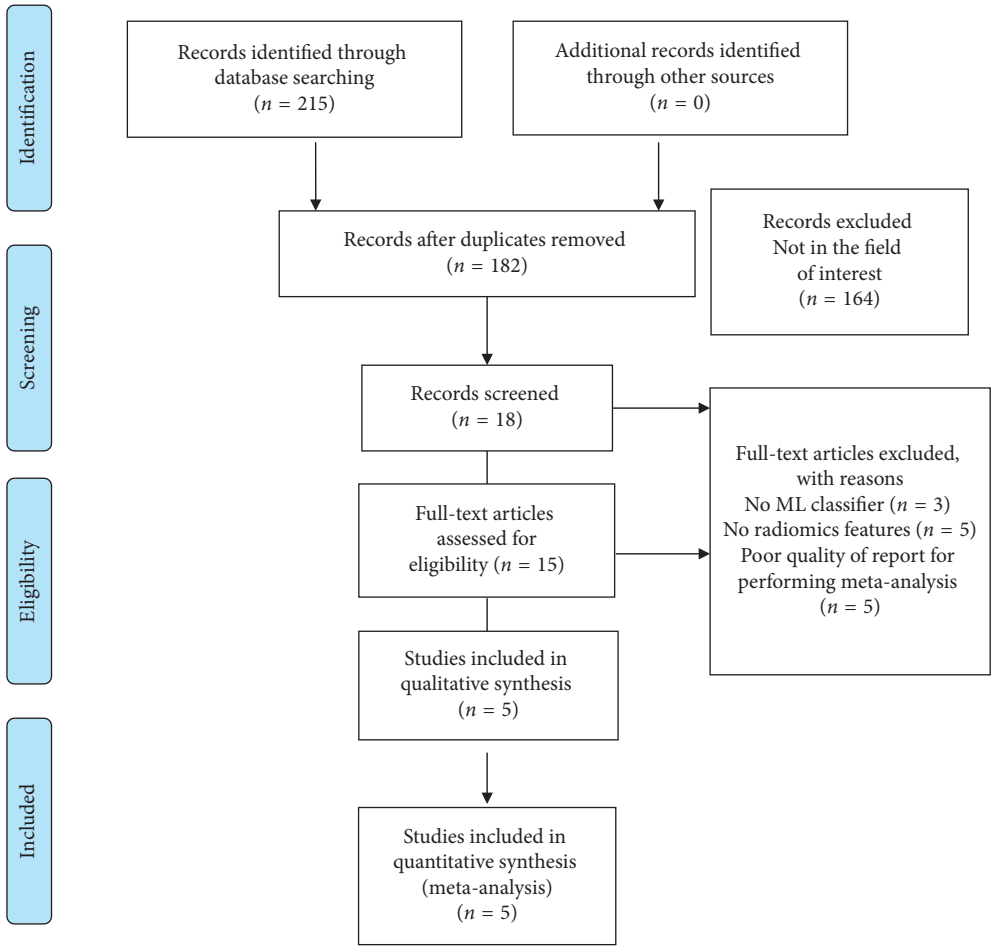


Figure 1: Search-strategy flowchart in accordance with the PRISMA guidelines used in the meta-analysis.

Table 1: Summary of QUADS-2 tool assessment of the literature used in the meta-analysis.

Authors and year	Risk of bias			
	Patient selection	Index test	Reference standard	Flow and timing
Cho et al. 2018	+	+	−	?
Tian et al. 2018	+	+	−	−
Hashido et al. 2018	+	−	−	?
Vamvakas et al. 2019	+	+	+	?
Zhao et al. 2020	+	−	−	?

Risk of bias: + = high, − = low, and ? = unclear risk.

The z-value for the regression coefficients for the sensitivities was significant for the sample size ($p \leq 0.01$) and feature number ($p \leq 0.01$). Therefore, the studies with sample sizes of more than 100 and a feature number smaller than the sample size exhibited better sensitivity, while the point estimate for the false positive rate did not indicate any effect. The z-value for the regression coefficient for the false-positive rate was significant ($p \leq 0.01$) in the single-image sequences group, which offered a higher false-positive rate than the multiparameter image group. The point estimate for sensitivity did not indicate any effect. The z-values for the regression coefficient for both the sensitivity and false-positive rate were insignificant ($p = 0.3$) and ($p = 0.4$), indicating no statistical difference whether the testing set was included or not.

3.8. Subgroup Analysis. The sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) were combined using a random effects model because of the heterogeneity across the reviewed studies in Table 5. In the subgroup analysis, the overall sensitivity of diagnosing HGG was higher (96% (95% CI, 0.93, 0.98)) than the specificity of diagnosing LGG (90% (95% CI, 0.85, 0.93)).

Similar to the result of metaregression, significant sensitivity, PLR, and NLR differences were found in the

Table 2: Detailed QUADS-2 tool assessment after the two reviewers reached a consensus.

Authors and year	Patient selection	Index test	Reference standard	Flow and timing
Cho et al. 2018	The experiment was designed to be a retrospective study and did not include a random sample. However, the study avoided inappropriate exclusion (high risk)	The index test results were interpreted knowing the results of the reference standard. It was unclear whether a prespecified threshold was used (high risk)	It was unclear whether the reference standard was likely to classify the target condition accurately. The reference standard results were interpreted knowing the results of the index test (low risk)	It was unclear whether there was an appropriate interval between the index test and reference standard and whether all patients were included in the analysis. All patients received a reference standard, but it is unclear whether it was the same reference standard (unclear)
Tian et al. 2018	The experiment was designed to be a retrospective study and did not include a random sample. However, the study avoided inappropriate exclusion (high risk)	The index test results were interpreted knowing the results of the reference standard. However, a pre specified threshold was used (high risk)	The reference standard was likely to classify the target condition accurately. The reference standard results were interpreted without knowing the results of the index test (low risk)	It was clear whether there was an appropriate interval between the index test and reference standard. All patients received a reference standard, but it was unclear whether it was the same reference standard. Not all patients were included in the analysis (low risk)
Hashido et al. 2018	The experiment was designed to be a retrospective study and did not include a random sample. However, the study avoided inappropriate exclusion (high risk)	The index test results were interpreted without knowing the results of the reference standard. Furthermore, a pre specified threshold was used (low risk)	The reference standard was likely to classify the target condition accurately. The reference standard results were interpreted without knowing the results of the index test (low risk)	It was unclear whether there was an appropriate interval between the index test and reference standard. All patients received a reference standard, but it was unclear whether it was the same reference standard. Not all patients were included in the analysis (unclear)
Vamvakas et al. 2019	The experiment was designed to be a retrospective study and did not include a random sample. In addition, it was unclear whether the study avoided inappropriate exclusion (high risk)	The index test results were interpreted knowing the results of the reference standard. It was unclear whether a prespecified threshold was used (high risk)	The reference standard was likely to classify the target condition accurately. The reference standard results were interpreted knowing the results of the index test (high risk)	It was unclear whether there was an appropriate interval between the index test and reference standard and included all patients in the analysis. All patients received a reference standard, but it was unclear whether it was the same reference standard (unclear)
Zhao et al. 2020	The experiment was designed to be a retrospective study and did not include a random sample. In addition, it was unclear whether the study avoided inappropriate exclusion (high risk)	The index test results were interpreted without knowing the results of the reference standard. However, it was unclear whether a prespecified threshold was used (low risk)	It was unclear whether the reference standard was likely to classify the target condition accurately. The reference standard results were interpreted without knowing the results of the index test (low risk)	It was unclear whether there was an appropriate interval between the index test and reference standard. All patients received a reference standard, but it was unclear whether it was the same reference standard. Not all patients were included in the analysis (unclear)

covariates of sample size and feature number; however, specificity was not detected. The studies using a sample size of more than 100 had higher sensitivity, PLR, and lower NLR than studies with a sample size of fewer than 100 (98 vs. 88), PLR (12.10 vs. 7.89), and NLR (0.03 vs. 0.14), but no difference in the specificity (90% vs. 90%). Furthermore, this result was aligned with the metaregression result, in which the studies using sample sizes of more than 100 had a lower false-positive rate.

In terms of feature engineering, the studies with selected feature numbers smaller than the total sample size had better sensitivity (97% vs. 85%), PLR (13.48 vs. 5.71), specificity

Table 3: Summary of the results evaluated in the reviewed studies.

Study and year	Method	Algorithm	Dataset/ HGG-LGG	MRI sequence	Best performance				Limitation
					AUC	DA (%)	Sen (%)	Spe (%)	
Cho et al. 2018	Classic machine learning	Multiple algorithms	WHO II–IV ($n = 285$)/210-75	T1, T1-C, T2, T2-FLAIR	0.94	92.92	97.86	79.11	No dataset separation information for training and testing cohort. Sample imbalance size between LGG and HGG.
Tian et al. 2018	Classic machine learning	SVM	WHO II–IV gliomas ($n = 153$)/111-42	Multiparametric	0.99	96.80	96.40	97.30	Sample imbalance sample size between LGG and HGG.
Hashido et al. 2018	Classic machine learning	Logistic regression	WHO II–IV ($n = 46$)/31-15	ASL, PWI (DSC)	0.96	NA	89.30	92.90	Small sample size. Small sample size used in the training set. Large feature number than the total sample size.
Vamvakas et al. 2019	Classic machine learning	SVM	WHO I–IV ($n = 40$) 20-20	Multiparametric	0.96	95.50	95	96	Small sample size.
Zhao et al. 2020	Classic machine learning	RF	WHO II-III gliomas ($n = 36$) 17-19	T1-C, T2- FLAIR	0.86	78.10	78.30	77.80	Small sample size. Large feature number compared to the total sample size.

Table 4: Summary of the methods used in the reviewed studies.

Study and year	Data source	External validation	Feature type	Feature extraction	Feature selection	Segmentation
Cho et al. 2018	Public	Training + testing	First-order and second-order (GLCM, ISZ)	486	5	ROI
Tian et al. 2018	Private	Training	First-order, second-order (GLCM, GLCGM)	510	30	VOI
Hashido et al. 2018	Private	Training (42) + testing (4)	First-order, second-order (GLCM, GLDM, GLRLM, GLSZM, and NGTDM)	91	75	Random forest-based semiautomatic tumor segmentation
Vamvakas et al. 2019	Private	Training	First-order and second-order texture (GLCM, GLRLM)	581	21	VOI
Zhao et al. 2020	Private	Training	First-order and second-order (GLCM, GLRLM, GLSZM, and GLDM)	1072	30	VOI

(90% vs. 85%), and lower NLR (0.03 vs. 0.18) than the studies that selected feature numbers close to or greater than the total sample size. Furthermore, the lower false-positive rate result from the metaregression when selecting a number of features fewer than the total sample size was observed. However, there was no significant statistical difference in sensitivity between the studies that extracted only second-order features or both first- and second-order features ($p = 0.3$), even with the higher sensitivity in the group that extracted both features (97% vs. 94%) in the subgroup analysis, but with lower specificity (84% vs. 95%), high PLR (11.19 vs. 7.23), and NLR (0.11 vs. 0.05).

No sensitivity difference was estimated in the group using single sequence and multiple sequences (96% vs. 96%), but the specificity was higher in the group using multiple sequences (97% vs. 81%) with a higher PLR (30.39 vs. 4.61) and lower NLR (0.04 vs. 0.09). Applying only the training set produced a higher sensitivity (94% vs. 97%), specificity (95% vs. 81%), PLR (12.91 vs. 5.32), and lower NLR (0.09 vs. 0.05) than the group that used both the training and testing sets. However, the result of metaregression indicated no statistical difference in either the sensitivity or false-positive rate. The overlap of the confidence interval in the sensitivity between the two groups was aligned with the result of metaregression. However, the higher specificity in the group that used only the training set indicated a better differentiation of LGG.

4. Discussion

Overall, the meta-analysis confirmed the source of heterogeneity from the covariates, including sensitivity, specificity,

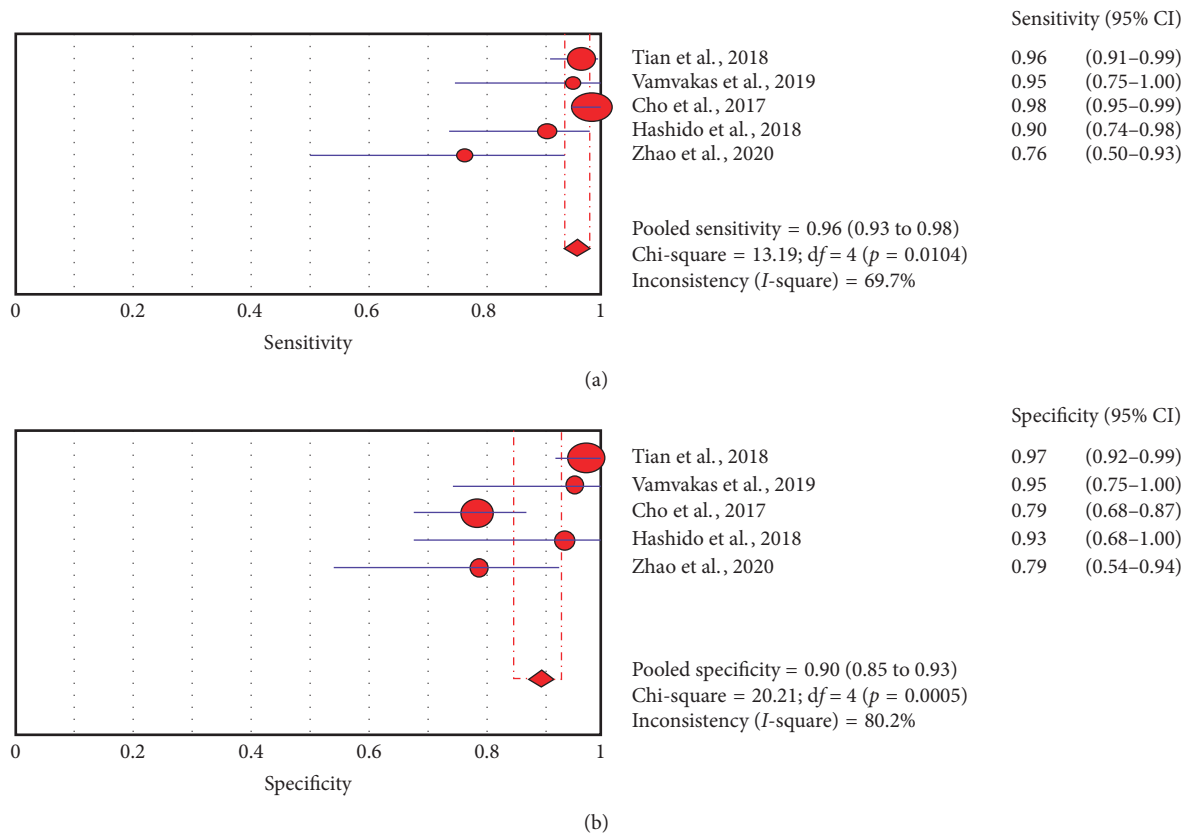


Figure 2: Pooled estimation of sensitivity and specificity for the diagnostic accuracy of radiomics using machine learning in differentiating HGG from LGG. Circles and horizontal lines represent the point estimate and 95% confidence intervals.

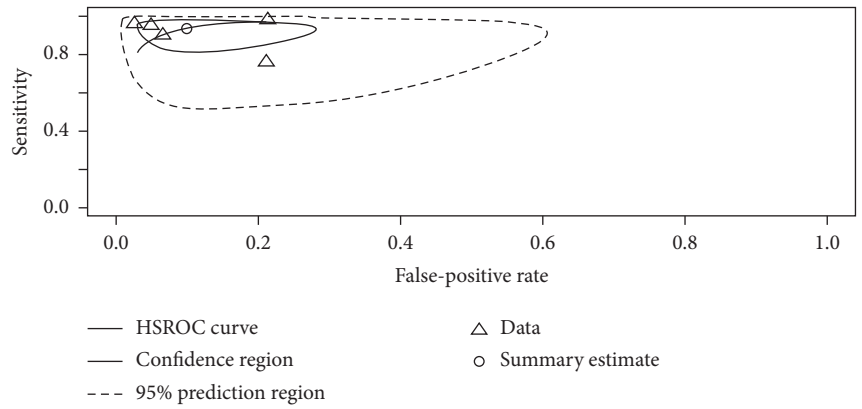


Figure 3: The HSROC curve displaying the diagnostic performance of ML-based radiomics in differentiating HGG from LGG. A large difference between the 95% confidence and prediction regions indicates a high possibility of heterogeneity across the reviewed studies.

sample size, imaging sequences, and data source. The main reason for greater heterogeneity could be attributed to the nature of the multiple steps included in the radiomics process of image acquisition, data source, segmentation, feature engineering, and ad hoc analysis [24].

The results of the meta-analysis indicated that a greater than 100 sample size positively affected the diagnostic performance of sensitivity (HGG), but not specificity (LGG). This result could be attributed to two causes. First, having a large sample size is essential to improve training [25] and to avoid the overfitting that occurs in ML-based research. Therefore, both statistical analysis and ML training may favor the result of the group with a greater than 100 sample size. The total sample size difference between the two groups (507 vs. 122) also contributed to the results. Second, the result could be attributed mainly to the imbalance in the HGG and LGG sample ratio. This is the main challenge in the medical dataset, where the ML-based classification

Table 5: Result of multiple subgroup analysis of machine learning-based radiomics for grading gliomas.

Subgroup	Study number	Patient number	Sensitivity	Specificity	PLR	NLR	Diagnostic odds ratio
All combined	5	629	0.96 (0.93–0.98)	0.90 (0.85–0.93)	9.53 (3.55–25.57)	0.07 (0.02–0.20)	153.85 (32.36–731.44)
Populations							
>100	2	507	0.98 (0.95–0.99)	0.90 (0.85–0.94)	12.099 (1.37–107.12)	0.03 (0.02–0.06)	393.81 (80.89–1917.3)_
<100	3	122	0.88 (0.78–0.95)	0.90 (0.77–0.96)	7.89 (2.21–28.15)	0.14 (0.05–0.39)	65.13 (7.84–540.95)
Sequence							
Single (CS or advanced)	2	262	0.96 (0.93–0.98)	0.81 (0.72–0.88)	4.61 (3.14–6.77)	0.09 (0.02–0.44)	66.75 (10.33–431.19)
Multiple (CS and advanced)	3	367	0.96 (0.91–0.99)	0.97 (0.92–0.99)	30.391 (11.585–79.726)	0.04 (0.017–0.09)	774.25 (202.54–2959.77)
Feature number							
≥Sample size	2	82	0.85 (0.72–0.94)	0.85 (0.69–0.95)	5.71 (1.39–23.46)	0.18 (0.07–0.52)	33.76 (3.36–339.14)
<Sample size	3	547	0.97 (0.95–0.99)	0.90 (0.85–0.94)	13.48 (2.56–71.12)	0.03 (0.02–0.06)	369.98 (19.68–6956.0)
Training and testing set							
Training set	3	331	0.94 (0.88–0.97)	0.95 (0.89–0.98)	12.91 (2.02–82.22)	0.09 (0.02–0.47)	154.56 (7.30–3276.9)
Training + testing set	2	298	0.97 (0.94–0.99)	0.81 (0.72–0.89)	5.32 (2.55–11.09)	0.05 (0.1–0.22)	176.99 (63.76–491.30)

method prefers a larger to a smaller sample during training [26, 27]. Therefore, the imbalanced ratio between LGG and HGG (total sample 240 vs. 389) may increase the sensitivity due to the large sample size of HGG without considering the distribution of data ratio between LGG and HGG, while reducing the specificity performance [19, 21]. In contrast, the overall equal sample ratio was found in the group with a sample size of less than 100, as summarized in Table 3 [14, 20, 22]. As a result, the specificity may increase in the small sample size group that did not consider the LGG sample as a minor group during the classifier training. Therefore, there was no difference in the specificity even with the higher sample size of LGG in the large sample size group than in the small sample size group (117 vs. 54). Both limited and imbalanced numbers of samples between LGG and HGG across the studies also affected the separation of the validation and testing datasets. The majority of studies did not include the testing set because of the small sample size. Furthermore, poor-quality reporting that did not include the subject number used in training and testing [19] and included only four subjects in the testing set [20] was observed. Therefore, subgroup analysis that includes the testing set may not be an appropriate criterion for this meta-analysis, even though it is necessary for the external validation of the model. However, a higher specificity was observed in the group that included only the training set [14, 21, 22]. This could be attributed to the balanced ratio between LGG and HGG in the group described in Table 3, not to the inclusion of the testing set. In contrast, imbalanced ratios of large HGG samples over LGG samples were observed in the group that included both the training and testing sets [19, 20]. In short, it is reasonable to assume that the labeled data balance plays a key role in grading gliomas.

Interestingly, there was a significant difference in the specificity but not in the sensitivity between the groups using a single MRI sequence [14, 19, 20] and multiparametric images [21, 22]. The nonsignificant sensitivity difference between the two groups can be attributed to two reasons. First, quantifying the heterogeneous spatial gray distribution including intratumoral spatial variation and intensity of the entire tumor in the second-order features allowed for the classification of the heterogeneous HGG over the first-order statistical TA [21, 28]. Therefore, the second-order features that were extracted in all reviewed studies could contribute to a statistically insignificant sensitivity difference between the two groups. Second, the performance of combining different MRI techniques in differentiating HGG from LGG is questionable because the combination of conventional and advanced MRI sequences, including PWI, DWI, and magnetic resonance spectroscopy, did not significantly increase the glioma grading performance [29, 30]. Furthermore, conventional MRI variables including enhancement and necrosis have been reported as the major predictors in differentiating HGG from LGG above the combination of conventional MRI with PWI, DWI, and MRS [31]. Therefore, the sequence differences between the two groups may be insignificant because of the contributory role played by the conventional MRI sequences that were included in each group. As a result, no significant difference in the sensitivity between the single MRI sequence and the multiparametric image was observed in this study.

The higher specificity in the multiparametric group can be attributed to the apparent diffusion coefficient (ADC) values extracted from DWI [21] and cerebral blood flow extracted by perfusion imaging [22]. Even though the ADC values in differentiating LGG from HGG varied from study to study [29], T1-CE, ADC slow, T2 WI, and CBF have been widely used to classify LGG. Among various biomarkers, the ADC value has demonstrated the feasibility of classifying LGG by the higher ADC values in LGG than HGG [32, 33], and the entropy of ADC values among various texture analysis software showed promising results [7, 9]. Therefore, it is reasonable to assume that advanced MRI sequences for estimating angiogenesis and blood perfusion play a role in classifying LGG. In short, the higher false-positive rate in the single-image sequence group and the higher specificity of differentiating LGG in the multiple sequences resulted from the use of multiple imaging sequences for grading gliomas.

All results including sensitivity, specificity, PLR, and NLR were higher in the group with fewer feature numbers [14, 19, 21] than those that were greater than or equal to the total sample size [20, 22]. This can be attributed to two causes. First, the sample size difference between the two groups may affect the performance of the group that selected fewer feature numbers than the total sample size (547 vs. 82), as described in the following section of this paper. Second, selecting a method to reduce the dimensionality of feature space in radiomics is an essential component of feature selection. In fact, reducing redundant features is important to avoid overfitting and improving data, even though there is no gold standard for the appropriate number of features. Therefore, a small number of features are recommended, either one-tenth of the total sample data [34] or the square root of the total sample data [35]. Furthermore, a high contribution of the gray-level gradient cooccurrence matrix (GLGCM) features when grading gliomas suggested the importance of the second-order feature numbers in the study [21]. In short, the group that selected a smaller number of quantitative features than the total sample size exhibited better performance in grading gliomas.

4.1. Limitation. Several limitations were observed in our study. Only five studies were included in the meta-analysis, while the recommended number of studies for meta-analysis, according to the Cochrane DTA handbook, is at least 30 for sufficient power [23].

Therefore, the assessment of a publication bias using the Deek funnel plot asymmetry test was excluded because the small number of meta-analyses could skew the result due to the number difference between the small and large studies [36, 37] and the heterogeneous sources in the meta-analysis [38].

The poor quality of the report that excluded crucial result information contributed to the scarcity of literature for performing a meta-analysis of the accuracy of this diagnostic test in radiomics studies. Overall, the poor quality of the reporting has limited the study of radiomics in neuro-oncology [39] because guidelines for reporting quantitative imaging results have not yet been established. Regardless of the scarcity of literature, the total sample size of 629 subjects may be sufficient to represent the predictive value of ML-based radiomics analysis in differentiating HGG from LGG.

4.2. Future. Several factors should be improved for future studies related to ML-based radiomics for grading glioma. First, all reviewed studies did not include the updated WHO 2016 glioma classification standard for combining molecular profiling with histopathological profiling. Second, the dataset size still plays a key role in grading gliomas. Therefore, enlarging the dataset should also be considered to overcome the imbalance caused by oversampling a small sample to improve classifier performance [21].

Furthermore, the enlarged dataset would lead to separate training and testing sets that allow for the external validation of ML classifier performance. Third, incorporating patient demographics or clinical history should be considered to improve the classification of ML models. Finally, the variation process included in the radiomics analysis that is based on the numerical extraction approach to image analysis could affect the result due to bias and variance, not underlying biologic effects. Therefore, standardization in image acquisition, segmentation, feature engineering, statistical analysis, and the reporting format should be established for reproducibility and the generalization of ML-based radiomics studies. Essential steps for standardization include optimizing the standard imaging acquisition process, fully automating the process for segmentation and feature engineering, reducing the redundancy of feature numbers, enhancing the reproducibility of radiomics features, and transparently reporting results. Therefore, the following guidelines suggested by the relevant professional societies, such as the Society of Nuclear Medicine and Molecular Imaging, the Quantitative Imaging Network, Radiology Society of North America, and the European Society of Radiology that lead the field in imaging methods, including radiomics, should be considered. Furthermore, it has been reported that the magnet strength, flip angles, number of excitations, and different scanner platforms could affect both first-order and second-order features [40]. For example, the gray-level gradient cooccurrence matrix (GLCM) could be invariant to magnetic strength but susceptible to flip angles. However, the first-order features of entropy that are considered the most stable features have high reproducibility [41]. Therefore, the reproducibility of information related to radiomics features should be considered depending on the image acquisition method. Apart from the radiomics feature type, a segmentation method should be considered for the reproducibility of radiomics features. A registration distortion between MRI sequences could cause the incorrect localization of the region of interest and could affect the feature extraction process in radiomics analysis [42]. To reduce individual variability, deep learning-based automated segmentation and feature engineering gained attention [43]. Therefore, a deep learning-based approach should also be considered in radiomics analysis. The SVM-based classifier with recursive feature elimination (RFE) was found to be superior to other 25 ML-based classifiers and 8 independent attribute selection methods in grading gliomas using the multiparameter approach [44].

Therefore, an SVM-based classifier with RFE could be the best method to reduce feature redundancy and improve the ML classifier performance of glioma grading [28].

5. Conclusion

ML with radiomics demonstrated excellent diagnostic performance in differentiating HGG from LGG. The results of the meta-analysis provide the following recommendations to perform grading gliomas with ML-based radiomics: (a) use a large sample with oversampling of a minor class to balance the sample ratio and include the external validation set; (b) employ a multiparameter approach to extract the second-order features from the T1-CE sequence and ADC entropy from DWI; (c) select features with a number smaller than the total sample size by combining clinical information; and (d) implement the SVM classifier with SVM-REF attribute selection.

We submit that a methodological standard to ensure the reproducibility of ML-based radiomics is warranted for clinical application.

Abbreviations

ASL:	Arterial spin labeling
HGG:	High-grade glioma
LGG:	Low-grade glioma
ML:	Machine learning
ADC:	Apparent diffusion coefficient
FA:	Fractional anisotropy
MRI:	Magnetic resonance imaging
DWI:	Diffusion-weighted imaging
DTI:	Diffusion tensor imaging
GBM:	Glioblastoma
SVM:	Support vector machine
SVM-RFE:	Recursive feature elimination
IDH:	Isocitrate dehydrogenase
MGMT:	O6-Methylguanine-DNA-methyltransferase
HSROC:	Hierarchical summary receiver-operating characteristic curve
AUC:	Area under curve
QUADAS-2:	Quality assessment of diagnostic accuracy studies-2
TP:	True-positive
TN:	True-negative
PLR:	Positive likelihood ratio
NLR:	Negative likelihood ratio
DOR:	Diagnostic odds ratio
RFE:	Recursive feature elimination
GLCM:	The gray-level gradient cooccurrence matrix.

Data Availability

The data used to support the findings of this study are available upon request to skgtohn@ucl.ac.uk.

Additional Points

Key Findings. (i) Machine learning with radiomics demonstrated excellent diagnostic performance in differentiating high-grade glioma from low-grade glioma. (ii) The lower false-positive rate and higher specificity for classifying lower-grade glioma favored the use of multiparametric imaging, even though a nonstatistically significant sensitivity difference between the single-image sequence and multiparameter imaging group was observed. (iii) Based on the results of the meta-analysis, several factors positively affected the performance of machine learning-based radiomics when grading gliomas: (a) a large sample size with oversampling of a minor class to balance the sample ratio; (b) use of the multiparametric approach to extract the second-order features from the gadolinium-based contrast-enhanced sequence and an apparent diffusion coefficient entropy from diffusion-weighted image; (c) selection of features with fewer numbers than the total sample size combined with clinical information; and (d) implementation of a support vector machine classifier with support vector machine-recursive feature elimination attribute selection. (iv) A methodological standard to ensure the reproducibility of ML-based radiomics is warranted for clinical application.

Consent

Written informed consent was not required because of the nature of study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Supplementary Materials

The supplementary file contains the supplementary information for key terms. (*Supplementary Materials*)

References

[1] G. Minniti, G. Lombardi, and S. Paolini, "Glioblastoma in elderly patients: current management and future perspectives," *Cancers*, vol. 11, no. 3, p. 336, 2019.

[2] S. R. Chandana, S. Movva, M. Arora, and T. Singh, "Primary brain tumors in adults," *American Family Physician*, vol. 77, no. 10, p. 1423, 2008.

[3] P. Wesseling and D. Capper, "WHO 2016 classification of gliomas," *Neuropathology and Applied Neurobiology*, vol. 44, no. 2, pp. 139–150, 2018.

[4] Q. T. Ostrom, H. Gittleman, P. Farah et al., "CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2006–2010," *Neuro-Oncology*, vol. 15, 2013.

[5] A. F. Tamimi and M. Juweid, "Epidemiology and outcome of glioblastoma," in *Glioblastoma*, S. De Vleeschouwer, Ed., Codon Publications, Brisbane, Australia, 2017.

[6] M. Ceccarelli, F. P. Barthel, T. M. Malta et al., "Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma," *Cell*, vol. 164, no. 3, pp. 550–563, 2016.

[7] N. Soni, S. Priya, and G. Bathla, "Texture analysis in cerebral gliomas: a review of the literature," *American Journal of Neuroradiology*, vol. 40, no. 6, pp. 928–934, 2019.

[8] J. Liang, D. Liu, P. Gao et al., "Diagnostic values of DCE-MRI and DSC-MRI for differentiation between high-grade and

low-grade gliomas: a comprehensive meta-analysis," *Academic Radiology*, vol. 25, no. 3, pp. 338–348, 2018.

[9] P. Svolos, E. Kousi, E. Kapsalaki et al., "The role of diffusion and perfusion weighted imaging in the differential diagnosis of cerebral tumors: a review and future perspectives," *Cancer Imaging*, vol. 14, p. 20, 2014.

[10] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: images are more than pictures, they are data," *Radiology*, vol. 278, no. 2, pp. 563–577, 2016.

[11] A. Chaddad, M. J. Kucharczyk, P. Daniel et al., "Radiomics in glioblastoma: current status and challenges facing clinical implementation," *Frontiers in Oncology*, vol. 9, p. 374, 2019.

[12] Y.-b. Xi, F. Guo, Z.-l. Xu et al., "Radiomics signature: a potential biomarker for the prediction of MGMT promoter methylation in glioblastoma," *Journal of Magnetic Resonance Imaging*, vol. 47, no. 5, pp. 1380–1387, 2018.

[13] B. Zhang, K. Chang, S. Ramkissoon et al., "Multimodal MRI features predict isocitrate dehydrogenase genotype in high-grade gliomas," *Neuro-Oncology*, vol. 19, no. 1, pp. 109–117, 2017.

[14] S.-S. Zhao, X.-L. Feng, Y.-C. Hu et al., "Better efficacy in differentiating WHO grade II from III oligodendrogliomas with machine-learning than radiologist's reading from conventional T1 contrast-enhanced and fluid attenuated inversion recovery images," *BMC Neurology*, vol. 20, no. 1, p. 48, 2020.

[15] C. J. Belden, P. A. Valdes, C. Ran et al., "Genetics of glioblastoma: a window into its imaging and histopathologic variability," *Radiographics*, vol. 31, no. 6, pp. 1717–1740, 2011.

[16] J. Wang and M. Leeflang, "Recommended software/packages for meta-analysis of diagnostic accuracy," *Journal of Laboratory and Precision Medicine*, vol. 4, p. 22, 2019.

[17] A. Liberati, D. G. Altman, J. Tetzlaff et al., "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration," *Journal of Clinical Epidemiology*, vol. 62, no. 10, pp. e1–e34, 2009.

[18] P. F. Whiting, A. W. Rutjes, M. E. Westwood et al., "QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies," *Annals of Internal Medicine*, vol. 155, no. 8, pp. 529–536, 2011.

[19] H.-h. Cho, S.-h. Lee, J. Kim, and H. Park, "Classification of the glioma grading using radiomics analysis," *PeerJ*, vol. 6, p. e5982, 2018.

[20] T. Hashido, S. Saito, and T. Ishida, "A radiomics-based comparative study on arterial spin labeling and dynamic susceptibility contrast perfusion-weighted imaging in gliomas," *Scientific Reports*, vol. 10, no. 1, p. 6121, 2020.

[21] Q. Tian, L.-F. Yan, X. Zhang et al., "Radiomics strategy for glioma grading using texture features from multiparametric MRI," *Journal of Magnetic Resonance Imaging*, vol. 48, no. 6, pp. 1518–1528, 2018.

[22] A. Vamvakas, S. C. Williams, K. Theodorou et al., "Imaging biomarker analysis of advanced multiparametric MRI for glioma grading," *Physica Medica*, vol. 60, pp. 188–198, 2019.

[23] J. P. T. Higgins and S. Green, *Cochrane Handbook for Systematic Reviews of Interventions*, Wiley, Hoboken, NJ, USA, 2019.

[24] V. Kumar, Y. Gu, S. Basu et al., "Radiomics: the process and the challenges," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1234–1248, 2012.

[25] D. Padmanabhan, S. Bhat, S. Shevade, and Y. Narahari, "Topic model based multi-label classification from the crowd," in *Proceedings of the IEEE 28th International Conference on Tools with Artificial Intelligence*, San Jose, CA, USA, 2016.

[26] R. Longadge, S. S. Dongre, and L. Malik, "Class imbalance problem in data mining: review," 2020, https://arxiv.org/ftp/arxiv/papers/1305/1305.1707.pdf.

[27] M. M. Rahman and D. N. Davis, "Addressing the class imbalance problem in medical datasets," *International Journal of Machine Learning and Computing*, vol. 3, no. 2, p. 224, 2013.

[28] Y. Yang, L.-F. Yan, X. Zhang et al., "Optimizing texture retrieving model for multimodal MR image-based support vector machine for classifying glioma," *Journal of Magnetic Resonance Imaging*, vol. 49, no. 5, pp. 1263–1274, 2019.

[29] M. A. Naveed, P. Goyal, A. Malhotra et al., "Grading of oligodendroglial tumors of the brain with apparent diffusion coefficient, magnetic resonance spectroscopy, and dynamic susceptibility contrast imaging," *The Neuroradiology Journal*, vol. 31, no. 4, pp. 379–385, 2018.

[30] J. H. Yoon, J.-h. Kim, W. J. Kang et al., "Grading of cerebral glioma with multiparametric MR imaging and 18F-FDG-PET: concordance and accuracy," *European Radiology*, vol. 24, no. 2, pp. 380–389, 2014.

[31] J. A. Guzmán-De-Villoria, J. M. Mateos-Pérez, P. Fernández-García, E. Castro, and M. Desco, "Added value of advanced over conventional magnetic resonance imaging in grading gliomas and other primary brain tumors," *Cancer Imaging*, vol. 14, no. 1, p. 35, 2014.

[32] Y. Lin, J. Li, Z. Zhang et al., "Comparison of intravoxel incoherent motiondiffusion-weighted MR imaging and arterial spin labeling MR imaging in gliomas," *Biomed Research International*, vol. 2015, Article ID 234245, 10 pages, 2015.

[33] Y. C. U, L. F. Yan, L. Wu et al., "Intravoxel incoherent motion diffusion-weighted MR imaging of gliomas: efficacy in preoperative grading," *Scientific Reports*, vol. 4, p. 7208, 2014.

[34] B. Koçak, E. Ş Durmaz, E. S. Durmaz, E. Ates, and O. Kilickesmez, "Radiomics with artificial intelligence: a practical guide for beginners," *Diagnostic and Interventional Radiology*, vol. 25, no. 6, pp. 485–495, 2019.

[35] J. Hua, Z. Xiong, J. Lowey, E. Suh, and E. R. Dougherty, "Optimal number of features as a function of sample size for various classification rules," *Bioinformatics*, vol. 21, no. 8, pp. 1509–1515, 2005.

[36] J. Berglund, J. P. A. Ioannidis, N. Terrin, C. H. Schmid, and I. Olkin, "The case of the misleading funnel plot," *BMJ*, vol. 333, no. 7568, pp. 597–600, 2006.

[37] A. Sotiriadis, S. I. Papatheodorou, and W. P. Martins, "Synthesizing evidence from diagnostic accuracy TEsts: the SEDATE guideline," *Ultrasound in Obstetrics & Gynecology*, vol. 47, no. 3, pp. 386–395, 2016.

[38] N. Terrin, C. H. Schmid, J. Lau, and I. Olkin, "Adjusting for publication bias in the presence of heterogeneity," *Statistics in Medicine*, vol. 22, no. 13, pp. 2113–2126, 2003.

[39] J. E. Park, H. S. Kim, D. Kim et al., "A systematic review reporting quality of radiomics research in neuro-oncology: toward clinical utility and quality improvement using high-dimensional imaging features," *BMC Cancer*, vol. 20, no. 1, p. 29, 2020.

[40] K. Buch, H. Kuno, M. M. Qureshi, B. Li, and O. Sakai, "Quantitative variations in texture analysis features dependent on MRI scanning parameters: a phantom model," *Journal of Applied Clinical Medical Physics*, vol. 19, no. 6, pp. 253–264, 2018.

[41] A. Traverso, L. Wee, A. Dekker, and R. Gillies, "Repeatability and reproducibility of radiomic features: a systematic review," *International Journal of Radiation Oncology, Biology, Physics*, vol. 102, no. 4, pp. 1143–1158, 2018.

[42] J. S. Lin, D. T. Fuentes, A. Chandler et al., "Performance assessment for brain MR imaging registration methods," *American Journal of Neuroradiology*, vol. 38, no. 5, pp. 973–980, 2017.

[43] S. Rizzo, F. Botta, S. Raimondi et al., "Radiomics: the facts and the challenges of image analysis," *European Radiology Experimental*, vol. 2, no. 1, p. 36, 2018.

[44] X. Zhang, L.-F. Yan, Y.-C. Hu et al., "Optimizing a machine learning based glioma grading system using multi-parametric MRI histogram and texture features," *Oncotarget*, vol. 8, no. 29, pp. 47816–47830, 2017.

Hindawi
Contrast Media & Molecular Imaging
Volume 2020, Article ID 8894213, 12 pages
https://doi.org/10.1155/2020/8894213

Review Article

Diagnostic Value of Structural and Functional Neuroimaging in Autoimmune Epilepsy

Limei Luo,[1] Nian Wei,[1] Jing Wang,[2] Yuemei Luo,[1] Fei Yang,[1] and Zucai Xu [1,3]

[1]*Department of Neurology, Affiliated Hospital of Zunyi Medical University, Zunyi, China*
[2]*Prevention and Health Care, Affiliated Hospital of Zunyi Medical University, Zunyi, China*
[3]*Key Laboratory of Brain Science, Zunyi Medical University, Zunyi, China0712*

Correspondence should be addressed to Zucai Xu; docxzc@126.com

Received 14 September 2020; Revised 14 November 2020; Accepted 2 December 2020; Published 14 December 2020

Academic Editor: Wei Li

Epilepsy is a common nervous system disease, which affects about 70 million people all over the world. In 2017, the International League Against Epilepsy (ILAE) considered immune factors as its independent cause, and the concept of autoimmune epilepsy (AE) was widely accepted. Early diagnosis and timely treatment can effectively improve the prognosis of the disease. However, due to the diversity of clinical manifestations, the expensive cost of autoantibody detection, and the increased prevalence in Western China, the difficulty for clinicians in early diagnosis and treatment has increased. Fortunately, convenient and fast imaging examinations are expected to help even more. The imaging manifestations of AE patients were characteristic, especially the combined application of structural and functional neuroimaging, which improved the diagnostic value of imaging. In this paper, several common autoantibodies associated with AE and their structure and function changes in neuroimaging were reviewed to provide help for neurologists to achieve the goal of precision medicine.

1. Introduction

Precision medicine is increasingly important in modern clinical medicine, as it aims to obtain an early and accurate diagnosis and reduce the subsequent treatment failure and intervention in disease development, which usually involves a highly individualized patient management and multidisciplinary cooperation [1–3]. For clinicians, the goal of achieving precision medicine for autoimmune epilepsy is fraught with challenges.

Epilepsy is a chronic neurological disorder characterized by recurrent abnormal discharge of neurons. Its etiology is complex, and a lot of studies showed that autoimmune factors may participate in its occurrence and development [4–6]. In 2017, ILAE regards immune factors as one of its six independent causes, and more attention has been paid to the research progress of AE [7, 8]. The understanding of AE can be divided into two types: one covers all epilepsy related to systemic autoimmune diseases and the other mainly includes epilepsy related to nervous system autoantibodies [9, 10]. Here, we mainly focus on the latter. Early and accurate diagnosis of AE is important because affected patients have seizures that are resistant to common antiepileptic therapy but usually respond to immunotherapy [6, 11, 12]. Antibody testing has always been essential for the diagnosis and evaluation of autoimmune diseases. There are still some situations in the laboratory examination of AE, such as sensitivity/specificity of antibody testing, inconsistent antibody titer between serum and cerebrospinal fluid (CSF), and lack of the rare antibody test and low popularity [13–15]. It should be noted that there was no difference in prevalence or incidence between autoimmune and infectious factors for inflammatory lesions of the central nervous system, and more than 50% of patients do not have specific autoantibodies [16, 17]. In patients with presumptive autoimmune encephalitis, there was no significant difference in the clinical manifestations of antibody-negative cases and confirmed cases, and the therapeutic response to immunotherapy was similar [18]. Therefore, the better application of convenient and fast imaging examination cannot be postponed.

In recent years, many studies have found that there is a unique value in the diagnosis and prognosis evaluation of structural and functional neuroimaging features in patients with AE. These imaging technologies, including magnetic resonance imaging (MRI), functional MRI (fMRI), positron emission tomography (PET), and single-photon emission computed tomography (SPECT), have opened a new way for the diagnosis and treatment of diseases [19–22]. Although the imaging changes of AE are relatively new in the field of radiology, the aim of this paper is to review the major antibody subtypes and imaging changes of AE, provide a framework for radiologists to understand the relevant neuroimmunology, and help clinicians to identify the causes of epilepsy for early and precise treatment (Table 1).

2. Neuroimaging Techniques

The etiologies of epilepsy are varied and multifactorial in most cases. Autoantibody testing takes a long time and has not been carried out in some hospitals, but immunotherapy usually must be based on clinical presentation and quasi-clinical results obtained during that time. And in these circumstances, the study of biomarkers may be further helpful for the early and accurate diagnosis of AE. In addition to CSF analysis, electroencephalogram, and physical examination, neuroimaging techniques are also included. Because of the special anatomy of the brain, here we focus on MRI, PET, and SPECT.

MRI is the most commonly used neuroimaging test for lesions of the brain parenchyma. The magnetic resonance signal is generated by the radiofrequency pulsations, and the selected pulse sequence will determine the appearance of the image. T1-weighted images have an advantage in the presentation of anatomic detail, but T2-weighted images are often needed to demonstrate pathology [23]. The purpose of MRI in epileptic patients includes etiology/differential diagnosis, follow-up observation, and preoperative evaluation [20]. Not only can T1 and T2-weighted be contrasted in conventional MRI images, fMRI can also improve the detection of pathological conditions [24, 25]. The fMRI is increasingly used to evaluate the relationship between brain activation and sensory/motor and cognitive activities, and its application in AE has also been reported [26–28]. The fMRI usually uses blood oxygen level-dependent (BOLD) contrast to locate brain function [29]. In this paper, fMRI not only refers to the application of BOLD technology but also includes magnetic resonance spectroscopy (MRS), diffusion tensor imaging (DTI), and so on.

PET is a nuclear medical imaging technology that can be used to investigate human metabolic processes. The 2-deoxy-2-18F-fluoro-*D*-glucose positron emission tomography/computer tomography (18F-FDG PET/CT) visualizes regional neuronal activity by measuring cerebral glucose, and 18F-FDG activity was projected to predefined surface pixels after stereotactic anatomic standardization [30]. In the field of inflammation and tumor, the use of FDG-PET imaging has been widely reported [31, 32]. In recent years, its application in AE is also on the rise.

SPECT is a high-resolution noninvasive imaging mode, which belongs to the category of functional imaging. The imaging principle is to image the gamma ray emitted from the patient's body and realize the imaging of body function and metabolism with the aid of single photon nuclide-labeled drugs [33, 34]. SPECT is not only widely used in the diagnosis and follow-up of cardiovascular diseases, tumors, and kidney diseases but also in epilepsy [35–38]. It can locate the active epileptic brain tissue according to the change of local cerebral blood flow and provide important basis in the preoperative evaluation of drug-resistant epilepsy [39, 40].

3. Structural and Functional Neuroimaging Imaging Features of AE

In the central nervous system (CNS), neural antigen-specific autoimmune diseases characterized by seizures and other symptoms have been identified. According to the existing reports, some patients with the CNS autoimmunity showed focal seizures alone or seizures as the most prominent clinical manifestation [41]. There is a tendency to classify autoimmune antibodies associated with central nervous system diseases into three categories: anti-NMDAR antibodies, limbic encephalitis- (LE-) related antibodies, and other antibodies [42]. They are organized as a review for those autoantibodies associated with AE and a description of the reported structural and functional imaging findings.

3.1. NMDAR Antibody-Related AE. The N-methyl-*D*-aspartate receptor (NMDAR) is a type of ligand-gated ion channel that mediates a major component of excitatory neurotransmission in the CNS, widely present in the brain. The anti-NMDAR encephalitis is the most common form of autoimmune antibody-mediated encephalitis, and the antibody caused a titer-dependent and reversible reduction of synaptic NMDAR through an approach of internalisation and crosslinking [43]. The specific binding of CSF antibodies to their cognate receptor leads to the functional decline and reversible reduction of NMDAR synaptic localization and surface density [44, 45]. For people with anti-NMDAR encephalitis, women are more likely to develop psychiatric disorders at the beginning, while male patients are more likely to have seizures initially [46, 47]. There are many types of epileptic seizures in NMDAR antibodies related to AE, including complex partial seizures, generalized tonic-clonic seizure, epilepticus state, and persistent intractable epilepsy, and some patients may have two or more types of seizures in the course of the disease [47, 48].

Although anti-NMDAR encephalitis has various symptoms and frequency of seizures, the MRI is usually normal in most cases. Previous studies have shown that the proportion of MRI abnormalities is less than 50%, with most of these abnormalities presenting T2 and fluid-attenuated inversion recovery (FLAIR) hyperintensity [49]. Lesions can be widely found in the cortex and subcortical white matter area; the most common is the temporal lobe, especially the medial temporal lobe (MTL), followed by the frontal lobe, and others include the thalamus, basal ganglia, cerebellum,

Table 1: Types and imaging characteristics of antibodies related to AE.

Antibody types	MRI		PET	SPECT
	Regular MRI	fMRI		
NMDAR antibody	T2/FLAIR hyperintensity in the cortex and subcortical white matter areas, including temporal lobe, cerebellum, thalamus, basal ganglia, etc.	Bilateral functional connectivity of hippocampus decreased. DTI revealed widespread changes in white matter. The decrease of NAA is related to clinical improvement	A high to low metabolic gradient from the frontal lobe to the occipital lobe	Hyperperfusion in basal ganglia and cortex, especially frontal cortex
Limbic encephalitis-related antibodies	T2/FLAIR hyperintensity in MTL. MTL and hippocampal volume from swelling to atrophy	Extensive damage to brain network connections. MRS showed that NAA decreased and lactate peak increased	MTL hypermetabolisma is the most common manifestation	Hypoperfusio-n in the frontal lobe, parietal lobe, thalamus, and cerebellum
GABAAR antibody	Multifocal cortical-subcortical T2/FLAIR abnormalities, predominantly involved temporal and frontal lobes but also basal ganglia and other regions	MRS showed elevated lactate signals and Lac/creatine ratio in the voxel of interest	—	—
CASPR2 antibody	T2/FLAIR hyperintensity in MTL and diffuse meningeal enhancement. Bilateral hippocampal and generalized cortical atrophy	—	Temporal hypermetabolism, temporomandibular, frontal and diffuse hypometabolism	—
GAD antibody	Acute/subacute lesions usually presented as temporal lobe encephalitis with high T2/FLAIR signal and swelling of unilateral or bilateral medial temporal structures. Hippocampal atrophy is associated with drug-resistant temporal lobe epilepsy	DTI showed wide range of effects in various regions of brain	Multiple hypermetabolism in brain tissue, mainly in the frontal or temporal lobes	—
Anti-Hu antibody	The most common abnormality on MRI was T2/FLAIR hyperintensity in the temporal lobe and showed multifocal subcortical/subcortical lesions in patients with SCLC	—	High metabolism in one or two temporal lobes, only a small number of brain MRI cases are related to PET	SPECT scan revealed asymmetric cortical activity, but distinct seizure focus could not be identified

and brainstem [50–53]. The only prognostic MRI finding in this type of encephalitis is progressive cerebellar atrophy. A long-term follow-up study found that some patients developed reversible diffuse cerebral atrophy and progressive cerebellar atrophy which is irreversible and is closely associated with a poor long-term prognosis [54]. Recurrence of encephalitis can be manifested as isolated atypical symptoms, suggesting involvements of the brainstem and cerebellum, but recurrence is not associated with abnormal MRI manifestations [55].

Some scholars confirmed that the functional connectivity of bilateral hippocampus decreased in resting state fMRI, and DTI revealed widespread changes in the white matter, especially in cingulate gyrus, which was related to the severity of the disease [28, 56]. In the brain functional activity analysis of 17 patients with anti-NMDAR encephalitis, the decrease of amplitude of low-frequency fluctuation values in the left precuneus, bilateral posterior cingulate gyrus, and cerebellum could be observed [57]. Leptomeningeal contrast enhancement was also observed in some patients [49]. A neurometabolic study showed that in a 31-year-old woman, MRS revealed a hypoglutamatergic state in the left prefrontal cortex, and the increased *N*-acetylaspartate (NAA) concentration was detected only in the left hemisphere with low metabolism [58]. The decrease of NAA concentration in the basal ganglia and thalamus was also observed on MRS, and the NAA signal returned to normal after the clinical symptoms subsided [22]. In a male patient who was admitted to hospital with epilepsy and was finally diagnosed with anti-NMDAR encephalitis, MRS indicated a decrease in NAA, and SPECT showed

hyperperfusion in the right temporooccipital territory [59]. These findings are progressive reversibility with clinical improvement [60]. The SPECT in patient with anti-NMDAR encephalitis also indicated hyperperfusion in the basal ganglia and cortex, especially in the frontal cortex [56].

The correlation between FDG-PET findings and epileptic activity is always direct. Compared with MRI with poor sensitivity, FDG-PET showed more evidence in detecting the progressive stages of anti-NMDAR encephalitis [61]. The metabolic changes on FDG-PET vary widely and involve all the cerebral lobes, including the temporal and occipital lobes, insular cortex, basal ganglia, hippocampi, striatum, caudate nuclei, cerebellum, and brainstem [51, 62–64]. The FDG-PET images of anti-NMDAR encephalitis-associated epilepsy showed a pattern of decreased metabolism from the front to the back, that is, high metabolism in the frontal lobe, temporal lobe, and basal ganglia and low metabolism in the parietal occipital lobe, and the metabolic pattern could change with disease progression, treatment, and follow-up [63]. During the acute and subacute phases, antibody levels were high in all patients, and FDG-PET indicated severe hypermetabolism in the frontal, temporal cortex, and basal ganglia and hypometabolism in bilateral occipital lobes; in the early stage of recovery, diffuse cortical metabolism was the main feature, and the antibody levels of these patients were weak and positive at the same time; during the recovery period, abnormal metabolism and antibody levels returned to normal [61, 63, 65]. The comparison of FDG uptake in patients and healthy probands showed the cortical anteroposterior gradient and increased uptake in the striatum [64]. During the treatment, the deterioration of brain metabolism occurred when the clinical condition deteriorated, which was accompanied by severe extensive cortical hypometabolism and basal ganglia hypermetabolism [66]. Focal hypermetabolism of the left temporal lobe can be observed in the context of decreased diffuse cortical uptake, as the patient was in a long-term epileptic state throughout the course of the disease [61]. It is important to note that the manifestations of FDG-PET could be almost normal in patients without obvious clinical abnormalities and negative antibody recovery after treatment, but it may become abnormal again with the recurrence of the patient's condition; these abnormal manifestations include previous or new ones, and dynamic monitoring of FDG-PET showed a parallel relationship between cerebral glucose metabolism and clinical improvement [58, 61, 66].

3.2. Limbic Encephalitis-Related Antibodies. Autoimmune limbic encephalitis is an inflammatory disease of the central nervous system that mainly involves the MTL [67]. Although the types and techniques of antibody testing have improved, there are still a number of patients who are antibody-negative and can get a certain degree of clinical improvement by immunotherapy [68]. In the absence of antibody test results or negative antibody detection, LE can be diagnosed by clinical symptoms and abnormal T2/FLAIR high signal intensity of bilateral brain parenchyma highly limited to MTL on MRI [69]. The related autoimmune antibodies mainly include AMPAR, LGI1, and GABABR [42].

3.2.1. AMPAR Antibody-Related AE. The α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor (AMPAR) is a type of excitatory ionotropic glutamatergic receptor, which participated in majority of rapid excitatory synaptic transmission activities in the brain. Anti-AMPAR encephalitis patients contain antibodies against GluR1/2, which changes synaptic localization and number of AMPARs [70]. Anti-AMPAR selectively eliminated the surface and synaptic AMPARs and caused the amplitude and frequency of the microexcitatory postsynaptic current in neurons induced, resulting in the steady-state decrease of inhibitory synaptic transmission and the enhancement of intrinsic excitability, which may be an important cause of memory impairment and epilepsy [71, 72].

AMPAR antibody-associated encephalitis is relatively rare in AE and is often associated with MTL abnormals on MRI [73, 74]. In a 66-year-old woman with anti-AMPAR encephalitis and diabetes mellitus, MRI showed only a few high-intensity punctate lesions in the white matter, but FDG-PET revealed a wide range of low metabolism areas including the frontal, parietal, and temporal parahippocampal areas [65]. Anti-AMPAR encephalitis with generalized seizures was closely associated with sustained hypermetabolism of left hippocampal FDG in FDG-PET [75]. In a pregnant woman with anti-AMPAR encephalitis, the initial brain MRI showed bilateral marginal encephalitis, but with clinical progression, rapid brain atrophy appeared on MRI, and extensive cortical cortex, caudate metabolism, and brain stem perfusion were observed on FDG-PET [74]. Analysis of FDG-PET was performed in 2 patients with anti-AMPAR encephalitis without seizures. One patient presented with bilateral cerebellar hypermetabolism and the other with total cortical hypometabolism [76]. The metabolism of FDG in the brain is correlated with clinical manifestations, and abnormal metabolism turned to normal after antiepileptic and immune treatment [65, 75]. Due to the low incidence of anti-AMPAR encephalitis, current functional imaging studies are limited to PET imaging, and further SPECT and fMRI studies are needed.

3.2.2. GABABR Antibody-Related AE. Gamma-aminobutyric acid (GABA) is an inhibitory neurotransmitter that exists in the CNS, which can decrease the excitability of neurons and plays a significant role in the regulation of muscle tone. The GABABR is a G protein-coupled receptor composed of two subunits, GABA-B1 and GABA-B2, both of which are essential for the receptor to perform its functions [77]. GABABR antibodies bind to the extracellular domain of the GABA-B1 subunit, which is an inhibitory receptor associated with seizure and memory dysfunction when disrupted [78]. GABABR antibody-associated encephalitis is characterized by epilepsy and is associated with other conditions such as opsoclonus-myoclonus syndrome, ataxia, and small-cell lung cancer [78, 79].

MRI findings in most cases of GABABR encephalities showed T2/FLAIR hyperintensity in medial temporal lobes [73, 78, 80]. Atrophy and hypointensity of the MTL were also found in rare cases [81]. In a prospective study of 15 patients, all patients developed seizures, and in 13 patients, the seizures were the presenting symptom. MRI indicated MTL T2/FLAIR hyperintensity in 10 patients, increased FLAIR signal in the corpus callosum in 1 patient, and normal in 4 patients [78]. MRI findings can reflect the progress of the disease to a certain extent. The involvement of the limbic system in the group with poor prognosis is more extensive than that in the group with good prognosis [80]. MTL hypermetabolism is the most common manifestation in FDG-PET [82]. A 55-year-old male presented with progressive seizures at 3 weeks with a high anti-GABABR antibody titer in CSF. FDG-PET showed significant MTL hypermetabolism and hypometabolism in other parts of the brain, but there were no related abnormal findings on MRI [83]. SPECT revealed that the hypoperfusion areas were consistent with the high expression area of GABABR, and the uptake of the motor area and left temporal lobe was increased, which may be related to convulsive seizures and tongue movement disorder, and all areas showed normal absorption following corticosteroid treatment and neurologic improvement [84]. Anti-GABABR encephalitis in MRS also suggested inflammatory changes, mainly manifested as decreased NAA and elevated lactate peaks [85].

3.2.3. LGI1 Antibody-Related AE. Leucine-rich glioma-inactivated 1 (LGI1) is a secreted neuronal protein that interacts with voltage-gated potassium channels Kv1.1 to perform its functions. The patient's antibodies destroy the LGI1 signal transduction around synapses, leading to neuronal overexcitement and reduced plasticity [86]. Anti-LGI1 encephalitis is the second most common type of autoimmune encephalitis known, which can lead to memory impairment and various forms of seizures, among which faciobrachial dystonic seizure (FBDS) was representative to a certain extent [87, 88]. LGI1 gene mutation is associated with autosomal dominant temporal lobe epilepsy that seizures can be well controlled by antiepileptic treatment [89].

MRI abnormalities in MTL at the early stage of the disease are an important basis for the diagnosis of anti-LGI1 encephalitis [90, 91]. MRI abnormalities in anti-LGI1 encephalitis are most common in MTL and basal ganglia with T2/FLAIR hyperintensity [90, 92–95]. Other manifestations may also involve extratemporal structures which include insula, thalamus, and frontal cortex; however, cortical involvement beyond the limbic region on MRI is rare [90, 94–96]. MRI findings are abnormal in patients with FBDS, usually located in the basal ganglia, in which T1 hyperintensity can be a useful biomarker for FBDS [97, 98]. Radiologic progression was also noted. Most patients showed T2/FLAIR hyperintensities on conventional MRI in the hippocampus during the acute phase of the disease [91]. Changes in MTL and hippocampal volume from swelling to atrophy were observed during the follow-up, and anti-LGI1 encephalitis can be considered as a potential cause of MTL sclerosis [92, 99, 100].

The results of the fMRI study on 27 sufferers with anti-LGI1 encephalitis showed that the disease had extensive damage to brain network connections, including the change of the brain default mode network, and it suggested that the hippocampal damage and the increase of brain default mode network connections might be a compensation mechanism for memory damage [91].

The PET-CT obtained in the acute disease stage often showed FDG hypermetabolism in the affected area. In a study of 18 anti-LGI1 encephalitis patients with seizures, abnormalities were found in 50% of patients and most commonly involved the middle temporal lobe [94]. Even if MRI indicated no structural changes in the brain, abnormal FDG uptake could be seen on PET [101]. The hypermetabolism of bilateral temporal lobes shown in the PET-CT of the studied patients corresponded to the patient's seizure pattern [99]. Hypermetabolism in the striatum and cerebellum was also observed [52, 102]. There was a significant correlation between anti-LGI1 encephalitis with FDBS and basal ganglia [100]. In a study, five out of the eight patients had hypermetabolic abnormalities in basal ganglia [101]. One case of anti LGI1 encephalitis complicated with FDBs showed hyperintense T1 signal in basal ganglia on MRI, while hypermetabolism was found in the same area on PET [98]. However, the location of the MRI results is not always consistent with that of the FDG-PET, and this is a hint that the LGI1 antibodies may affect sugar metabolism and the hippocampus structure through two different steps [94].

3.3. Other Antibodies

3.3.1. GABAAR Antibody-Related AE. GABAARs are a class of ligand-gated ion channels, and its main epitope targets were the $\alpha 1/\beta 3$ subunits of the GABAAR [103, 104]. The antibodies caused a decrease in synaptic GABAAR selectivity, and high antibody titers in CSF and serum are associated with brain parenchymal lesions with seizures and/or intractable status epilepticus [105]. Compared with adults, children were more likely to have generalized seizures in GABAAR antibody-associated encephalitis; this disorder is severe, but most patients respond to treatment [104].

MRI abnormalities in most cases of anti-GABAAR encephalities showed not only multifocal cortical-subcortical T2/FLAIR abnormalities and predominantly involved temporal and frontal lobes but also basal ganglia and other regions [104, 106]. Epileptic persistence accompanied with extensive cortical-subcortical MRI abnormalities and limbic involvement occurs and is often accompanied by stiff-person syndrome [105, 106]. MRI pathologies are associated with disease progression and can be resolved completely after early immunoregulatory therapy [106]. In febrile infection-associated epileptic syndrome caused by GABAAR antibody with refractory status epilepticus, MRI remains consistently negative over the course of the disease, despite the epileptic discharge shown by electroencephalogram [107]. In a 67-year-old woman with anti-GABAaR encephalitis, MRI demonstrated a multifocal cortical-subcortical lesion, and MRS showed elevated lactate signals and Lac/creatine ratio in the voxel of interest [108].

3.3.2. CASPR2 Antibody-Related AE. Contactin-associated protein 2 (CASPR2) is a transmembrane axonal protein localized at the juxtaparanodes of myelinated axons, a specialized region between axons and myelinating glial cells, and contributes to the jump conduction of action potential [109, 110]. CASPR2 autoantibodies are predominantly IgG4, which target multiple epitopes on the extracellular domain of the protein and destroy the combination of CASPR2 with contactin-2; it may interfere with the accumulation of of voltage-gated potassium channel (VGKC) at juxtaparanodes and leading to hyperexcitability of peripheral nerves [110, 111]. CASPR2 antibodies can also bind to hippocampal inhibitory interneurons at the presynaptic level and have a disruptive effect on inhibitory synapses [112]. Seizure is the first symptom, and sometimes, it is the only clinical manifestation in some CASPR2 antibodies positive patients; immunotherapy has a good effect on clinical improvement [113]. Patients can have a variety of seizures, including generalized tonic-clonic, which is the most common form of seizures, and the rest also includes simple partial seizures, complex partial seizures, and even epileptic persistence [113, 114].

MRI is a highly sensitive and specific predictor for CASPR2 encephalitis [115]. MRI abnormalities in anti-CASPR2 encephalitis are most common in MTL with T2/FLAIR hyperintensity [114, 116, 117]. Such abnormal changes in T2 can be restored to normal after immunotherapy [115]. Bilateral hippocampal atrophy was observed in the first MRI analysis in 13.3% of patients [114]. Generalized cortical atrophy and diffuse meningeal enhancement were also found in a small number of patients [113, 117].

In a retrospective study, FDG-PET was performed in 35 patients with CASPR2 autoantibodies-related diseases, of which 85.7% was abnormal, which commonly included temporomandibular hypometabolism (36.7%), frontal hypometabolism (20%), temporal hypermetabolism (16.7%), and diffuse hypometabolism (10%) [118]. FDG-PET showed reduced FDG uptake in one case of anti-CASPR2 encephalitis with generalized seizures, especially in orbitofrontal regions bilaterally, as well as in bilateral anterolateral temporal and left medial temporal regions [119]. A 72-year-old man with positive CASPR2 antibodies presented with hallucinations and seizures, on FDG-PET, revealed hypometabolism in the left temporal and occipital cortex [120]. The imaging findings of FDG-PET in two patients with anti-CASPR2 encephalitis were studied retrospectively: one patient had hypometabolism in association cortices and hypermetabolism of striata and the other showed normal [52]. There are no SPECT and fMRI studies on CASPR2 antibodies associated with AE, and further functional imaging studies are needed.

3.3.3. GAD Antibody-Related AE. Glutamic acid decarboxylase (GAD) is an intracellular enzyme expressed in GABAergic neurons, which catalyzes the transformation of glutamate into GABA [121, 122]. GAD antibodies interfere with the endocytosis of GABAergic neuron vesicles and have been proved to be related to immune response [123, 124]. Acute seizures and chronic epilepsy with temporal lobe onset have been reported in patients with GAD encephalitis [125, 126]. Seizure can be the only clinical symptom of it [127]. A study on the etiologies of temporal lobe epilepsy shows that GAD antibodies were positive in 21.7% of the unknown etiology group, and epilepsy in patients with high antibody titers is often drug-resistant and has been linked to depression, memory disorders, and other autoimmune diseases [128].

The MRI manifestations of GAD antibodies associated to encephalitis involve a wide range, including the thalamus, insulae, parietal lobe, and brain stem, in addition to the most common temporal lobes [127, 129]. Patients in the acute/subacute setting often present with temporal encephalitis evidenced by T2/FLAIR hyperintensity and swelling of temporal structures [127, 130]. Hippocampal atrophy has been found in patients with GAD positive drug-resistant temporal lobe epilepsy [131]. Compared with the patients with positive anti-NMDAR and anti-VGKC antibodies, anti-GAD encephalitis showed higher FLAIR intensity in hippocampus on postprocessed images [130]. A volumetric analysis of serial MRIs indicated that the amygdala volume was increased obviously within the first 12 months after the onset of GAD encephalitis and and tended to be normal during the follow-up period; the increase of hippocampal volume showed no significant difference from the control group [132]. It is worth noting that although early immunotherapy is helpful to avoid brain injury, MRI abnormalities may not be visible in patients with anti-GAD in the early stages of acute immunoactivation [133].

DTI was used to study the changes of white matter in anti-GAD encephalitis, and the results showed that there was a wide range of effects in various regions of the brain, including wide changes of fractional anisotropy and all diffusivity parameters, and lesions with a trend toward a negative correlation of figural memory performance with diffusivity parameters were mainly appeared in the right temporal lobe [134].

The relevant FDG-PET results showed that high metabolism corresponds to early swelling of the lesion parenchyma, and low metabolism corresponds to atrophy at the later stage of the lesion process [127]. In patients with anti-GAD encephalitis, FDG-PET showed multiple hypermetabolism in brain tissue, mainly in the frontal or temporal lobes [127, 135]. When patient presents with cognitive decline, FDG-PET indicated bifrontal hypometabolism and hypoperfusion [136].

3.3.4. Anti-Hu Antibody-Related AE. Anti-Hu antibody is a kind of antinuclear antibody that is related to a variety of tumors, including neuroendocrine tumor of the duodenum, neuroblastoma, and small cell lung cancer (SCLC) [137–139]. The paraneoplastic neurological syndromes associated to anti-Hu are severe and have no effective treatment. Its pathogenicity is believed to be related to nerve cells death and T cell immune response [140, 141]. Epilepsia partialis continua and intractable epilepsy are associated

with Hu-ab [138, 142, 143]. The management of epilepsy was difficult in those epileptics without cancer-received antiepileptic drugs and immunotherapy [144].

When epilepsy occurs in patients with anti-Hu encephalitis, the most common abnormality on MRI was T2/FLAIR hyperintensity in the temporal lobe [19, 138, 142, 144]. In a boy with anti-Hu encephalitis, his MRI had no abnormality at the time of paroxysmal ataxia at first, but T2/FLAIR hyperintensity in the temporal lobe appeared after intractable epilepsy, and this change was consistent with electroencephalogram [142]. In patients' combination with SCLC, MRI showed multifocal subcortical/subcortical lesions with T2/FLAIR hyperintensity without any contrast enhancement in T1-weighted images, and with the development of the disease, brain atrophy and ventricular enlargement may occur [145]. Abnormalities in T2 may represent the sequelae of recurrent seizures, and changes from focal to multifocal may be observed in the course of the disease [146].

In the case of paraneoplastic limbic encephalitis, FDG-PET usually shows high metabolism in one or two temporal lobes, but only a small number of brain MRI cases are related to FDG-PET [147]. FDG-PET is particularly useful for diagnosis, recurrence, and evolution of tumors for anti-Hu-related AE [145, 147]. When patients with anti-Hu paraneoplastic syndrome developed partial status epilepticus, SPECT scan revealed asymmetric cortical activity, but could not identify obvious epileptic foci [148].

There are also some autoantibodies related to AE, such as those related to recombinant dipeptidyl peptidase 6, Ma2, mGluR5, and so on [7]. However, there are few reports or no specific manifestations on neuroimaging; so, we will not list them in this review.

4. Conclusions and Future Perspectives

The current diagnosis of AE relies too much on antibody detection and immunotherapy response. However, many institutions are not easy to carry out antibody detection, and it takes some time to obtain the test results, and it is not easy to obtain the information of immunotherapy response in the early stage. When the status of autoantibodies is not clear, clinical syndrome and imaging findings can determine the diagnosis of probable or definite autoimmune encephalitis [69]. Therefore, the value of convenient imaging in the diagnosis of AE should be paid more attention. MRI and PET are important imaging methods for detecting parenchymal lesions, and they have their own advantages and disadvantages. MRI is more readily available and is essential for preoperative evaluation of epileptic lesion resection [20]. Due to the wide popularity of magnetic resonance, many large sample data can be obtained, and sometimes, PET does not seem particularly important. One study showed that anti-LGI1 LE affects a wide range of brain regions, including the medial temporal lobe and basal ganglia, and these changes can be detected by head MRI without the need for PET/CT [96]. Multiple studies have found that PET is more sensitive than MRI because it can be abnormal in patients with normal MRI, and it is a trend toward that PET could be better used as an early biomarker for AE, so that treatment can start in the early phase [30]. Region-specific changes in brain FDG uptake occurred throughout childhood; so, age-specific adjustments were necessary in the statistical analysis of studies comparing FDG images of children's brains [149]. Considering the characteristics of the two technologies, simultaneous PET/MRI combines metabolic information of PET to localize the abnormality with high-resolution structural and functional information of MRI and holds the dual advantage of providing PET and MRI in single temporal as well as spatial domain [150].

At present, there are still some deficiencies in the research of functional imaging. On the one hand, there is a lack of large-scale prospective research on the causal relationship between brain dysfunction and autoimmune epilepsy. On the other hand, for some antibody types, due to the lack of functional imaging data, AE related to AMPAR, GABAAR, and CASPR antibodies, there is still a lack of SPECT- and fMRI-related research, so specific brain functional imaging changes cannot be obtained. Therefore, further large-scale clinical imaging research is needed in the future.

Precision medicine is extremely important in modern clinical medicine. It is an ideal goal that involves early accurate diagnosis of disease and customizes the optimal treatment plan. Because delayed immunotherapy is associated with poorer prognosis and higher mortality, the diagnosis of AE requires consideration of multiple factors. Antibody status as the only criterion for early diagnosis is clearly unrealistic. Convenient and fast neuroimaging can be used as an essential reference index for the diagnosis of AE. Both structural and functional neuroimaging techniques are particularly important in diagnosing and assessing disease progression. According to the current research, there is a tendency to combine the two to make better clinical decisions.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] K. S. Ramos, "Precision medicine: a wider definition," *Journal of the American Geriatrics Society*, vol. 63, no. 9, pp. 1971-1972, 2015.

[2] D. Aletaha, "Precision medicine and management of rheumatoid arthritis," *Journal of Autoimmunity*, vol. 110, p. 102405, 2020.

[3] S. Lambin, R. T. H. LeijenaarT. M. Deist et al., Radiomics: the bridge between medical imaging and personalized medicine," *Nature Reviews Clinical Oncology*, vol. 14, no. 12, pp. 749–762, 2017.

[4] G. Gozubatik-Celik, C. Ozkara, C. Ulusoy et al., "Anti-neuronal autoantibodies in both drug responsive and resistant focal seizures with unknown cause," *Epilepsy Research*, vol. 135, pp. 131–136, 2017.

[5] J. B. Lilleker, V. Biswas, and R. Mohanraj, "Glutamic acid decarboxylase (GAD) antibodies in epilepsy: diagnostic yield and therapeutic implications," *Seizure*, vol. 23, no. 8, pp. 598–602, 2014.

[6] D. Dubey, A. Alqallaf, R. Hays et al., "Neurological autoantibody prevalence in epilepsy of unknown etiology," *JAMA Neurology*, vol. 74, no. 4, pp. 397–402, 2017.

[7] C. Geis, J. Planagumà, M. Carreño, F. Graus, and J. Dalmau, "Autoimmune seizures and epilepsy," *Journal of Clinical Investigation*, vol. 129, no. 3, pp. 926–940, 2019.

[8] I. E. Scheffer, S. Berkovic, G. Capovilla et al., "ILAE classification of the epilepsies: position paper of the ILAE commission for classification and terminology," *Epilepsia*, vol. 58, no. 4, pp. 512–521, 2017.

[9] M.-S. Ong, I. S. Kohane, T. Cai, M. P. Gorman, and K. D. Mandl, "Population-level evidence for an autoimmune etiology of epilepsy," *JAMA Neurology*, vol. 71, no. 5, pp. 569–574, 2014.

[10] J. Britton, "Autoimmune epilepsy," *Handbook of Clinical Neurology*, vol. 133, pp. 219–245, 2016.

[11] M. Toledano, J. W. Britton, A. McKeon et al., "Utility of an immunotherapy trial in evaluating patients with presumed autoimmune epilepsy," *Neurology*, vol. 82, no. 18, pp. 1578–1586, 2014.

[12] A. M. L. Quek and O. O'Toole, "Autoimmune epilepsy: the evolving science of neural autoimmunity and its impact on epilepsy management," *Seminars in Neurology*, vol. 38, no. 3, pp. 290–302, 2018.

[13] S. Wu, H. Li, Y. Lian et al., "Anti-N-methyl-D-aspartate receptor encephalitis: a prospective study focused on cerebrospinal fluid and clinical symptoms," *Neurological Sciences*, vol. 41, no. 11, pp. 3255–3263, 2020.

[14] Y. Gu, M. Zhong, L. He et al., "Epidemiology of antibody-positive autoimmune encephalitis in southwest China: a multicenter study," *Frontiers in Immunology*, vol. 10, p. 2611, 2019.

[15] H.-Z. Guan, H.-T. Ren, and L.-Y. Cui, "Autoimmune encephalitis," *Chinese Medical Journal*, vol. 129, no. 9, pp. 1122–1127, 2016.

[16] D. Dubey, S. J. Pittock, C. R. Kelly et al., "Autoimmune encephalitis epidemiology and a comparison to infectious encephalitis," *Annals of Neurology*, vol. 83, no. 1, pp. 166–177, 2018.

[17] J. Granerod, H. E. Ambrose, N. W. Davies et al., "Causes of encephalitis and differences in their clinical presentations in England: a multicentre, population-based prospective study," *The Lancet Infectious Diseases*, vol. 10, no. 12, pp. 835–844, 2010.

[18] S. Pradhan, A. Das, A. Das, and M. Mulmuley, "Antibody negative autoimmune encephalitis- does it differ from definite one," *Annals of Indian Academy of Neurology*, vol. 22, no. 4, pp. 401–408, 2019.

[19] J. Aupy, N. Collongues, F. Blanc, C. Tranchant, E. Hirsch, and J. De Seze, "Encéphalites dysimmunitaires, données cliniques, radiologiques et immunologiques," *Revue Neurologique*, vol. 169, no. 2, pp. 142–153, 2013.

[20] T. Rüber, B. David, and C. E. Elger, "MRI in epilepsy: clinical standard and evolution," *Current Opinion in Neurology*, vol. 31, no. 2, pp. 223–231, 2018.

[21] F. Graus and J. Dalmau, "Role of (18)F-FDG-PET imaging in the diagnosis of autoimmune encephalitis—authors' reply," *The Lancet Neurology*, vol. 15, no. 10, p. 1010, 2016.

[22] H. Kataoka, J. Dalmau, T. Taoka, and S. Ueno, "Reduced N-acetylaspartate in the basal ganglia of a patient with anti-NMDA receptor encephalitis," *Movement Disorders*, vol. 24, no. 5, pp. 784–786, 2009.

[23] L. L. Seeger, "Physical principles of magnetic resonance imaging," *Clinical Orthopaedics & Related Research*, vol. 244, pp. 7–16, 1989.

[24] R. Bammer, S. Skare, R. Newbould et al., "Foundations of advanced magnetic resonance imaging," *NeuroRx*, vol. 2, no. 2, pp. 167–196, 2005.

[25] S. J. Holdsworth and R. Bammer, "Magnetic resonance imaging techniques: fMRI, DWI, and PWI," *Seminars in Neurology*, vol. 28, no. 4, pp. 395–406, 2008.

[26] K. R. Thulborn, "MRI in the management of cerebrovascular disease to prevent stroke," *Neurologic Clinics*, vol. 26, no. 4, pp. 897–921, 2008.

[27] I. Becerra-Laparra, D. Cortez-Conradis, H. G. Garcia-Lazaro, M. Martinez-Lopez, and E. Roldan-Valadez, "Radial diffusivity is the best global biomarker able to discriminate healthy elders, mild cognitive impairment, and Alzheimer's disease: a diagnostic study of DTI-derived data," *Neurology India*, vol. 68, no. 2, pp. 427–434, 2020.

[28] C. Finke, U. A. Kopp, M. Scheel et al., "Functional and structural brain changes in anti-N-methyl-D-aspartate receptor encephalitis," *Annals of Neurology*, vol. 74, no. 2, pp. 284–296, 2013.

[29] Huettel, A. Scott, A. W. Song, and G. Mccarthy, *Functional Magnetic Resonance Imaging*, Sinauer Associates, Sunderland, MA, USA, 2nd edition, 2009.

[30] L. B. Solnes, K. M. Jones, S. P. Rowe et al., "Diagnostic value of 18F-FDG PET/CT versus MRI in the setting of antibody-specific autoimmune encephalitis," *Journal of Nuclear Medicine*, vol. 58, no. 8, pp. 1307–1313, 2017.

[31] F. Jamar, J. Buscombe, A. Chiti et al., "EANM/SNMMI guideline for 18F-FDG use in inflammation and infection," *Journal of Nuclear Medicine*, vol. 54, no. 4, pp. 647–658, 2013.

[32] T. Mochizuki, E. Tsukamoto, Y. Kuge et al., "FDG uptake and glucose transporter subtype expressions in experimental tumor and inflammation models," *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine*, vol. 42, no. 10, pp. 1551–1555, 2001.

[33] D. J. Brooks, "Positron emission tomography and single-photon emission computed tomography in central nervous system drug development," *NeuroRx*, vol. 2, no. 2, pp. 226–236, 2005.

[34] R. J. Jaszczak, R. E. Coleman, and C. B. Lim, "SPECT: single photon emission computed tomography," *IEEE Transactions on Nuclear Science*, vol. 27, no. 3, pp. 1137–1153, 1980.

[35] H. S. Sachdev, B. Patel, M. McManis, M. Lee, and D. F. Clarke, "Comparing single-photon emission computed tomography (SPECT), electroencephalography (EEG), and magneto-encephalography (MEG) seizure localizations in pediatric cases of laser ablation," *Journal of Child Neurology*, vol. 34, no. 6, pp. 303–308, 2019.

[36] A. D. Cesare, D. V. Giuseppe, G. Stefano et al., "Single-photon-emission computed tomography (SPECT) with technetium-99m sestamibi in the diagnosis of small breast cancer and axillary lymph node involvement," *World Journal of Surgery*, vol. 35, no. 12, pp. 2668–2672, 2011.

[37] B. Sabha, H. Abdul, D. Sunitha et al., "Prognostic value of regadenoson myocardial single-photon emission computed tomography in patients with different degrees of renal dysfunction," *European Heart Journal-Cardiovascular Imaging*, vol. 15, no. 8, pp. 933–940, 2014.

[38] J. P. Greenwood, N. Maredia, J. F. Younger et al., "Cardiovascular magnetic resonance and single-photon emission computed tomography for diagnosis of coronary heart

disease (CE-MARC): a prospective trial," *The Lancet*, vol. 379, no. 9814, pp. 453–460, 2012.

[39] C. Stamoulis, J. Connolly, E. Axeen et al., "Noninvasive seizure localization with single-photon emission computed tomography is impacted by preictal/early ictal network dynamics," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 7, pp. 1863–1871, 2019.

[40] A. Desai, K. Bekelis, V. M. Thadani et al., "Interictal PET and ictal subtraction SPECT: sensitivity in the detection of seizure foci in patients with medically intractable epilepsy," *Epilepsia*, vol. 54, no. 2, pp. 341–350, 2013.

[41] A. M. L. Quek, J. W. Britton, A. McKeon et al., "Autoimmune epilepsy," *Archives of Neurology*, vol. 69, no. 5, pp. 582–593, 2012.

[42] F. Leypoldt, T. Armangue, and J. Dalmau, "Autoimmune encephalopathies," *Annals of the New York Academy of Sciences*, vol. 1338, no. 1, pp. 94–114, 2015.

[43] J. Dalmau, E. Lancaster, E. Martinez-Hernandez, M. R. Rosenfeld, and R. Balice-Gordon, "Clinical experience and laboratory investigations in patients with anti-NMDAR encephalitis," *The Lancet Neurology*, vol. 10, no. 1, pp. 63–74, 2011.

[44] J. Planagumà, F. Leypoldt, F. Mannara et al., "Human N-methyl D-aspartate receptor antibodies alter memory and behaviour in mice," *Brain*, vol. 138, no. 1, pp. 94–109, 2015.

[45] E. G. Hughes, X. Peng, A. J. Gleichman et al., "Cellular and synaptic mechanisms of anti-NMDA receptor encephalitis," *Journal of Neuroscience*, vol. 30, no. 17, pp. 5866–5875, 2010.

[46] A. Viaccoz, V. Desestret, F. Ducray et al., "Clinical specificities of adult male patients with NMDA receptor antibodies encephalitis," *Neurology*, vol. 82, no. 7, pp. 556–563, 2014.

[47] W. Wang, J. M. Li, F. Y. Hu et al., "Anti- NMDA receptor encephalitis: clinical characteristics, predictors of outcome and the knowledge gap in southwest C hina," *European Journal of Neurology*, vol. 23, no. 3, pp. 621–629, 2016.

[48] X.-p. Qu, J. Vidaurre, X.-l. Peng, L. Jiang, M. Zhong, and Y. Hu, "Seizure characteristics, outcome, and risk of epilepsy in pediatric anti-N-Methyl-d-Aspartate receptor encephalitis," *Pediatric Neurology*, vol. 105, pp. 35–40, 2020.

[49] S. Bacchi, K. Franke, D. Wewegama, E. Needham, S. Patel, and D. Menon, "Magnetic resonance imaging and positron emission tomography in anti-NMDA receptor encephalitis: a systematic review," *Journal of Clinical Neuroscience*, vol. 52, pp. 54–59, 2018.

[50] H. Qianyi, X. Yue, H. Zhiping, and T. Xiangqi, "Anti-N-methyl-D-aspartate receptor encephalitis: a review of pathogenic mechanisms, treatment, prognosis," *Brain Research*, vol. 15, no. 1727, Article ID 146549, 2020.

[51] W. O. Tobin, E. A. Strand, H. M. Clark, V. J. Lowe, C. E. Robertson, and S. J. Pittock, "NMDA receptor encephalitis causing reversible caudate changes on MRI and PET imaging," *Neurology: Clinical Practice*, vol. 4, no. 6, pp. 470–473, 2014.

[52] A. Baumgartner, S. Rauer, I. Mader, and P. T. Meyer, "Cerebral FDG-PET and MRI findings in autoimmune limbic encephalitis: correlation with autoantibody types," *Journal of Neurology*, vol. 260, no. 11, pp. 2744–2753, 2013.

[53] J. Zhang, T. Ji, Q. Chen et al., "Pediatric autoimmune encephalitis: case series from two Chinese tertiary pediatric Neurology centers," *Frontiers in Neurology*, vol. 10, p. 906, 2019.

[54] T. Iizuka, J. Kaneko, N. Tominaga et al., "Association of progressive cerebellar atrophy with long-term outcome in patients with anti-N-Methyl-d-Aspartate receptor encephalitis," *JAMA Neurology*, vol. 73, no. 6, pp. 706–713, 2016.

[55] I. Gabilondo, A. Saiz, L. Galán et al., "Analysis of relapses in anti-NMDAR encephalitis," *Neurology*, vol. 77, no. 10, pp. 996–999, 2011.

[56] V. Llorens, I. Gabilondo, J. C. Gómez-Esteban et al., "Abnormal multifocal cerebral blood flow on Tc-99m HMPAO SPECT in a patient with anti-NMDA-receptor encephalitis," *Journal of Neurology*, vol. 257, no. 9, pp. 1568-1569, 2010.

[57] L. Cai, Y. Liang, H. Huang, X. Zhou, and J. Zheng, "Cerebral functional activity and connectivity changes in anti-N-methyl-D-aspartate receptor encephalitis: a resting-state fMRI study," *NeuroImage: Clinical*, vol. 25, p. 102189, 2020.

[58] E. Dominique, E. Perlov, S. Rauer et al., "Hypoglutamatergic state is associated with reduced cerebral glucose metabolism in anti-NMDA receptor encephalitis: a case report," *BMC Psychiatry*, vol. 15, no. 1, 2015.

[59] S. Yamamoto, Y. Koide, M. Fujiwara, K. Nakazawa, Y. Takahashi, and H. Hara, "Subacute encephalitis associated with anti-glutamate receptor antibodies: serial studies of MRI, 1H-MRS and SPECT," *Rinsho Shinkeigaku*, vol. 48, no. 3, pp. 196–201, 2008.

[60] H. Kasahara, M. Sato, S. Nagamine, K. Makioka, K. Tanaka, and Y. Ikeda, "Temporal changes on 123I-iomazenil and cerebral blood flow single-photon emission computed tomography in a patient with anti-N-methyl-D-aspartate receptor encephalitis," *Internal Medicine*, vol. 58, no. 10, pp. 1501–1505, 2019.

[61] J. Yuan, H. Guan, X. Zhou et al., "Changing brain metabolism patterns in patients with ANMDARE," *Clinical Nuclear Medicine*, vol. 41, no. 5, pp. 366–370, 2016.

[62] J. Guerin, R. E. Watson, C. M. Carr, G. B. Liebo, and A. L. Kotsenas, "Autoimmune epilepsy: findings on MRI and FDG-PET," *The British Journal of Radiology*, vol. 92, no. 1093, Article ID 20170869, 2019.

[63] F. Leypoldt, R. Buchert, I. Kleiter et al., "Fluorodeoxyglucose positron emission tomography in anti-N-methyl-D-aspartate receptor encephalitis: distinct pattern of disease," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 83, no. 7, pp. 681–686, 2012.

[64] J. Novy, G. Allenbach, C. G. Bien, E. Guedj, J. O. Prior, and A. O. Rossetti, "FDG-PET hyperactivity pattern in anti-NMDAr encephalitis," *Journal of Neuroimmunology*, vol. 297, pp. 156–158, 2016.

[65] Y. C. Wei, J. R. Tseng, C. L. Wu et al., "Different FDG-PET metabolic patterns of anti-AMPAR and anti-NMDAR encephalitis: case report and literature review," *Brain and Behavior*, vol. 10, no. 3, Article ID e01540, 2020.

[66] S. Lagarde, A. Lepine, E. Caietta et al., "Cerebral 18FluoroDeoxy-glucose positron emission tomography in paediatric anti N-methyl-d-aspartate receptor encephalitis: a case series," *Brain and Development*, vol. 38, no. 5, pp. 461–470, 2016.

[67] A. Budhram, A. Leung, M. W. Nicolle, and J. G. Burneo, "Diagnosing autoimmune limbic encephalitis," *Canadian Medical Association Journal*, vol. 191, no. 19, pp. E529–E534, 2019.

[68] F. Graus, D. Escudero, L. Oleaga et al., "Syndrome and outcome of antibody-negative limbic encephalitis," *European Journal of Neurology*, vol. 25, no. 8, pp. 1011–1016, 2018.

[69] F. Graus, M. J. Titulaer, R. Balu et al., "A clinical approach to diagnosis of autoimmune encephalitis," *The Lancet Neurology*, vol. 15, no. 4, pp. 391–404, 2016.

[70] M. Lai, E. G. Hughes, X. Peng et al., "AMPA receptor antibodies in limbic encephalitis alter synaptic receptor location," *Annals of Neurology*, vol. 65, no. 4, pp. 424–434, 2009.

[71] A. J. Gleichman, J. A. Panzer, B. H. Baumann, J. Dalmau, and D. R. Lynch, "Antigenic and mechanistic characterization of anti-AMPA receptor encephalitis," *Annals of Clinical and Translational Neurology*, vol. 1, no. 3, pp. 180–189, 2014.

[72] X. Peng, E. G. Hughes, E. H. Moscato, T. D. Parsons, J. Dalmau, and R. J. Balice-Gordon, "Cellular plasticity induced by anti-α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor encephalitis antibodies," *Annals of Neurology*, vol. 77, no. 3, pp. 381–398, 2015.

[73] M. Dogan Onugoren, D. Deuretzbacher, C. A. Haensch et al., "Limbic encephalitis due to GABABand AMPA receptor antibodies: a case series," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 86, no. 9, pp. 965–972, 2015.

[74] Y. C. Wei, C. H. Liu, J. J. Lin et al., "Rapid progression and brain atrophy in anti-AMPA receptor encephalitis," *Journal of Neuroimmunology*, vol. 261, no. 1-2, pp. 129–133, 2013.

[75] M. Spatola, V. Stojanova, J. O. Prior, J. Dalmau, and A. O. Rossetti, "Serial brain ^{18}FDG-PET in anti-AMPA receptor limbic encephalitis," *Journal of Neuroimmunology*, vol. 271, no. 1-2, pp. 53–55, 2014.

[76] O. Laurido-Soto, M. R. Brier, L. E. Simon, A. McCullough, R. C. Bucelli, and G. S. Day, "Patient characteristics and outcome associations in AMPA receptor encephalitis," *Journal of Neurology*, vol. 266, no. 2, pp. 450–460, 2019.

[77] A. Pinard, R. Seddik, and B. Bettler, "GABAB receptors: physiological functions and mechanisms of diversity," *GABABReceptor Pharmacology—A Tribute to Norman Bowery*, vol. 58, pp. 231–255, 2010.

[78] E. Lancaster, M. Lai, X. Peng et al., "Antibodies to the GABAB receptor in limbic encephalitis with seizures: case series and characterisation of the antigen," *The Lancet Neurology*, vol. 9, no. 1, pp. 67–76, 2010.

[79] R. Höftberger, M. J. Titulaer, L. Sabater et al., "Encephalitis and GABAB receptor antibodies: novel findings in a new case series of 20 patients," *Neurology*, vol. 81, no. 17, pp. 1500–1506, 2013.

[80] X. Zhang, Y. Lang, L. Sun, W. Zhang, W. Lin, and L. Cui, "Clinical characteristics and prognostic analysis of anti-gamma-aminobutyric acid-B (GABA-B) receptor encephalitis in Northeast China," *BMC Neurology*, vol. 20, no. 1, p. 1, 2020.

[81] M. H. van Coevorden-Hameete, M. A. A. M. de Bruijn, E. de Graaff et al., "The expanded clinical spectrum of anti-GABABR encephalitis and added value of KCTD16 autoantibodies," *Brain*, vol. 142, no. 6, pp. 1631–1643, 2019.

[82] T. J. Kim, S. T. Lee, J. W. Shin et al., "Clinical manifestations and outcomes of the treatment of patients with GABAB encephalitis," *Journal of Neuroimmunology*, vol. 270, no. 1-2, pp. 45–50, 2014.

[83] M. Su, D. Xu, and R. Tian, "18F-FDG PET/CT and MRI findings in a patient with anti-GABAB receptor encephalitis," *Clinical Nuclear Medicine*, vol. 40, no. 6, pp. 515–517, 2015.

[84] K. Ohta, M. Seki, J. Dalmau, and Y. Shinohara, "Perfusion123IMP-SPECT shows reversible abnormalities in GABABreceptor antibody associated encephalitis with normal MRI," *Brain and Behavior*, vol. 1, no. 2, pp. 70–72, 2011.

[85] S. Qiao, Y.-X. Zhang, B.-J. Zhang et al., "Clinical, imaging, and follow-up observations of patients with anti-GABAB receptor encephalitis," *International Journal of Neuroscience*, vol. 127, no. 5, pp. 379–385, 2017.

[86] M. Petit-Pedrol, J. Sell, J. Planagumà et al., "LGI1 antibodies alter Kv1.1 and AMPA receptors changing synaptic excitability, plasticity and memory," *Brain*, vol. 141, no. 11, pp. 3144–3159, 2018.

[87] J. Dalmau, C. Geis, and F. Graus, "Autoantibodies to synaptic receptors and neuronal cell surface proteins in autoimmune diseases of the central nervous system," *Physiological Reviews*, vol. 97, no. 2, pp. 839–887, 2017.

[88] M. Spatola and J. Dalmau, "Seizures and risk of epilepsy in autoimmune and other inflammatory encephalitis," *Current Opinion in Neurology*, vol. 30, no. 3, pp. 345–353, 2017.

[89] C. Di Bonaventura, F. F. Operto, G. Busolin et al., "Low penetrance and effect on protein secretion of LGI1 mutations causing autosomal dominant lateral temporal epilepsy," *Epilepsia*, vol. 52, no. 7, pp. 1258–1264, 2011.

[90] X. Yang, A.-N. Li, X.-H. Zhao, X.-W. Liu, and S.-J. Wang, "Clinical features of patients with anti-leucine-rich glioma inactivated-1 protein associated encephalitis: a Chinese case series," *International Journal of Neuroscience*, vol. 129, no. 8, pp. 754–761, 2019.

[91] J. Heine, H. Prüss, U. A. Kopp et al., "Beyond the limbic system: disruption and functional compensation of large-scale brain networks in patients with anti-LGI1 encephalitis," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 89, no. 11, pp. 1191–1199, 2018.

[92] A. L. Kotsenas, R. E. Watson, S. J. Pittock et al., "MRI findings in autoimmune voltage-gated potassium channel complex encephalitis with seizures: one potential etiology for mesial temporal sclerosis," *American Journal of Neuroradiology*, vol. 35, no. 1, pp. 84–89, 2014.

[93] Y. Li, F. Song, W. Liu, and Y. Wang, "Clinical features of nine cases of leucine-rich glioma inactivated 1 protein antibody-associated encephalitis," *Acta Neurologica Belgica*, 2020.

[94] C. Chen, X. Wang, C. Zhang et al., "Seizure semiology in leucine-rich glioma-inactivated protein 1 antibody-associated limbic encephalitis," *Epilepsy & Behavior*, vol. 77, pp. 90–95, 2017.

[95] Z. Li, T. Cui, W. Shi, and Q. Wang, "Clinical analysis of leucine-rich glioma inactivated-1 protein antibody associated with limbic encephalitis onset with seizures," *Medicine*, vol. 95, no. 28, p. e4244, 2016.

[96] J. Yu, X. Yu, S. Fang, Y. Zhang, and W. Lin, "The treatment and follow-up of anti-LGI1 limbic encephalitis," *European Neurology*, vol. 75, no. 1-2, pp. 5–11, 2016.

[97] P. Flanagan Eoin, L. Kotsenas Amy, W. Britton Jeffrey et al., "Basal ganglia T1 hyperintensity in LGI1-autoantibody faciobrachial dystonic seizures," *Neurology(R) Neuroimmunology & Neuroinflammation.*vol. 2, no. 6, 2015.

[98] A. S. López Chiriboga, J. L. Siegel, W. O. Tatum, J. J. Shih, and E. P. Flanagan, "Striking basal ganglia imaging abnormalities in LGI1 ab faciobrachial dystonic seizures," *Neurology—Neuroimmunology Neuroinflammation*, vol. 4, no. 3, p. e336, 2017.

[99] S. Fauser, J. Talazko, K. Wagner et al., "FDG-PET and MRI in potassium channel antibody-associated non-paraneoplastic limbic encephalitis: correlation with clinical course and neuropsychology," *Acta Neurologica Scandinavica*, vol. 111, no. 5, 2005.

[100] K. K. Kamaleshwaran, R. S. Iyer, J. Joppy Antony, E. K. Radhakrishnan, and A. Shinto, "18F-FDG PET/CT findings in voltage-gated potassium channel limbic

encephalitis," *Clinical Nuclear Medicine*, vol. 38, no. 5, pp. 392–394, 2013.

[101] S. R. Irani, A. W. Michell, B. Lang et al., "Faciobrachial dystonic seizures precede Lgi1 antibody limbic encephalitis," *Annals of Neurology*, vol. 69, no. 5, pp. 892–900, 2011.

[102] P. Moloney, R. Boylan, M. Elamin, S. O'Riordan, R. Killeen, and C. McGuigan, "Semi-quantitative analysis of cerebral FDG-PET reveals striatal hypermetabolism and normal cortical metabolism in a case of VGKCC limbic encephalitis," *The Neuroradiology Journal*, vol. 30, no. 2, pp. 160–163, 2017.

[103] H. C. Chua and M. Chebib, "GABA A receptors and the diversity in their structure and pharmacology," *Advances in Pharmacology*, vol. 79, pp. 1–34, 2017.

[104] M. Spatola, M. Petit-Pedrol, M. M. Simabukuro et al., "Investigations in GABAA receptor antibody-associated encephalitis," *Neurology*, vol. 88, no. 11, pp. 1012–1020, 2017.

[105] M. Petit-Pedrol, T. Armangue, X. Peng et al., "Encephalitis with refractory seizures, status epilepticus, and antibodies to the GABAA receptor: a case series, characterisation of the antigen, and analysis of the effects of antibodies," *The Lancet Neurology*, vol. 13, no. 3, pp. 276–286, 2014.

[106] M. Nikolaus, E. Knierim, C. Meisel et al., "Severe GABA A receptor encephalitis without seizures: a paediatric case successfully treated with early immunomodulation," *European Journal of Paediatric Neurology*, vol. 22, no. 3, pp. 558–562, 2018.

[107] D. Caputo, R. Iorio, F. Vigevano, and L. Fusco, "Febrile infection-related epilepsy syndrome (FIRES) with super-refractory status epilepticus revealing autoimmune encephalitis due to GABA A R antibodies," *European Journal of Paediatric Neurology*, vol. 22, no. 1, pp. 182–185, 2018.

[108] H. Ueno, T. Iizuka, Y. Tagane et al., "Focal hyperperfusion and elevated lactate in the cerebral lesions with anti-GABAaR encephalitis: a serial MRI study," *Journal of Neuroradiology*, vol. 47, no. 3, pp. 243–246, 2020.

[109] S. Poliak, L. Gollan, D. Salomon et al., "Localization of Caspr2 in myelinated nerves depends on axon-glia interactions and the generation of barriers along the axon," *The Journal of Neuroscience*, vol. 21, no. 19, pp. 7568–7575, 2001.

[110] K. R. Patterson, J. Dalmau, and E. Lancaster, "Mechanisms of Caspr2 antibodies in autoimmune encephalitis and neuromyotonia," *Annals of Neurology*, vol. 83, no. 1, 2017.

[111] A. L. Olsen, Y. Lai, J. Dalmau, S. S. Scherer, and E. Lancaster, "Caspr2 autoantibodies target multiple epitopes," *Neurology—Neuroimmunology Neuroinflammation*, vol. 2, no. 4, p. e127, 2015.

[112] P. Delphine, H. Bruno, B. José et al., "Inhibitory axons are targeted in hippocampal cell culture by anti-caspr2 autoantibodies associated with limbic encephalitis," *Frontiers in Cellular Neuroscience*, vol. 9, 2015.

[113] J.-S. Sunwoo, S.-T. Lee, J.-I. Byun et al., "Clinical manifestations of patients with CASPR2 antibodies," *Journal of Neuroimmunology*, vol. 281, pp. 17–22, 2015.

[114] B. Joubert, M. Saint-Martin, N. Noraz et al., "Characterization of a subtype of autoimmune encephalitis with anti-contactin-associated protein-like 2 antibodies in the cerebrospinal fluid, prominent limbic symptoms, and seizures," *JAMA Neurology*, vol. 73, no. 9, pp. 1115–1124, 2016.

[115] C. G. Bien, Z. Mirzadjanova, C. Baumgartner et al., "Anti-contactin-associated protein-2 encephalitis: relevance of antibody titres, presentation and outcome," *European Journal of Neurology*, vol. 24, no. 1, pp. 175–186, 2017.

[116] A. van Sonderen, H. Ariño, M. Petit-Pedrol et al., "The clinical spectrum of Caspr2 antibody-associated disease," *Neurology*, vol. 87, no. 5, pp. 521–528, 2016.

[117] E. Lancaster, M. G. M. Huijbers, V. Bar et al., "Investigations of caspr2, an autoantigen of encephalitis and neuromyotonia," *Annals of Neurology*, vol. 69, no. 2, pp. 303–311, 2011.

[118] M. Boyko, K. L. K. Au, C. Casault, P. de Robles, and G. Pfeffer, "Systematic review of the clinical spectrum of CASPR2 antibody syndrome," *Journal of Neurology*, vol. 267, no. 8, 2020.

[119] A. T. Toosy, S. E. Burbridge, M. Pitkanen et al., "Functional imaging correlates of fronto-temporal dysfunction in Morvan's syndrome," *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 79, no. 6, pp. 734-735, 2008.

[120] Y. Chen, X.-W. Xing, J.-T. Zhang et al., "Autoimmune encephalitis mimicking sporadic Creutzfeldt-Jakob disease: a retrospective study," *Journal of Neuroimmunology*, vol. 295-296, pp. 1–8, 2016.

[121] C. J. Martyniuk, R. Awad, R. Hurley, T. E. Finger, and V. L. Trudeau, "Glutamic acid decarboxylase 65, 67, and GABA-transaminase mRNA expression and total enzyme activity in the goldfish (Carassius auratus) brain," *Brain Research*, vol. 1147, pp. 154–166, 2007.

[122] C. S. Pinal and A. J. Tobin, "Uniqueness and redundancy in GABA production," *Perspectives on Developmental Neurobiology*, vol. 5, no. 2-3, pp. 109–118, 1998.

[123] C. Werner, M. Pauli, S. Doose et al., "Human autoantibodies to amphiphysin induce defective presynaptic vesicle dynamics and composition," *Brain*, vol. 139, no. 2, pp. 365–379, 2016.

[124] C. G. Bien, A. Vincent, M. H. Barnett et al., "Immunopathology of autoantibody-associated encephalitides: clues for pathogenesis," *Brain*, vol. 135, no. 5, pp. 1622–1638, 2012.

[125] M. Esclapez and C. R. Houser, "Up-regulation of GAD65 and GAD67 in remaining hippocampal GABA neurons in a model of temporal lobe epilepsy," *The Journal of Comparative Neurology*, vol. 412, no. 3, pp. 488–505, 1999.

[126] L. Errichiello, S. Striano, F. Zara, and P. Striano, "Temporal lobe epilepsy and anti glutamic acid decarboxylase autoimmunity," *Neurological Sciences*, vol. 32, no. 4, pp. 547–550, 2011.

[127] M. P. Malter, C. Helmstaedter, H. Urbach, A. Vincent, and C. G. Bien, "Antibodies to glutamic acid decarboxylase define a form of limbic encephalitis," *Annals of Neurology*, vol. 67, no. 4, pp. 470–478, 2010.

[128] M. Falip, M. Carreño, J. Miró et al., "Prevalence and immunological spectrum of temporal lobe epilepsy with glutamic acid decarboxylase antibodies," *European Journal of Neurology*, vol. 19, no. 6, pp. 827–833, 2012.

[129] A. Daif, R. V. Lukas, N. P. Issa et al., "Antiglutamic acid decarboxylase 65 (GAD65) antibody-associated epilepsy," *Epilepsy & Behavior*, vol. 80, pp. 331–336, 2018.

[130] J. Wagner, J.-C. Schoene-Bake, M. P. Malter et al., "Quantitative FLAIR analysis indicates predominant affection of the amygdala in antibody-associated limbic encephalitis," *Epilepsia*, vol. 54, no. 9, pp. 1679–1687, 2013.

[131] N. Hansen, L. Ernst, T. Rüber et al., "Pre- and long-term postoperative courses of hippocampus-associated memory impairment in epilepsy patients with antibody-associated limbic encephalitis and selective amygdalohippocampectomy," *Epilepsy & Behavior*, vol. 79, pp. 93–99, 2018.

[132] J. Wagner, J.-A. Witt, C. Helmstaedter, M. P. Malter, B. Weber, and C. E. Elger, "Automated volumetry of the

mesiotemporal structures in antibody-associated limbic encephalitis," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 86, no. 7, pp. 735–742, 2015.

[133] K. M. Mäkelä, A. Hietaharju, A. Brander, and J. Peltola, "Clinical management of epilepsy with glutamic acid decarboxylase antibody positivity: the interplay between immunotherapy and anti-epileptic drugs," *Frontiers in Neurology*, vol. 9, p. 579, 2018.

[134] J. Wagner, J.-C. Schoene-Bake, J.-A. Witt et al., "Distinct white matter integrity in glutamic acid decarboxylase and voltage-gated potassium channel-complex antibody-associated limbic encephalitis," *Epilepsia*, vol. 57, no. 3, pp. 475–483, 2016.

[135] G. Kojima, M. Inaba, and M. K. Bruno, "PET-positive extralimbic presentation of anti-glutamic acid decarboxylase antibody-associated encephalitis," *Epileptic Disorders*, vol. 16, no. 3, pp. 358–361, 2014.

[136] M. Takagi, H. Yamasaki, K. Endo et al., "Cognitive decline in a patient with anti-glutamic acid decarboxylase autoimmunity; case report," *BMC Neurology*, vol. 11, p. 156, 2011.

[137] L. Nappi, L. Formisano, V. Damiano, E. Matano, R. Bianco, and G. Tortora, "Paraneoplastic sensitive neuropathy associated with anti-hu antibodies in a neuroendocrine tumor of duodenum: a case report," *International Journal of Immunopathology and Pharmacology*, vol. 23, no. 4, pp. 1281–1285, 2010.

[138] M. Sweeney, M. Sweney, M. M. P. Soldán, and S. L. Clardy, "Antineuronal nuclear autoantibody type 1/anti-hu-associated opsoclonus myoclonus and epilepsia partialis continua: case report and literature review," *Pediatric Neurology*, vol. 65, pp. 86–89, 2016.

[139] F. Graus, F. Keime-Guibert, R. Reñe et al., "Anti-Hu-associated paraneoplastic encephalomyelitis: analysis of 200 patients," *Brain*, vol. 124, no. 6, pp. 1138–1148, 2001.

[140] J. W. de Beukelaar and P. A. S. Smitt, "Managing paraneoplastic neurological disorders," *The Oncologist*, vol. 11, no. 3, pp. 292–305, 2006.

[141] J. E. Greenlee, S. A. Clawson, K. E. Hill et al., "Neuronal uptake of anti-Hu antibody, but not anti-Ri antibody, leads to cell death in brain slice cultures," *Journal of Neuroinflammation*, vol. 11, p. 160, 2014.

[142] J. E. Langer, M. B. S. Lopes, N. B. Fountain et al., "An unusual presentation of anti-Hu-associated paraneoplastic limbic encephalitis," *Developmental Medicine & Child Neurology*, vol. 54, no. 9, pp. 863–866, 2012.

[143] Y. B. Shavit, F. Graus, A. Probst, R. Rene, and A. J. Steck, "Epilepsia partialis continua: a new manifestation of anti-Hu-associated paraneoplastic encephalomyelitis," *Annals of Neurology*, vol. 45, no. 2, pp. 255–258, 1999.

[144] J. Honnorat, A. Didelot, E. Karantoni et al., "Autoimmune limbic encephalopathy and anti-Hu antibodies in children without cancer," *Neurology*, vol. 80, no. 24, pp. 2226–2232, 2013.

[145] M. Mut, D. Schiff, and J. Dalmau, "Paraneoplastic recurrent multifocal encephalitis presenting with epilepsia partialis continua," *Journal of Neuro-Oncology*, vol. 72, no. 1, pp. 63–66, 2005.

[146] F. Nahab, A. Heller, and S. M. Laroche, "Focal cortical resection for complex partial status epilepticus due to a paraneoplastic encephalitis," *The Neurologist*, vol. 14, no. 1, pp. 56–59, 2008.

[147] B. Sandip and A. Abass, "Role of FDG-PET in the clinical management of paraneoplastic neurological syndrome: detection of the underlying malignancy and the brain PET-MRI correlates," *Molecular Imaging and Biology*, vol. 10, no. 3, 2008.

[148] D. A. Jacobs, K. M. Fung, N. M. Cook, W. W. Schalepfer, H. I. Goldberg, and M. M. Stecker, "Complex partial status epilepticus associated with anti-Hu paraneoplastic syndrome," *Journal of the Neurological Sciences*, vol. 213, no. 1-2, pp. 77–82, 2003.

[149] K. London and R. Howman-Giles, "Voxel-based analysis of normal cerebral [18F]FDG uptake during childhood using statistical parametric mapping," *Neuroimage*, vol. 106, pp. 264–271, 2015.

[150] S. Taneja, V. Suri, A. Ahuja, and A. Jena, "Simultaneous 18F-FDG PET/MRI in autoimmune limbic encephalitis," *Indian Journal of Nuclear Medicine*, vol. 33, no. 2, pp. 174–176, 2018.

Hindawi
Contrast Media & Molecular Imaging
Volume 2020, Article ID 6058159, 8 pages
https://doi.org/10.1155/2020/6058159

Review Article

Current and Potential Applications of Artificial Intelligence in Gastrointestinal Stromal Tumor Imaging

Cai-Wei Yang,[1] **Xi-Jiao Liu**,[1] **Si-Yun Liu**,[2] **Shang Wan**,[1] **Zheng Ye**,[1] **and Bin Song**[1]

[1]*Department of Radiology, West China Hospital, Sichuan University, Chengdu 610041, Sichuan Province, China*
[2]*GE Healthcare (China), Beijing 100176, China*

Correspondence should be addressed to Bin Song; songlab_radiology@163.com

Cai-Wei Yang and Xi-Jiao Liu contributed equally to this work.

Received 30 July 2020; Revised 18 October 2020; Accepted 31 October 2020; Published 26 November 2020

Academic Editor: Jian-Guo Zhou

The most common mesenchymal tumors are gastrointestinal stromal tumors (GISTs), which have malignant potential and can occur anywhere along the gastrointestinal system. Imaging methods are important and indispensable of GISTs in diagnosis, risk staging, therapy, and follow-up. The recommended imaging method for staging and follow-up is computed tomography (CT) according to current guidelines. Artificial intelligence (AI) applies and elaborates theses, procedures, modes, and utilization systems for simulating, enlarging, and stretching the intellectual capacity of humans. Recently, researchers have done a few studies to explore AI applications in GIST imaging. This article reviews the present AI studies in GISTs imaging, including preoperative diagnosis, risk stratification and prediction of prognosis, gene mutation, and targeted therapy response.

1. Introduction

The most frequent mesenchymal tumors from the gastrointestinal system are gastrointestinal stromal tumors (GISTs), with a prevalence of 14–20 cases per million [1]. GISTs can occur anywhere in the gastrointestinal system, with 50–60% located in the stomach, 30–35% sited in the small intestine, 5% originated in the colon and rectum, and less than 1% in the esophagus [2]. The surgical operation is the first treatment way for GISTs with malignant potential.

Imaging methods are important and indispensable of GISTs in diagnosis, staging, follow-up, and surveilling adjuvant therapy response [3]. The recommended imaging method for GISTs classification is computed tomography (CT) according to current guidelines [4], while magnetic resonance imaging (MRI) or enhanced endoscopic ultrasonography (EUS) could be replacements for iodine allergic or pregnant patients, [18F]-fluorodeoxyglucose positron emission tomography (PET)–CT can be conducive for early phase monitoring of tumor response to tyrosine kinase inhibitor (TKI) therapy [4].

At present, the clinical images practice mainly depends on the subjective interpretation by radiologists of morphological signs such as the location, margin, contour, size, attenuation, growth type, and enhancement degree. With the application and popularization of high-end multislice spiral CT, high-quality images containing rich digital information are available and prevalent, promoting artificial intelligence (AI) techniques to mine and process the big data deep in the images. Recently, an explosion of AI research emerged, particularly in the medicine field.

Recently, researchers have reported a few studies exploring the AI applications in GISTs imaging, including preoperative diagnosis, risk stratification and prediction of prognosis, gene mutation, and targeted therapy response. The current article aims to review the AI imaging studies in GISTs in relation to these four aspects.

2. Artificial Intelligence

As an information science, AI applies and elaborates theses, procedures, modes, and utilization systems for simulating, enlarging, and stretching intellectual capacity of humans [5].

Lambin et al. [6] initially proposed the notion of radiomics in 2012, which consisted of a computer-aided operating instrument derived from a great number of features from radiographic images. This technique as a new imaging technology, which can provide objective image information that cannot be recognized by the naked eye, is more detailed than the personal image interpretation by radiologists' vision. And texture analysis can quantitatively evaluate and extract the characteristics of tumors and can assess tumor heterogeneousness related to histopathological components in tumor tissues and mainly influenced by tumor neovascularization (vascularization formation and vascular permeability), tumor cellular structure, tumor cell density, and microcirculation deformation. Such quantitative-feature-based method could be of clinical associations of tumor diagnosis, staging, prognosis, and therapy.

In detail, radiomic texture is mainly composed of statistical texture, morphology-based texture, and transform-based texture. The statistical texture is formed on assessing texture as a measurement of the gray levels statistical properties based on processing the region of interest (ROI). It mainly includes (1) first-order statistical features, depicting distributed pixels in an image, such as histogram analysis; (2) second-order statistical features, as well as texture features, represent spatial relations between pixels and corresponding pair ratios, including gray-level cooccurrence matrix (GLCM), gray-level difference matrix (GLDM), gray-level run-length (GLRLM), gray-level size zone matrix (GLSZM), and neighborhood gray-tone difference matrix (NGTDM); (3) higher-order statistical features refer to the features extracted after applying filters or mathematical transformations for images, such as first-order and textural features extracted from the wavelet-filtered, Laplacian of Gaussian (LoG)-filtered, or local binary pattern (LBP) filtered images. The transform-based analysis includes texture characteristic extraction based on wave spectral statistical properties of and characterization of the global periodicity of gray level by high-energy apices and their varied types in the spectrum.

Morphology-basing method incorporates the decomposition of an image into basic units and the determination of the rules required to assemble a given image based on these basic units. All of the above methods consist in various descriptors. The detailed descriptions are presented in Table 1. In a simplistic way, a representative radiomics workflow is composed of four tasks: image attainment, image segmentation, parameter extraction, and statistical analysis (Figure 1).

Radiogenomics, as well as an encouraging novel exemplification, has the potential to extend and expand traditional radiographic images into the field of molecular and genomic imaging [7]. It aims to correlate image features with patterns of gene expressions, gene mutations, and any other genes associated traits, promoting a deeper level explanation of tumor heterogeneity and the development of imaging biomarkers [8].

Deep learning is a group of machine learning algorithms extracting deep features of the input image via multiple hidden layers [9]. Such multilayered computational models can progressively learn representations of data during multilevel abstraction [10]. A neural network is an embranchment of machine learning that organizes the basic structure of a deep learning network [11]. The models of deep learning algorithms used in medical imaging processing include Sparse Autoencoder, Convolutional Neural Network (CNN), Deep Belief Network, Restricted Boltzmann Machine, and Residual Neural Network (ResNet) [10]. Among various deep learning networks, CNN is the most popular architecture, and a further improved neural network included more computational layers.

3. Diagnosis and Differential Diagnosis

GISTs represent a distinct histopathological group of subepithelial tumors. A broad range of other mesenchymal tumors can also manifest similar imaging features with GISTs, while the two groups have distinguished prognosis and treatment. Previous studies have differentiated GISTs from other mesenchymal tumors based on tumor location, margin, contour, size, attenuation, growth type, enhancement degree, and necrosis [12–17]. However, it is still difficult to discriminate GISTs with a diameter less than 5 cm from other mesenchymal tumors, only counting on subjective imaging interpretations. There is a vacancy of AI research in this area, we look forward to more AI researches to dig new data in this field.

Clinically, the preoperative diagnosis of GISTs around the periampullary area poses a dilemma in conventional imaging performance. Rather, pancreatic ductal adenocarcinomas (PDACs), duodenal adenocarcinomas (DACs), and GISTs differed in surgery procedures and prognosis [15–17]. Recently, Lu et al. [18] retrospectively studied 74 patients with duodenal tumors around the periampullary area: 26 DACs, 20 DACs, and 28 GISTs. Volumetric histogram analysis was performed on enhanced multidetector CT images based on tumor heterogeneity. They concluded that some parameters of CT histogram analysis of periampullary tumors could be valuable for diagnostic differentiating DACs, PDACs, and GISTs arising from the periampullary area. However, the sample size and tumors type involved in this article are limited. Further researches with more sample capacity and various kinds of tumors will reinforce AI application in GISTs diagnosis.

4. Prediction of Risk Stratification and Prognosis

Several risk assessment systems for postoperative recurrence of GISTs have been proposed and evolved over the years, including the National Institute of Health (NIH) criteria, Armed Forces Institute of Pathology (AFIP) standard, and National Comprehensive Cancer Network (NCCN) risk classification. In 2008, modifications of the NIH criteria were

Table 1: Feature metrics extracted in the radiomic analysis of images.

Texture metric	Method (s)	Descriptors
First-order (statistical)	Histogram analysis	Mean, median, kurtosis, skewness, quartiles, minimum, maximum, energy (uniformity), entropy, standard deviation
Second-order (statistical)	GLCM, GLDM, NGTDM, GLRLM, GLSZM	Homogeneity, contrast, autocorrelation, prominence, maximum probability, difference variance, dissimilarity, inverse difference moment, sum entropy, sum variance, sum average, inertia, coarseness, busyness, complexity, texture strength, short run emphasis, long run emphasis, gray-level nonuniformity, run-length nonuniformity, intensity variability, run-length variability, long-zone emphasis, short-zone emphasis, intensity nonuniformity, intensity, zone percentage, variability, size zone variability
Transform (statistical)	Fourier, wavelets, discrete cosine, Gabor, law, LoG, LBP	Metrics assessing magnitude, phase, direction, edge, noise, and other descriptors
Structural analysis	Fractal analysis	Hurst component, mean fractal dimension, standard deviation, lacunarity

Note. GLCM = gray-level cooccurrence matrix, GLDM = gray-level difference matrix, NGTDM = Neighborhood gray-tone difference matrix, GLRLM = gray-level run-length, GLSZM = gray-level size zone matrix, LoG = Laplacian of Gaussian, LBP = local binary pattern.

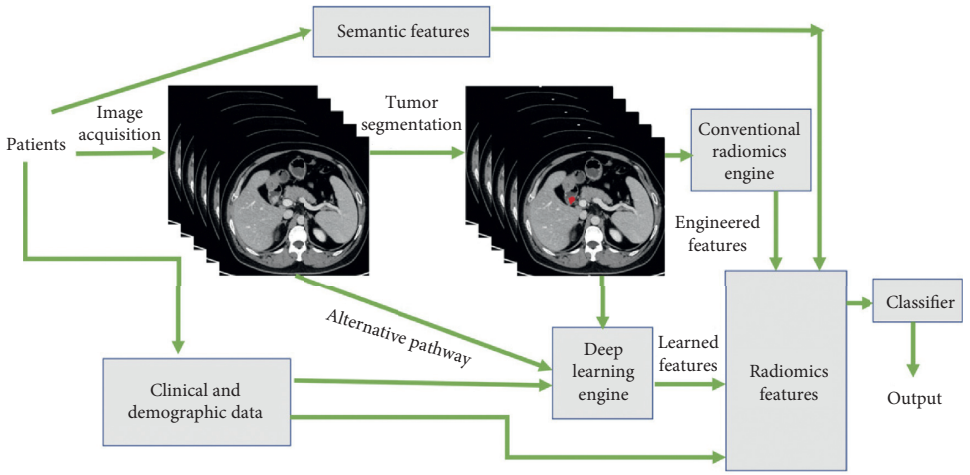


Figure 1: A representative radiomics workflow is composed of four tasks: image acquisition, tumor segmentation, features extraction, and subsequent statistical analysis.The patient in Figure 1 had a small gastrointestinal stromal tumor in the duodenum.

proposed, which incorporated tumor location, size, mitotic count, and tumor rupture. The criteria of the recurrence risk categorized into four groups (including very-low-risk group, low-risk group, intermediate-risk group, and high-risk group) and is accepted worldwide [19]. Imaging can provide more findings related to the risk stratification of GISTs. According to previous studies [20–24], tumor growth mode is related to the risk, and the risk level of GISTs with exophytic or mixed growth mode is high. It has also been suggested that the enhancement type, boundary, enlarged blood vessels, necrosis, calcification, and invasions to adjacent organs are connected to the tumor risk stratification.

The differences among the observers of subjective evaluations urged researchers to find more stable and more objective parameters and indicators. The texture analysis could extract more information hidden from medical images, which cannot be identified by subjective visual interpretation. In theory, the judgment efficiency of texture analysis of GISTs risk stratification is better than the conventional imaging [24, 25]. Nine studies have researched the performance of CT-derived radiomic signature for risk stratification [24–32], and one study evaluated EUS-derived texture [33] associated with risk stratification. The details are summarized in Table 2. In CT-derived analysis, four studies have applied NIH criterion or modified NIH criterion for GISTs malignant risk classification [25–28], while three studies were determined on NCCN guideline [24, 29, 30] and one study without clear guideline [33] and one study used Ki-67 expression standard [32]. Two of the four NIH studies based on NIH risk classification only evaluated CT textural parameters [26, 27]. The remaining two studies combined and compared conventional visual CT findings and clinical indexes models [25, 28].

In 2018, Feng et al. [26] retrospectively reviewed 90 intestinal GISTs patients. GISTs risk levels were evaluated by CT-derived histogram features that were compared according to modified NIH risk classification. They believe that volumetric CT texture features show the feasibility to be biomarkers for distinguishing low-risk, intermediate-risk, and high-risk intestinal GISTs (area under the curve (AUC) = 0.830, $P < 0.001$). However, some studies have reported contradictory results with the present study [29, 34–36]. We speculate that the differences in ROI delineation methods, and differences between enhanced and

Table 2: Details of 10 articles on artificial intelligence in the prediction of GISTs' risk stratification and prognosis.

Author	Year	Nation	Study design	Sample size	Extracted features of AI	Software
Feng C et al. [26].	2018	China	Retrospective	90	First-order statistics: Mean attenuation; 10th, 25th, 50th, 75th, and 90th percentile attenuation; skewness; kurtosis; entropy	CT kinetics
Wang C et al. [27].	2019	China	Retrospective	333 Training cohort = 233 Validation cohort = 100	First-order (histogram), haralick features, GLCM, GLRLM	AK
Chen T et al. [25].	2019	China	Retrospective	222 Training cohort = 130 Validation cohort = 92	GLV, GLRLM, GLSZM, NGTDM, GLSZM	MATLAB
Yan J et al. [28].	2018	China	Retrospective	213	First-order (histogram) gradient features, GLCM, GLRLM	MaZda
Liu S et al. [29].	2018	China	Retrospective	78	First-order (histogram)	Image analyzer
Zhang L et al. [30].	2020	China	Retrospective	140 Training cohort = 100 Validation cohort = 40	First-order features, shape and size features, second-order features (GLCM, GLRLM, GLSZM) features, and haralick features	AK
Choi I et al. [24].	2019	Korea	Retrospective	145	First-order statistics: Mean SD of mean, entropy, MPP, skewness, and kurtosis. Geometry with Gaussian filtration	MATLAB
Ning Z et al. [31].	2018	China	Retrospective	231 Training cohort = 130 Validation cohort = 101	First-order, second-order (GLCM, GLRLM, GLSZM, and NGTDM) features	MATLAB PYTHON
Zhang Q et al. [32].	2020	China	Retrospective	339 Training cohort = 148 |Internal validation cohort = 41 External validation cohort = 150	First-order statistics, features of shape, second-order features (GLCM, GLRLM, GLSZM)	PYTHON
Li X et al. [33]	2020	China	Retrospective	915 Training cohort = 680 Validation cohort = 54 Testing cohort = 181	First-order (histogram), second-order (GLCM, GLRLM, GLSZM, NGTDM) and wavelet-filtered features	MATLAB

Note. GLCM = gray-level cooccurrence matrix, GLRLM = gray-level run-length matrix, GLV = gray-level variance, GLSZM = gray-level size-zone matrix, NGTDM = Neighborhood gray-tone difference matrix.

unenhanced CT-derived texture features might be associated. In this study, the numbers of some risk groups of intestinal GISTs were limited and the author combined some groups. Moreover, this study only applied first-order statistical radiomics features. It will require further studies to explain the controversy.

Another research [27] constructed a radiomics model using multiple-order statistical radiomics features based on contrast-enhanced CT to noninvasively predict malignant-transformation potential and mitotic indexes of GISTs. In this research, the patients were classified as low- (including very-low-risk GISTs, low-risk GISTs, and intermediate-risk GISTs) and high-malignant-transformation-potential group (high-risk GISTs) based on the NIH criterion, and the sample size is enlarged with 333 numbers in total (training cohort = 233 and validation cohort = 100). The radiomics model showed a good predictive performance in differentiating high-from low-malignant-transformation-potential GISTs with an AUC value of 0.882 in the training group and 0.920 in the validation group.

The above two studies have only constructed radiomics model, and a single radiomics model could not utilize and compare the performance of conventional image findings and clinical information in GISTs' risk stratification. The next two

studies [25, 28] compared the accuracy of CT-derived textural parameters, subjective CT parameters, and clinical index models in predicting risk stratification. Yan et al. [28] included 213 intestinal GISTs patients to assess the predictive effect of clinical and subjective imaging findings and multidetector CT texture findings on preoperative risk stratification. They reported that an AUC of the model combining clinical and conventional imaging findings and multidetector CT texture features was 0.943. They deduced that CT texture may be a useful integrated tool for preoperative risk stratification of intestinal GISTs. In 2019, Chen et al. [25] constructed a radiomics nomogram for predicting GISTs malignancy potential. In comparison to conventional CT parameters and clinical indexes, the radiomics model could discriminate low-from high-malignant-transformation-potential group GISTs with a higher AUC value of 0.858. Besides, the generated radiomic nomogram model achieved the highest diagnostic performance, which showed an AUC of 0.867 and 0.847 in the internal and external cohort.

The same predicament for the only usage of radiomics model and limited sample size applied to these studies using NCCN guidelines [29, 30]. Liu et al. [29] found meaningful texture parameters from various phases in differentiating malignancy risks GISTs based on NCCN risk stratification, which was consistent with a previous study [25]. But the sample size is small as no more than 100 patients, and this study only applied first-order statistical radiomics features. With a larger included sample size (total number = 140, training cohort = 100 and validation cohort = 40) and various statistical radiomics features, Zhang et al. [30] highlighted discriminative performance with an AUC value of 0.935 and an accuracy value of 90.2% in the validation set for advanced from nonadvanced GISTs. Further, the radiomics indicated satisfied discriminative performance for four groups of GISTs risk stratification with an AUC value of 0.809 and an accuracy value of 67.5% in the validation set. Nevertheless, these studies did not conduct a direct or indirect correlation among radiomics features, subjective imaging findings, and pathological results.

So then, Choi et al. [24] evaluated and compared the diagnostic performance of CT radiomics parameters and visual CT inspection to predict malignancy grade and mitosis index of GISTs. They found the diagnostic accuracy of special radiomics features was better than visual inspection.

However, the previous studies independently used radiomics methods for pattern classification, without regard to relatively global artificially predefined parameters. Researchers also start to explore the GISTs classification efficiency of deep features obtained by deep learning networks. In 2019, Ning et al. [31] introduced an integrated structure including various features applied to a radiomics model and deep convolutional models and incorporated these features to engage in GISTs categorization. The hybrid structure with the combination of radiomics and CNNs features exhibited better performance with an AUC of 0.882 than that of the conventional CT features model (AUC = 0.774), radiomics model (global features) (AUC = 0.807), and CNN model (local features) (AUC = 0.826). As far as we can tell, this is the initial and exclusive study to apply radiomics model and CNNs for GISTs risk stratification, in which the radiomics parameters are derived from a three-dimensional universal section and deep convolutional features derived from a regional section were combined. This integrated structure enhances not merely model robustness but classifier efficiency as well.

In addition, the risk-related molecules were also predicted by using radiomics methods. The ki-67 index is an important marker related to cell proliferation and tumor heterogeneity [37]. Ki-67 is signified in the majority of the reproducing cells in high level expression, besides G0 cells, and Ki-67 is deemed as a global risk marker of malignant potential in GISTs [38]. Previous literature has also demonstrated that expression of high level Ki-67 indexes is an unrelated risk marker for high-malignancy GISTs [39–41]. A multicenter study [32] has also demonstrated a nomogram that consisted of CT-based radiomics features combined with tumor size indicated significant performance in predicting Ki-67 indexes expression in GISTs, with respective AUCs of 0.801, 0.828, and 0.784 in the training, internal validation, and external validation cohort, respectively. This proved that the Ki-67 indexes expression rate in GISTs was potentially connected with the CT textural signature.

Radiomics methods extended its applicability to various imaging modalities. For EUS-based radiomics, Li et al. [33] performed a EUS-derived radiomics model to differentiate GISTs of the higher-risk classification (intermediate-risk and high-risk) from the lower-risk classification (very-low-risk and low-risk). This model can promote the preoperative diagnosis and supply a beneficial reference for clinicians.

All of the above results show that radiomics is superior to traditional imaging description in predicting the risk stratification of GISTs, which built a foundation for the application of radiomics in the future. However, the existing studies remained insufficient. Present studies only evaluated CT-derived texture. The MRI-derived texture analysis may be more potential to dig hidden information, and quantitative imaging modalities may be useful in precise medical improvement. It should also be noted that at present, the sample sizes of most studies were limited. The inconsistency of scanning parameters, scanners, image acquisition protocol, lesion segmentation, the delineation of ROI, and statistical modeling is also presented. Selection bias of texture parameters extraction also manifested in the statistics of the levy, which leads to the consequence that duplication of research results be questioned. In addition, the conclusions of small samples also brought about poor generalization ability in specific clinical applications. Furthermore, CNN based on deep learning may substantially supplement and extend the applicability of radiomics, in the aspects of feature library or the prediction accuracy, but its effectiveness still remains to be verified.

5. Prediction of Gene Mutation

GISTs grow up in the interstitial Cajal cells from the gastrointestinal system [42], and 90% express CD117 antigen (C-KIT) [43], a tyrosinase kinase growth factor receptor [44]. GISTs with KIT exon 11 mutated genes are more

responsive to imatinib therapy [45–47], while other molecular mutations respond more poorly to imatinib. In addition, GISTs with KIT exon 9 mutations are more responsive to sunitinib. The connection between CT findings and GISTs genotype has been investigated [48]. GISTs with KIT exon 9 mutation classification have significant linkages with tumor size more than 10 cm, a stronger enhancement grade and greater area of tumor necrosis when compared to those of the KIT exon 11 mutation classification ($P < 0.05$).

In 2018, Xu et al. [49] reported a radiogenomic study on GISTs. They included enhanced CT images of 86 GISTs and performed texture analysis. They found that texture analysis could be of use to discriminate GISTs without KIT exon 11 mutated gene group from those with KIT exon 11 mutated gene group. In addition, the nongastric orientation, lower CD34 staining, and higher radiogenomic signature values were connected with GISTs without the KIT exon 11 mutated gene, which achieved satisfactory diagnostic efficiency in the validation group (AUC = 0.904 – 0.962). However, the sample size of the training cohort and validation cohort was 69 and 17 cases, respectively, and there were only four cases of tumors without KIT 11 exon mutation in the validation cohort, which may have affected the accuracy of the results. Hence, a study with a large number of patients is required to validate these conclusions.

6. Response Evaluation of Targeted Therapy

Adjuvant TKIs therapy is suggested for patients with a high recurrence risk of GISTs, and enhanced CT is the recommend imaging method for evaluating treatment response.

The study in [50] constructed and confirmed a predictive nomogram for recurrence-free survival (RFS) of GISTs after surgery without aid treatment based on deep learning (ResNet model). The ResNet nomogram was investigated on enhanced CT and clinicopathological factors including mitotic index of tumor, tumor location, and size. Both the ResNet nomogram and model manifested significant prognostic capabilities in 3- and 5-year RFS in receiver operating characteristic curves. They suggested that ResNet nomogram was supreme to the existing risk stratification standards and clinicopathological nomogram majority of the probability of exceeding reasonable threshold probabilities.

For metastatic GISTs undergoing TKI therapy, Ekert et al. [51] identified 25 GISTs patients with KIT and PDGFR mutations. All patients underwent first-line imatinib therapy and different TKI therapies after disease progression. CT texture features were extracted and associated with response categories according to the modified Choi criterion. They came to the conclusion that some of the CT texture features (GLCM inverse difference, GLCM inverse difference normalized, GLRLM, and NGTDM) correlated with prognosis, progressive-free survival, gene mutations, and treatment regimens.

7. Conclusions

Previous studies had some limitations. First, all of the above studies were retrospective. Most of them were the single center and the sample sizes were limited. The restricted number of samples not only limited the setting of imaging radiomics threshold standard, but also imposed restrictions on the training of the models [52]. Second, several image acquisition scanners and parameters were used in the same study, which might reduce the reliability and reproducibility of potential findings. Third, all the studies evaluated CT-derived texture. MRI-derived texture analysis might have more potential to uncover hidden information, and quantitative imaging modalities may be useful for improving precision medicine. In the end, most of the significant texture semantics are statistical terms, which lacked explainable correlations to the specific clinicopathological significance and biological characteristics directly and limited the interpretation of AI in repeatable research and clinical application.

The present studies demonstrated that AI methods including radiomics or deep learning have clinical value for GISTs and built a foundation for future application. Considering the limitations, prospective multicenter studies with large samples are needed. Besides, further standardization of inspection techniques and in-depth excavation of detailed signs will deepen our understanding of GIST imaging. The development of AI imaging in PET-CT and MRI will broaden our exploration. In the future, more AI studies and applications are expected in preoperative prediction of various gene mutations and evaluation of the efficacy of targeted therapies to make continuous progress towards the goal of individualized and accurate treatment.

Abbreviations

GISTs: Gastrointestinal stromal tumors
CT: Computed tomography
MRI: Magnetic resonance imaging
PET-CT: Positron emission tomography-CT
TKI: Tyrosine kinase inhibitor
EUS: Endoscopic ultrasonography
AI: Artificial intelligence
ROI: Region of interest
GLCM: Gray-level cooccurrence matrix
GLDM: Gray-level difference matrix
NGTDM: Neighborhood gray-tone difference matrix
GLRLM: Gray-level run-length
GLSZM: Gray-level size zone matrix
LoG: Laplacian of Gaussian
LBP: Local binary pattern
CNN: Convolutional neural network
ResNet: Residual neural network
PDAC: Pancreatic ductal adenocarcinoma
DAC: Duodenal adenocarcinoma
NIH: National institute of health
AFIP: Armed forces institute of pathology
NCCN: National comprehensive cancer network
AUC: Area under the curve
RFS: Recurrence-free survival.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Cai-Wei Yang, Shang Wan, and Zheng Ye collected materials. Cai-Wei Yang and Xi-Jiao Liu wrote the paper. Xi-Jiao Liu, Si-Yun Liu, and Bin Song revised the paper. Cai-Wei Yang and Xi-Jiao Liu contributed to this work equally.

Acknowledgments

This work was supported by Research Grant of the National Natural Science Foundation of China, No. 82001810, and No. 81971571.

References

[1] H. Joensuu, P. Hohenberger, and C. L. Corless, "Gastrointestinal stromal tumour," *The Lancet*, vol. 382, no. 9896, pp. 973–983, 2013.

[2] H. Joensuu, A. Vehtari, J. Riihimäki et al., "Risk of recurrence of gastrointestinal stromal tumour after surgery: an analysis of pooled population-based cohorts," *The Lancet Oncology*, vol. 13, no. 3, pp. 265–274, 2012.

[3] H. Choi, "Imaging modalities of gastrointestinal stromal tumors," *Journal of Surgical Oncology*, vol. 104, no. 8, pp. 907–914, 2011.

[4] I. Judson, R. Bulusu, B. Seddon et al., "UK clinical practice guidelines for the management of gastrointestinal stromal tumours (GIST)," *Clinical Sarcoma Research*, vol. 7, p. 6, 2017.

[5] N. Peek, C. Combi, R. Marin, and R. Bellazzi, "Thirty years of artificial intelligence in medicine (AIME) conferences: a review of research themes," *Artificial Intelligence in Medicine*, vol. 65, no. 1, pp. 61–73, 2015.

[6] P. Lambin, E. Rios-Velazquez, R. Leijenaar et al., "Radiomics: extracting more information from medical images using advanced feature analysis," *European Journal of Cancer*, vol. 48, no. 4, pp. 441–446, 2012.

[7] M. D. Kuo and N. Jamshidi, "Behind the numbers: decoding molecular phenotypes with radiogenomics-guiding principles and technical considerations," *Radiology*, vol. 270, no. 2, pp. 320–325, 2014.

[8] K. Pinker, F. Shitano, E. Sala et al., "Background, current role, and potential applications of radiogenomics," *Journal of Magnetic Resonance Imaging*, vol. 47, no. 3, pp. 604–620, 2018.

[9] M. P. McBee, O. A. Awan, A. T. Colucci et al., "Deep learning in radiology," *Academic Radiology*, vol. 25, no. 11, pp. 1472–1480, 2018.

[10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[11] G. Litjens, T. Kooi, B. E. Bejnordi et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[12] Y. R. Choi, S. H. Kim, S.-A. Kim et al., "Differentiation of large (≥5cm) gastrointestinal stromal tumors from benign subepithelial tumors in the stomach: radiologists' performance using CT," *European Journal of Radiology*, vol. 83, no. 2, pp. 250–260, 2014.

[13] C. Zhou, X. Duan, X. Zhang, H. Hu, D. Wang, and J. Shen, "Predictive features of CT for risk stratifications in patients with primary gastrointestinal stromal tumour," *European Radiology*, vol. 26, no. 9, pp. 3086–3093, 2016.

[14] J. S. Kim, H. J. Kim, S. H. Park, J. S. Lee, A. Y. Kim, and H. K. Ha, "Computed tomography features and predictive findings of ruptured gastrointestinal stromal tumours," *European Radiology*, vol. 27, no. 6, pp. 2583–2590, 2017.

[15] F. Yang, C. Jin, Z. Du et al., "Duodenal gastrointestinal stromal tumor: clinicopathological characteristics, surgical outcomes, long term survival and predictors for adverse outcomes," *The American Journal of Surgery*, vol. 206, no. 3, pp. 360–367, 2013.

[16] A. El Nakeeb, M. El Shobary, M. El Dosoky et al., "Prognostic factors affecting survival after pancreaticoduodenectomy for pancreatic adenocarcinoma (single center experience)," *Hepatogastroenterology*, vol. 61, no. 133, pp. 1426–1438, 2014.

[17] M. Zenali, M. J. Overman, A. Rashid et al., "Clinicopathologic features and prognosis of duodenal adenocarcinoma and comparison with ampullary and pancreatic ductal adenocarcinoma," *Human Pathology*, vol. 44, no. 12, pp. 2792–2798, 2013.

[18] J. Lu, D. Hu, H. Tang et al., "Assessment of tumor heterogeneity: differentiation of periampullary neoplasms based on CT whole-lesion histogram analysis," *European Journal of Radiology*, vol. 115, pp. 1–9, 2019.

[19] Y.-M. Lin, N.-C. Chiu, A. F.-Y. Li, C.-A. Liu, Y.-H. Chou, and Y.-Y. Chiou, "Unusual gastric tumors and tumor-like lesions: radiological with pathological correlation and literature review," *World Journal of Gastroenterology*, vol. 23, no. 14, pp. 2493–2504, 2017.

[20] H. Joensuu, "Risk stratification of patients diagnosed with gastrointestinal stromal tumor," *Human Pathology*, vol. 39, no. 10, pp. 1411–1419, 2008.

[21] F. J. Maldonado, S. P. Sheedy, V. R. Iyer et al., "Reproducible imaging features of biologically aggressive gastrointestinal stromal tumors of the small bowel," *Abdominal Radiology*, vol. 43, no. 7, pp. 1567–1574, 2018.

[22] H. Li, G. Ren, R. Cai, J. Chen, X. Wu, and J. Zhao, "A correlation research of Ki67 index, CT features, and risk stratification in gastrointestinal stromal tumor," *Cancer Medicine*, vol. 7, no. 9, pp. 4467–4474, 2018.

[23] T. Chen, L. Xu, X. Dong et al., "The roles of CT and EUS in the preoperative evaluation of gastric gastrointestinal stromal tumors larger than 2 cm," *European Radiology*, vol. 29, no. 5, pp. 2481–2489, 2019.

[24] I. Y. Choi, S. K. Yeom, J. Cha et al., "Feasibility of using computed tomography texture analysis parameters as imaging biomarkers for predicting risk grade of gastrointestinal stromal tumors: comparison with visual inspection," *Abdominal Radiology*, vol. 44, no. 7, pp. 2346–2356, 2019.

[25] T. Chen, Z. Ning, L. Xu et al., "Radiomics nomogram for predicting the malignant potential of gastrointestinal stromal tumours preoperatively," *European Radiology*, vol. 29, no. 3, pp. 1074–1082, 2019.

[26] C. Feng, F. Lu, Y. Shen et al., "Tumor heterogeneity in gastrointestinal stromal tumors of the small bowel: volumetric CT texture analysis as a potential biomarker for risk stratification," *Cancer Imaging*, vol. 18, no. 1, p. 46, 2018.

[27] C. Wang, H. Li, Y. Jiaerken et al., "Building CT radiomics-based models for preoperatively predicting malignant potential and mitotic count of gastrointestinal stromal tumors," *Translational Oncology*, vol. 12, no. 9, pp. 1229–1236, 2019.

[28] J. Yan, X. Zhao, S. Han, T. Wang, and F. Miao, "Evaluation of clinical plus imaging features and multidetector computed tomography texture analysis in preoperative risk grade prediction of small bowel gastrointestinal stromal tumors," *Journal of Computer Assisted Tomography*, vol. 42, no. 5, pp. 714–720, 2018.

[29] S. Liu, X. Pan, R. Liu et al., "Texture analysis of CT images in predicting malignancy risk of gastrointestinal stromal tumours," *Clinical Radiology*, vol. 73, no. 3, pp. 266–274, 2018.

[30] L. Zhang, L. Kang, G. Li et al., "Computed tomography-based radiomics model for discriminating the risk stratification of gastrointestinal stromal tumors," *La Radiologia Medica*, vol. 125, no. 5, pp. 465–473, 2020.

[31] Z. Ning, J. Luo, Y. Li et al., "Pattern classification for gastrointestinal stromal tumors by integration of radiomics and deep convolutional features," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 3, pp. 1181–1191, 2019.

[32] Q. W. Zhang, Y. J. Gao, R. Y. Zhang et al., "Personalized CT-based radiomics nomogram preoperative predicting Ki-67 expression in gastrointestinal stromal tumors: a multicenter development and validation cohort," *Clinical and Translational Medicine*, vol. 9, no. 1, p. 12, 2020.

[33] X. Li, F. Jiang, Y. Guo, Z. Jin, and Y. Wang, "Computer-aided diagnosis of gastrointestinal stromal tumors: a radiomics method on endoscopic ultrasound image," *International Journal of Computer Assisted Radiology and Surgery*, vol. 14, no. 10, pp. 1635–1645, 2019.

[34] C. Yip, D. Landau, R. Kozarski et al., "Primary esophageal cancer: heterogeneity as potential prognostic biomarker in patients treated with definitive chemotherapy and radiation therapy," *Radiology*, vol. 270, no. 1, pp. 141–148, 2014.

[35] F. Ng, B. Ganeshan, R. Kozarski, K. A. Miles, and V. Goh, "Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis: contrast-enhanced CT texture as a biomarker of 5-year survival," *Radiology*, vol. 266, no. 1, pp. 177–184, 2013.

[36] G.-M.-Y. Zhang, H. Sun, B. Shi, Z.-Y. Jin, and H.-D. Xue, "Quantitative CT texture analysis for evaluating histologic grade of urothelial carcinoma," *Abdominal Radiology*, vol. 42, no. 2, pp. 561–568, 2017.

[37] J. R. A. Hutchins, Y. Toyoda, B. Hegemann et al., "Systematic analysis of human protein complexes identifies chromosome segregation proteins," *Science*, vol. 328, no. 5978, pp. 593–599, 2010.

[38] N. Nakamura, H. Yamamoto, T. Yao et al., "Prognostic significance of expressions of cell-cycle regulatory proteins in gastrointestinal stromal tumor and the relevance of the risk grade," *Human Pathology*, vol. 36, no. 7, pp. 828–837, 2005.

[39] J. Jiang, M. S. Jin, J. Suo et al., "Evaluation of malignancy using Ki-67, p53, EGFR and COX-2 expressions in gastrointestinal stromal tumors," *World Journal of Gastroenterology*, vol. 18, no. 20, pp. 2569–2575, 2012.

[40] X. Wang, I. Mori, W. Tang et al., "Helpful parameter for malignant potential of gastrointestinal stromal tumors (GIST)," *Japanese Journal of Clinical Oncology*, vol. 32, no. 9, pp. 347–351, 2002.

[41] A. Nanding, L. Tang, L. Cai et al., "Low ING4 protein expression detected by paraffin-section immunohistochemistry is associated with poor prognosis in untreated patients with gastrointestinal stromal tumors," *Gastric Cancer*, vol. 17, no. 1, pp. 87–96, 2014.

[42] L. G. Kindblom, H. E. Remotti, F. Aldenborg, and J. M Meis-Kindblom, "Gastrointestinal pacemaker cell tumor (GIPACT): gastrointestinal stromal tumors show phenotypic characteristics of the interstitial cells of Cajal," *The American Journal of Pathology*, vol. 152, no. 5, pp. 1259–1269, 1998.

[43] S. Hirota, K. Isozaki, Y. Moriyama et al., "Gain-of-function mutations of c-kit in human gastrointestinal stromal tumors," *Science*, vol. 279, no. 5350, pp. 577–580, 1998.

[44] C. D. M. Fletcher, J. J. Berman, C. Corless et al., "Diagnosis of gastrointestinal stromal tumors: a consensus approach," *Human Pathology*, vol. 33, no. 5, pp. 459–465, 2002.

[45] M. C. Heinrich, K. Owzar, C. L. Corless et al., "Correlation of kinase genotype and clinical outcome in the north American intergroup phase III trial of imatinib mesylate for treatment of advanced gastrointestinal stromal tumor: CALGB 150105 study by cancer and leukemia group B and southwest oncology group," *Journal of Clinical Oncology*, vol. 26, no. 33, pp. 5360–5367, 2008.

[46] A. Patrikidou, J. Domont, S. Chabaud et al., "Long-term outcome of molecular subgroups of GIST patients treated with standard-dose imatinib in the BFR14 trial of the French Sarcoma Group," *European Journal of Cancer*, vol. 52, pp. 173–180, 2016.

[47] C. L. Corless, K. V. Ballman, C. R. Antonescu et al., "Pathologic and molecular features correlate with long-term outcome after adjuvant therapy of resected primary GI stromal tumor: the ACOSOG Z9001 trial," *Journal of Clinical Oncology*, vol. 32, no. 15, pp. 1563–1570, 2014.

[48] Y. Q. Yin, C. J. Liu, B. Zhang et al., "Association between CT imaging features and KIT mutations in small intestinal gastrointestinal stromal tumors," *Scientific Reports*, vol. 9, no. 1, p. 7257, 2019 10.

[49] F. Xu, X. Ma, Y. Wang et al., "CT texture analysis can be a potential tool to differentiate gastrointestinal stromal tumors without KIT exon 11 mutation," *European Journal of Radiology*, vol. 107, pp. 90–97, 2018.

[50] T. Chen, S. Liu, Y. Li et al., "Developed and validated a prognostic nomogram for recurrence-free survival after complete surgical resection of local primary gastrointestinal stromal tumors based on deep learning," *EBioMedicine*, vol. 39, pp. 272–279, 2019.

[51] K. Ekert, C. Hinterleitner, and M. Horger, "Prognosis assessment in metastatic gastrointestinal stromal tumors treated with tyrosine kinase inhibitors based on CT-texture analysis," *European Journal of Radiology*, vol. 116, pp. 98–105, 2019.

[52] C. M. Park, "Can artificial intelligence fix the reproducibility problem of radiomics?" *Radiology*, vol. 292, no. 2, pp. 374-375, 2019.

Hindawi
Contrast Media & Molecular Imaging
Volume 2020, Article ID 3725353, 6 pages
https://doi.org/10.1155/2020/3725353

Research Article

Contrast-Enhanced Ultrasound Evaluation of Mifepristone for Treatment of Low-Risk Cesarean Scar Pregnancy

Xi Xiong,[1] **Chun-yan Gao**,[1] **De-mei Ying**,[1] **Ping Yan**,[1] **Zhi-jia Zhang**,[2] **Na Kuang**,[1] **Hong-ju Tian**,[1] **Li Luo**,[1] **Shu-yu Long**,[1] **and Zheng-qiong Chen**[1]

[1]*Department of Obstetrics and Gynecology, Second Clinical Medical College of Army Medical University, Chongqing 400037, China*
[2]*Department of Clinical Laboratory, Second Clinical Medical College of Army Medical University, Chongqing 400037, China*

Correspondence should be addressed to Zheng-qiong Chen; chenzhengqiong75@163.com

Received 22 June 2020; Revised 9 October 2020; Accepted 18 October 2020; Published 31 October 2020

Academic Editor: Wei Li

Purpose. The effect of mifepristone for treatment of low-risk cesarean scar pregnancy (CSP) was monitored by contrast-enhanced ultrasound (CEUS). *Methods*. Data were collected from 23 CSP patients with a 10-point risk score <5 (low-risk CSP) and from 23 intrauterine pregnancy (IUP) patients with a scar from a previous cesarean delivery. All patients were prescribed 75 mg mifepristone daily for 2 days and underwent transvaginal CEUS before and after administration of mifepristone. On the third day, uterine curettage was performed after transvaginal CEUS. Arrival time (AT), peak intensity (PI), and area under the curve (AUC) around the gestational sac were monitored by CEUS before and after application of mifepristone, and the rate of effective treatment was compared between the two patient groups. *Results*. No patients experienced side effects from either the CEUS procedure or the mifepristone treatment. Changes in AT, PI, and AUC index from before vs. after mifepristone treatment did not differ significantly between the two groups (all p values >0.05). There was also no significant difference in the rate of effective treatment between the two groups (95.65% in the CSP group vs. 100% in the IUP group; $p > 0.05$). *Conclusions*. Based on monitoring by CEUS, the effect of mifepristone in low-risk CSP was comparable to that in IUP.

1. Introduction

Cesarean scar pregnancy (CSP), or implantation of the gestational sac in a hysterotomy scar, is a rare but serious complication that can occur in a subsequent pregnancy after cesarean delivery [1]; it is especially concerning in China [2]. Prenatal diagnosis of CSP is based on the presence of a gestational sac at the site of the previous uterine incision and the presence of an empty uterine cavity and cervix and thin myometrium adjacent to the bladder [3]. The severity of CSP has been found to correlate with clinical and sonographic characteristics including the implantation site, blood flow around the gestational sac, timing within gestation, and number of previous cesarean deliveries [4, 5]. Numerous management options for CSP have been evaluated based on case series, including laparoscopy, uterine artery embolization (UAE), and high-intensity focused ultrasound (HIFU) [6–8]. However, no standardized diagnostic or management guidelines have been published [9].

Our group has developed and validated a scoring system to rate the severity of CSP on a 10-point scale based on clinical indicators including thickness of the myometrium at uterine incision, grading of blood flow, fetal heartbeat, location of the gestational sac, maximal diameter of the gestational sac, and number of previous cesarean sections [10]. To validate this scoring system, patients were assigned a risk score based on these indicators, and treatment modalities employed were then assessed in relation to risk scores. Results showed that patients with CSP risk scores lower than 5 were significantly less likely to need invasive salvage treatments compared to higher-risk patients [10].

Mifepristone is the most commonly used progesterone antagonist [11]. In addition to producing prostaglandins to accelerate the degeneration and necrosis of villi, mifepristone can also reduce the vascular endothelial growth factor in decidual tissue, thereby reducing blood supply to the embryo and bringing about termination of the pregnancy [12]. A case report showed that the approach with mifepristone for treatment of CSP may be a safer and less invasive method [13]. By contrast, another study suggested that mifepristone is not very effective in the treatment of CSP [14]. Therefore, there are no commonly accepted clinical management guidelines on the use of mifepristone for CSP.

Contrast-enhanced ultrasound (CEUS) is a safe, widely available, and relatively inexpensive imaging technique that uses dedicated imaging ultrasound sequences and FDA-approved contrast microbubbles, permitting high diagnostic accuracy [15]. CEUS is a convenient method for diagnosis of CSP, has excellent spatial and temporal resolution, and can be used for quantitative assessment of microcirculation perfusion of the gestational sac [16]. Building on findings from our previous work, we sought to quantitatively analyze changes in microcirculation around the gestational sac using CEUS in order to evaluate the efficacy of mifepristone in the treatment of low-risk CSP.

2. Materials and Methods

Data were prospectively collected between July 2018 and March 2019 from patients seen in the department of obstetrics and gynecology of the Second Clinical Medical College of Army Medical University of China. Participants included a group of patients with low-risk CSP (10-point risk score <5, $n = 23$) and an equal-sized control group of patients having intrauterine pregnancy (IUP) with scar and electing to terminate pregnancy. Patients were excluded from the study if they had serious diseases of vital organs such as the heart, kidney, and lungs. Diagnosis of CSP was confirmed by review of sonographic images. In accordance with the "2013 revision of the Declaration of Helsinki," all study participants gave written informed consent regarding study procedures and treatment modalities after the procedures had been fully explained to them.

We assessed relevant demographic and clinical characteristics including age, parity, gestational age, BMI, and remnant myometrial thickness. All patients were prescribed 75 mg mifepristone daily for 2 days and underwent transvaginal CEUS before and after administration of mifepristone (Figures 1 and 2). On the third day, uterine curettage was performed after transvaginal CEUS.

2.1. CEUS Examination. All patients were examined by two obstetric ultrasound technicians with at least 5 years of experience. All ultrasound examinations were conducted using a Philips IU-22 system (Philips Electronics N.V., Amsterdam, Netherlands) with a 5–9 MHz transvaginal transducer. A 21G trocar was used to puncture the cubital vein and establish a venous channel. Next, 2.5 mL of the contrast agent was injected, and 5 mL of 0.9% normal saline was used for tube washing. When the contrast agent was injected, the patient began holding their breath and took a shallow breath when required or alternatively continued with slow shallow breathing. All patients were trained in the required breathing regime before the contrast process. Two minutes of ultrasound data were recorded and saved for analysis. The lesion area with the most evident enhancement was identified as the region of interest (ROI), with ROIs set as 5 mm diameter circles and remaining unchanged. The ROIs were located at the embryo decidua basalis. Related parameters obtained through the time intensity curve (TIC) included arrival time (AT, the time from injection of the agent to the point when the first contrast bubbles appeared in the gestational sac), peak intensity (PI, the maximal intensity of the TIC), and area under the curve around the gestational sac (AUC, the area under the TIC) [17].

Operations were ceased if vaginal bleeding exceeded 300 mL, in which case uterine balloons were employed for temporary hemostasis. It is recommended that UAE be performed if bleeding exceeds 500 mL. All study patients were followed up for at least one month following the study procedure; assessments included the serum β-hCG level and presence of abdominal pain, vaginal bleeding, and fever every month.

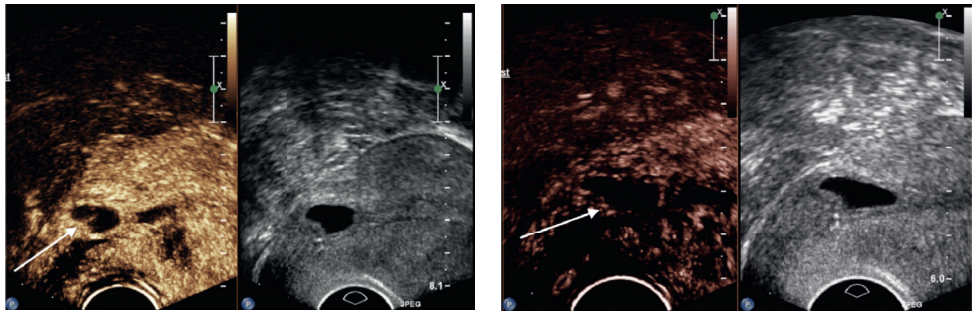
2.2. Evaluation of Curative Effects. Curative effects were assessed at three months following treatment and were rated as follows:

2.2.1. Excellent Curative Effect. Ultrasound showed no residual gestational tissue; the patient had no abdominal pain, vaginal bleeding, or fever, and the serum β-hCG level decreased and returned to normal in three months.

2.2.2. Moderate Curative Effect. Ultrasound results showed residual gestational tissue; the patient had abdominal pain, vaginal bleeding, or fever, and the serum β-hCG level was decreased. After recurettage or pharmaceutical treatment, ultrasound showed no residual gestational tissue, there were no symptoms such as abdominal pain, vaginal bleeding, or fever, and the serum β-hCG level returned to normal within three months.

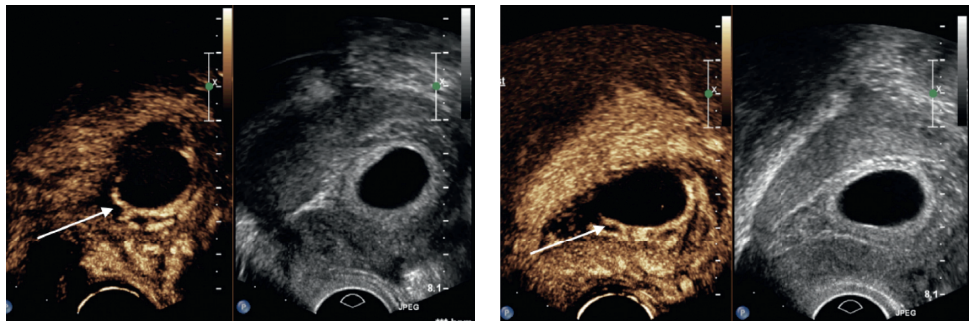
2.2.3. Poor Curative Effect. Ultrasound showed residual gestational tissue; the patient had abdominal pain, vaginal bleeding, or other symptoms, and the serum β-hCG level may have increased or decreased but did not decrease to normal levels. After recurettage or pharmaceutical treatment, ultrasound showed that the residual gestational tissue had persisted or grown. Patients had persistent vaginal bleeding or abdominal pain and needed further treatment such as laparoscopic surgery or UAE.

The total effective treatment rate was defined as the number of patients for whom treatment was rated as excellent or moderate divided by the total number of patients in each study group and expressed as a percentage.



(a) (b)

Figure 1: Findings from transvaginal contrast-enhanced ultrasonography before and after mifepristone administration in patients with low-risk cesarean scar pregnancy. (a) Before mifepristone treatment. (b) After mifepristone treatment.



(a) (b)

Figure 2: Findings from transvaginal contrast-enhanced ultrasonography before and after mifepristone administration in patients with intrauterine pregnancy. (a) Before mifepristone treatment. (b) After mifepristone treatment.

2.3. Statistical Analysis. Analyses were performed using SPSS software version 16.0 (IBM, Armonk, NY, USA). All variables are presented as mean ± standard deviation. The paired *t*-test was used to compare patient characteristics between the CSP and IUP groups, to compare perfusion parameters as assessed by CEUS before vs. after mifepristone treatment, and to compare changes in perfusion parameters and curative effects between the two study groups. The chi-squared test was used to verify the efficacy of mifepristone. A p value of <0.05 was considered statistically significant.

3. Results

3.1. Patient Demographics and Clinical Characteristics. Patient characteristics for the CSP and IUP groups are shown in Table 1. Mean thickness of the lower uterine segment was significantly higher among women with IUP (5.72 ± 1.65 mm vs. 2.60 ± 1.20, $p \leq 0.001$). There were no other significant differences between the two patient groups.

There were 16 patients with fetal heart activity in the CSP group and 15 patients in the IUP group. Two days after misoprostol administration, there were 9 patients without fetal heart activity detected in the CSP group and 7 patients without fetal heart activity in the IUP group (difference between the two groups not statistically significant, $\chi^2 = 2.000$; $p = 0.368$).

Three patients in the CSP group reported pain without bleeding, 7 patients presented with bleeding but did not report pain, and 7 patients presented with both pain and bleeding. In the IUP group, 4 patients reported pain with no bleeding, 6 patients presented with bleeding but did not report pain, and 6 patients presented with both pain and bleeding.

3.2. CEUS Findings. CEUS perfusion quantification values for the two study groups before and after mifepristone treatment are shown in Table 2. In the CSP group, before mifepristone treatment, AT is 18.42 ± 3.38 (s), PI is 17.68 ± 2.84 (dB), and AUC is 1011.03 ± 194.53, and after mifepristone treatment, AT is 13.39 ± 1.98 (s), PI is 14.48 ± 2.81 (dB), and AUC is 800.33 ± 109.41. In IUP group, before mifepristone treatment, AT is 18.71 ± 2.01 (s), PI is 17.85 ± 2.61 (dB), and AUC is 1041.76 ± 168.14, and after mifepristone treatment, AT is 14.06 ± 2.85 (s), PI is 15.47 ± 2.44 (dB), and AUC is 878.49 ± 162.23. Based on TIC analysis, AT, PI, and AUC around the gestational sac were significantly lower in both study groups after mifepristone treatment than before ($p < 0.05$; Table 2).

Forty-eight hours following administration of mifepristone, in the CSP group, AT, PI, and AUC changes in blood flow around the gestational sac are 5.03 ± 2.97 (s),

Table 1: Patient characteristics.

Characteristic	CSP ($n = 23$)	IUP ($n = 23$)	t	p
Maternal age (years)	32.26 ± 3.99	32.43 ± 3.78	0.152	0.880
BMI	23.12 ± 2.92	22.37 ± 3.96	−0.726	0.471
Gravidity	4.29 ± 1.43	4.61 ± 1.47	0.507	0.615
Parity	1.47 ± 0.51	1.52 ± 0.51	0.289	0.774
Diameter of gestational sac (mm)	20.40 ± 7.90	21.37 ± 11.64	0.331	0.742
Previous cesarean deliveries (times)	1.47 ± 0.51	1.52 ± 0.51	0.289	0.774
Thickness of the lower uterine segment (mm)	2.60 ± 1.20	5.72 ± 1.65	7.341	≤0.001

CSP, cesarean scar pregnancy; IUP, intrauterine pregnancy.

Table 2: CEUS perfusion quantification before and after mifepristone treatment in CSP and IUP patients.

Group	Parameter	Before mifepristone treatment	After mifepristone treatment	t	p
CSP ($n = 23$)	AT (s)	18.42 ± 3.38	13.39 ± 1.98	8.117	≤0.001
	PI (dB)	17.68 ± 2.84	14.48 ± 2.81	6.446	≤0.001
	AUC	1011.03 ± 194.53	800.33 ± 109.41	8.341	≤0.001
IUP ($n = 23$)	AT (s)	18.71 ± 2.01	14.06 ± 2.85	7.208	≤0.001
	PI (dB)	17.85 ± 2.61	15.47 ± 2.44	4.411	≤0.001
	AUC	1041.76 ± 168.14	878.49 ± 162.23	10.961	≤0.001

AT, arrival time; AUC, the area under the time intensity curve; CEUS, contrast-enhanced ultrasonography; CSP, cesarean scar pregnancy; IUP, intrauterine pregnancy; PI, peak intensity.

0.01 ± 1.78 (dB), and 210.69 ± 121.14; in the IUP group, AT, PI, and AUC changes in blood flow around the gestational sac are 4.65 ± 3.09 (s), 0.15 ± 1.53 (dB), and 169.26 ± 74.06. Based on TIC analysis, changes in blood flow around the gestational sac, including AT, PI, and AUC, did not differ significantly between the two groups (Table 3).

Table 3: Changes in blood flow around the gestational sac before vs. after mifepristone treatment in CSP and IUP patients.

Parameter	CSP ($n = 23$)	IUP ($n = 23$)	t	p
AT (s)	5.03 ± 2.97	4.65 ± 3.09	−0.424	0.674
PI (dB)	0.01 ± 1.78	0.15 ± 1.53	−1.115	0.271
AUC	210.69 ± 121.14	169.26 ± 74.06	−1.399	0.169

AT, arrival time; AUC, the area under the time intensity curve; CSP, cesarean scar pregnancy; IUP, intrauterine pregnancy; PI, peak intensity.

3.3. Evaluation of Curative Effects. In the IUP group, no patients had vaginal bleeding >500 mL during uterine curettage. One week after uterine curettage, ultrasound showed no residual gestational tissue and no abdominal pain, vaginal bleeding, or fever, and serum β-hCG levels returned to normal in three months in all patients. Thus, the total effective treatment rate was 100% in the IUP group.

In the CSP group, two patients had vaginal bleeding >500 mL during uterine curettage; bleeding was significantly reduced following administration of 1 mL oxytocin during the operation, and no further treatment was needed. One week following the operation, 6 patients had residual gestational tissue, minimal vaginal bleeding, and no abdominal pain. These patients were instructed to take mifepristone orally. Among them, 2 patients underwent recurettage because there had been no evident reduction of gestational tissue. After recurettage, ultrasound showed no residual gestational tissue, there was no vaginal bleeding, and serum β-hCG levels returned to normal. One patient underwent laparoscopic resection after conservative treatment for 1 month because of persistent vaginal bleeding, continuous enlargement of residual gestational tissue, and disappearance of local muscular layer (Figure 2); there was no significant decrease in the serum β-hCG level in this patient. In the remaining 3 CSP patients with no abdominal pain, residual gestational tissue gradually shrank and returned to normal within one month, and serum β-hCG levels returned to normal within three months. The total effective treatment rate among CSP patients was thus 95.65% (Table 4). The effective treatment rate did not differ significantly between the two study groups ($\chi^2 = 4.000$; $p = 0.261$).

4. Discussion

With continuing advances in research on CSP, it is now understood that the risk posed by CSP is affected by many factors, including the number of previous cesarean deliveries, the position of implantation of the gestational sac, and the timing within gestation [4, 5]. While numerous management options for CSP have been identified and evaluated, no standardized diagnostic or management guidelines have been developed [8, 9, 18]. Therefore, it is crucial to make an accurate diagnosis and to provide prompt therapy to avoid potentially catastrophic complications.

In our previous study evaluating the utility of a CSP risk scoring system to predict appropriate treatment, we found mifepristone combined with uterine curettage to be the optimal treatment for low-risk patients (those with a risk score <5) [10]. In the present study, we found no significant differences in maternal age, BMI, gravidity, parity, or gestational days between CSP and IUP patients, but average muscle layer thickness in the CSP group was less than half of that of the normal pregnancy group. Our finding of low

Table 4: Comparison of curative effects between CSP and IUP patients.

Group	n	Excellent curative effect	Moderate curative effect	Poor curative effect	Effective treatment rate (%)
CSP	23	16	6	1	95.65
IUP	23	23	0	0	100

CSP, cesarean scar pregnancy; IUP, intrauterine pregnancy.

remnant myometrial thickness in CSP is consistent with results from previous studies [9] and may stem from erosion of the muscular layer of the gestational sac when it is implanted in the scar, resulting in thinning of the gestational sac.

The effectiveness of high-dose mifepristone for abortion has been well-established [19–21]. Mifepristone influences the human endometrium during the luteal phase by reducing stromal edema, increasing venular diameter, and causing erythrocyte and leukocyte diapedesis and focal hemorrhage and degeneration of the stromal extracellular matrix. Through these mechanisms, eventual degradation of the endometrium is initiated, leading to termination of pregnancy.

CEUS has become a widely available and well-accepted imaging modality in recent years. By overcoming some of the limitations of conventional ultrasonography, CEUS creates a significant opportunity for visualization of the microcirculation [22]. Findings from the present study show that perfusion parameters around the gestational sac were significantly reduced following mifepristone treatment in both low-risk CSP patients and women with IUP. Accordingly, it appears that mifepristone brings about medical abortion in part through reducing microcirculation of the gestational sac by acting on endometrial vessels. In addition, changes in microcirculation of the gestational sac did not differ significantly between the two study groups, suggesting that mifepristone has the same effect on pregnancy termination in low-risk CSP as in normal pregnancy. Furthermore, both groups achieved similar curative effects through mifepristone combined with curettage. Accordingly, more aggressive treatments such as laparoscopy, hysteroscopy, and UAE can be avoided through the use of mifepristone combined with curettage in low-risk CSP patients. This conclusion is consistent with findings from Fu et al. [23] and suggests that personalizing treatment options based on the patient's condition can reduce the physical and mental impact of treatment on patients while also reducing the cost of their care.

Several limitations to our study should be acknowledged. First, this study was conducted at a single center, and our results should be confirmed in larger multicenter studies before being applied more widely in clinical practice. Second, we did not look at long-term outcomes such as recurrent ectopic pregnancy or subsequent fertility. We recommend that future studies examine these outcomes in order to develop a richer understanding of the long-term safety and risks of mifepristone for treatment of low-risk CSP.

In conclusion, based on monitoring by CEUS, the effect of mifepristone in low-risk CSP was comparable to that in IUP, and combined with uterine curettage, we found this treatment was safe and effective in patients with low-risk CSP. In this patient population, such a treatment course can be used to avoid more aggressive treatments such as laparoscopy, hysteroscopy, and UAE.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Xi Xiong, Chun-yan Gao, and De-mei Ying have contributed equally to this paper.

Acknowledgments

The authors thank LetPub (http://www.letpub.com) for its linguistic assistance and scientific consultation during the preparation of this manuscript. Funding for this research was provided by the Clinical Research Foundation of the Second Clinical Medical College of Army Medical University (2016YLC23), and the Technological Innovation and Application Development Project of the Chongqing Science and Technology Bureau (stc2019jscx-msxmX0269), with key project support from the Second Affiliated Hospital of Army Medical University (2019JSLC011).

References

[1] I. E. Timor-Tritsch, A. Monteagudo, R. Santos, T. Tsymbal, G. Pineda, and A. A. Arslan, "The diagnosis, treatment, and follow-up of cesarean scar pregnancy," *American Journal of Obstetrics and Gynecology*, vol. 207, p. 44, 2012.

[2] L. Luo, X. Ruan, C. Li, S. Chen, Q. Hu, and A. O. Mueck, "Early clinical features and risk factors for cesarean scar pregnancy: a retrospective case-control study," *Gynecological Endocrinology*, vol. 35, no. 4, pp. 337–341, 2019.

[3] F. D'Antonio, J. Palacios-Jaraquemada, P. S. Lim et al., "Counseling in fetal medicine: evidence-based answers to clinical questions on morbidly adherent placenta," *Ultrasound in Obstetrics & Gynecology*, vol. 47, no. 3, pp. 290–301, 2016.

[4] X. Xiong, X. M. Zhang, P. Yan, C. Y. Gao, and Q. L. Sun, "The value of ultrasound for early classification in cesarean scar pregnancy," *International Journal of Clinical and Experimental Medicine*, pp. 17267–17273, 2016.

[5] O. Grechukhina, U. Deshmukh, L. Fan et al., "Cesarean scar pregnancy, incidence, and recurrence," *Obstetrics & Gynecology*, vol. 132, no. 5, pp. 1285–1295, 2018.

[6] S. Tanigaki, C. Nagata, K. Ueno et al., "Successful treatment of caesarean scar pregnancies by local treatment only," *Obstetrics*

and Gynecology International, vol. 2017, Article ID 9543570, 2017.

[7] I. E. Timor-Tritsch, N. Khatib, A. Monteagudo, J. Ramos, R. Berg, and S. Kovács, "Cesarean scar pregnancies," *Journal of Ultrasound in Medicine*, vol. 34, no. 4, pp. 601–610, 2015.

[8] J. Xiao, Z. Shi, J. Zhou et al., "Cesarean scar pregnancy: comparing the efficacy and tolerability of treatment with high-intensity focused ultrasound and uterine artery embolization," *Ultrasound in Medicine & Biology*, vol. 43, no. 3, pp. 640–647, 2017.

[9] P. K. Birch, E. Hoffmann, R. C. Larsen, and H. S. Nielsen, "Cesarean scar pregnancy: a systematic review of treatment studies," *Fertility and Sterility*, vol. 105, no. 4, pp. 958–967, 2016.

[10] Q.-L. Sun, X.-H. Wu, L. Luo, D.-M. Ying, Y. Yang, and Z.-Q. Chen, "Characteristics of women with mixed mass formation after evacuation following uterine artery chemoembolization for cesarean scar pregnancy," *Archives of Gynecology and Obstetrics*, vol. 297, no. 4, pp. 1059–1066, 2018.

[11] C. Devane, R. M. Renner, S. Munro et al., "Implementation of mifepristone medical abortion in Canada: pilot and feasibility testing of a survey to assess facilitators and barriers," *Pilot and Feasibility Studies*, vol. 8, no. 5, p. 126, 2019.

[12] D. Ghosh, P. Kumar, and J. Sengupta, "Effect of early luteal phase administration of mifepristone (RU486) on leukaemia inhibitory factor, transforming growth factor beta and vascular endothelial growth factor in the implantation stage endometrium of the rhesus monkey," *Journal of Endocrinology*, vol. 157, no. 1, pp. 115–125, 1998.

[13] E. Kalampokas, I. Boutas, K. Panoulis, and T. Kalampokas, "Novel medical therapy of cesarean scar pregnancy with a viable embryo combining multidose methotrexate and mifepristone: a case report," *Medicine*, vol. 94, no. 41, Article ID e1697, 2015.

[14] Y. Y. Li, Z. Y. Yin, S. Li et al., "Comparison of transvaginal surgery and methotrexate/mifepristone-combined transcervical resection in the treatment of cesarean scar pregnancy," *European Review for Medical and Pharmacological Sciences*, vol. 21, no. 12, pp. 2957–2963, 2017.

[15] I. Durot, S. R. Wilson, and J. K. Willmann, "Contrast-enhanced ultrasound of malignant liver lesions," *Abdominal Radiology*, vol. 43, no. 4, pp. 819–847, 2018.

[16] X. Xiong, P. Yan, C. Gao, Q. Sun, and F. Xu, "The value of contrast-enhanced ultrasound in the diagnosis of cesarean scar pregnancy," *BioMed Research International*, vol. 2016, Article ID 4762785, 2016.

[17] S. J. Zhang, J. Du, R. H. Tian, S. W. Xie, F. H. Li, and Z. Li, "Assessment of the use of contrast enhanced ultrasound in guiding microdissection testicular sperm extraction in non-obstructive azoospermia," *BMC Urology*, vol. 18, no. 1, p. 48, 2018.

[18] S. Y. Kim, S. R. Yoon, M. J. Kim, J. H. Chung, M. Y. Kim, and S. W. Lee, "Cesarean scar pregnancy; Diagnosis and management between 2003 and 2015 in a single center," *Taiwanese Journal of Obstetrics and Gynecology*, vol. 57, no. 5, pp. 688–691, 2018.

[19] World Health Organization Task Force on Post-ovulatory Methods for Fertility Regulation, "Lowering the doses of mifepristone and gameprost for early abortion: a randomised controlled trial," *BJOG*, vol. 108, pp. 738–742, 2001.

[20] C. Papp, F. Schatz, G. Krikun, V. Hausknecht, and C. J. Lockwood, "Biological mechanisms underlying the clinical effects of mifepristone (RU 486) on the endometrium," *Early Pregnancy*, vol. 4, no. 4, pp. 230–239, 2000.

[21] L.-P. Song, S.-Y. Tang, C.-L. Li, L.-J.-G.-Y.-K. Zhou, and X.-T. Mo, "Early medical abortion with self-administered low-dose mifepristone in combination with misoprostol," *Journal of Obstetrics and Gynaecology Research*, vol. 44, no. 9, pp. 1705–1711, 2018.

[22] M. Sperandeo, G. Rea, M. A. Grimaldi, F. Trovato, L. M. C. Dimitri, and V. Carnevale, "Contrast-enhanced ultrasound does not discriminate between community acquired pneumonia and lung cancer," *Thorax*, vol. 72, no. 2, pp. 178–180, 2017.

[23] L. P. Fu, "Therapeutic approach for the cesarean scar pregnancy," *Medicine*, vol. 97, no. 18, Article ID e0476, 2018.

Hindawi
Contrast Media & Molecular Imaging
Volume 2020, Article ID 5418364, 10 pages
https://doi.org/10.1155/2020/5418364

Research Article

A Nomogram Based on Radiomics with Mammography Texture Analysis for the Prognostic Prediction in Patients with Triple-Negative Breast Cancer

Xian Jiang,[1,2] **Xiuhe Zou**,[3] **Jing Sun**,[4] **Aiping Zheng**,[5] **and Chao Su**[6]

[1]*Clinical Research Center for Breast, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu, China*
[2]*Laboratory of Tumor Targeted and Immune Therapy, Clinical Research Center for Breast, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University and Collaborative Innovation Center, Chengdu, China*
[3]*Department of Thyroid Surgery, West China Hospital, Sichuan University, Chengdu, China*
[4]*Department of Integrated Chinese and Western Medicine, Qingdao Central Hospital, Qingdao University, Qingdao, Shandong, China*
[5]*West China School of Medicine, Sichuan University, Chengdu, China*
[6]*State Key Laboratory of Biotherapy and Cancer Center/Collaborative Innovation Center for Biotherapy, West China Hospital, Sichuan University, Chengdu, China*

Correspondence should be addressed to Chao Su; schbhw@scu.edu.cn

Received 8 June 2020; Accepted 6 July 2020; Published 25 August 2020

Guest Editor: Jianguo Zhou

Objectives. To develop and validate a radiomics-based nomogram with texture features from mammography for the prognostic prediction in patients with early-stage triple-negative breast cancer (TNBC). *Methods*. The study included 200 consecutive patients with TNBC (training cohort: $n = 133$, validation cohort: $n = 67$). A total of 136 mammography-derived textural features were extracted, and LASSO (least absolute shrinkage and selection operator) was applied to select features for building the radiomics score (Rad-score). After univariate and multivariate logistic regression, a radiomics-based nomogram was constructed with independent prognostic factors. The discrimination and calibration power were assessed, and further the clinical applicability of the nomograms was evaluated. *Results*. Among the 136 mammography-derived textural features, fourteen were used to build the Rad-score after LASSO regression. A radiomics nomogram that incorporates Rad-score and pN stage was constructed. This nomogram achieved a C-index of 0.873 (95% CI: 0.758–0.989) for predicting iDFS (invasive disease-free survival), which outperformed the clinical model. Moreover, it is feasible to stratify patients into high-risk and low-risk groups based on the optimal cut-off point of Rad-score. The validations of the nomogram confirmed favorable discrimination and considerable predictive efficiency. *Conclusions*. The radiomics nomogram that incorporates Rad-score and pN stage exhibited favorable performance in the prediction of iDFS in patients with early-stage TNBCs.

1. Introduction

Breast cancer, the most frequently diagnosed malignancy, is the leading cause of cancer-related deaths among women [1]. Triple-negative breast cancer (TNBC), accounting for 15–20% of breast cancers [2], does not benefit from endocrine therapy or classic targeted therapy due to the absence of estrogen receptor (ER), progesterone receptor (PR), and the human epidermal growth factor receptor (HER-2) gene amplification. Despite the attempts of novel therapeutic agents such as immune checkpoint inhibitors [3] and poly (ADP-ribose) polymerase inhibitors [4], traditional cytotoxic chemotherapy is still the mainstream systemic treatment option for TNBC [5], while the clinical outcomes remain the poorest among all molecular subtypes of breast cancers.

Nomogram is a feasible and efficacious statistical predictive tool that incorporates multiple variables of value [6, 7]. Seeking new prognostic factors and embedding them into nomograms is an important research method in the prediction of clinical outcomes. To date, nomogram predictions for the survival of TNBCs merely considered conventional clinical and pathological risk factors [8–10]. While efforts have been made in adding novel prognostic factors such as the expression of HIF-1α and c-myc to nomograms [11], imaging features have not yet been analyzed. Nowadays, the advent of deep-learning-based "radiomics" technology has allowed for the high-throughput extraction of quantitative imaging features from images, thus enhancing the accuracy of diagnosis and prognostic prediction, especially for malignancies [12]. Radiomics-based methods are being widely applied in discrimination of confusing lesions on the images [13, 14]. Mammography, ultrasound, and MRI are the most important diagnostic imaging modalities in the management of breast cancers. And imaging-based radiomics nomograms have been constructed in the prediction of axillary lymph node metastasis in early-stage breast cancer [12, 15, 16].

Mammography, mainly displayed in craniocaudal (CC) and mediolateral oblique (MLO) views, has long been a routine screening method for early detection of breast cancers, typically through detection of characteristic masses, microcalcifications, and/or architectural distortions. Apart from distinct biological properties and clinical activation, TNBCs might be distinguished from non-TNBCs with radiomics features based on mammography [17]. More importantly, mammographic features can further help in differentiating basal-like and normal-like subtypes of TNBCs [18]. These facts suggested that mammographic radiomics features might be a potential prognostic factor for TNBCs.

So far, there has been no radiomics-based study for the prognostic prediction of TNBC to the best of our knowledge. Therefore, this study analyzes the prognostic value of mammography textures using deep-learning strategies and constructs an optimized nomogram for the prognostic prediction of TNBCs.

2. Methods

This study was approved by the West China Hospital Research Ethics Committee (No. 2019[887]). The study only involved retrospective analysis of anonymous data, and in consequence the requirement for informed consent was waived.

2.1. Patient Selection. Between April 14, 2010, and April 17, 2017, a total of 200 consecutive patients with TNBC who were treated at West China Hospital of Sichuan University were retrospectively identified from hospital database. The inclusion criteria were pathologically diagnosed TNBCs with mammography performed within 3 months before surgery. In the determination of TNBCs, statuses of ER, PR, and HER-2 were tested with immunostaining. Uncertain status of HER-2 amplification was confirmed by fluorescence in situ hybridization. The exclusion criteria were as follows: a. ductal carcinoma in situ (DCIS) or Paget's disease without invasive elements; *b*. excisional biopsy prior to mammography; *c*. neoadjuvant therapy prior to mammography; *d*. nonmass lesions, i.e., abnormities visible on either mammographic views, which could not be characterized as a distinct mass because of lack of a conspicuous margin or shape [19], including (1) calcification without clear boundary, (2) architectural distortion, and (3) focal asymmetric density; *e*. patients with negative mammography; *f*. recurrent or metastatic diseases.

The included patients were randomly divided into the training ($n = 133$) and validation ($n = 67$) datasets at a ratio of 2 : 1 using random number table (Supplementary Table 1). The clinicopathological features and treatment strategies of the patients were retrieved from medical records, including age, clinical stage, WHO classification, pathological type, Ki-67 expression, type of surgery, radiation therapy, and chemotherapy.

2.2. Follow-Up. The primary end point of this study was invasive disease-free survival (iDFS). After surgeries, patients were routinely followed up every six months for the first five years and annually ever after. IDFS was defined as the period from the date of diagnosis until the date of ipsilateral invasive breast cancer recurrence, ipsilateral locoregional invasive breast cancer recurrence, contralateral invasive breast cancer, distant recurrence/metastasis, death from any cause, and/or the date of last follow-up.

2.3. Image Acquisition and Texture Feature Extraction. Mammography images were acquired through Mammomat Novation DR systems (SIEMENS, German). Meanwhile, craniocaudal (CC) and mediolateral oblique (MLO) projections of both breasts were acquired for each patient.

Two participants (XJ and JY) who were blinded to patient information independently extracted texture features of mammography images using Local Image Feature Extraction (LIFEx) software (http://www.lifexsoft.org, version 5.10) and were supervised by a senior breast radiologist in case of controversies [20]. Following the software instructions (supplementary data), the regions of interests (ROIs) on both CC and MLO views were carefully drawn along the edge of lesions (Figure 1). Sixty-eight features of each mammographic view were automatically extracted, and 136 features of both CC and MLO projections were used to form the radiomic statistical dataset for subsequent machine-learning analysis. The extracted texture features by LIFEx include the grey level co-occurrence matrix (GLCM), the neighborhood grey-level different matrix (NGLDM), the grey-level run length matrix (GLRLM), and the grey-level zone length matrix (GLZLM). Original data of texture features extracted from 200 patients' mammography images is provided in the Supplementary Table 2. Detailed description of features is provided in the supplementary data.

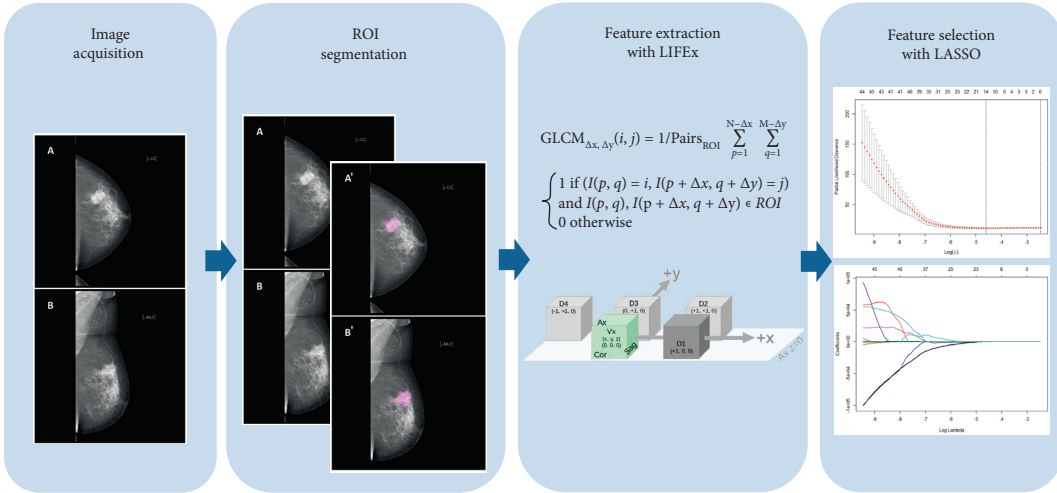


Figure 1: Workflow of radiomic signature building. Abbreviation: ROI = region of interest; LIFEx = Local Image Feature Extraction; LASSO = the least absolute shrinkage and selection operator.

2.4. Selection of Radiomics Signatures. A total of 136 textural features were extracted for each patient. The logistic regression model with least absolute shrinkage and selection operator (LASSO) was adopted to select value features for clinical outcomes with nonzero coefficients (Figure 1) [21]. The calculation formula of Rad-score was subsequently constructed with a linear combination of the selected features of value that were weighted by their respective coefficients [22, 23].

2.5. Construction, Assessment, and Validation of the Radiomics Nomogram. Univariate and multivariate analyses were conducted with Cox proportional hazards regression model. Firstly, a univariate analysis of Rad-score and clinicopathological features for iDFS prediction was performed within the training dataset. For each parameter, hazard ratios (HRs) and 95% confidence intervals (95% CIs) were calculated, and significant variables ($p < 0.05$) in univariate analysis were then tested in backward stepwise selection in the multivariate logistic regression model. Upon the basis of the multivariate regression model, the independent predictive factors ($p < 0.05$) of iDFS were incorporated in the ultimate nomogram, through which a risk score was calculated for each patient.

In order to assess the predictive efficacy of nomogram, the calibration curve was drawn to evaluate the calibration of nomogram and the concordance index (C-index) was applied to further assess its performance.

The internal validation of the nomogram was performed with the validation dataset. Each patient in the validation cohort received a Rad-score with the established formula. The calibration curve and C-index calculation were performed subsequently. Moreover, Kaplan–Meier (K–M) survival curve analysis of iDFS based on the median value of the radiomics nomogram was performed to stratify patients into high- and low-risk subgroups.

2.6. Clinical Utility Evaluation of Radiomics Nomogram. Decision curve analysis (DCA) was conducted to evaluate the clinical significance of radiomics nomogram in predicting iDFS in TNBC patients. More specifically, the net benefits at ranges of threshold probabilities were calculated in the combined training and validation cohorts.

2.7. Statistical Analysis. The comparisons of clinicopathological features between training and validation cohorts were assessed by Student's *t*-test or Mann–Whitney *U* test for continuous variables and Pearson's chi-squared test or Fisher's exact test for categorical variables. The survival curves were displayed with Kaplan–Meier method and differences in survival were examined using the log-rank test. All statistical analysis were performed using *R* software (version 3.5.2). The *R* packages implemented included glmnet, psych, rms, Hmisc, survival, survminer, grid, Lattice, Formula, ggplot2, nomogramEx, tidyverse, dplyr, tidyr, rmda, devtools, rmda, and MASS. A two-tailed $p < 0.05$ was considered to be statistically significant.

3. Result

3.1. Patient Characteristics and iDFS. A total of 200 TNBC patients were included for data analysis and patient characteristics in training and validation cohorts are summarized in Table 1. There were no statistically significant differences in the follow-up duration, clinical-pathological characteristics, or treatment strategies between the two cohorts.

As of the last follow-up, 17 patients (8.50%) had experienced disease relapse or death. The mean iDFS was 17.58 months and the median iDFS was 16.23 (3.10 to 36.43) months. The 1-, 2-, 5-, and 8-year cumulative iDFS of all patients were 2.50% (5/200), 7.00% (14/200), 8.50% (17/200), and 8.50% (17/200), respectively.

3.2. Construction of Radiomics Score. Among the 136 mammography-derived textural features, fourteen were used to build the Rad-score after LASSO regression (Figure 2). The equation of Rad-score was as follows:

Table 1: Clinical and pathological characteristics of patients in training and validation cohorts.

	Training cohort ($n = 133$)	Validation cohort ($n = 67$)	p value
Age (years, mean ± SD)	49.22 ± 9.87	47.79 ± 10.10	0.196
BMI (kg/m [2], mean ± SD)	23.59 ± 3.32	22.88 ± 2.63	0.232
Follow-up (months, mean ± SD)	54.98 ± 21.72	54.39 ± 22.63	0.701
Menopausal status			0.889
Premenopausal	72	39	
Postmenopausal	60	29	
T Stage			0.651
T1	38	23	
T2-4	92	42	
Tx	3	2	
N stage			0.452
N0	95	44	
N1-3	37	23	
WHO classification			0.276
1	0	0	
2	6	6	
3	108	47	
NA	19	14	
Ki-67 status			0.710
<14%	7	2	
≥14%	126	65	
Type of breast surgery			0.928
Mastectomy	120	60	
Lumpectomy	6	4	
NA	7	3	
Neoadjuvant chemotherapy			1.000
Yes	13	7	
No	120	60	
Adjuvant chemotherapy			1.000
Yes	127	64	
No	6	3	
Adjuvant radiotherapy			0.468
Yes	30	19	
No	103	48	

SD: standard deviation; BMI: body mass index; NA: not available.

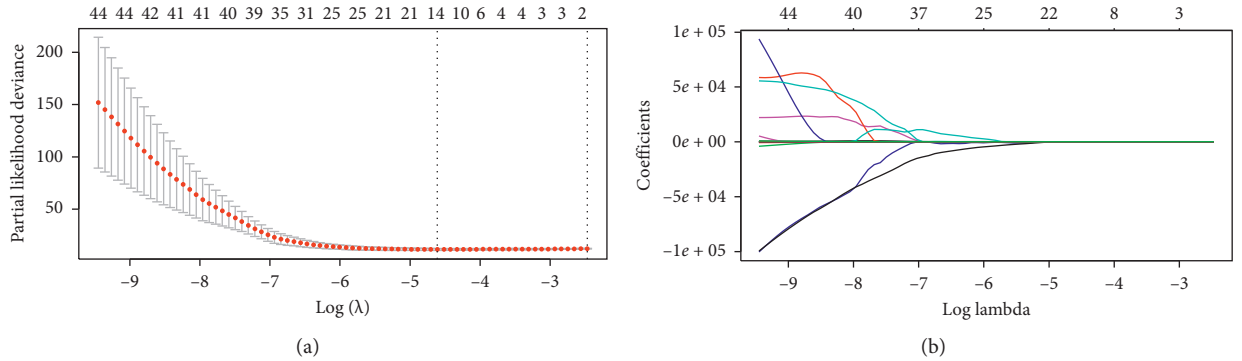


Figure 2: LASSO selection and the predictive efficacy of radiomics features. (a). Tuning parameter (λ) selection with minimum criteria-based 10-fold cross-validation in the LASSO model. Binomial deviances (y-axis) were plotted as a function of log (λ) (lower x-axis), and the upper x-axis represents the average number of predictors. The dotted vertical lines were drawn at the optimal values of λ and the value that gave the minimum average binomial deviance was used to select radiomics features. The optimal λ value of 0.01 (log (λ) = −4.610) was selected. (b) LASSO coefficient profiles of the 136 texture features. Each colored curve represents the trajectory of the change of an independent variable. At the value selected using 10-fold cross-validation, the optimal λ resulted in fourteen coefficients.

$$\begin{aligned}\text{Rad} - \text{score} = {} & \text{CONVENTIONAL\_\#std\_}C^{*}\,(-0.00132048)\\ & + \text{SHAPE\_Volume}\,(mL)\_CC^{*}\,(0.004541706)\\ & + \text{SHAPE\_Volume}\,(\#vx)\_CC^{*}\,(2.016478E-19)\\ & + \text{GLRLM\_SRE\_}CC^{*}\,(-74.75245)\\ & + \text{NGLDM\_Contrast\_}CC^{*}\,(-5.214166)\\ & + \text{NGLDM\_Busyness\_}CC^{*}\,(7.177503)\\ & + \text{CONVENTIONAL\_\# min }\_MLO^{*}\,(0.0006480479)\\ & + \text{HISTO\_Skewness\_}MLO^{*}\,(-0.4479433)\\ & + \text{GLCM\_Correlation\_}MLO^{*}\,(-0.7861222)\\ & + \text{GLCM\_Entropy\_log10\_}MLO^{*}\,(0.4937121)\\ & + \text{GLCM\_Entropy\_log2 (=Joint entropy)\_}MLO^{*}\,(0.0000005908434)\\ & + \text{GLRLM\_GLNU\_}MLO^{*}\,(-0.0004946444)\\ & + \text{GLZLM\_SZE\_}MLO^{*}\,(20.57717)\\ & + \text{GLZLM\_SZHGE\_}MLO^{*}\,(0.00003251064).\end{aligned} \tag{1}$$

3.3. Development and Validation of Nomogram. Univariate Cox regression model analysis indicated that pN stage (HR: 3.964, 95% CI: 1.258–12.490, $p = 0.019$) and Rad-score (HR: 3.071, 95% CI: 1.949–4.840, $p < 0.001$) were associated with iDFS in TNBC patients (Table 2). Subsequently, multivariate Cox regression analysis confirmed pN stage (HR: 3.898, 95% CI: 1.181–12.867, $p = 0.026$) and Rad-score (HR: 3.052, 95% CI: 1.868–4.985, $p < 0.001$) as independent risk factors for iDFS. Accordingly, the nomogram was constructed to quantify probability 1-, 2-, 5-, and 8-year survival (Figure 3(a)). As assessed by the calibration curves, the nomogram revealed good calibration in the prediction of iDFS (Figures 3(b)–3(e)). The C-index for the radiomics nomogram was 0.873 (95% CI: 0.758–0.989) in the training cohort and 0.944 (95% CI: 0.883–1.004) in the validation cohort, while for the clinical nomogram (*N* stage), it was 0.668 (95% CI: 0.531–0.805) in the training and 0.761 (95% CI: 0.608–0.913) in the validation cohort, indicating a better efficacy of radiomics nomogram than clinical nomogram in predicting iDFS.

We were able to stratify patients into high-risk and low-risk groups based on the optimal cut-off point obtained by the "surv_cutpoint" function of the "survminer" *R* package [24, 25]. Patients with a total score higher than or equal to -51.53 were identified as high-risk patients ($n = 21$), and those with a total score less than -51.53 were classified as low-risk patients ($n = 179$). The verification with K-M survival curves showed that the iDFS in the high-risk group was much lower than that in the low-risk group in both the training ($p < 0.0001$) and validation ($p < 0.0001$) cohorts (Figure 4).

3.4. Clinical Utility. The DCAs for the radiomics nomogram and clinical nomogram are presented in Figure 5. The radiomics nomogram adds more net benefit than the "treat all" or "treat none" strategies without limitation on the threshold probability.

4. Discussion

Due to an advanced histological grade, a more aggressive behavior, and the lack of effective therapeutic targets, the clinical outcomes of TNBCs remain the poorest among all molecular subtypes of breast cancers [26]. The recurrence pattern of TNBCs is distinct from non-TNBCs. The risk of disease relapse and death steadily continues for seventeen years after diagnosis in non-TNBCs [27]. However, in patients with TNBCs, the risk of recurrence reached its peak in the first three years after diagnosis and declines thereafter. Recurrence is unlikely to occur in patients who remain disease-free over eight years after diagnosis [27]. Moreover, TNBCs are highly heterogeneous and they respond variously to standard chemo-regimens, resulting in a comparatively wide range of survival [28]. Therefore, a model that validly predicts the survival of TNBCs has significant clinical value. In this study, we assessed the value of radiomics features of mammography in the prognostic prediction of TNBCs, and the results revealed a desired effect. Accordingly, we further established and validated a radiomics-based nomogram to accurately predict the iDFS in TNBC patients. The nomogram contains two indicators, namely, Rad-score and pN stage. With the addition of mammography-derived radiomic score, the nomogram significantly improved the predictive efficiency compared to the existing predictive models.

Mammography has been adopted as a screening modality since 1960s and is currently accepted as the most effective screening approach for breast cancer [29, 30]. There are several discriminative mammographic findings on TNBCs from non-TNBCs. The majorities of TNBCs appear as a mass on mammograms [31]. In a retrospective study with 198 premenopausal patients with breast cancer, all TNBCs were associated with a mass ($n = 33$) while 55% of HER-2+ cancers and 48% of ER + cancers were related [32]. Moreover, TNBCs tend to be less frequently associated with calcifications on mammography compared to non-TNBCs

Table 2: Univariate and multivariate analysis of clinicopathological-radiomics characteristics for prognostic prediction in training cohort.

Variables	Univariate regression		Multivariate regression	
	HR (95% CI)	p value	HR (95% CI)	p value
Age (years)	1.007 (0.951–1.066)	0.812		
BMI (kg/m^2)	0.950 (0.793–1.138)	0.575		
Menopausal status	0.373 (0.101–1.377)	0.139		
pT stage	2.199 (0.482–10.040)	0.309		
pN stage	3.964 (1.258–12.490)	0.019*	3.898 (1.181–12.867)	0.026*
Histological type	3.886×10^{-8} (0-Inf)	0.998		
WHO classification	0.506 (0.065–3.956)	0.516		
Ki-67 status	2.643×10^{7} (0-inf)	0.998		
Type of breast surgery	1.750 (0.226–13.560)	0.592		
Neoadjuvant chemotherapy	1.774 (0.389–8.096)	0.459		
Adjuvant chemotherapy	0.534 (0.069–4.138)	0.548		
Adjuvant radiotherapy	1.117 (0.302–4.127)	0.868		
Radiomics score	3.071 (1.949–4.840)	1.331×10^{-6}***	3.052 (1.868–4.985)	8.380×10^{-6}***

* $p < 0.05$; ** $p < 0.001$. HR: hazard ratio; BMI: body mass index.

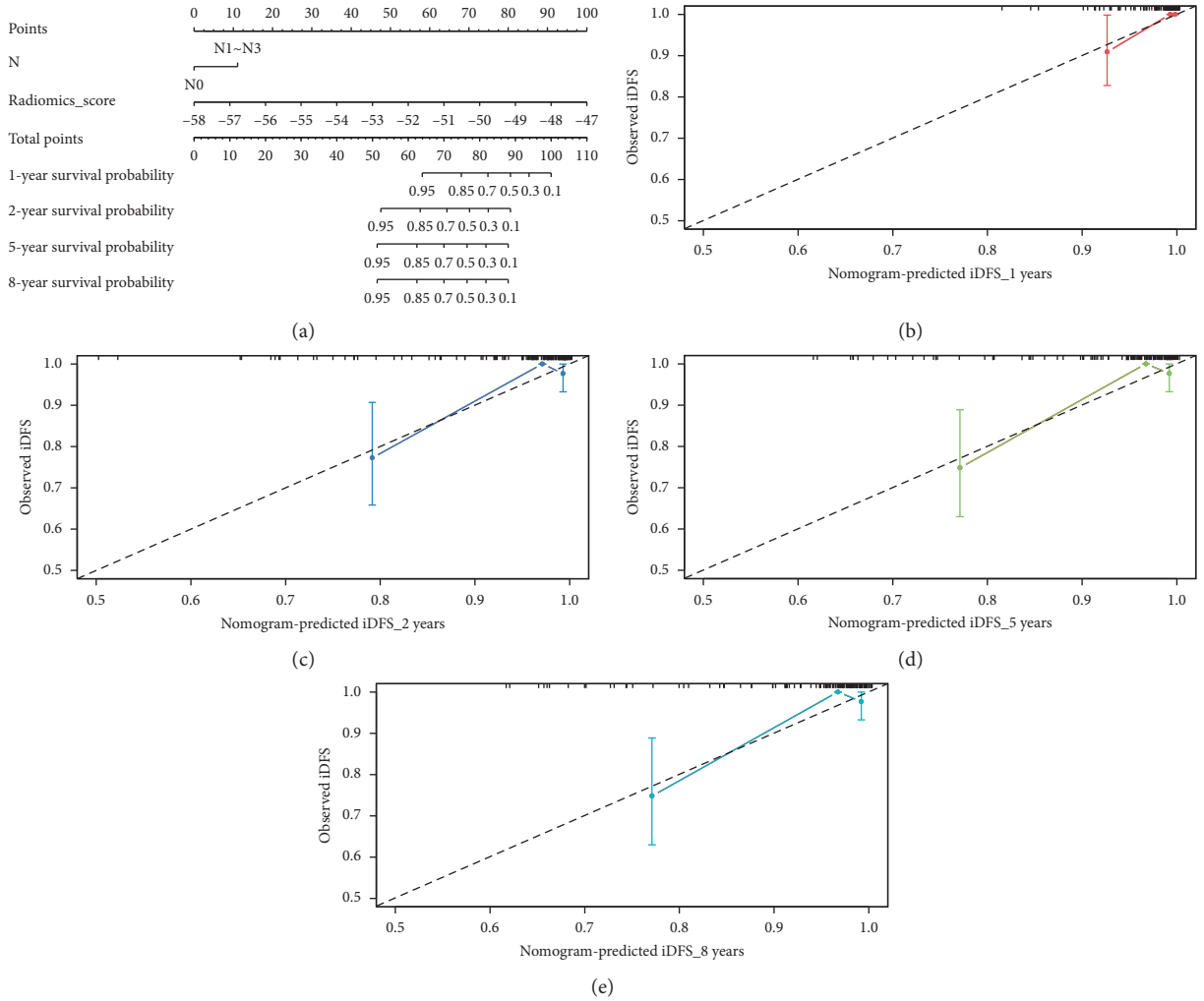


Figure 3: Radiomics nomogram to estimate iDFS for patients with triple-negative breast cancers and its discrimination performance. (a) The radiomics nomogram was developed by incorporating pN stage and radiomics score. b–e. Calibration curves of the nomogram for the estimation of 1-year (b), 2-year (c), 5-year (d), and 8-year (e) iDFS in the training cohort. The diagonal line represents a perfect match between the predicted (x-axis) and actual (y-axis) probabilities, and the colored line represents the predictive performance of the nomogram. The closeness between the two lines indicates the predictive accuracy of the nomogram.

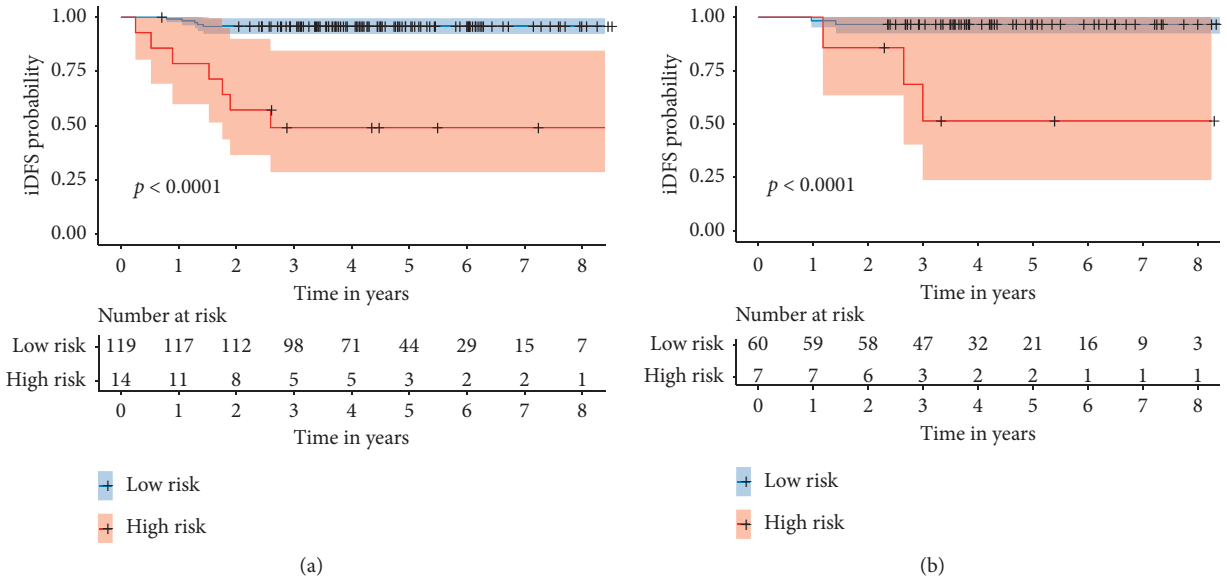


Figure 4: Kaplan–Meier survival analyses of high-risk and low-risk patients in the training (a) and validation cohort (b).

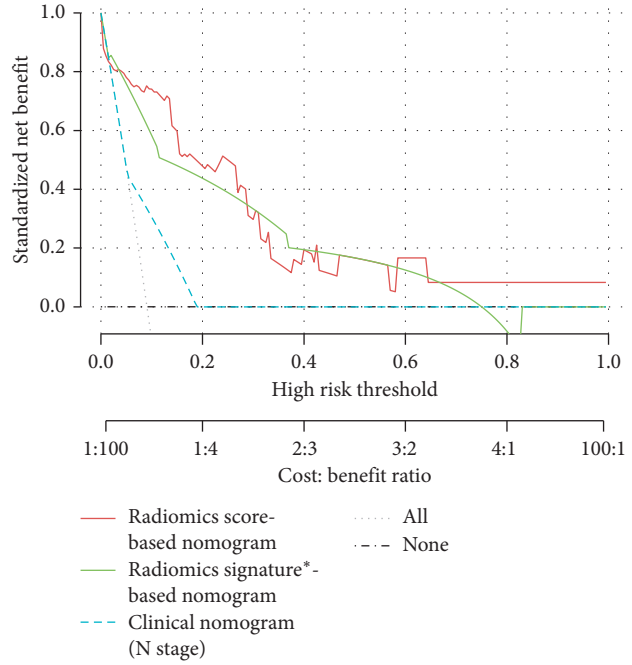


Figure 5: Decision curve analysis for each model in survival prediction in patients with triple-negative breast cancers (TNBCs). The x-axis represents the threshold probability and the y-axis represents the net benefit. The grey line represents the assumption that all patients had experienced events of invasive disease. The black horizontal line represents the assumption that no patient had invasive disease. The decision curves show that using the radiomics nomograms to predict survival adds more benefit for TNBC patients than all other models. *Radiomics score binarily classified as high- and low-risk groups.

[32], potentially because TNBCs progress rapidly to invasive disease without a period of precancerous disease with in situ components that is sufficient to allow calcifications to form [31]. In a study involving 91 TNBC patients, analyses of mammographic features suggested that mass margins were significantly different between basal-like TNBCs and normal-like TNBCs. More specifically, margins of basal-like TNBCs were microlobulated or speculated, whereas those normal-like TNBCs were more likely to be microlobulated [18]. These results suggest that the mammographic presentation of tumors may reflect the histological characteristics and biological behavior of the tumors with sophisticated mechanism.

Despite the wide popularity and the huge number of examinations, which makes mammography a potential data reservoir in big data medicine, the in-depth information

hidden in the images has not been taken advantage of. Nowadays, the advent of deep learning-based radiomics technology has improved the accuracy of disease diagnosis and prognostic prediction [12]. Radiomics is the process of high-throughput extraction of a large number of image features, which converts traditional medical images into high-dimensional data that can be mined, and allows the subsequent quantitative analysis of these data [33]. It helps in the identification of tumor types, noninvasively and quantitatively evaluates tumor biological heterogeneity [34], and therefore optimizes disease detection, diagnosis, treatment response prediction, and prognosis evaluation to promote clinical decision-making. At present, radiomics is mainly applied into the management of malignant tumors, such as liver cancer [35, 36], lung cancer [37], and glioblastoma [38].

In breast cancers, MRI-based radiomics has been evaluated in the prediction of response to neoadjuvant therapy [39–41], prediction of sentinel lymph node metastasis [42], and recognition of molecular subtypes [43]. Moreover, ultrasound-based radiomics has been applied in the prediction of axillary lymph node metastasis [12] and the differential diagnosis between TNBC and fibroadenoma [44]. Mammography-based radiomics has been adopted in breast cancer diagnosis [45]. However, few studies have applied radiomics in prognostic prediction of patients with breast cancer. This study analyzed the prognostic value of mammography textures and constructed an optimized nomogram that incorporated Rad-score and clinical features (pN stage) for the prognostic prediction of TNBCs. According to the results of univariate and multivariate regression, Rad-score rather than *T* stage remained an independent prognostic factor apart from pN stage. The Rad-score was constructed with parameters of the grey level co-occurrence matrix (GLCM), the neighborhood grey-level different matrix (NGLDM), the grey-level run length matrix (GLRLM), and the grey-level zone length matrix (GLZLM). These parameters take into account number, distance, angle, etc. (supplementary data). The Rad-score and the according nomogram exhibited favorable discrimination and considerable predictive efficiency. Based on the Rad-score and the according nomogram, we can predict the prognosis of patients according to the texture features of tumor combined with pN stage and optimize the surgical mode and chemotherapy accordingly. Rad-score successfully identified high-risk patients with poor survival outcomes who need more intensive treatment. Compared to long-term outcome overall survival, iDFS is an end point that avoids extended follow-up and enables early adjustment of treatment. Therefore, our study may provide a more effective tool for making early personalized treatment.

There are several limitations of our study. First, the retrospective nature of the study design might inevitably bias the patient selection. Second, this is a single-center study and the sample size was relatively small. Although the incidence of breast cancers remains the highest among malignancies of all sites in women, the mortality is comparatively lower. In 2019, breast cancer accounts for 30% of the estimated new cases of cancers in the United States, while the estimated deaths only account for 15% [46]. As a consequence in our situation, the number of iDFS events was even smaller which may weaken the statistical power of the current model. Thus, the nomogram should be further validated in prospective studies with larger sample size and longer follow-up. Third, we only included TNBC patients with a definite mass on mammogram to allow a ROI to be drawn. Therefore, our nomogram is not applied to TNBCs with negative mammograms and those with lesions appeared as distortions or calcifications with no clear boundaries.

In conclusion, we established a novel nomogram that can effectively predict the iDFS in TNBC patients by incorporating mammography-based radiomics features into clinicopathological variables. This nomogram mainly benefits primary TNBCs with a mass-like lesion on mammography. Still, future work is required to evaluate its reliability as a routine clinical tool.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Xian Jiang and Xiuhe Zou have contributed equally to this work.

Supplementary Materials

Supplementary data: image feature extraction process, description of 14 mammography-derived textural features. Supplementary Table 1: random number table. Supplementary Table 2: original data of texture features extracted from 200 patients' mammography images. (*Supplementary Materials*)

References

[1] F Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[2] K.-L. Lee, Y.-C. Kuo, Y.-S. Ho, and Y.-H. Huang, "Triple-negative breast cancer: Current understanding and future therapeutic breakthrough targeting cancer stemness," *Cancers*, vol. 11, no. 9, p. 1334, 2019.

[3] S. Adams, M. E. Gatti-Mays, K. Kalinsky et al., "Current landscape of immunotherapy in breast cancer," *JAMA Oncology*, vol. 5, no. 8, p. 1205, 2019.

[4] M. Papadimitriou, G. Mountzios, and C. A. Papadimitriou, "The role of PARP inhibition in triple-negative breast cancer: Unraveling the wide spectrum of synthetic lethality," *Cancer Treatment Reviews*, vol. 67, pp. 34–44, 2018.

[5] A. Naik, A. M. Monjazeb, and J. Decock, "The obesity paradox in cancer, tumor immunology and immunotherapy: Potential

therapeutic implications in triple negative breast cancer," *Frontiers in Immunology*, vol. 10, p. 1940, 2019.

[6] S.-H Chen, Q.-S Wan, D. Zhou et al., "A simple-to-use nomogram for predicting the survival of early hepatocellular carcinoma patients," *Frontiers in Oncology*, vol. 9, p. 584, 2019.

[7] M. Wu, X. Li, T. Zhang, Z. Liu, and Y. Zhao, "Identification of a nine-gene signature and establishment of a prognostic nomogram predicting overall survival of pancreatic cancer," *Frontiers in Oncology*, vol. 9, p. 996, 2019.

[8] T. Wadasadawala, S. Kannan, S. Gudi et al., "Predicting locoregional recurrence risk in T1, T2 breast cancer with 1–3 positive axillary nodes postmastectomy: Development of a predictive nomogram," *Indian Journal of Cancer*, vol. 54, no. 1, pp. 352–357, 2017.

[9] Y. Lin, F. Fu, S. Lin et al., "A nomogram prediction for the survival of patients with triple negative breast cancer," *Oncotarget*, vol. 9, no. 63, pp. 32108–32118, 2018.

[10] D. Dai, H. Jin, and X. Wang, "Nomogram for predicting survival in triple-negative breast cancer patients with histology of infiltrating duct carcinoma: A population-based study," *American Journal of Cancer Research*, vol. 8, no. 8, p. 1576, 2018.

[11] J. Cui and H. Jiang, "Prediction of postoperative survival of triple-negative breast cancer based on nomogram model combined with expression of HIF-1α and c-myc," *Medicine*, vol. 98, no. 40, Article ID e17370, 2019.

[12] F.-H. Yu, J.-X. Wang, X.-H. Ye, J. Deng, J. Hang, and B. Yang, "Ultrasound-based radiomics nomogram: A potential biomarker to predict axillary lymph node metastasis in early-stage invasive breast cancer," *European Journal of Radiology*, vol. 119, Article ID 108658, 2019.

[13] X. Ma, H. Xu, W. Guo et al., "Three-dimensional texture analysis based on PET/CT images to distinguish hepatocellular carcinoma and hepatic lymphoma," *Frontiers in Oncology*, vol. 9, p. 844, 2019.

[14] J. Yang, X. Guo, X. Ou, W. Zhang, and X. Ma, "Discrimination of pancreatic serous cystadenomas from mucinous cystadenomas with CT textural features: Based on machine learning," *Frontiers in Oncology*, vol. 9, p. 494, 2019.

[15] L. Han, Y. Zhu, Z. Liu et al., "Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer," *European Radiology*, vol. 29, no. 7, pp. 3820–3829, 2019.

[16] J. Yang, T. Wang, L. Yang et al., "Preoperative prediction of axillary lymph node metastasis in breast cancer using mammography-based radiomics method," *Scientific Reports*, vol. 9, no. 1, p. 4429, 2019.

[17] H.-X. Zhang, Z.-Q. Sun, Y.-G. Cheng, and G.-Q. Mao, "A pilot study of radiomics technology based on X-ray mammography in patients with triple-negative breast cancer," *Journal of X-Ray Science and Technology*, vol. 27, no. 3, pp. 485–492, 2019.

[18] Z. Zeng, C. J. Hou, Q. H. Hu et al., "Mammography and ultrasound effective features in differentiating basal-like and normal-like subtypes of triple negative breast cancer," *Oncotarget*, vol. 8, no. 45, pp. 79670–79679, 2017.

[19] K.-H. Ko, H.-H. Hsu, J.-C. Yu et al., "Non-mass-like breast lesions at ultrasonography: Feature analysis and BI-RADS assessment," *European Journal of Radiology*, vol. 84, no. 1, pp. 77–85, 2015.

[20] C. Nioche, F. Orlhac, S. Boughdad et al., "LIFEx: A freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity," *Cancer Research*, vol. 78, no. 16, pp. 4786–4789, 2018.

[21] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[22] W. Cai, B. He, M. Hu et al., "A radiomics-based nomogram for the preoperative prediction of posthepatectomy liver failure in patients with hepatocellular carcinoma," *Surgical Oncology*, vol. 28, pp. 78–85, 2019.

[23] S. Wu, J. Zheng, Y. Li et al., "A radiomics nomogram for the preoperative prediction of lymph node metastasis in bladder cancer," *Clinical Cancer Research*, vol. 23, no. 22, pp. 6904–6911, 2017.

[24] X. Li, Y. Yuan, J. Ren, Y. Shi, and X. Tao, "Incremental prognostic value of apparent diffusion coefficient histogram analysis in head and neck squamous cell carcinoma," *Academic Radiology*, vol. 25, no. 11, pp. 1433–1438, 2018.

[25] A. Kassambara, C. Gourzones-Dmitriev, S. Sahota et al., "A DNA repair pathway score predicts survival in human multiple myeloma: the potential for therapeutic strategy," *Oncotarget*, vol. 5, no. 9, pp. 2487–2498, 2014.

[26] J. Reis-Filho and A.N.J. Tutt, "Triple negative tumours: A critical review," *Histopathology*, vol. 52, no. 1, pp. 108–118, 2008.

[27] R. Dent, M. Trudeau, K. I. Pritchard et al., "Triple-negative breast cancer: Clinical features and patterns of recurrence," *Clinical Cancer Research*, vol. 13, no. 15, pp. 4429–4434, 2007.

[28] H. Na, J. Han, N.-L. Ka et al., "High expression of NR1D1 is associated with good prognosis in triple-negative breast cancer patients treated with chemotherapy," *Breast Cancer Research*, vol. 21, no. 1, p. 127, 2019.

[29] A. B. Miller, "The role of screening mammography in the era of modern breast cancer treatment," *Climacteric*, vol. 21, no. 3, pp. 204–208, 2018.

[30] S. Shapiro, W. Venet, P. Strax, L. Venet, and R. Roeser, "Ten-to fourteen-year effect of screening on breast cancer mortality," *Journal of the National Cancer Institute*, vol. 69, no. 2, pp. 349–355, 1982.

[31] Y. Kojima and H. Tsunoda, "Mammography and ultrasound features of triple-negative breast cancer," *Breast Cancer*, vol. 18, no. 3, pp. 146–151, 2011.

[32] W.-T. Yang, M. Dryden, K. Broglio et al., "Mammographic features of triple receptor-negative primary breast cancers in young premenopausal women," *Breast Cancer Research and Treatment*, vol. 111, no. 3, pp. 405–410, 2008.

[33] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: Images are more than pictures, they are data," *Radiology*, vol. 278, no. 2, pp. 563–577, 2015.

[34] H. J. Aerts, E. R. Velazquez, R. T. Leijenaar et al., "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nature Communications*, vol. 5, no. 1, p. 4006, 2014.

[35] B.-H. Zheng, L.-Z. Liu, Z.-Z. Zhang et al., "Radiomics score: A potential prognostic imaging feature for postoperative survival of solitary HCC patients," *BMC Cancer*, vol. 18, no. 1, p. 1148, 2018.

[36] C. Yuan, Z. Wang, D. Gu et al., "Prediction early recurrence of hepatocellular carcinoma eligible for curative ablation using a Radiomics nomogram," *Cancer Imaging*, vol. 19, no. 1, p. 21, 2019.

[37] G. Lee, H. Y. Lee, H. Park et al., "Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art," *European Journal of Radiology*, vol. 86, pp. 297–307, 2017.

[38] A. Kotrotsou, P. O. Zinn, and R. R. Colen, "Radiomics in brain tumors," *Magnetic Resonance Imaging Clinics of North America*, vol. 24, no. 4, pp. 719–729, 2016.

[39] J. Zhou, J. Lu, C. Gao et al., "Predicting the response to neoadjuvant chemotherapy for breast cancer: wavelet transforming radiomics in MRI," *BMC Cancer*, vol. 20, no. 1, p. 100, 2020.

[40] R. W. Y. Granzier, T. J. A van Nijnatten, H. C Woodruff, M. L. Smidt, and M. B. I Lobbes, "Exploring breast cancer response prediction to neoadjuvant systemic therapy using MRI-based radiomics: A systematic review," *European Journal of Radiology*, vol. 121, Article ID 108736, 2019.

[41] N. Braman, P. Prasanna, J. Whitney et al., "Association of peritumoral radiomics with tumor biology and pathologic response to preoperative targeted therapy for HER2 (ERBB2)–positive breast cancer," *JAMA Network Open*, vol. 2, no. 4, Article ID e192561, 2019.

[42] J. Liu, D. Sun, L. Chen et al., "Radiomics analysis of dynamic contrast-enhanced magnetic resonance imaging for the prediction of sentinel lymph node metastasis in breast cancer," *Frontiers in Oncology*, vol. 9, p. 980, 2019.

[43] W. Li, K. Yu, C. Feng, and D. Zhao, "Molecular subtypes recognition of breast cancer in dynamic contrast-enhanced breast magnetic resonance imaging phenotypes from radiomics data," *Computational and Mathematical Methods in Medicinea*, vol. 2019, 2019.

[44] S. E. Lee, K. Han, J. Y. Kwak, E. Lee, and E.-K. Kim, "Radiomics of US texture features in differential diagnosis between triple-negative breast cancer and fibroadenoma," *Scientific Reports*, vol. 8, no. 1, pp. 1–8, 2018.

[45] N. Mao, P. Yin, Q. Wang et al., "Added value of radiomics on mammography for breast cancer diagnosis: A feasibility study," *Journal of the American College of Radiology*, vol. 16, no. 4, pp. 485–491, 2019.

[46] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," *CA: A Cancer Journal for Clinicians*, vol. 69, no. 1, pp. 7–34, 2019.