



# Radiomics analysis of contrast-enhanced computerized tomography for differentiation of gastric schwannomas from gastric gastrointestinal stromal tumors

Cui Zhang<sup>1</sup> · Chongwei Wang<sup>2</sup> · Guoqun Mao<sup>1</sup> · Guohua Cheng<sup>3</sup> · Hongli Ji<sup>3</sup> · Linyang He<sup>3</sup> · Yang Yang<sup>4</sup> · Hongjie Hu<sup>5</sup> · Jian Wang<sup>1</sup>

Received: 16 August 2023 / Accepted: 20 November 2023 / Published online: 9 February 2024

© The Author(s) 2024

## Abstract

**Purpose** To assess the performance of radiomics-based analysis of contrast-enhanced computerized tomography (CE-CT) images for distinguishing GS from gastric GIST.

**Methods** Forty-nine patients with GS and two hundred fifty-three with GIST were enrolled in this retrospective study. CT features were evaluated by two associate chief radiologists. Radiomics features were extracted from portal venous phase images using Pyradiomics software. A non-radiomics dataset (combination of clinical characteristics and radiologist-determined CT features) and a radiomics dataset were used to build stepwise logistic regression and least absolute shrinkage and selection operator (LASSO) logistic regression models, respectively. Model performance was evaluated according to sensitivity, specificity, accuracy, and receiver operating characteristic (ROC) curve, and Delong's test was applied to compare the area under the curve (AUC) between different models.

**Results** A total of 1223 radiomics features were extracted from portal venous phase images. After reducing dimensions by calculating Pearson correlation coefficients (PCCs), 20 radiomics features, 20 clinical characteristics + CT features were used to build the models, respectively. The AUC values for the models using radiomics features and those using clinical features were more than 0.900 for both the training and validation groups. There were no significant differences in predictive performance between the radiomic and clinical data models according to Delong's test.

**Conclusion** A radiomics-based model applied to CE-CT images showed comparable predictive performance to senior physicians in the differentiation of GS from GIST.

**Keywords** Radiomics · Schwannoma · Gastrointestinal stromal tumor · Computed tomography

✉ Jian Wang  
wangj\_0309@163.com

Cui Zhang  
zc553217018@163.com

Chongwei Wang  
cadillac\_luck@zju.edu.cn

Guoqun Mao  
maoguoqun123@163.com

Guohua Cheng  
cheng.guohua@jianpeicn.com

Hongli Ji  
ji.hongli@jianpeicn.com

Linyang He  
he.linyang@jianpeicn.com

Yang Yang  
1107019018@qq.com

Hongjie Hu  
hongjiehu@zju.edu.cn

<sup>1</sup> Department of Radiology, TongDe Hospital of ZheJiang Province, No. 234, Gucui Road, Hangzhou 310013, Zhejiang, China

<sup>2</sup> Department of Pathology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

<sup>3</sup> Jianpei Technology, Hangzhou, Zhejiang, China

<sup>4</sup> Department of Radiology, The First Affiliated Hospital of Bengbu Medical College, Bengbu, Anhui, China

<sup>5</sup> Department of Radiology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

## Abbreviations

AUC	Area under the curve
CE-CT	Contrast-enhanced computed tomography
CTU/CTA/CTV	The CT attenuation value in unenhanced phase, arterial phase, and portal venous phase
DEAP/DEPP	Degree of enhancement on arterial phase and portal venous phase
GIST	Gastrointestinal stromal tumor
GLCM	Gray level co-occurrence matrix
GLDM	Gray level dependence matrix
GLRLM	Gray level run length matrix
GLSZM	Gray level size zone matrix
GS	Gastric schwannoma
ICC	Inter-and intra-class correlation coefficient
LASSO	Least absolute shrinkage and selection operator
LD	Long diameter
LD/SD	Ratio of long diameter to short diameter
LN	Lymph node
NGTDM	Neighboring gray tone difference matrices matrix
ROI	Region of interest
PCC	Pearson correlation coefficient
ROC	Receiver operating characteristics
SD	Short diameter

## Introduction

Gastric schwannoma (GS) is usually a benign, neurogenic, and mesenchymal neoplasm derived from the Schwann cells of the Auerbachs nerve plexus, with an incidence rate of 2–17% (Choi et al. 2012). Gastric gastrointestinal tumors (GISTs) originating from the interstitial cells of Cajal have the highest incidence rate among mesenchymal tumors, accounting for 60–70% of occurrences (Gao et al. 2017; Tsuboi et al. 2021). GSs and GISTs share a similar affected population, clinical symptoms, and even computed tomography (CT) imaging characteristics, particularly in large (> 5 cm) tumors (Hong et al. 2008; Ji et al. 2015; Levy et al. 2003; Wang et al. 2019a, b). However, the treatment strategies and prognoses for these two tumors differ substantially. GS, which is almost always a benign tumor, has an excellent prognosis after surgery (Cai et al. 2016). By contrast, GISTs appear potentially malignant and require complete excision. Furthermore, high-risk GISTs should receive imatinib treatment as adjuvant or neoadjuvant therapy (Casali et al. 2022; Li et al. 2017). It is of great clinical

significance to accurately distinguish these two tumors preoperatively.

Numerous studies on the identification of GISTs and GSs on contrast-enhanced (CE)-CT images have been reported (Choi et al. 2012, 2014; He et al. 2017; Liu et al. 2017; Wang et al. 2021; Xu et al. 2022). We previously evaluated five machine learning models to identify the two tumor types on CT image analysis and found that logistic regression classification with six CT features provided the best differentiation of GSs from GISTs, with a diagnostic efficiency greater than traditional multivariate analysis (Wang et al. 2021). Nevertheless, although the abovementioned studies were of value in distinguishing the two kinds of tumors, they all focused on conventional image features recognized by the naked eye and lacked radiomics analysis, thereby relying heavily on the observer's professional level.

Radiomics features of tumors, which are related to the histopathology and physiology of tumors, can quantitatively and objectively evaluate tumor heterogeneity (Liu et al. 2019). Compared with traditional image evaluation, radiomics features can more comprehensively, objectively, and accurately reflect the nature of a lesion, and permit the establishment of a radiomics database for in-depth quantitative research (Gillies et al. 2016). Radiomics-based analysis has been widely applied to various diseases including the differentiation of benign and malignant lung nodules (Zhuo et al. 2021), classification of liver lesions (Wei et al. 2020), gastric cancer staging (Wang et al. 2020; Liu et al. 2020), determination of malignancy (Zhang et al. 2020; Chen et al. 2019, 2021; Wang et al. 2019a), recurrence (Ao et al. 2021), and Ki-67 expression prediction (Feng et al. 2022) of GISTs.

In this study, CT radiomics features together with clinical data and conventional image features were used to build two kinds of models to distinguish GSs from GISTs: a stepwise logistic regression model and a least absolute shrinkage and selection operator (LASSO) logistic regression model.

## Methods

### Patients

This retrospective study received institutional review board and ethics committee approval from Tongde Hospital of Zhejiang Province (Approval No. 2021-040) and informed consent was waived by the ethics review boards. Patients from two centers (Center 1: Tongde Hospital of Zhejiang Province, Center 2: Sir Run Run Shaw Hospital) with either of the two tumor types between January 2015 and August 2022 were identified. The inclusion criteria were: (1) operation-confirmed diagnosis of solitary GIST or schwannoma of the stomach; (2) available detailed clinical data and CE-CT

images (with unenhanced, arterial, and portal phases); (3) an interval time within 15 days between CT imaging and surgery; and (4) a lesion larger than 1 cm and smaller than 10 cm in the long diameter. The exclusion criteria were as follows: (1) lack of CT scan or uncompleted CT data; (2) unavailable clinical data; (3) an interval time more than 15 days between CT scan and surgery; (4) the lesion smaller than 1 cm or larger than 10 cm in the long diameter; and (5) multiple lesions. Process of the patients enrolling is shown in Fig. 1. The final study series consisted of 49 patients with GS (16 men and 33 women; mean age,  $56.51 \pm 10.34$  years) and 253 patients with GIST (131 men and 122 women; mean age,  $59.56 \pm 11.64$  years). From these, 211 patients from Center 1 were assigned to a training cohort and 91 patients from Center 2 to a validation cohort. The data for each patient included non-radiomics dataset (consisting of

clinical baseline characteristics and CT features) and radiomics dataset.

### CT scan protocols

All patients were requested to drink 800–1000 mL of water on an empty stomach to attain sufficient gastric distension before CT examination. Two 64-slice spiral CT scanners (Siemens Healthineers, Forchheim, Germany; or Philips Medical Systems, Cleveland, OH, USA) were used for the abdominal CE-CT examinations. The CT parameters were: tube voltage, 120 kV; tube current, 150–250 mA; tube rotation time, 0.5 s; detector collimation,  $64 \times 0.625$  mm; field of view,  $350 \times 350$  mm; section thickness, 5 mm; and reconstruction interval, 1–1.5 mm. After a routine unenhanced scan, contrast material was injected at a dose of 1.0 mL/kg

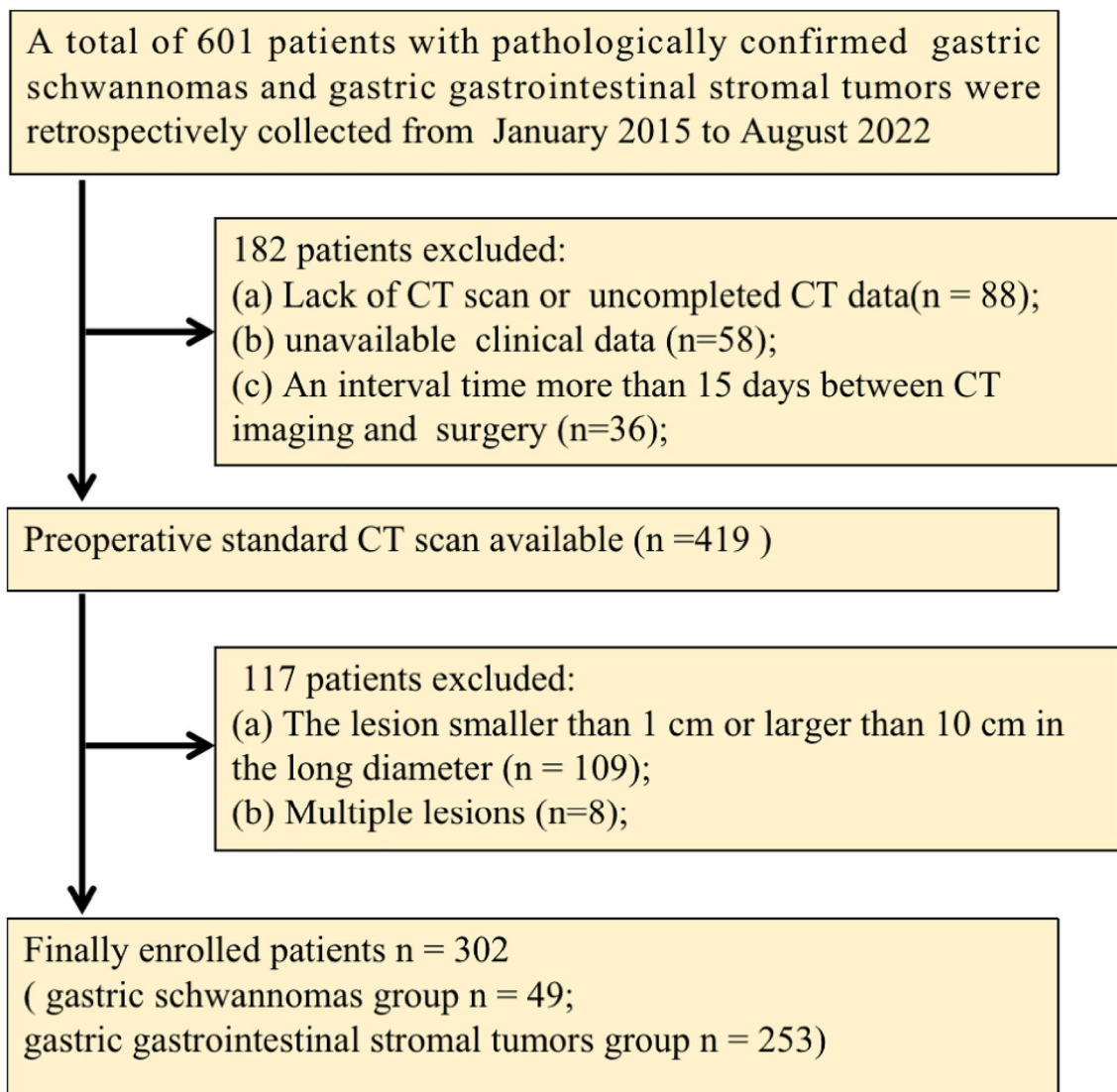


Fig. 1 Process of the patients enrolling

body weight at a rate of 3–4 mL/s, and arterial and portal venous phase imaging were acquired at 30–40 s and 60–70 s after injection.

## Image analysis

All CT images were independently retrospectively reviewed and evaluated by two associate chief radiologists (Radiologist 1 with 13 and Radiologist 2 with 15 years of experience in abdominal imaging). The two physicians were blinded to the surgical and pathological data of each patient, and the final conclusions were achieved through consensus decisions. The evaluated CT features consisted of qualitative and quantitative variables. Qualitative variables included location (cardia, fundus, body, or antrum), grow pattern (endophytic, exophytic, or mixed), contour (round, oval, or irregular), necrosis, calcification, surface ulceration, lymph node (LN), hemorrhage, peritumoral exudation, necrosis under the tumor wall, and intratumoral vessel. Quantitative variables included the CT attenuation value on unenhanced phase (CTU), arterial phase (CTA), and portal venous phase (CTV) imaging, the degree of enhancement on arterial phase (DEAP;  $DEAP = CTA - CTU$ ) and portal venous phase (DEPP;  $DEPP = CTV - CTU$ ) imaging, long diameter (LD), short diameter (SD), and the ratio of long diameter to short diameter (LD/SD). Lymph node was defined as positive when the short diameter of the lymph node was larger than 10 mm.

## Image preprocessing

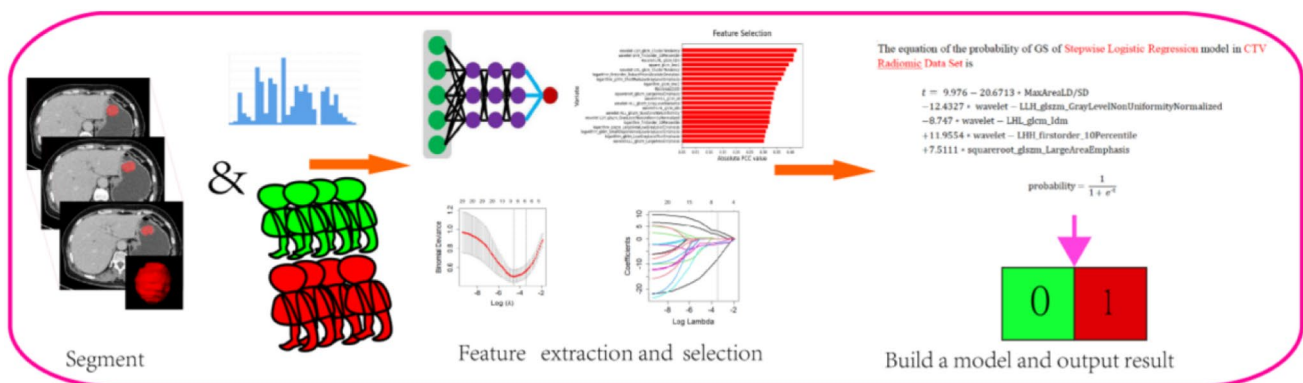
Before image segmentation and feature extraction, it is necessary to preprocess images to reduce the data heterogeneity collected by different CT devices. First, the voxel size should be resampled to  $1 \times 1 \times 1 \text{ mm}^3$ , and the anisotropic images should be homogenized to minimize the dependence of radiomics features on the image voxel size. The voxel

intensity values were discretized using a fixed bin width of 25 HU with the aim of reducing image noise and standardizing the intensity to achieve a stable intensity resolution in all images. The images were normalized and the signal intensity was normalized to 1 to 500 HU to reduce the difference in signal intensity of images acquired by different machines. Z score was used to standardize the image gray value to reduce the influence of the inconsistency of image parameters on the variation of radiomics features.

## Image segmentation and radiomics feature extraction

Figure 2 depicts the workflow of this study. Open-source ITK-SNAP software (version 2.2.0; <http://www.itk-snap.org>) was used for image segmentation. Portal venous phase images were used for image segmentation in this study since the portal venous phase had a better performance than unenhanced and arterial phases in identifying the lesion from the surrounding normal tissue. Region of interests (ROIs) were drawn manually along the margin of tumors in a slice-by-slice manner by two radiologists (with 2 and 4 years of experience in abdominal imaging). The radiomics characteristics were then extracted using Pyradiomics (version 3.0.0; <http://www.radiomics.io/pyradiomics.html>).

A total of 1223 radiomics features were extracted from the tumor region of interests (ROIs) drawn on the portal venous phase imaging: (1) 19 first-order features; (2) 26 size- and shape-related features; (3) 24 Gy level co-occurrence matrix (GLCM) features; (4) 14 Gy level dependence matrix (GLDM) features; (5) 16 Gy level run length matrix (GLRLM) features; (6) 16 Gy level size zone matrix (GLSZM) features; (7) 5 neighboring gray tone difference matrix (NGTDM) features; and (8) 1103 features after transformations including square, square root, exponential, logarithm, wavelet, and Laplacian of Gaussian.



**Fig. 2** Flow diagram of this study. CT images were segmented on portal venous phase images. After preprocessing and feature selection, radiomics features, or clinical characteristics and CT features,

were used to construct stepwise logistic regression and least absolute shrinkage and selection operator (LASSO) logistic regression models for differentiation of GSs from GISTs

To assess inter- and intra-observer reproducibility, 40 ROIs of 2 tumors delineated by 2 radiologists were chosen at random. Inter- and intra-class correlation coefficients (ICCs) were used to evaluate the stability and reproducibility in radiomics feature extraction. The intra-observer ICCs were calculated based on the twice feature extractions by one radiologist. The inter-observer ICCs were evaluated based on the first extracted features by one radiologist and those by another radiologist. The feature with  $ICC > 0.75$  was deemed to have a good repeatability or reproducibility. After inter- and intra-observer analyses, a total of 968 features with  $ICC > 0.75$  were applied for subsequent feature selection.

### Feature selection and model building

First, feature dimensionality reduction was performed before model building. The Pearson correlation coefficient (PCC) between two features was calculated for each pair of features in the non-radiomics and radiomics datasets. These two features were considered to be highly correlated if the PCC larger than 0.8, and one of them was selected for deletion. Then we calculated the PCCs between features and response variables (GSs or GISTs) and selected the 20 features with the largest PCCs.

These 20 features selected from the clinical and radiomics features were used to build 2 kinds of models: a stepwise logistic regression model and a LASSO-logistic regression model. The two models were constructed using data from the training dataset and were tested using the test dataset.

#### Stepwise logistic regression

Logistic regression models were plotted using the R package “glm”. Initially, all 20 variables were input into the model and the variables with the largest  $p$  value were eliminated in each cycle, with the variables with  $p$  values less than 0.1 being continually entered into the univariate analysis until the variables no longer changed.

#### LASSO-logistic regression

The LASSO algorithm was applied with the R package “glmnet”. The most useful predictive features were selected using the LASSO method. Briefly, the optimized hyperparameter  $\lambda$  was determined using tenfold cross-validation via the minimum criteria. Then the features with non-zero coefficients were selected based on the determined optimal  $\lambda$ .

**Statistical analysis** All statistical analyses were performed using R software (version 3.6.3; <http://www.Rproject.org>). Continuous variables are presented as mean  $\pm$  standard deviation. Student’s  $t$  test or the Mann–Whitney  $U$  test

was used to compare continuous variables. Categorical variables are expressed as frequency (percentage) and were analyzed using the Chi-square or Fisher’s exact test. Sensitivity, specificity, and accuracy were calculated using the ‘caret’ package in R. ROC curve analysis was performed with the ‘pROC’ package to evaluate the diagnostic efficacy of the models. Delong’s test was used to compare the AUC values between the different methods. A two-sided  $p$  value of  $< 0.05$  was considered statistically significant.

## Results

### Clinical baseline characteristics and CT findings

Two hundred forty-three patients with tumors (forty-three GSs and two hundred GISTs) from Center 1 and fifty-nine patients with tumors (six GSs and fifty-three GISTs) from Center 2 were included in our series. The clinical baseline characteristics and CT findings of the 302 patients in the training ( $n = 211$ ) and validation ( $n = 91$ ) cohorts are listed in Table 1. For both training and validation sets, the number of underlying diseases, lesion location, necrosis, and necrosis under the tumor wall were significant variables in the univariate analysis between the two tumor types, while growth pattern was significantly different in the training cohort. Although patients with GIST were older than those with GS both in training and validation cohorts, the differences were not statistically significant. There were no significant differences in any of the clinical and CT parameters between the two datasets (all  $p > 0.05$ ), which indicates that the patient allocation to the two datasets was well balanced. The detailed results are summarized in Table 1.

### Feature selection and model building

The clinical features, CT features, and radiomics features were reduced in dimension by eliminating those features with PCC values of more than 0.8 with another feature. The PCC values (between features and GSs) in the 235 remaining radiomics features and 24 clinical characteristics and CT features were then calculated, and the 20 features with the largest PCC values were selected. The 20 features selected in the 2 datasets are shown in Fig. 3. Among the selected radiomics features, wavelet-LLH\_glcm\_ClusterTendency had the largest PCC, and wavelet-LLL\_glszm\_LargeAreaEmphasis the smallest PCC. Among the clinical features, LD/SD showed the largest PCC and tumor marker the lowest PCC.

#### Stepwise logistic regression model

Table 2 presents the five radiomics features (and their coefficients) with  $p < 0.05$  selected from the training dataset

**Table 1** Clinical baseline characteristics and CT findings in training and validation cohorts

	Training cohort(n = 211)			Validation cohort (n91)			
	GS(n = 34)	GIST(n = 177)	P1 value <sup>a</sup>	GS(n = 15)	GIST(n = 76)	P2 value <sup>b</sup>	P3 value <sup>c</sup>
<b>Clinical characteristics</b>							
Gender (male/female)	(11/23)	(91/86)	<b>0.042</b>	(5/10)	(40/36)	0.259	0.900
Age (Mean ± SD)	56.38 ± 10.05	59.70 ± 11.89	0.129	56.80 ± 11.32	59.22 ± 11.12	0.444	0.813
Symptom	(17/59)	(17/118)	0.079	(6/9)	(53/23)	<b>0.039</b>	0.968
Tumor makers	(8/46)	(26/131)	0.833	(4/11)	(20/56)	1.000	1.000
Number of underlying diseases (0/1/2/≥ 3)	(28/4/2/0)	(106/46/18/7)	<b>0.002</b>	(12/3/0/0)	(43/23/8/2)	<b>0.004</b>	0.546
<b>CT features</b>							
Location (cardia/fundus/body/antrum)	(0/2/20/12)	(6/57/97/17)	<b>0.000</b>	(0/0/11/4)	(3/25/42/6)	<b>0.009</b>	0.911
Growth pattern (endophytic/ exophytic/ mixed)	(4/18/12)	(66/67/44)	<b>0.015</b>	(3/7/5)	(30/33/13)	0.209	0.466
Contour (round/oval/irregular)	(9/16/16)	(51/46/80)	0.064	(7/6/2)	(30/16/30)	0.092	0.307
Necrosis (yes/no)	(3/31)	(95/82)	<b>0.000</b>	(1/14)	(33/43)	<b>0.000</b>	0.802
Calcification (yes/no)	(0/34)	(28/149)	<b>0.027</b>	(0/15)	(13/63)	0.185	0.855
Surface ulceration (yes/no)	(3/31)	(37/140)	0.149	(2/13)	(16/60)	0.740	0.874
LN (yes/no)	(6/28)	(15/162)	0.186	(2/13)	(6/70)	0.614	0.834
Hemorrhage (yes/no)	(0/34)	(2/175)	1.000	(0/15)	(1/75)	1.000	1.000
Peritumoral exudation (yes/no)	(0/34)	(5/172)	1.000	(0/15)	(1/75)	1.000	0.782
Necrosis under the tumor wall (yes/no)	(2/32)	(69/108)	<b>0.000</b>	(2/13)	(29/47)	<b>0.017</b>	1.000
Intratumoral vessel (yes/no)	(5/29)	(35/142)	0.635	(2/32)	(16/60)	0.740	0.874
CTU(HU)	34.80 ± 5.00	34.40 ± 7.10	0.755	32.33 ± 4.34	35.43 ± 8.70	0.183	0.619
CTA(HU)	56.47 ± 12.85	58.51 ± 17.31	0.514	55.71 ± 13.94	57.37 ± 13.10	0.672	0.270
CTV(HU)	71.20 ± 15.63	71.90 ± 18.10	0.826	68.45 ± 17.85	75.04 ± 13.31	0.179	0.307
DEAP(HU)	21.67 ± 10.80	24.11 ± 16.61	0.278	20.28 ± 12.48	25.05 ± 12.05	0.178	0.158
DEPP(HU)	36.38 ± 14.39	37.51 ± 18.41	0.735	33.03 ± 17.57	42.71 ± 11.37	0.054	0.221
LD (mm)	28.79 ± 13.74	44.83 ± 34.94	<b>0.000</b>	24.40 ± 10.83	50.87 ± 40.60	<b>0.000</b>	0.346
SD (mm)	28.50 ± 12.23	35.26 ± 24.42	<b>0.000</b>	25.67 ± 13.32	37.80 ± 28.41	<b>0.015</b>	0.594
LD/SD	1.02 ± 0.21	1.26 ± 0.25	<b>0.000</b>	1.00 ± 0.18	1.28 ± 0.24	<b>0.000</b>	0.813

GIST gastrointestinal stromal tumor; GS gastric schwannoma; CTU/CTA/CTV the CT attenuation value of unenhancement phase/arterial phase/portal venous phase; DEAP the CT attenuation value of arterial phase-unenhanced phase; DEPP CT attenuation value of portal venous phase-unenhanced phase; LD long diameter; SD short diameter

P1 value<sup>a</sup>: comparison of GSs and GISTs in training cohort

P2 value<sup>b</sup>: comparison of GSs and GISTs in validation cohort

P3 value<sup>c</sup>: comparison of patients in training cohort and patients in validation cohort

P value written in bold indicates a significant difference

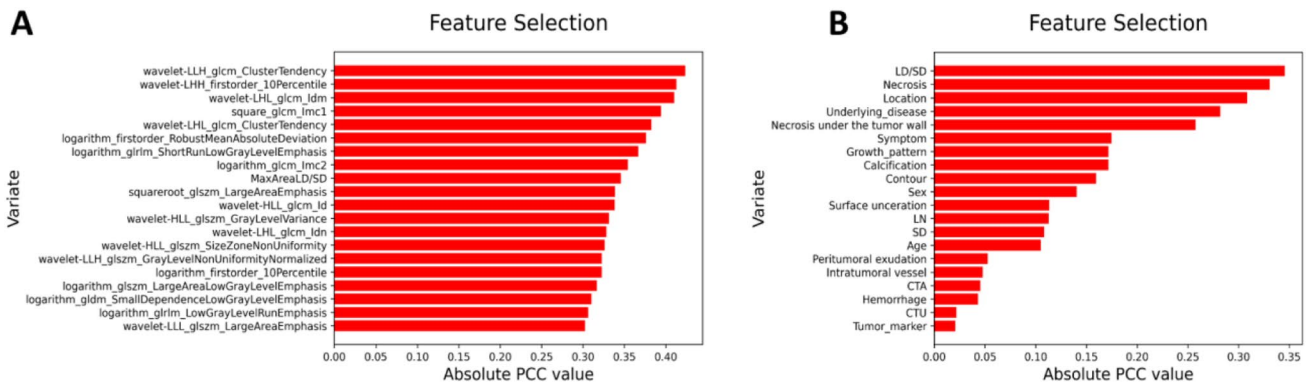
according to the stepwise logistic regression model. These features included one from size- and shape-related features, two from GLSZM, and one from GLCM. The wavelet-LHH\_firstorder\_10Percentile with the largest β (SE) value (11.9554) increased the likelihood of GSs diagnosis rather than GISTs. However, MaxAreaLD/SD with a β (SE) value of -20.6713 provided the greatest contribution to the prediction of GIST.

Six clinical and CT features (sex, LD/SD, location, growth pattern, necrosis, and underlying disease) were selected by the stepwise logistic regression model, as shown in Table 3. A location in the body and antrum had the highest association with a diagnosis of GS, while a large LD/SD value was most associated with a finding of GIST.

The equation for calculating the probability of GS in the stepwise logistic regression model for radiomics features was:

where

$$\begin{aligned}
 \text{probability} = & \frac{1}{1 + e^{-t}} \quad t = 9.976 - 20.6713 * \text{MaxAreaLD/SD} \\
 & - 12.4327 * \text{wavelet} \\
 & - \text{LLH\_glszm\_GrayLevelNonUniformityNormalized} \\
 & - 8.747 * \text{wavelet} - \text{LHL\_glcm\_Idm} \\
 & + 11.9554 * \text{wavelet} - \text{LHH\_firstorder\_10Percentile} \\
 & + 7.5111 * \text{squareroot\_glszm\_LargeAreaEmphasis}
 \end{aligned}$$



**Fig. 3** Feature selection by calculation of the Pearson correlation coefficients (PCCs) before model building. The 20 features with the largest PCC were selected as input indexes to build models using features from non-radiomics dataset (**A**) and radiomics dataset (**B**)

**Table 2** The selected features and relevant coefficients of the stepwise logistic regression model in radiomics dataset

Variate	$\beta$ (SE)	$\beta$ (SE)[0.025	0.975]	<i>P</i> value
Constant	9.976	3.728	16.224	<b>0.002</b>
MaxAreaLD/SD	- 20.6713	- 28.626	- 12.717	<b>0.000</b>
Wavelet-LLH_glszm_GrayLevelNonUniformityNormalized	- 12.4327	- 19.886	- 4.979	<b>0.001</b>
wavelet-LHL_glcM_Idm	- 8.747	- 14.74	- 2.754	<b>0.004</b>
wavelet-LHH_firstorder_10Percentile	11.9554	6.996	16.915	<b>0.000</b>
squareroot_glszm_LargeAreaEmphasis	7.5111	3.151	11.871	<b>0.001</b>

*P* value written in bold indicates a significant difference

**Table 3** The selected features and relevant coefficients of the Stepwise Logistic Regression model in non-radiomic dataset

Variate	$\beta$ (SE)	$\beta$ (SE)[0.025	0.975]	<i>P</i> value
Constant	- 1.5608	- 4.398	1.276	0.281
Sex	- 1.1934	- 2.343	- 0.044	<b>0.042</b>
LD/SD	- 8.3933	- 14.564	- 2.223	<b>0.008</b>
Location	5.1244	1.993	8.256	<b>0.001</b>
Growth pattern	2.1954	0.704	3.687	<b>0.004</b>
Necrosis	- 2.7189	- 4.11	- 1.328	<b>0.000</b>
Underlying disease	- 5.1407	- 8.399	- 1.883	<b>0.002</b>

*LD*: long diameter; *SD*: short diameter

*P* value written in bold indicates a significant difference

For the non-radiomics dataset, the equation for calculating the probability of GS was

$$\text{probability} = \frac{1}{1 + e^{-t}}$$

Where

$$t = -1.5608 - 1.1934 * \text{Sex} - 8.3933 * \frac{LD}{SD} + 5.1244 * \text{Location} + 2.1954 * \text{Growthpattern} - 2.7189 * \text{Necrosis} - 5.1407 * \text{Underlyingdisease}$$

**LASSO-logistic regression model**

The results of the LASSO-logistic regression using features from non-radiomics dataset or radiomics dataset are summarized in Tables 4 and 5 and Fig. 4. The LASSO method selected nine radiomics features and eight clinical and CT features for input into the regression algorithm. High MaxAreaLD/SD and LD/SD yielded the largest coefficients supporting the identification of GISTs in the two datasets. Supplemental Table S1 presents the five selected radiomics features with *p* < 0.05 and compares them between GSs and GISTs in the training dataset.

The probability of GS according to the radiomics datasets was equal to

**Table 4** The selected features and relevant coefficients of LASSO-logistic regression model in radiomics dataset

Variate	$\beta$ (SE)	$\beta$ (SE)[0.025	0.975]	P value
Constant	8.5564	- 0.725	17.838	0.071
MaxAreaLD/SD	- 17.0994	- 24.064	- 10.135	<b>0.000</b>
wavelet-LHL_glcm_ClusterTendency	- 0.8538	- 6.172	4.465	0.753
wavelet-LHL_glcm_Idm	- 3.8902	- 12.601	4.82	0.381
wavelet-LHH_firstorder_10Percentile	5.6835	2.267	9.1	<b>0.001</b>
square_glcm_Imc1	- 3.2946	- 10.537	3.948	0.373
squareroot_glszm_LargeAreaEmphasis	3.886	0.014	7.758	<b>0.049</b>
logarithm_firstorder_RobustMeanAbsoluteDeviation	- 1.238	- 7.659	5.183	0.706
logarithm_glrIm_ShortRunLowGrayLevel Emphasis	- 2.8064	- 9.113	3.5	0.383
logarithm_glszm_LargeAreaLowGrayLevelEmphasis	1.4232	- 3.513	6.359	0.572

P value written in bold indicates a significant difference

**Table 5** The selected features and relevant coefficients of LASSO-logistic regression in non-radiomics dataset

Variate	$\beta$ (SE)	$\beta$ (SE)[0.025	0.975]	P value
Constant	- 1.6349	- 4.569	1.299	0.275
Sex	- 1.3814	- 2.613	- 0.149	<b>0.028</b>
LD/SD	- 8.0618	- 14.016	- 2.108	<b>0.008</b>
Location	5.2375	2.028	8.447	<b>0.001</b>
Growth pattern	2.0087	0.489	3.529	<b>0.010</b>
Necrosis	- 1.9847	- 3.893	- 0.077	<b>0.041</b>
LN	1.9776	- 0.094	4.049	0.061
Necrosis under the tumor wall	- 1.5074	- 3.805	0.79	0.199
Underlying disease	- 5.2904	- 8.755	- 1.826	<b>0.003</b>

LD/SD: the ratio of long diameter to short diameter; LN: lymph node  
P value written in bold indicates a significant difference

$$\text{probability} = \frac{1}{1 + e^{-t}}$$

where

$$t = 8.5564 - 17.0994 * \text{MaxAreaLD/SD} - 0.8538 * \text{wavelet-LHL\_glcm\_ClusterTendency} - 3.8902 * \text{wavelet-LHL\_glcm\_Idm} + 5.6835 * \text{wavelet-LHH\_firstorder\_10Percentile} - 3.2946 * \text{square\_glcm\_Imc1} + 3.886 * \text{squareroot\_glszm\_LargeAreaEmphasis} - 1.238 * \text{logarithm\_firstorder\_RobustMeanAbsoluteDeviation} - 2.8064 * \text{logarithm\_glrIm\_ShortRunLowGrayLevelEmphasis} + 1.4232 * \text{logarithm\_glszm\_LargeAreaLowGrayLevelEmphasis}$$

For the non-radiomics dataset, the equation for the probability of GS was

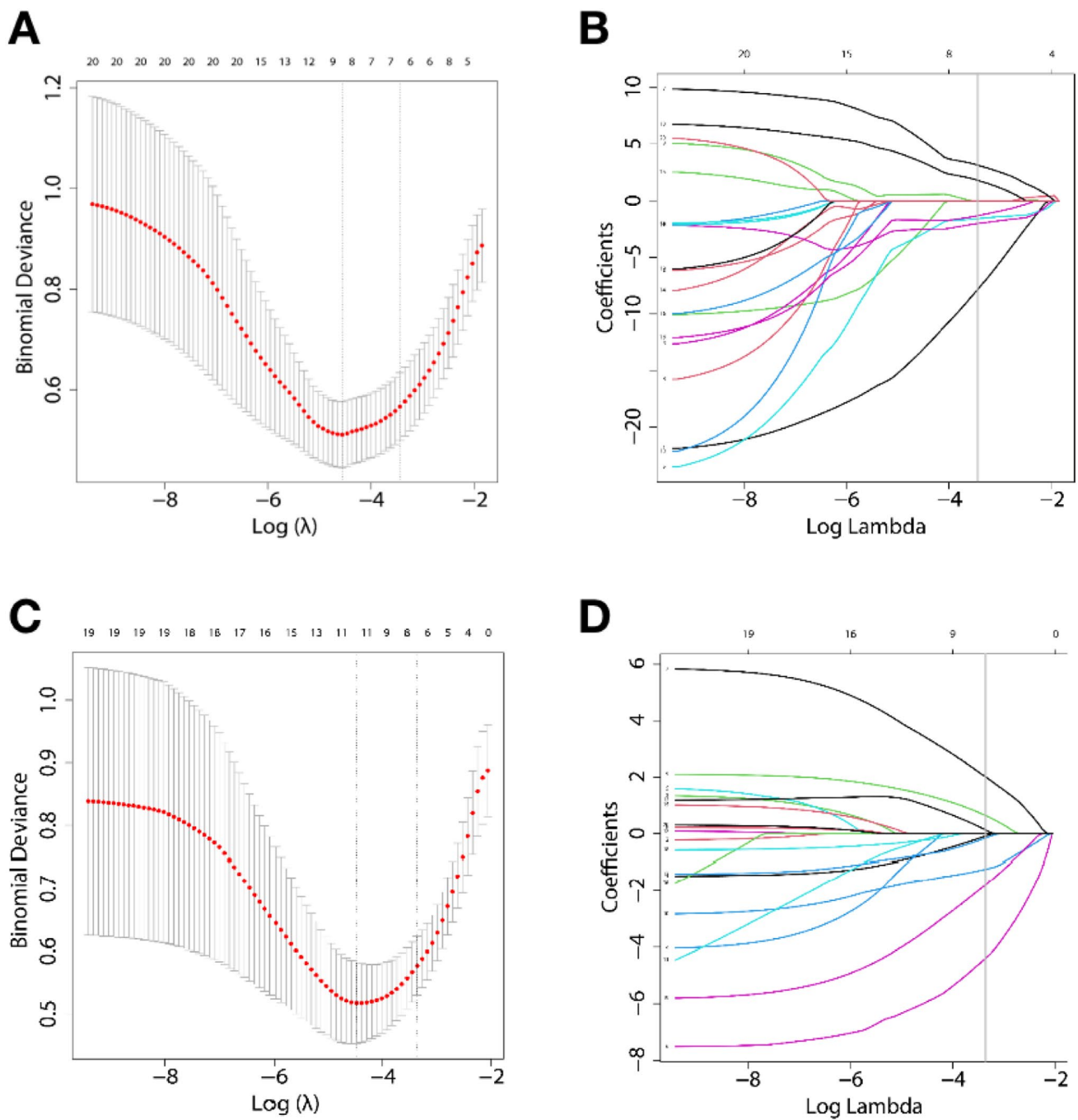
$$\text{probability} = \frac{1}{1 + e^{-t}}$$

$$t = - 1.6349 - 1.3814 * \text{Sex} - 8.0618 * \text{LD/SD} + 5.2375 * \text{Location} + 2.0087 * \text{Growth pattern} - 1.9847 * \text{Necrosis} + 1.9776 * \text{LN} - 1.5074 * \text{Necrosis under the tumor wall} - 5.2904 * \text{Underlying disease}$$

### Diagnostic performance analysis

The diagnostic efficacy of the two radiomics feature models in the training and validation sets is summarized in Table 6. The stepwise logistic regression model applied to the training dataset yielded sensitivity, specificity, accuracy, and AUC of 94.1%, 85.3%, 86.7%, and 0.955, respectively, whereas the





**Fig. 4** Feature selection via LASSO-logistic regression. Radiomics feature selection (**A**, **B**). Selection of the tuning parameter ( $\lambda$ ) using tenfold cross-validation and a minimum criterion. A  $\lambda$  value of 0.011 with  $\log(\lambda) = -4.50$  was selected. A coefficient profile is plotted against the  $\log(\lambda)$  sequence using tenfold cross-validation. A vertical

line is drawn at the value selected, which resulted in ten non-zero coefficients. Clinical and CT features selection (**C**, **D**). The optimal  $\lambda$  value of 0.037 with  $\log(\lambda) = -3.31$  was retained using tenfold cross-validation. Eight non-zero coefficients were selected finally

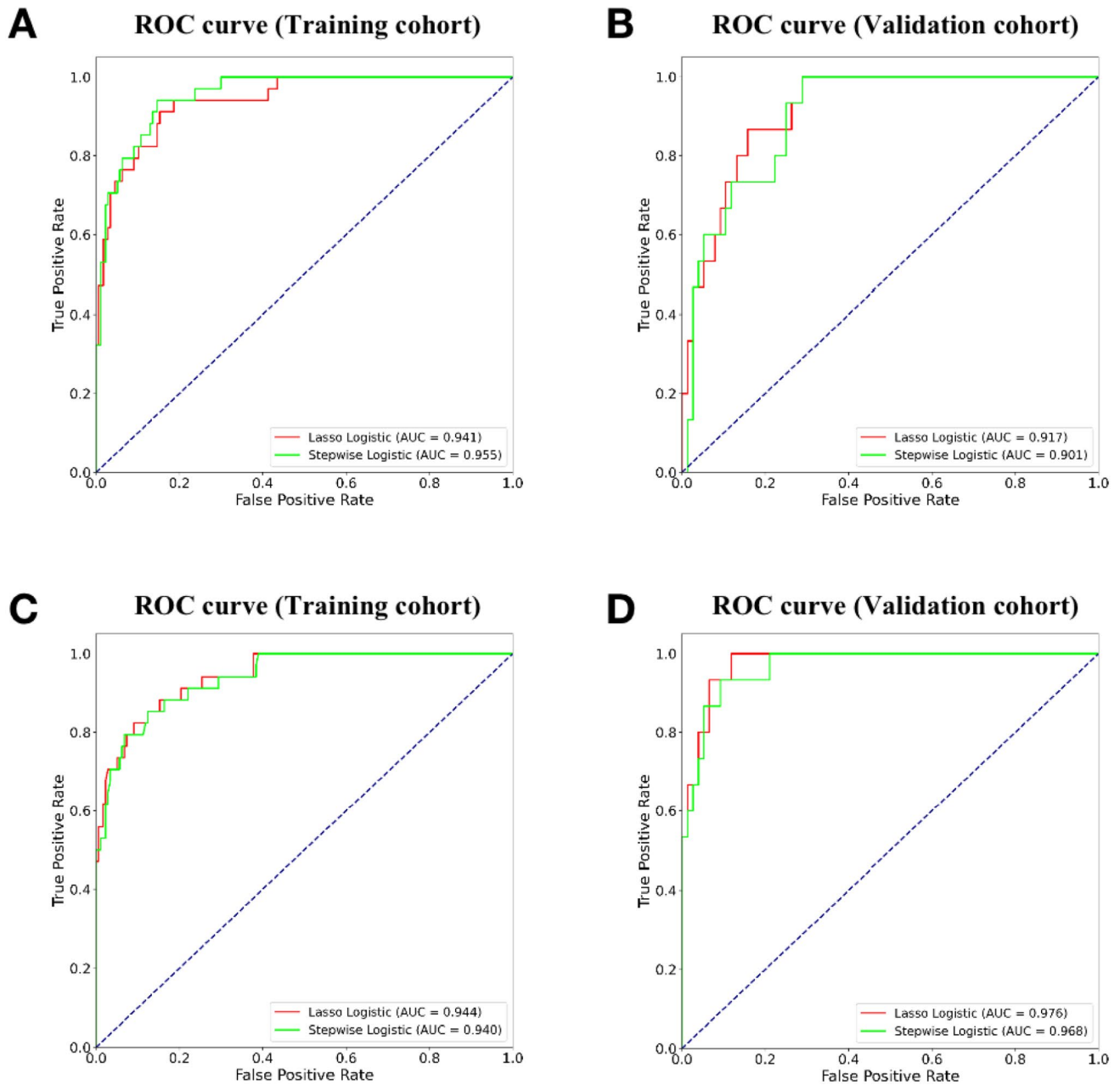
LASSO-logistic regression model yielded sensitivity, specificity, accuracy, and AUC of 91.2%, 84.7%, 85.8%, and 0.941. Supplemental Table S2 shows the diagnostic performance

results for the features from non-radiomics dataset. The ROC curves of the two models applied to the training and validation datasets are plotted in Fig. 5. We found that the AUC values

**Table 6** The diagnostic performance analysis in radiomics dataset

Cohort	Model	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC
Training( $n = 211$ )	Stepwise logistic regression	94.1	85.3	86.7	0.955
	LASSO-logistic regression	91.2	84.7	85.8	0.941
Validation( $n = 91$ )	Stepwise logistic regression	93.0	71.1	75.8	0.901
	LASSO-logistic regression	95.6	71.1	72.8	0.917

AUC: area under the curve



**Fig. 5** Receiver operating characteristic (ROC) curves of two models for differentiating GS and GIST. All the AUCs of the two models using features from radiomics dataset (**A, B**) and non-radiomics dataset (**C, D**) were above 0.900

**Table 7** The Delong's test analysis of the two models in different dataset

Cohort	Model	Dataset	AUC	<i>P</i> value
Training ( <i>n</i> = 211)	Stepwise logistic regression	Radiomics dataset	0.955	0.493
		Combination dataset	0.940	
	LASSO-logistic regression	Radiomics dataset	0.941	0.902
		Combination dataset	0.944	
Validation ( <i>n</i> = 91)	Stepwise logistic regression	Radiomics dataset	0.901	0.089
		Combination dataset	0.968	
	LASSO-logistic regression	Radiomics dataset	0.917	0.079
		Combination dataset	0.976	

*AUC*: area under the curve

of the two kinds of models were larger than 0.900 for both datasets.

### Delong's test analysis

Delong's test was independently implemented in the training and validation datasets. We compared the diagnostic efficiency differences between the two data sets within the same model (stepwise logistic regression or LASSO-logistic regression). Table 7 presents the Delong's test analysis results. There were no significant differences in predictive performance between features from non-radiomics and radiomics datasets, regardless of which model was used (all  $p > 0.05$ ).

Supplemental Tables S3 and S4 separately summarize the Delong's test results between the stepwise logistic regression model and LASSO-logistic regression model in the two datasets.

### Discussion

While a number of studies have investigated the differentiation of GSs from gastric stromal tumors using CT images, we believe this study is the first attempt to apply radiomics to this problem. In our study, 211 patients with either GISTs or GSs from Center 1 were assigned to training cohort and 91 patients from Center 2 to validation cohort. Univariate analysis showed that the clinical and CT parameters were well balanced between the two cohorts.

In our study, 1223 extracted radiomics features and 24 clinical and CT characteristics were used to build stepwise logistic regression and LASSO-logistic regression models. A total of eight models using two kinds of algorithms were developed, and the AUC values attained were above 0.9 for both datasets with both radiomics features and clinical and CT features. Although a Delong's test revealed no significant differences in predictive performance between the non-radiomics dataset and radiomics dataset, we still suggest that

the radiomics-based model shows promise, with comparable predictive performance to senior physicians in the differentiation of GS from GIST.

Five radiomics features were selected by the stepwise logistic regression model as useful parameters to differentiate GS from GIST, and three of these features were picked in the LASSO-logistic regression.

MaxAreaLD/SD is a shape-based feature, equal to the ratio of the major axis length to the least axis length, which is similar to the LD/SD feature manually extracted from the CT imaging. In this study, the MaxAreaLD/SD and LD/SD values were higher for GISTs than for GSs. High MaxAreaLD/SD from the radiomics features and LD/SD clinical values were of greater benefit to identification of GISTs rather than GSs, which reveals that GISTs, as a potentially malignant entity, tended to grow at different rates in all directions and had a more irregular shape than GSs.

The features wavelet-LLH\_glszm\_GrayLevelNonUniformityNormalized, wavelet-LHL\_glcm\_Idm, and wavelet-LHH\_firstorder\_10Percentile are wavelet texture features. Wavelet features are high-order features that not only reflect the distribution characteristics of space and frequency, but also reflect the heterogeneity of the tumor's microenvironment, and they are known to improve the diagnostic performance of radiomics models (Zhou et al. 2020). GrayLevelNonUniformityNormalized and glszm\_LargeAreaEmphasis are both GLSZM characteristics. GLSZM features are used to calculate the number of connections of voxels with the same gray value in the image. Wavelet\_HHL\_glszm\_GrayLevelNonUniformity refers to the non-uniformity of the gray levels; the lower the value is, the more uniform the gray level and the lower the heterogeneity of the image. GSs showed lower wavelet\_HHL\_glszm\_GrayLevelNonUniformity values than GISTs, which indicated that GSs had more homogeneous parenchyma. Firstorder\_10Percentile describes the distribution of voxel intensities in low-density regions. The lower values found in GISTs may be due to their susceptibility to necrosis and cysts. GLCM features can reflect the

distribution of the gray levels of two pixels in a specific direction and distance, and are most widely used to evaluate tumor heterogeneity. The *glcm\_Idm* feature reflects local changes in the image texture. If different areas of the image texture are more uniform, the inverse difference will be larger, whereas otherwise it will be smaller. Large *glcm\_Idm* values in GSs indicate a more uniform nature.

Among the clinical and CT features, sex, LD/SD, location, growth pattern, necrosis, and number of underlying diseases were found to be significant features for the differentiation of GSs from GISTs. GS was more predominantly found in female patients than GIST, which is consistent with the previous studies of Ji et al. (2015), but different to the finding of Xu et al. (2022). We found that GSs tended to grow in the gastric body and antrum, whereas gastric GISTs were often seen in the body and the fundus, which is similar to the findings in Xu et al. (2022) and our previous reports (Wang et al. 2021). An exophytic or mixed growth pattern found in GS in several previous studies (Choi et al. 2012; He et al. 2017; Xu et al. 2022) is also in agreement with our findings. Our study also revealed necrosis to be a significant CT feature suggesting GIST rather than GS, which is in line with the study of He et al. (He et al. 2017). The lack of necrosis in schwannomas may be a result of the slower growth rate of GS compared with GIST (Choi et al. 2014). To our surprise, patients with GIST tended to have more underlying diseases than those with GS, which has not been reported before. The underlying diseases included hypertension, diabetes, cardiovascular disease, cerebrovascular disease, chronic obstructive pulmonary disease, and chronic kidney diseases. The reason for this may be related to the fact that our patients with GIST were older than those with GS, and the possibility of suffering from underlying diseases may increase accordingly. We also speculate that patients with underlying diseases may be more likely to be affected by malignant tumors than benign ones, although this hypothesis still requires more cases and studies to confirm it.

This study has several limitations. First, the sample size is relatively small, especially for GSs. Our series was collected from two hospitals and was acquired using two different CT scanners, which may have resulted in data heterogeneity. Second, the radiomics features were only extracted from portal venous phase images. Radiomics information from unenhanced and arterial phase images was not applied to optimize the performance of the models. Hence, we intend to build radiomics models using features from all three phases in future work. Finally, the models built from CT features assessed by two associate chief radiologists showed good performance, but we did not make a comparison with predictive performance assessed by junior radiologists. It is possible that poorer performance may be shown assessed by junior radiologists, highlighting the excellence of the radiomics methods.

## Conclusion

This preliminary study built radiomics-based models to explore associations between radiomics features and GSs and GISTs. Our study showed that a radiomics model applied to CE-CT images had comparable predictive performance to senior radiologists in the differentiation of GSs from GISTs.

**Electronic Supplementary Material** The online version of this article (<https://doi.org/10.1007/s00432-023-05545-w>) contains supplementary material, which is available to authorized users.

**Acknowledgements** We thank Liwen Bianji (Edanz) ([www.liwenbianji.cn](http://www.liwenbianji.cn)) for editing the language of a draft of this manuscript.

**Author contributions** JW, GQM, and CZ participated in the design of this study, and they all performed the statistical analysis. CWW, HJH, and YY carried out the study and collected important background information. CXC and HLJ collected raw data. LYH performed data processing on deep learning software. CZ drafted the manuscript. All authors read and approved the final manuscript.

**Funding** This research was supported by Zhejiang Provincial Natural Science Foundation under Grant No.LGF21H030004.

**Data availability** The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Conflict of interest** The authors declare no conflict of interest.

**Ethics approval and consent to participate** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the Declaration of Helsinki. For retrospective study, formal consent is not required. Informed consent was obtained from all individual participants included in the study. This study has been approved by the Ethics Committee of the TongDe Hospital (Approval No. 2021–040).

**Consent for publication** All authors consent to publish this article.

**Institutional review board statement** This study was reviewed and approved by the institutional review board and ethics committee of TongDe Hospital of ZheJiang Province.

**Informed consent statement** Patients were not required to give informed consent to the study because the analysis used anonymous clinical data that were obtained after each patient agreed to treatment by written consent.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not

permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ao W, Cheng G, Lin B, Yang R, Liu X, Zhou S, Wang W, Fang Z, Tian F, Yang G, Wang J (2021) A novel CT-based radiomic nomogram for predicting the recurrence and metastasis of gastric stromal tumors. *Cancer Res* 11(6):3123–3134
- Cai MY, Xu JX, Zhou PH, Xu MD, Chen SY, Hou J, Zhong YS, Zhang YQ, Ma LL (2016) Endoscopic resection for gastric schwannoma with long-term outcomes. *Surg Endosc* 30(9):3994–4000
- Casali PG, Blay JY, Abecassis N, Bajpai J, Bauer S, Biagini R, Bielack S, Bonvalot S, Boukovinas I, Bovee JV, Boye K, Brodowicz T (2022) Gastrointestinal stromal tumours: ESMO-EURACAN-GENTURIS Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 33(1):20–33
- Chen T, Ning Z, Xu L, Feng X, Han S, Roth HR, Xiong W, Zhao X, Hu Y, Liu H, Yu J, Zhang Y (2019) Radiomics nomogram for predicting the malignant potential of gastrointestinal stromal tumours preoperatively. *Eur Radiol* 29(3):1074–1082
- Chen Z, Xu L, Zhang C, Huang C, Wang M, Feng Z, Xiong Y (2021) CT radiomics model for discriminating the risk stratification of gastrointestinal stromal tumors: a multi-class classification and multi-center study. *Front Oncol* 11:654114
- Choi JW, Choi D, Kim KM, Sohn TS, Lee JH, Kim HJ, Lee SJ (2012) Small submucosal tumors of the stomach: differentiation of gastric schwannoma from gastrointestinal stromal tumor with CT. *Korean J Radiol* 13(4):425–433
- Choi YR, Kim SH, Kim SA, Shin CI, Kim HJ, Kim SH, Han JK, Choi BI (2014) Differentiation of large ( $\geq 5$  cm) gastrointestinal stromal tumors from benign subepithelial tumors in the stomach: Radiologists' performance using CT. *Eur J Radiol* 83(2):250–260
- Feng Q, Tang B, Zhang Y, Liu X (2022) Prediction of the Ki-67 expression level and prognosis of gastrointestinal stromal tumors based on CT radiomics nomogram. *Int J Comput Assist Radiol Surg* 17(6):1167–1175
- Gao Z, Lin T, Cao H, Shen L, Chinese Society Of Clinical Oncology CSCO Expert Committee On Gastrointestinal Stromal Tumor (2017) Chinese consensus guidelines for diagnosis and management of gastrointestinal stromal tumor. *Chin J Cancer Res* 29(4):281–293
- Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. *Radiology* 278(2):563–577
- He MY, Zhang R, Peng Z, Li Y, Xu L, Jiang M, Li ZP, Feng ST (2017) Differentiation between gastrointestinal schwannomas and gastrointestinal stromal tumors by computed tomography. *Oncol Lett* 13(5):3746–3752
- Hong HS, Ha HK, Won HJ, Byun JH, Shin YM, Kim AY, Kim PN, Lee MG, Lee GH, Kim MJ (2008) Gastric schwannomas: Radiological features with endoscopic and pathological correlation. *Clin Radiol* 63:536–542
- Ji JS, Lu CY, Mao WB, Wang ZF, Xu M (2015) Gastric schwannoma: CT findings and clinicopathologic correlation. *Abdom Imaging* 40(5):1164–1169
- Levy AD, Remotti HE, Thompson WM, Sobin LH, Miettinen M (2003) Gastrointestinal stromal tumors: radiologic features with pathologic correlation. *Radiographics* 23(283–304):456
- Li J, Ye Y, Wang J, Zhang B, Qin S, Shi Y, He Y, Liang X, Liu X, Zhou Y, Wu X, Zhang X, Wang M, Liu J, Chai Y, Zhou J, Dong C, Zhang W, Liu B (2017) Spectral computed tomography imaging of gastric schwannoma and gastric stromal tumor. *J Comput Assist Tomogr* 41(3):417–421
- Liu J, Chai Y, Zhou J, Dong C, Zhang W, Spectral Liu B (2017) Computed Tomography Imaging of Gastric Schwannoma and Gastric Stromal Tumor. *J Comput Assist Tomogr* 41(3):417–421
- Liu Z, Wang S, Dong D, Wei J, Fang C, Zhou X, Sun K, Li L, Li B, Wang M, Tian J (2019) The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges. *Theranostics* 9(5):1303–1322
- Liu S, He J, Liu S, Ji C, Guan W, Chen L, Guan Y, Yang X, Zhou Z (2020) Radiomics analysis using contrast-enhanced CT for preoperative prediction of occult peritoneal metastasis in advanced gastric cancer. *Eur Radiol* 30(1):239–246
- Tsuboi K, Yano F, Omura N, Misawa T, Kashiwagi H (2021) Reduced-port surgery with the cowboy technique for a gastric submucosal tumor. *Asian J Endosc Surg* 14(1):154–157
- Wang C, Li H, Jiaerken Y, Huang P, Sun L, Dong F, Huang Y, Dong D, Tian J, Zhang M (2019a) Building CT radiomics-based models for preoperatively predicting malignant potential and mitotic count of gastrointestinal stromal tumors. *Transl Oncol* 12(9):1229–1236
- Wang J, Zhang WM, Zhou XX, Xu JL, Hu HJ (2019b) Simple analysis of the computed tomography features of gastric schwannoma. *Can Assoc Radiol J* 70(3):246–253
- Wang Y, Liu W, Yu Y, Liu JJ, Xue HD, Qi YF, Lei J, Yu JC, Jin ZY (2020) CT radiomics nomogram for the preoperative prediction of lymph node metastasis in gastric cancer. *Eur Radiol* 30(2):976–986
- Wang J, Xie Z, Zhu X, Niu Z, Ji H, He L, Hu Q, Zhang C (2021) Differentiation of gastric schwannomas from gastrointestinal stromal tumors by CT using machine learning. *Abdom Radiol (NY)* 46(5):1773–1782
- Wei J, Jiang H, Gu D, Niu M, Fu F, Han Y, Song B, Tian J (2020) Radiomics in liver diseases: current progress and future opportunities. *Liver Int* 40(9):2050–2063
- Xu JX, Yu JN, Wang XJ, Xiong YX, Lu YF, Zhou JP, Zhou QM, Yang XY, Shi D, Huang XS, Fan SF, Yu RS (2022) A radiologic diagnostic scoring model based on CT features for differentiating gastric schwannoma from gastric gastrointestinal stromal tumors. *Am J Cancer Res* 12(1):303–314
- Zhang QW, Zhou XX, Zhang RY, Chen SL, Liu Q, Wang J, Zhang Y, Lin J, Xu JR, Gao YJ, Ge ZZ (2020) Comparison of malignancy-prediction efficiency between contrast and non-contrast CT based radiomics features in gastrointestinal stromal tumours: a multi-center study. *Clin Transl Med* 10(3):e291
- Zhou J, Lu J, Gao C, Zeng J, Zhou C, Lai X, Cai W, Xu M (2020) Predicting the response to neoadjuvant chemotherapy for breast cancer: wavelet transforming radiomics in MRI. *BMC Cancer* 20(1):100
- Zhuo Y, Zhan Y, Zhang Z, Shan F, Shen J, Wang D, Yu M (2021) Clinical and CT radiomics nomogram for preoperative differentiation of pulmonary adenocarcinoma from tuberculoma in solitary solid nodule. *Front Oncol* 11:701598

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.