



Application of radiomics in predicting the preoperative risk stratification of gastric stromal tumors

Li Yang 
Chong-Fei Ma 
Yang Li 
Chun-Ran Zhang 
Jia-Liang Ren 
Gao-Feng Shi 

PURPOSE

The stomach is the most common site of gastrointestinal stromal tumors. In this study, clinical model, radiomics models, and nomogram were constructed to compare and assess the clinical value of each model in predicting the preoperative risk stratification of gastric stromal tumors (GSTs).

METHODS

In total, 180 patients with GSTs confirmed postoperatively pathologically were included. 70% of patients was randomly selected from each category as the training group (n = 126), and the remaining 30% was stratified as the testing group (n = 54). The image features and texture characteristics of each patient were analyzed, and predictive model was constructed. The image features and the rad-score of the optimal radiomics model were used to establish the nomogram. The clinical application value of these models was assessed by the receiver operating characteristic curve and decision curve analysis. The calibration of each model was evaluated by the calibration curve.

RESULTS

The area under the curve (AUC) value of the nomogram was 0.930 (95% confidence interval [CI]: 0.886-0.973) in the training group and 0.931 (95% CI: 0.869-0.993) in the testing group. The AUC values of the training group and the testing group calculated by the radiomics model were 0.874 (95% CI: 0.814-0.935) and 0.863 (95% CI: 0.765-0.960), respectively; the AUC values calculated by the clinical model were 0.871 (95% CI: 0.811-0.931) and 0.854 (95% CI: 0.760-0.947).

CONCLUSION

The proposed nomogram can accurately predict the malignant potential of GSTs and can be used as repeatable imaging markers for decision-making to predict the risk stratification of GSTs non-invasively and effectively before surgery.

Approximately 1%-3% of all gastrointestinal tumors are gastrointestinal stromal tumors (GISTs), but GISTs are the most common gastrointestinal mesenchymal tumors.¹ It is considered that all GISTs have some malignant tendency.² Surgical resection is the preferred treatment for GISTs. Decision-making, including the choice of preoperative targeted therapy and postoperative adjuvant treatment, depends on the risk classification of GISTs.³ Preoperative prediction of stromal tumor risk classification has important significance for decision-making regarding treatment. According to 2008 National Institute for Integrative Healthcare standard, the risk of invasive clinical procedures can be stratified according to tumor size, mitotic count, and pathogenic site. Histological specimens are usually obtained through surgery, but surgery has the disadvantages of invasiveness and treatment delay. Furthermore, patients with surgical contraindications require a biopsy to determine the pathology, which may lead to tumor rupture and peritoneal metastasis, aggravating the risk of recurrence.⁴ The stomach is one of the most common GIST sites, accounting for 70% of the total.¹ Although the risk of recurrence and the rate of progression and metastasis of gastric stromal tumors (GSTs) are lower than those of non-GSTs of the same size and mitotic count, the risk of GSTs rupture is significantly increased,² and biopsy is likely to aggravate the risk of tumor seeding.³ Therefore, it is of great clinical value to

From the Hebei Medical University Fourth Affiliated Hospital and Hebei Provincial Tumor Hospital (L.Y., C.-F.M., Y.L., C.-R.Z., G.-F.S. ✉ gaofengs62@sina.com), Shijiazhuang, China and GE Healthcare (J.L.R.), China.

Received 20 January 2021; revision requested 22 February 2021; last revision received 21 May 2021; accepted 2 June 2021.

Publication date: 1 December 2022.

DOI: 10.5152/dir.2022.21033

You may cite this article as: Yang L, Ma C, Li Y, Zhang C, Ren J, Shi G. Application of radiomics in predicting the preoperative risk stratification of gastric stromal tumors. *Diagn Interv Radiol.* 2022;28(6):532-539.

develop a non-invasive and highly repeatable method that can be used for preoperative risk grading assessments of GSTs.

The characteristics of traditional computed tomography (CT) images, such as the location, size, growth pattern of the lesion, distant metastasis and cystic necrosis, are helpful to initially predict the degree of malignancy of GSTs.⁶⁻⁸ Radiomics, which is a state-of-the-art method, can quantitatively extract and analyze the pixel distribution and mutual relationship of the lesion area through postprocessing technology on medical images. It has been widely used in the judgment of the malignancy of tumors and in the evaluation of treatment response, which can reflect the potential biological characteristics and heterogeneity of tumors more objectively.^{9,10} Previous studies have shown that compared with clinical factors, the radiomic characteristics revealed by CT can be used to predict the risk classification of GSTs.¹¹⁻¹⁵ Feng et al.¹⁶ explored the clinical value of texture analysis of small bowel stromal tumors as potential biomarkers for risk stratification. However, it is still unknown whether radiomics can be used for preoperative GST malignancy assessments. The clinical model, radiomics model, and nomogram used in this work were established to evaluate the predictive value for the risk classification of GSTs. We further compare each model that has the best clinical diagnostic performance to support clinical treatment decisions.

Methods

Patients

This retrospective study was approved by our institutional review board (2016MEC111). The requirement for

informed consent was waived for this retrospective study. This retrospective study enrolled pathologically confirmed GSTs from June 2016 to October 2020. The inclusion criteria were as follows: (1) patients who underwent radical resection (laparoscopy or open surgery), (2) patients who underwent standard contrast-enhanced CT before surgical resection, and (3) patients who did not receive neoadjuvant therapy. The exclusion criteria were as follows: (1) minor stromal tumor (tumors with a diameter of <10 mm), which is insufficient to contain a region of interest (ROI), (2) poor filling of the gastric which is not sufficient to indicate a lesion during CT examinations, and (3) non-GSTs. The patient recruitment flowchart is presented in Figure 1. Finally, 180 patients (78 females and 102 males; mean age, 61.5 ± 9.2 years; age range, 34-84 years) with histopathologically proven primary GSTs were included. To make sure that the ratio of high-malignant risk and low-malignant risk GSTs of training group and testing group is nearly equal, 70% was randomly selected from each category as the training group (n = 126), and the remaining 30% was stratified as the testing group (n = 54). The main clinical symptoms were abdominal pain (15 cases), gastrointestinal bleeding (45 cases), abdominal distension (24 cases), abdominal discomfort (24 cases), acid reflux/heartburn (8 cases), choking (2 cases), and diarrhea (1 case). The other 61 patients were asymptomatic, and their tumors were unintentionally discovered during physical examination.

Computed tomography image acquisition

GE REVOLUTION Spiral CT (GE Healthcare), the second generation SOMATOM Definition Flash CT scanner of Siemens, and the SOMATOM Sensation Open CT scanner were used for CT scanning. Patients were instructed to fast 6-8 h before the scan and to drink 1000 mL of warm water 0.5 h before the scan. Routine plain scanning was done followed by enhanced scanning. Using a MEDTRON high-pressure syringe, the patient was administered an intravenous injection of 1.5 mL of the contrast agent iodohydrin per kg of body weight, and the injection rate was controlled at 3.0 mL/s. The contrast medium was injected for a 35 s arterial phase scan and 70 s venous phase scan. The relevant parameters of REVOLUTION GE spiral CT are as follows: layer spacing: 5 mm, layer thickness: 5 mm, pitch 0.993 : 1, smart mA, 100 or 120 kV, image reconstruction using a 1 mm thin layer, multiplanar reconstruction, etc. The scanning parameters of the second-generation SOMATOM Definition Flash CT scanner and SOMATOM Sensation Open CT scanner of Siemens (Siemens Healthcare) are as follows: the tube voltage is 120 kV, the tube current is 110 mA, the pitch is 0.2, the scanning matrix is set at 512 × 512, the normal scanning layer thickness is 5.0 mm, standard algorithm reconstruction is performed, and the reconstructed layer thickness is 1.0-1.25 mm. The mediastinum window width is 350 Hounsfield units (HU), and the window level is 40 HU.

Main point

- The proposed nomogram can accurately predict the malignant potential of gastric stromal tumor (GST) and can be used as an effective tool to guide preoperative clinical decisions.
- Above three machine learning algorithms, logistic regression model has shown best performance in the prediction of the risk classification of GST.
- The size of the tumors and ulceration might be helpful to predict the risk classification of GST.
- Different from previous studies, whose subjects were gastrointestinal tract, the research object of this study is the stomach.

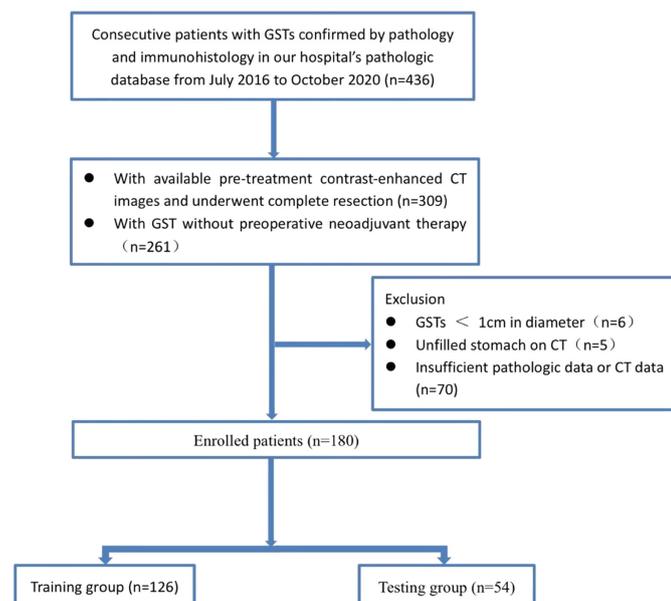


Figure 1. Flowchart of the study population. CT, computed tomography; GST, gastric stromal tumors.

Risk classification standard

MKI67, c-kit (CD117), and NU-4A1 (CD34) expression was evaluated in all lesions. According to a revised version of the National Institutes of Health (NIH) standard proposed by Joensuu in 2008,² 2 pathologists classified the tumor as very low risk, low, medium, and high, based on its size (maximum diameter), mitosis rate (mitosis per 50 high-magnification fields), tumor site, and tumor rupture. Risk stratification was performed according to the NIH risk classification of GISTs and invasive diagnostic criteria for grading postoperative pathological changes. The classification criteria of GISTs are shown in Table 1. In this study, no tumor rupture was observed during the operation. The very low- and low-risk groups were combined and regarded as the low-malignant-potential group, and the moderate- and high-risk groups were combined and regarded as the high-malignant-potential group.

Clinical model establishment

The imaging information for each patient was extracted from the Picture Archiving and Communications System. The clinical information obtained from the electronic medical record included the patient's name, age, sex, clinical symptoms, and presence of gastrointestinal bleeding. CT-imaging features included the maximum diameter of the primary lesion, growth pattern, margin, and shape; presence of ulceration, calcification, necrosis, or cystic degeneration; and pattern of enhancement,

degree of enhancement, and presence of enlarged vessels feeding or draining the mass (EVFDM, present or absent). The maximum diameter of the tumor was measured in sagittal, coronal, and cross-sections by multiplanar reconstruction. The margin of the lesion is considered to be either poorly demarcated or well demarcated. The morphology of the tumor was regarded as irregular and regular; a regular shape was defined as round or oval, and an irregular shape was defined as a lobular appearance. The growth mode was divided into endoluminal, exophytic, and mixed. Intracavitary growth was considered if the mass was confined to the gastrointestinal cavity under the mucosa without breaking through the gastric contour and growing outside the cavity. Exophytic growth was considered if the mass had invaded past the cavity but had not exhibited expansive growth into the cavity. Mixed growth was considered if the mass spanned the stomach wall and exhibited growth into and out of the cavity at the same time, similar to a dumbbell. Ulceration was defined when there are localized defects on the mucosal surface of tumor. If a low attenuation area was observed within the tumor in the portal venous phase, corresponding to a HU between 0 and 30 HU, and it did not show any changes between scans or any associated increase (up to 5 HU), then it was considered to be an area of tumor necrosis and was classified as follows: does not exist, mild (<50% of the tumor necrosis), or severe (>50% of tumor necrosis). Cystic

degeneration refers to cystic degeneration within the mass, which is defined as a completely liquefied cystic area with clear and sharp borders; solid portions can be observed with a density close to that of water (CT value of 0-20 HU). The enhancement pattern was divided subjectively into homogeneous enhancement and heterogeneous enhancement. A total of 180 patients were observed and recorded by an imaging physician with over 10 years of clinical experience and an imaging physician with 3 years of clinical experience. When there was disagreement, a consensus was reached through consultation.

Segmentation of gastrointestinal stromal tumors

The ROI was delineated using ITK-SNAP software (version 3.8.0, <https://www.itk-snap.org>). All image processing was performed for venous phase CT images. ROIs were manually segmented to encompass as much of the lesion as possible into 2 consecutive slices of portal phase images while leaving a 1-mm peripheral margin outside the ROI to exclude adjacent blood vessels, the gastric wall, and gastric contents. Finally, the ROI of all layers was fused. Segmentation of GISTs in ITK-SNAP software is shown in Figure 2. The same physician randomly selected 30 cases for the second delineation of the target area to evaluate the repetitive consistency of the delineation of the target area.

Image preprocessing and radiomics model construction

Texture feature extraction was carried out in 3-dimensional gray space. To avoid the influence of tumor orientation dependence and different layer thicknesses, isotropic voxel (1 × 1 × 1 mm³) resampling was performed by the linear interpolation method. Image filtering included the Laplace of Gaussian (LoG) and wavelet transforms. The filter parameters of LoG σ were 3, 5, and 7. Wavelet transformation included LLL, LLH, LHL, HLL, LHH, HLH, HHL, and HHH.

Texture features extraction and calculation were carried out automatically by using the PyRadiomics platform.¹⁷ In total, 1130 features were extracted from all original GIST images and filtered images and included (1) image intensity histogram characteristics, (2) shape and size characteristics, (3) gray-level cooccurrence matrix (GLCM), (4) gray-level run-length matrix, (5) gray-level size zone matrix, (6) gray-level

Table 1. Clinical and demographic characteristics and some CT features of the patients

	Low-malignant-potential GSTs (n = 86)	High-malignant-potential GSTs (n = 94)	Total (n = 180)	P
Sex				.6021
Female	47 (54.7%)	55 (58.5%)	102 (56.7%)	
Male	39 (45.3%)	39 (41.5%)	78 (43.3%)	
Age				.6092
Mean (SD)	59.733 (9.307)	60.330 (9.793)	60.044 (9.542)	
Size (%)				<.001 ¹
≤2 cm	18 (20.9%)	6 (6.4%)	24 (13.3%)	
2.1-5 cm	68 (79.1%)	25 (26.6%)	93 (51.7%)	
5.1-10 cm	0 (0.0%)	52 (55.3%)	52 (28.9%)	
>10 cm	0 (0.0%)	11 (11.7%)	11 (6.1%)	
Ulceration (%)				<.001 ²
–	59 (68.6%)	29 (30.9%)	88 (48.9%)	
+	27 (31.4%)	65 (69.1%)	92 (51.1%)	

CT, computed tomography; GSTs, gastric stromal tumors; SD, standard deviation.

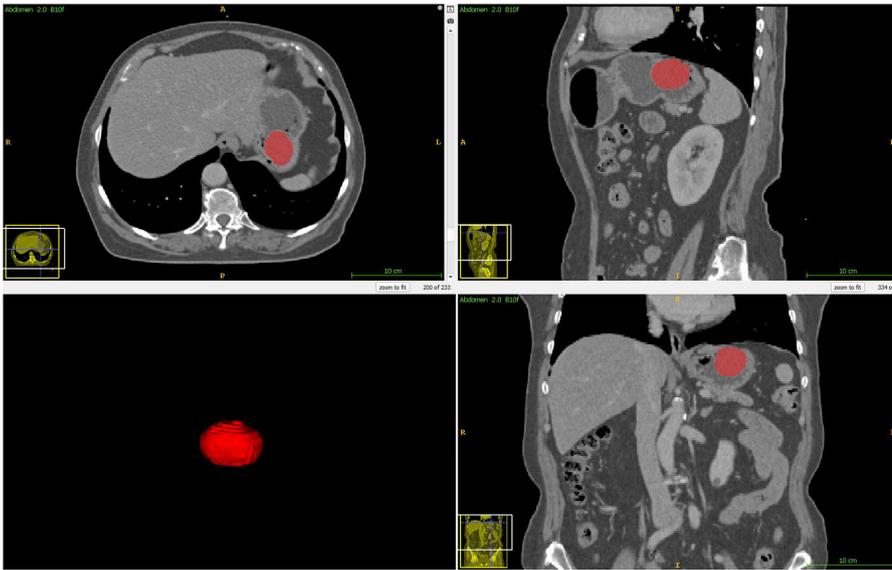


Figure 2. Segmentation of GSTs in ITK-SNAP software. There is manual segmentation of the sagittal, coronary, axial slices, and 3-dimensional volumetric reconstruction. GST, gastric stromal tumors.

dependency matrix, and (7) neighborhood gray-tone difference matrix.

To avoid model overfitting, the following methods were used to select the

most important features: (1) the Wilcoxon rank-sum test and Mann-Whitney U test were performed on retention characteristics, with $P < .001$ indicating significance

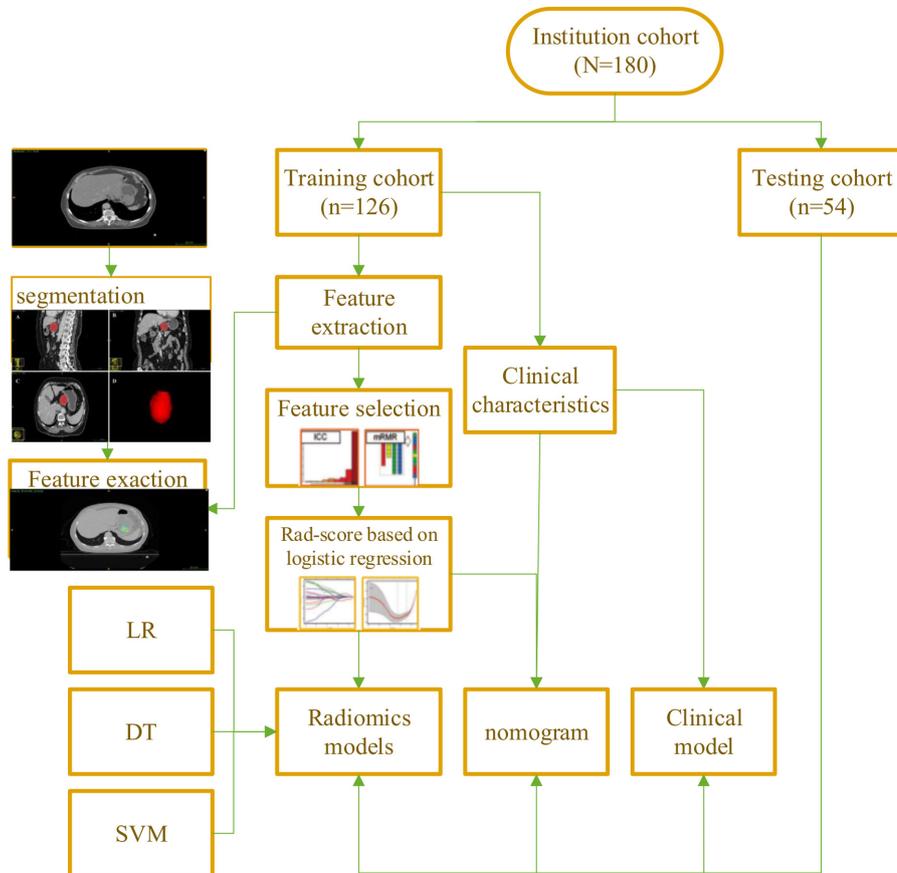


Figure 3. Schematic diagram of the construction proposed models. Based on the malignant potential profile, tumor area segmentation and feature extraction were performed. GIST patients were categorized into training group and testing group. GIST, gastrointestinal stromal tumor; LR, logistic regression; DT, decision tree; SVM, support vector machine.

(Bonferroni-adjusted P value). (2) Spearman rank correlation analysis was used to select the independent impact factors of the characteristics (excluding $R^2 > 0.90$). (3) The least absolute shrinkage and selection operator (LASSO) method was used for feature selection.¹⁸

Radiomics models included logistic regression, support vector machine (SVM) classifiers, and decision tree (DT).¹⁹ The rad-score of each patient was calculated by the probability of model output. Schematic diagram of the proposed workflow is presented in Figure 3.

To predict the risk stratification of GST patients quantitatively, this study developed a nomogram that incorporated the most important predictive indicators of clinical models and the rad-score of the optimal radiomics model.²⁰ The area under the receiver operating characteristic (ROC) curve (AUC) value, accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were obtained to quantitatively evaluate the fitting degree of the prediction model. The clinical efficacy of each model was evaluated by decision curve analysis (DCA) through the net benefit to the patient at different threshold probabilities. The reliability of each prediction model was evaluated by calibration curves.

Statistical analysis

All the statistical analyses were performed in R software (version 3.6.1, <https://www.rproject.org>). The LASSO, logistic, SVM, DT classifier, and ROC curve analysis were conducted based on “glmnet,” “base,” “e1071,” “rpart,” and “pROC” packages, respectively. Independent sample t or Mann-Whitney U tests were applied appropriately for continuous variables across the groups, and χ^2 or Fisher exact tests were used to assess the differences in patient categorical variables. Intra-class correlation coefficient (ICC) was used to evaluate the features’ consistency. Features with ICC value greater than 0.75 indicated good consistency. Delong test was used to evaluate the difference in AUC values among the various models. Consistency between the observers was evaluated by calculating the ICC. A 2-tailed P value of less than .05 indicates a significant statistical difference.

Results

In this study, there were 180 patients with GSTs. In the dichotomous prediction

Table 2. Diagnostic efficiency of the different models		
Models	Training group	Testing group
LR		
AUC	0.874 (0.814-0.935)	0.863 (0.765-0.960)
ACC	0.810 (0.730-0.874)	0.778 (0.644-0.880)
Sensitivity	0.746 (0.603-0.873)	0.806 (0.516-0.968)
Specificity	0.873 (0.714-0.953)	0.739 (0.522-0.914)
Positive predictive value	0.855 (0.826-0.873)	0.806 (0.727-0.833)
Negative predictive value	0.775 (0.738-0.790)	0.739 (0.667-0.778)
DT		
AUC	0.888 (0.839-0.938)	0.816 (0.701-0.930)
ACC	0.810 (0.730-0.874)	0.778 (0.644-0.880)
Sensitivity	0.667 (0.543-0.770)	0.710 (0.271-0.848)
Specificity	0.952 (0.787-1.000)	0.870 (0.609-1.000)
Positive predictive value	0.933 (0.919-0.942)	0.880 (0.737-0.898)
Negative predictive value	0.741 (0.702-0.750)	0.690 (0.609-0.719)
SVM		
AUC	0.874 (0.812-0.935)	0.857 (0.757-0.957)
ACC	0.817 (0.739-0.881)	0.778 (0.644-0.880)
Sensitivity	0.714 (0.524-0.841)	0.742 (0.452-0.968)
Specificity	0.921 (0.746-0.984)	0.826 (0.609-0.958)
Positive predictive value	0.900 (0.868-0.914)	0.852 (0.778-0.882)
Negative predictive value	0.763 (0.723-0.775)	0.704 (0.636-0.734)
Clinical model		
AUC	0.871 (0.811-0.931)	0.854 (0.760-0.947)
ACC	0.810 (0.730-0.874)	0.796 (0.665-0.894)
Sensitivity	0.635 (0.095-0.746)	0.667 (0.485-0.818)
Specificity	0.984 (0.863-1.000)	1.000 (0.685-1.000)
Positive predictive value	0.976 (0.857-0.979)	1.000 (1.000-1.000)
Negative predictive value	0.729 (0.703-0.733)	0.656 (0.567-0.656)
Nomogram		
AUC	0.930 (0.886-0.973)	0.931 (0.869-0.993)
ACC	0.865 (0.793-0.919)	0.870 (0.751-0.946)
Sensitivity	0.857 (0.714-0.937)	0.839 (0.581-0.968)
Specificity	0.873 (0.587-0.968)	0.913 (0.478-1.000)
Positive predictive value	0.871 (0.849-0.881)	0.929 (0.900-0.937)
Negative predictive value	0.859 (0.804-0.871)	0.808 (0.688-0.821)

AUC, area under curve; ACC, accuracy; DT, decision tree; LR, logistic regression; SVM, support vector machine.

(0.871, 95% CI: 0.811-0.931) and in the testing group (0.854, 95% CI: 0.760-0.947) for low-malignant-potential and high-malignant-potential GSTs, respectively.

After feature selection, the last 2 texture features were retained; (1) LoG filter image with 5.0 mm informational measure of correlation belongs to GLCM (log5_glcm_IMC) and (2) LoG filter image with 5.0 mm gray-level non-uniformity belongs to GLSZM (log5_glszm_GLNU). Moreover, the same radiologist had good consistency in delineating the target area before and after feature selection, with ICCs of 0.98 and 0.79, respectively. The features were incorporated into the radiomics models. The logistic regression model showed excellent and stable clinical diagnostic performance compared with SVM and DT.

The lesion size, the presence of the ulceration, and the rad-score of the LR model were included in the establishment of the nomogram, as shown in Figure 4. Moreover, the rad-score was obtained from logistic regression model by the weighted linear combination of all key features. The rad-score equation was given as follows:

$$\text{rad-score} = 0.0788 - 1.578 \times \log_5_{\text{glcm_IMC}} + 1.51 \times \log_5_{\text{glszm_GLNU}} \quad (1)$$

The univariate analysis of those independent predictors is shown in Table 3. All predictors have statistical significance difference between low malignant and high malignant in both groups.

ROC curves, calibration curves, and DCA curves assessed the clinical model, radiomics model, and nomogram. Figure 5 shows that the nomogram is superior to the clinical model and radiomics model with a good calibration effect in the calibration curve and has higher net benefit in the DCA curve. The sensitivity, specificity, and accuracy of the training group and testing group of each model were shown in Table 4. As shown in Figure 5 and Table 4, the AUC values of the training group and the testing group obtained in the nomogram were 0.930 (95% CI: 0.886-0.973) and 0.931 (95% CI: 0.869-0.993), respectively, which were higher than those obtained in the radiomics model and the clinical model. However, there was no significant difference in the AUCs between the clinical model and the radiomics model ($P = .90$). Thus, the nomogram has high accuracy and stable fitting effect.

experiment, the very low-risk and low-risk groups were combined and regarded as the low-malignant-potential group (86 cases), and the moderate-risk and high-risk groups were combined and regarded as the high-malignant-potential group (94 cases).

A total of 20 clinical features were obtained by combining clinical data with CT-imaging features. Multivariate logistic

regression analysis revealed that only the maximum tumor diameter and presence of ulceration remained independent risk factors for malignant GSTs, both of which were incorporated in the clinical model and nomogram. The clinical and demographic characteristics and CT features used to establish the clinical model are shown in Table 2. The clinical model predicted the AUCs and 95% CI in the training group

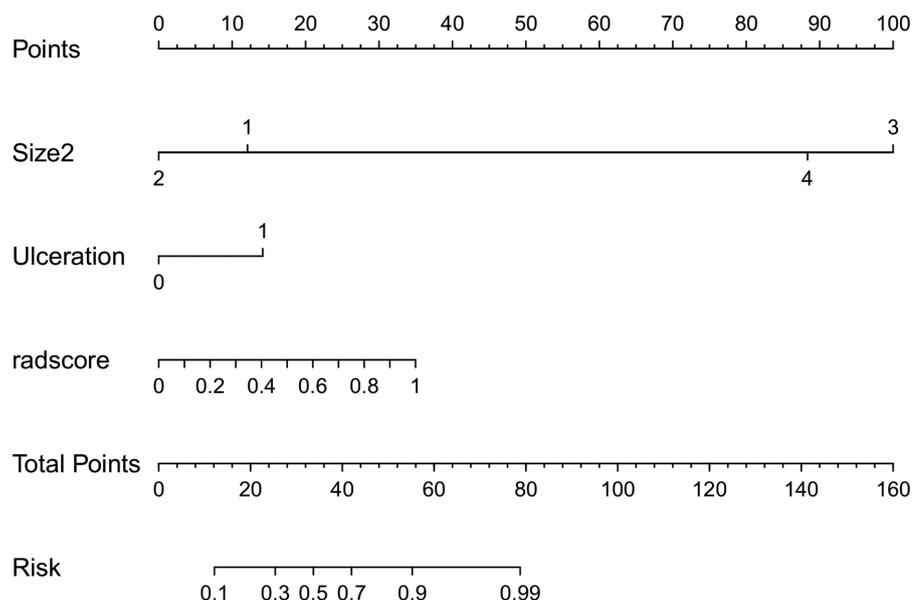


Figure 4. Image of nomogram. The rad-score, tumor diameter, and ulceration were used to construct the nomogram. The probability of each predictor can be converted into scores according to the first scale "Points" at the top of the nomogram. Add up the scores of these predictors in "Total Points." Then the corresponding prediction probability at the bottom of the nomogram reflects the malignancy of the tumor. The cut-off point of our nomogram is 0.486. Those total prediction probability beyond the cut-off point is diagnosed as having high malignant potential.

Discussion

A clinical model, a radiomics model, and a nomogram were developed in this study as repeatable imaging markers for decision-making to predict the risk stratification of GSTs before surgery. The nomogram showed excellent clinical diagnostic efficacy and calibration ability.

A previous study suggested that image signs are predictive factors of GIST risk classification.⁸ Zhou et al.⁸ incorporated tumor

lesion size, growth mode, and EVFDM into the logistic regression model, and further, ROC curve showed that the AUC of the obtained logistic regression model was 0.806 (95% CI: 0.727-0.885). Tumor size and mucous ulcers were included in the logistic regression model in this study, and ROC analysis showed that the AUCs were 0.871 (95% CI: 0.811-0.931) and 0.854 (95% CI: 0.760-0.947) in the logistic regression model for the training and testing groups, respectively. Yang et al.²¹ and Kim et al.²² noted a

relationship between mucosal ulceration and a high mitotic index. Mucosal ulcers can increase the odds of GIST recurrence and indicate a poor prognosis. Joensuu et al.² also proposed that tumor ulceration is related to a poor prognosis, which is consistent with the results of this study. Based on the texture analysis, the radiomics model does not rely only on subjective judgment of CT features. The relevance of the informational measure (informational measure of correlation) and gray level non-uniformity is incorporated into the radiomics model. ROC analysis showed that in the radiomics model, the AUCs were 0.874 (95% CI: 0.814-0.935) and 0.863 (95% CI: 0.765-0.960) for the training and testing groups, respectively. There was no significant difference in AUCs by the Delong test between the clinical model and the radiomics model ($P = .54$ in training group, 0.18 in testing group). However, the radiomics model excludes the subjective factors of the imaging physicians and factors affecting the difference in experience levels.

In this study, we developed a nomogram for the discrimination of malignant GSTs. This model has achieved satisfactory predictive efficacy and can likely be used as repeatable imaging markers for decision-making to predict the risk stratification of GSTs non-invasively and effectively before surgery. Together, clinical factors and conventional radiological examination information can provide comprehensive evidence for clinical decisions. The tumor size, mucosal ulceration, and rad-score of the groups

Table 3. Univariate analysis of predictors in both groups

	Training group				Testing group		
	Low-malignant-potential GSTs (n=63)	High-malignant-potential GSTs (n=63)	Cut-off	P	Low-malignant-potential GSTs (n=23)	High-malignant-potential GSTs (n=31)	P
Size (%)			-	<.001			<.001
≤2 cm	14 (22.2)	4 (6.3)			4 (17.4)	2 (6.5)	
2.1-5 cm	49 (77.8)	18 (28.6)			19 (82.6)	7 (22.6)	
5.1-10 cm	0 (0.0)	34 (54.0)			0 (0.0)	18 (58.1)	
>10 cm	0 (0.0)	7 (11.1)			0 (0.0)	4 (12.9)	
Ulceration (%)			-	<.001			.002
-	42 (66.7)	19 (30.2)			17 (73.9)	10 (32.3)	
+	21 (33.3)	44 (69.8)			6 (26.1)	21 (67.7)	
Log5_glcm_IMC (mean ± SD)	0.533 (1.021)	-0.533 (0.629)	0.168	<.001	0.636 (1.209)	-0.476 (0.850)	<.001
Log5_glszm_GLNU (mean ± SD)	-0.431 (0.521)	0.431 (1.169)	-0.118	<.001	-0.453 (0.406)	0.690 (1.679)	<.001
Rad-score (mean ± SD)	-1.414 (1.675)	1.572 (2.186)	0.343	<.001	-1.608 (2.181)	1.871 (3.079)	<.001

GST, gastric stromal tumor; glcm, gray-level cooccurrence matrix; IMC, informational measure of correlation; SD, standard deviation; glszm, gray-level size zone matrix; GLNU, gray-level non-uniformity belongs.

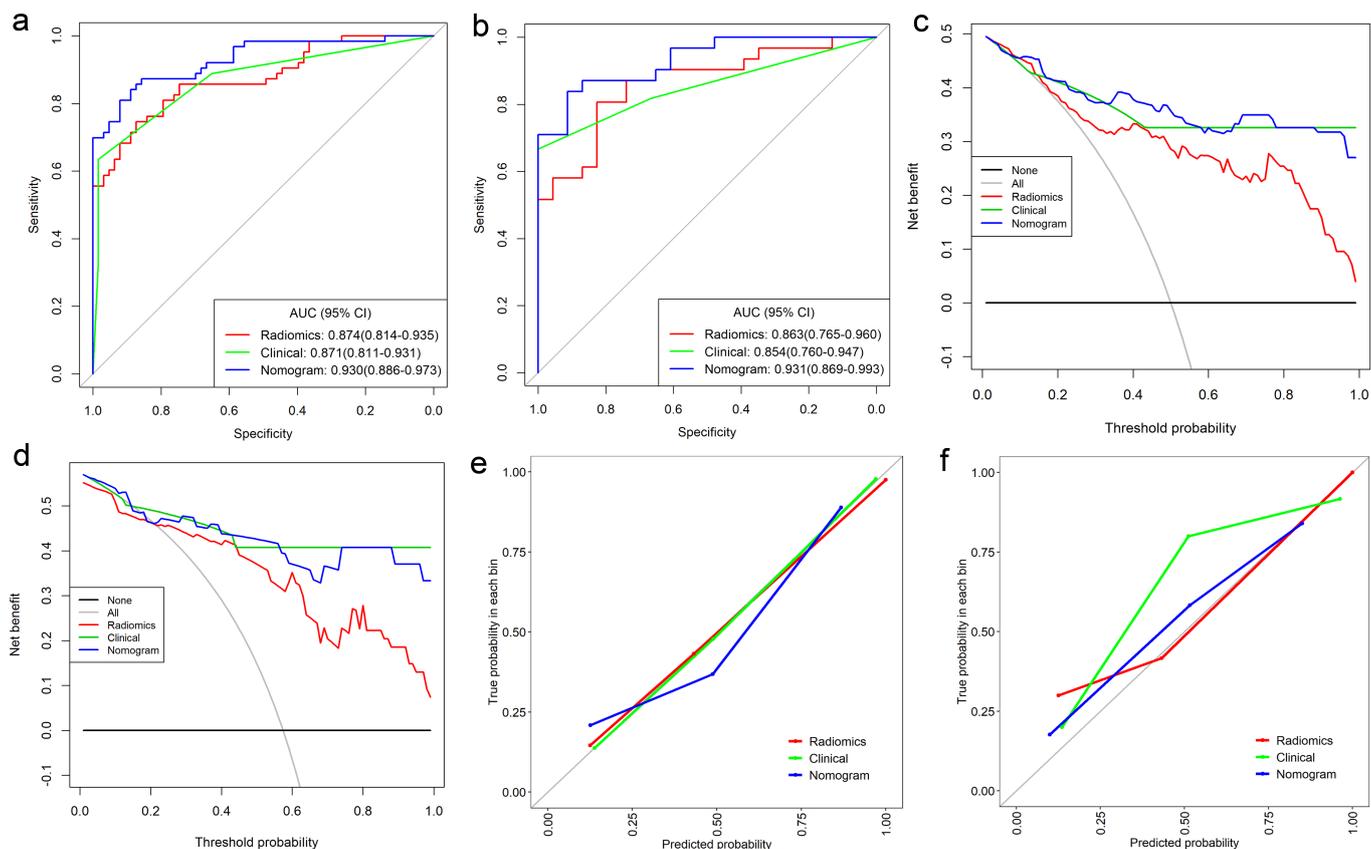


Figure 5. (a, b). The receiver operating characteristic (ROC) curve of each model (a—training group, b—testing group) in evaluating the clinical efficacy of each model quantitatively. (c, d). Decision curve analysis (DCA) of each model (c—training group, d—testing group) in evaluating the clinical efficacy of each model by quantifying the net benefit to the patient under different threshold probabilities. (e, f). Calibration curves of each model (E—training group, F—testing group) in assessing the reliability and accuracy of each prediction model. The red line, green line, and blue line represent the clinical model, radiomics model, and nomogram, respectively.

Table 4. Diagnostic efficiency of different models		
Models	Training group	Testing group
LR		
AUC	0.874 (0.814-0.935)	0.863 (0.765-0.960)
ACC	0.810 (0.730-0.874)	0.778 (0.644-0.880)
Sensitivity	0.746 (0.603-0.873)	0.806 (0.516-0.968)
Specificity	0.873 (0.714-0.953)	0.739 (0.522-0.914)
Positive predictive value	0.855 (0.826-0.873)	0.806 (0.727-0.833)
Negative predictive value	0.775 (0.738-0.790)	0.739 (0.667-0.778)
DT		
AUC	0.888 (0.839-0.938)	0.816 (0.701-0.930)
ACC	0.810 (0.730-0.874)	0.778 (0.644-0.880)
Sensitivity	0.667 (0.543-0.770)	0.710 (0.271-0.848)
Specificity	0.952 (0.787-1.000)	0.870 (0.609-1.000)
Positive predictive value	0.933 (0.919-0.942)	0.880 (0.737-0.898)
Negative predictive value	0.741 (0.702-0.750)	0.690 (0.609-0.719)
SVM		
AUC	0.874 (0.812-0.935)	0.857 (0.757-0.957)
ACC	0.817 (0.739-0.881)	0.778 (0.644-0.880)
Sensitivity	0.714 (0.524-0.841)	0.742 (0.452-0.968)

(Continued)

were included in the radiomics model and the nomogram to establish a combined nomogram model, which can obtain AUCs higher than those obtained by the clinical model and can be used for the radiomics model of AUC values. The Delong test showed that the differences in the AUC values among the nomogram, clinical model, and radiomics model were statistically significant, indicating that the performance of the nomogram used for GST risk classification prediction is superior to that of the clinical model and that of the radiomics model. In this study, the proposed nomogram incorporated the rad-score of the radiomics model and integrated the texture features of radiomics analyses. For example, in the present study, 1 high-malignant-potential case with a 1.8-cm-sized tumor was diagnosed accurately by the nomogram, while it was classified as the low-malignant-potential GST and underestimated by the radiomics model and clinical model. The integration of tumor texture information and traditional imaging information allows

Table 4. Diagnostic efficiency of different models (Continued)

Models	Training group	Testing group
Specificity	0.921 (0.746-0.984)	0.826 (0.609-0.958)
Positive predictive value	0.900 (0.868-0.914)	0.852 (0.778-0.882)
Negative predictive value	0.763 (0.723-0.775)	0.704 (0.636-0.734)
Clinical model		
AUC	0.871 (0.811-0.931)	0.854 (0.760-0.947)
ACC	0.810 (0.730-0.874)	0.796 (0.665-0.894)
Sensitivity	0.635 (0.095-0.746)	0.667 (0.485-0.818)
Specificity	0.984 (0.863-1.000)	1.000 (0.685-1.000)
Positive predictive value	0.976 (0.857-0.979)	1.000 (1.000-1.000)
Negative predictive value	0.729 (0.703-0.733)	0.656 (0.567-0.656)
Nomogram		
AUC	0.930 (0.886-0.973)	0.931 (0.869-0.993)
ACC	0.865 (0.793-0.919)	0.870 (0.751-0.946)
Sensitivity	0.857 (0.714-0.937)	0.839 (0.581-0.968)
Specificity	0.873 (0.587-0.968)	0.913 (0.478-1.000)
Positive predictive value	0.871 (0.849-0.881)	0.929 (0.900-0.937)
Negative predictive value	0.859 (0.804-0.871)	0.808 (0.688-0.821)

AUC, area under curve; ACC, accuracy; LR, logistic regression; DT, decision tree; SVM, support vector machine.

for a better quantitative analysis of diagnostic efficacy than the previous nomograms incorporating single texture features.

This study still has some limitations. First, similar to previous studies, this study is a retrospective study, which makes it difficult to avoid selection bias, so further prospective studies are needed. In addition, the CT images were obtained with different CT examination equipment, so there are certain confounding factors. We preprocessed the images, which greatly reduced these confounding factors. Second, this study lacks external validation; we plan to carry out further multicenter research in the future. Third, the prediction model does not consider genetic mutations. Further studies are needed to explore the relationship between gene mutations and radiographic characteristics.

Conclusion

In this study, a predictive nomogram based on image signs and texture analysis was developed and verified. The proposed nomogram is superior to the radiomics model and clinical model and can be conveniently and accurately applied for prediction of the risk stratification of GST preoperatively.

References

- Joensuu H, Vehtari A, Riihimäki J, et al. Risk of recurrence of gastrointestinal stromal tumor after surgery: an analysis of pooled population-based cohorts. *Lancet Oncol.* 2012;13(3):265-274. [CrossRef]
- Joensuu H. Risk stratification of patients diagnosed with gastrointestinal stromal tumor. *Hum Pathol.* 2008;39(10):1411-1419. [CrossRef]
- Li J, Ye Y, Wang J, et al. Chinese consensus guidelines for diagnosis and management of gastrointestinal stromal tumor. *Chin J Cancer Res.* 2017;29(4):281-293. [CrossRef]
- Demetri GD, Von MM, Antonescu CR, et al. NCCN Task Force report: update on the management of patients with gastrointestinal stromal tumors. *J Natl Compr Canc Netw.* 2010;8(suppl 2):S1-41.
- Wei S-C, Xu L, Li W-H, et al. Risk stratification in GIST: shape quantification with CT is a predictive factor. *Eur Radiol.* 2020;30(4):1856-1865. [CrossRef]
- Miettinen M, Lasota J. Gastrointestinal stromal tumors (GISTs): definition, occurrence, pathology, differential diagnosis and molecular genetics. *Pol J Pathol.* 2003;54(1):3-24.
- Mazzei MA, Cioffi Squitieri N, Vindigni C, et al. Gastrointestinal stromal tumors (GIST): a proposal of a "CT-based predictive model of Miettinen index" in predicting the risk of malignancy. *Abdom Radiol (NY).* 2020;45(10):2989-2996. [CrossRef]
- Zhou C, Duan X, Zhang X, Hu H, Wang D, Shen J. Predictive features of CT for risk stratifications in patients with primary gastrointestinal stromal tumor. *Eur Radiol.* 2016;26(9):3086-3093. [CrossRef]

- Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* 2012;48(4):441-446. [CrossRef]
- Liu LH, Liu YH, Xu L, et al. Application of texture analysis based on apparent diffusion coefficient maps in discriminating different stages of rectal cancer. *J Magn Reson Imaging.* 2017;45(6):1798-1808. [CrossRef]
- Liu S, Pan X, Liu R, et al. Texture analysis of CT images in predicting malignancy risk of gastrointestinal stromal tumors. *Clin Radiol.* 2018;73(3):266-274. [CrossRef]
- Chen T, Ning Z, Xu L, et al. Radiomics nomogram for predicting the malignant potential of gastrointestinal stromal tumors preoperatively. *Eur Radiol.* 2019;29(3):1074-1082. [CrossRef]
- Ren C, Wang S, Zhang S. Development and validation of a nomogram based on CT images and 3D texture analysis for preoperative prediction of the malignant potential in gastrointestinal stromal tumors. *Cancer Imaging.* 2020;20(1):5. [CrossRef]
- Choi IY, Yeom SK, Cha J, 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. *Abdom Radiol (NY).* 2019;44(7):2346-2356. [CrossRef]
- Zhang QW, Zhou XX, Zhang RY, et al. Comparison of malignancy prediction efficiency between contrast and noncontrast CT-based radiomics features in gastrointestinal stromal tumors: a multicenter study. *Clin Transl Med.* 2020;10(3):e291. [CrossRef]
- Feng C, Lu F, Shen Y, 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.* 2018;18(1):46. [CrossRef]
- van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 2017;77(21):e104-e107. [CrossRef]
- Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med.* 2007;26(30):5512-5528. [CrossRef]
- Noble WS. What is a support vector machine? *Nat Biotechnol.* 2006;24(12):1565-1567. [CrossRef]
- Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol.* 2008;26(8):1364-1370. [CrossRef]
- Yang TH, Hwang JI, Yang MS, et al. Gastrointestinal stromal tumors: computed tomographic features and prediction of malignant risk from computed tomographic imaging. *J Chin Med Assoc.* 2007;70(9):367-373. [CrossRef]
- Kim HC, Lee JM, Kim KW, et al. Gastrointestinal stromal tumors of the stomach: CT findings and prediction of malignancy. *AJR Am J Roentgenol.* 2004;183(4):893-898. [CrossRef]