



Utility of radiomics based on contrast-enhanced CT and clinical data in the differentiation of benign and malignant gallbladder polypoid lesions

Xiaodong Yang¹ · Yi Liu¹ · Yan Guo² · Ruimei Chai¹ · Meng Niu¹ · Ke Xu¹

Published online: 12 March 2020

© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Purpose To develop and validate a novel method based on radiomics for the preoperative differentiation of benign and malignant gallbladder polypoid lesions (PLG).

Patients and methods A total of 145 patients with pathological proven gallbladder polypoid lesions ≥ 1 cm were included in this retrospective study. All the patients underwent abdominal contrast-enhanced computed tomography (CT) examinations 3 weeks before cholecystectomy from January 2013 to January 2019. Seventy percent of the cases were randomly selected for the training dataset, and 30% of the cases were independently used for testing. Radiomics features extracted from portal venous-phase CT of the PLG and clinical features were analyzed, and the LASSO regression algorithm was used for data dimension reduction. Multivariable logistic regression was used to generate radiomics signatures, clinical signatures, and combination signatures. The receiver operating characteristic (ROC) curve and decision curve were plotted to assess the differentiating performance of the three signatures.

Results The area under the ROC curve (AUC) of the radiomics signature and clinical signature was 0.924 and 0.861 in the testing dataset, respectively. For the radiomics signature, the accuracy was 88.6%, with 88.0% specificity and 89.5% sensitivity. When combined, the AUC was 0.931, the specificity was 84.0%, and the sensitivity was 89.5%. The differences between the AUC values of the two sole models and the combination model were statistically nonsignificant.

Conclusion Radiomics based on CT images can be helpful to differentiate benign and malignant gallbladder polyps ≥ 1 cm in size.

Keywords Radiomics · Polypoid lesions of the gallbladder · CT · US

Abbreviations

3D	Three-dimensional
AUC	Area under the curve
CI	Confidence interval
CT	Computed tomography
DICOM	Digital Imaging and Communications in Medicine
PLG	Gallbladder polypoid lesions
ROC	Receiver operating characteristic
ROI	Region of interests

Introduction

The prevalence of polypoid lesions of the gallbladder (PLG), as a common clinical gallbladder disease, is approximately 5% in the adult population around the world [1], and with the widespread use of abdominal imaging modalities in modern clinical practice, an increasing number of PLG are being detected. PLG is a morphological concept named on the basis of imaging manifestations and refers to a collection of lesions that elevate from the gallbladder wall and grow like a polyp into the gallbladder cavity. Clinically, most gallbladder polyps are benign, and malignant polyps are only a minority [2]. Unfortunately, although the proportion of malignant polyps is small, it presents late in diagnosis and carries a dismal prognosis, and the 5-year survival rate is less than 5% [3]. Therefore, it is of great importance in differentiating benign and malignant PLG at an early stage.

✉ Ke Xu
kxu@cmu.edu.cn

¹ Department of Radiology, The First Affiliated Hospital of China Medical University, No.155 Nanjing Road Heping Area, Shenyang 110000, Liaoning Province, China

² GE Health, Shanghai 200000, China

At present, the imaging diagnostic equipment for gallbladder polyps mainly includes ultrasound (US), computed tomography (CT), and magnetic resonance imaging (MRI). However, studies have shown that it is often difficult to differentiate between benign and malignant PLG using these modalities [4–6].

Radiomics is an emerging and rapidly developing discipline whose primary goal is to use existing medical images to dig up the potential high-dimensional information behind these images to aid clinical decision-making, and radiomics has been widely used in research of lung cancer [7], liver cancer [8], colorectal cancer [9], head and neck tumor [10], etc. Unfortunately, until now, there have been no reports of making use of radiomics to differentiate the nature of PLG, whether it is based on CT, US, or MRI. Therefore, the aim of this study was to develop a new method to differentiate the benignity and malignancy of PLG by radiomics based on CT imaging for the early diagnosis of malignant PLG.

Methods

Patients

Ethical approval was obtained for this retrospective analysis, and the informed consent requirement was waived. A total of 145 patients with pathological proven gallbladder polypoid lesions ≥ 1 cm were ultimately included in this

retrospective study (Fig. 1). All the patients underwent abdominal contrast-enhanced CT examination and cholecystectomy from January 2013 to January 2019. The inclusion criteria were as follows: (1) an abdominal enhanced CT scan performed in our hospital within 3 weeks before the operation, (2) the maximum diameter of PLG ≥ 1 cm, and (3) availability of the pathologic report from the cholecystectomy. The exclusion criteria were as follows. (1) The lesion had obviously invaded the surrounding tissues, such as liver, hepatic hilar lymph node. (2) Patients underwent some operation and treatment before surgery, such as percutaneous transhepatic gallbladder drainage and radiochemotherapy. (3) The PLG could not be clearly displayed because of acute cholecystitis that caused gallbladder wall edema accompanied by a large amount of exudate, the obstruction by very large gallbladder stones, respiratory movement artifacts. (4) Most or all of the gallbladder cavity was occupied by PLG. All the patients were randomly divided into a training group and a testing group according to the ratio of 7:3.

Imaging

This study involved three CT scanners, Aquilion One CT (Toshiba Corporation, Tokyo, Japan), Somatom Definition Flash CT (Siemens Healthineers, Erlangen, Germany), and Brilliance I CT (Royal Philips Electronics, Amsterdam, Nederlande). The scanning parameters are shown in Table 1. For CT scanning, the patients were in a conventional supine

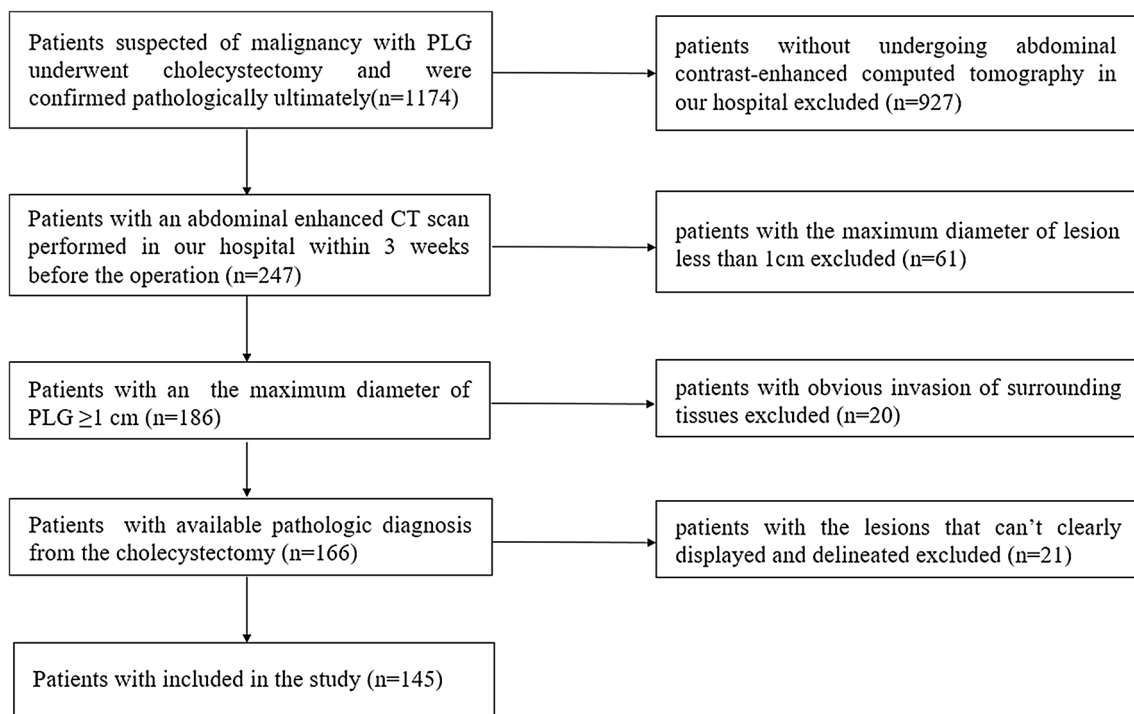


Fig. 1 Flowchart for selecting patients

Table 1 CT image acquisition parameters

	Toshiba aquilion one	Siemens somatom	Philips brilliance I
Tube voltage (kV)	120	120	120
Tube current (mA)	Auto	Auto	Auto
Width of the collimator (mm)	128×0.625	128×0.625	128×0.625
Pitch	0.870	0.993	0.993
Field of view (FOV) (mm)	300–400	350–500	300–400
Matrix	512×512	512×512	512×512
Rotation time (s)	0.5	0.5–1	0.5
Slice thickness (mm)	0.8–1.5	0.8–1.5	0.8–1.5

position, received a plain scan first, and then an enhanced scan. Nonionic contrast agent was used in the enhanced scan, i.e., 80 mL of iohexol and 20 mL of normal saline were injected into the antecubital vein through a double-cylinder high-pressure syringe at a speed of 3.0 mL/s. An enhanced CT scan was performed at 25 to 30 s (arterial phase), 60 to 70 s (portal venous phase), and 160 to 180 s (delayed phase) after the start of the injection.

Image preprocessing

Studies have shown that images acquired from different scanner modalities may affect the extraction and analysis of radiomics features due to differences in scanning parameters and reconstruction methods [11]. Therefore, we resampled the original image before extraction of the radiomic features. First, a linear interpolation algorithm was used to generate new data points within the range of the known data points; thus, all of the original image data were resampled to a common voxel spacing of $0.700 \times 0.700 \times 0.700 \text{ mm}^3$. Consequently, the images after data preprocessing were isotropic and used for the next segmentation.

Segmentation

The preprocessed portal-phase CT images were imported to ITK-SNAP software (<http://www.itksnap.org>, version 3.6.0), and the segmentation of the 3D volume of interest (VOI) based on portal-phase CT images were manually delineated by one radiologist with the years of experience in the same window width and window level (150 HU, 80 HU) without knowing the pathological diagnosis or any clinical information about the patients. Then, the segmentations were reviewed by another senior radiologist, with 20 years of experience, in all patients.

Radiomics feature extraction and selection

The preprocessed portal-phase CT image and the segmented VOI were imported into the AK platform. Radiomic features were automatically generated based on lesion morphology,

first-order histogram, and high-order texture features. Then, the selection of radiomics features was carried out in the training group.

First, all the features were standardized by using the formula: $\text{features} = (f_i - \mu) / \text{std}$. Second, the abnormal and missing values of all radiomics features were replaced by median values. Then, the least absolute shrinkage and selection operator (LASSO) method with tenfold cross-validation (repeated 100 times) was used to iteratively select the most robust features in the training group until the feature coefficients were not zero.

Clinical features

Clinical features of every patient were retrospectively collected and recorded from medical records, including age, sex, pathological diagnosis, CA199, CA125, carcinoembryonic antigen (CEA), complicated symptoms or not (such as abdominal pain, nausea, vomiting, cutaneous or sclera icterus), and CT imaging features measured by experienced radiologists that included the PLG maximum diameter, location, basal width, single or multiple lesions, consolidated stone or not, CT value of the plain scan, and each enhanced phase.

The clinical biomarker was recorded as positive when it was higher than normal; otherwise, it was considered negative. If any of the symptoms mentioned above existed, they were recorded as positive. When the patient had multiple CT examination records, the most recent was selected. In addition, the size of the lesion was measured at the slice of the largest diameter of the PLG. The location of the lesion was strictly divided according to the definition of gallbladder anatomy, which was recorded as gallbladder fundus, body, or neck. The basal width of the PLG was defined as follows in reference to Yamada's classification: the wide base refers to the angle between protuberance of the PLG and the basement mucosa $> 90^\circ$, and the narrow base refers to the angle $< 90^\circ$. If there was only one PLG, it was defined as single; otherwise, it was defined as multiple. And when there were multiple polyps, the polyp with the largest diameter was regarded as the target polyp. The measurement of

the CT value in each period was selected at the level of the maximum diameter of the PLG while avoiding the necrotic area. When the CT value cannot be measured, it is recorded as the missing value.

Signature building and validation

Regarding the above selected radiomics and clinical features as independent variables and the pathological diagnosis of each patient's PLG as dependent variables, the backward stepwise method was used to construct a multivariate logistic regression model. As shown in the following formula:

$$\text{Logit}(P) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

$x = \{x_i, i = 1, 2, \dots, n\}$ indicates the selected omics feature and $\beta = \{\beta_i, i = 0, 1, \dots, n\}$ indicates the regression coefficient.

After the models were constructed in the training group, based on the principle of maximum Youden's index, the cutoff points were confirmed, and then the test group data were substituted into the models to validate the diagnostic efficiency and accuracy of the model. The receiver operating characteristic (ROC) curve was used to evaluate the predictive ability of the models for PLG benignity and malignancy in this study. Then, the sensitivity (SEN), specificity

(SPE), positive predictive rate (PPV), negative predictive rate (NPV), accuracy (ACC), and area under the ROC curve (AUC) were calculated under the optimal diagnostic threshold. The whole process is shown in Fig. 2.

Statistical analysis

All statistical analyses were completed by R language software (<https://www.r-project.org>, version 3.5.2). The packages named "verification," "pROC," "rms," and "glmnet" were used. The ROC curve was plotted to evaluate the prediction efficiency of the models for benign and malignant PLG. $p < 0.05$ was considered statistically significant. The Delong test was used to compare the performance of the ROC curves. Clinical characteristics of the two groups were compared using Student's t tests and χ^2 tests, and $p < 0.05$ was considered statistically significant. The median was used to replace the missing value.

Results

Clinical characteristics

Patient characteristics in the training and testing cohorts are given in Table 2. Among 145 patients, 63 cases were

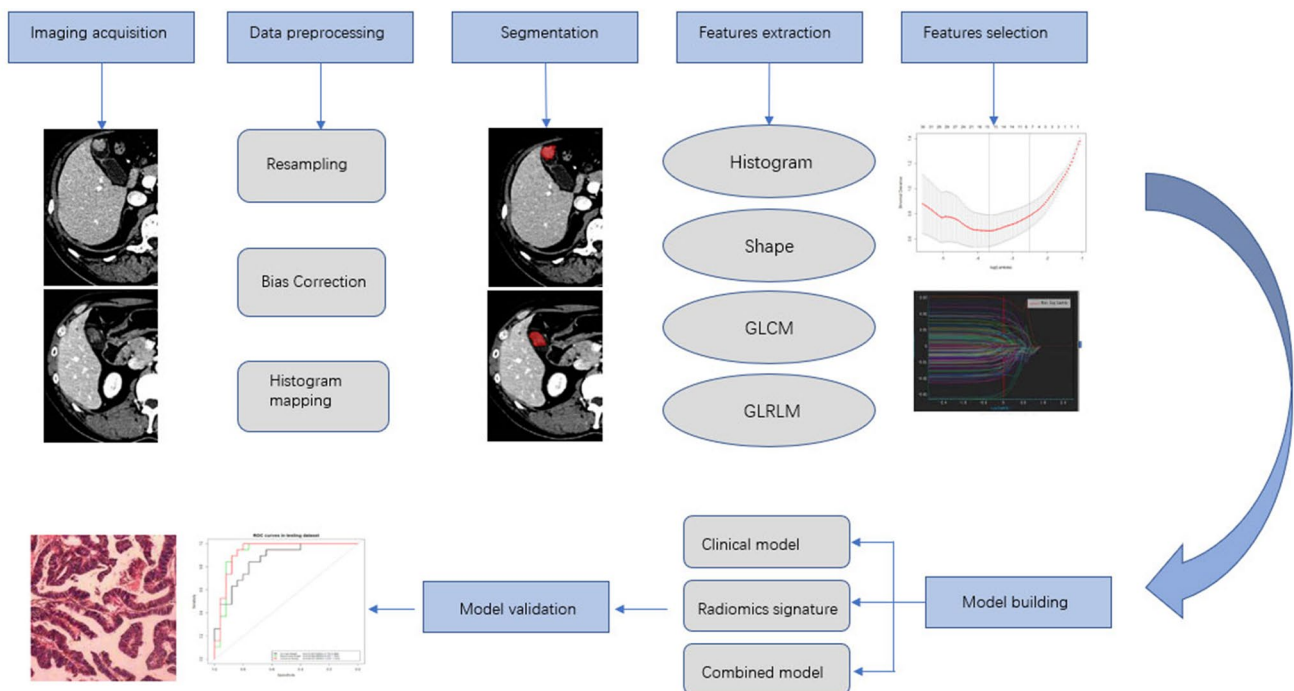


Fig. 2 Flowchart illustrating the image data analysis in the study. The two original traverse CT images were obtained from the patient with the PLG, and after preprocessing, 3D regions of interest (ROIs) were manually segmented and reconstructed, as shown in the figure. Radi-

omics features were automatically generated and selected by the "R" language. After the model was established, the ROC curve was used to evaluate the predictive ability of the models, and pathologic diagnosis was regarded as the gold standard to verify

Table 2 Characteristics of patients in the training and testing cohorts

Characteristics	Training group (N= 101)		P	Testing group (44)		P
	Malignant (n=44)	Benign (n=57)		Malignant (n=19)	Benign (n=25)	
Age (years)/Mean ± SD	62.80 ± 10.18	52.07 ± 11.28	<0.001*	65.53 ± 9.70	59.28 ± 11.40	0.057
Sex/No. (%)	Male	15 (34.1)	0.958	5 (26.3)	10 (40.0)	0.530
	Female	29 (65.9)		14 (73.7)	15 (60.0)	
Symptoms/No. (%)	(+)	33 (75.0)	<0.001*	14 (73.7)	9 (36.0)	0.029*
	(-)	11 (25.0)		5 (26.3)	16 (64.0)	
Diameter/No. (%)	2.65 ± 1.27	1.61 ± 0.62	<0.001*	2.55 ± 0.70	1.70 ± 0.59	<0.001*
Number	Single	23 (52.3)	0.068	10 (52.6)	16 (64.0)	0.653
	Multiple	21 (47.7)		9 (47.4)	9 (36.0)	
Base	Wide	34 (77.3)	<0.001*	12 (63.2)	10 (40.0)	0.223
	Narrow	10 (22.7)		7 (36.8)	15 (60.0)	
Location/No. (%)	Bottom,	16 (36.4)	0.075	9 (47.4)	10 (40.0)	0.443
	Body	16 (36.4)		5 (26.3)	11 (44.0)	
	Neck	12 (27.3)		5 (26.3)	4 (16)	
Calcification/No. (%)	(+)	6 (13.6)	0.648	3 (15.8)	7 (28.0)	0.552
	(-)	38 (86.4)		16 (84.2)	18 (72.0)	
CT value (HU) /Mean ± SD	Arterial	63.41 ± 21.34	0.026*	62.00 ± 23.49	52.56 ± 15.35	0.217
	Portal	80.5 ± 21.89	0.227	86.58 ± 30.69	70.32 ± 10.36	0.024*
	Delay	74.93 ± 16.00	0.018*	78.37 ± 21.11	63.16 ± 14.86	0.005*
Rad-Score/Median (25%, 75%)	5.24 (2.35, 9.47)	-5.94 (-8.69, -3.06)	<0.001*	3.41 (1.38, 5.72)	-5.17 (-8.91, -2.93)	<0.001*

*Significant difference ($p < 0.05$)

malignant, including 57 cases of adenocarcinoma (6 cases of local adenocarcinomas in the background of adenoma), 3 cases of neuroendocrine carcinoma, 1 case of adenosquamous carcinoma, and 82 cases were benign, including 44 cases of adenomatous polyps, 28 cases of cholesterol polyps, 8 cases of inflammatory polyps, and 2 cases of gallbladder adenomyomatosis (Fig. 3). There is no difference in the clinical characteristics between the training and testing groups. In the process of collecting data, nearly one-third of the patients did not have measured tumor markers, such as CA199, CA125, and CEA before the operation which resulted in too many missing values. Therefore, this study discarded these characteristics. Univariate correlation analysis was conducted between the remaining 11 features and the nature of the PLG. The results are shown in Fig. 4.

After statistical analysis, there were five characteristics with significant differences in the training group: max diameter, age, symptom, base, and arterial CT value. Regarding these characteristics as independent variables and benignity and malignancy of the PLG as dependent variables, a multivariate logistic regression model was constructed with a backward stepwise method, and the details are shown in Table 3. $\text{Logit}(P) = -10.26 + 0.0867 \times \text{age} + 1.3307 \times \text{symptom} + 1.5592 \times \text{base} + 0.0230 \times \text{arterial CT value} + 1.0357 \times \text{maximum diameter}$.

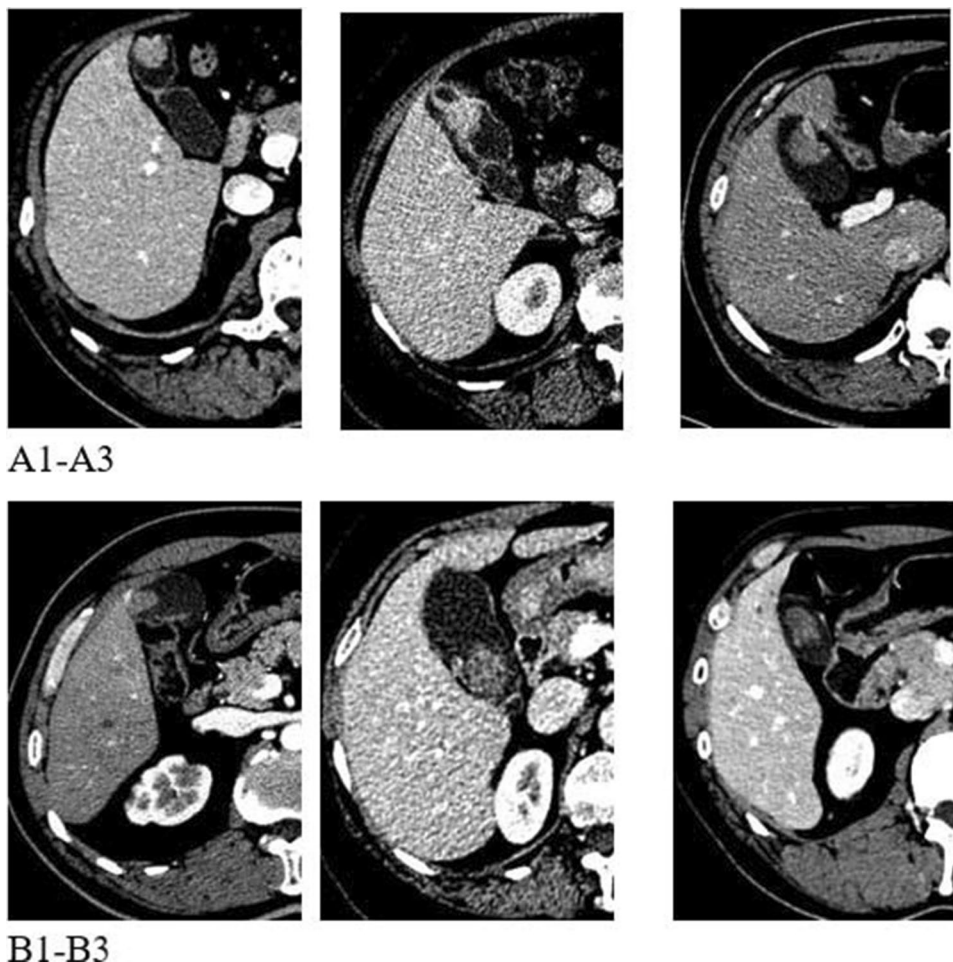
Using pathological diagnosis as the gold standard and $\text{Logit}(P)$ as the predictive value, ROC curves were drawn

in the training group and the testing group. The results are shown in Fig. 5. The diagnostic efficacy was validated with 0.474 as the cutoff value. The specificity and sensitivity of the test group were 0.680 and 0.895, respectively, with an $\text{AUC} = 0.861$, which showed that the prediction model had a certain level of accuracy.

Radiomics signature

After importing the VOI files and original images into the AK analysis platform, 396 radiomic features of every patient were extracted, including 9 morphological features, 42 first-order histogram features, 154 GLCM features, 180 GLRLM features, and 11 GLSZM features, which were extracted from each patient. The steps of the above GLCM and GLRLM features were 1/4/7, and the angles were 0 degrees, 45 degrees, 90 degrees, and 135 degrees. A total of 396 radiomics features were reduced to 15 potential predictors on the basis of 101 patients in the training group (Fig. 6), and these 15 predictors were features with nonzero coefficients in the LASSO logistic regression model. Finally, there were seven features involved in the construction of the radiomics signature. These features are presented in the Rad-score calculation formula as

Fig. 3 Portal CT images of some patients with PLG in our study. A1-A3: Pathological diagnosis confirmed adenocarcinoma with different differentiation. A1: Male, 59 years old, lesion was 1.8 cm located at the fundus of gallbladder; A2: Female, 66 years old, lesion was 1.7 cm located at the body of gallbladder; A3: Male, 62 years old, lesion was 3.0 cm located at the fundus of gallbladder. B1-B3, Pathological diagnosis confirmed adenoma. B1: Male, 56 years old, lesion was 1.7 cm located at the bottom of gallbladder; B2: Female, 65 years old, lesion was 2.8 cm located at the neck of gallbladder; B3: Male, 55 years old, lesion was 2.1 cm located at the fundus of gallbladder



shown. Specific parameters of the radiomics model are given in Table 4.

$$\begin{aligned} \text{Rad-score} = & -12.3 - 9.55 \times \beta_1 + 1.45 \times 10^8 \\ & \times \beta_2 + 15200 \times \beta_3 + 6.38 \times 10^{-9} \\ & \times \beta_4 - 2.11 \times 10^{-10} \times \beta_5 + 0.000401 \\ & \times \beta_6 + 0.781 \times \beta_7 \end{aligned}$$

β_1 : skewness, β_2 : Correlation_AllDirection_offset4_SD, β_3 : Correlation_angle45_offset7, β_4 : HaralickCorrelation_angle135_offset7, β_5 : LongRunHighGreyLevelEmphasis_AllDirection_offset4_SD, β_6 : HighIntensitySmallAreaEmphasis, β_7 : ZonePercentage.

There was a significant difference in Rad-score between benign and malignant PLG in the training group ($p < 0.0001$), which was then confirmed in the testing group ($p < 0.0001$). ROC curves were drawn in the training group and the testing group, and the results are shown

in Fig. 5. After the Delong test, there was no significant difference in the ROC curve between the two groups ($p = 0.2$), which suggested that the radiomics model was stable.

The diagnostic efficacy was validated with 0.370 as the cutoff value. The specificity and sensitivity of the test group were 0.880 and 0.895, respectively, with an AUC = 0.924, which showed that the prediction model had high accuracy. The results are shown in Table 5.

Combinations

As mentioned above, we incorporated clinical features and radiomics signatures of the training set into the model analysis and obtain the combined logistic regression model. The calculation formula was as follows:

$$\text{Logit}(P) = -9.1697 + 0.1577 \times \text{age} + 1.3724 \times \text{radscore}$$

Fig. 4 The spearman correlation heatmap of the remained 11 features and the group, which was the gold standard that defined the benign or malignancy of the PLG. If the correlation between two features is positive, the block would be colored blue, otherwise colored red. The stronger the correlation is, the block would be bigger and then the color would be darker. As shown, the feature named “max diameter” has the strongest correlation with the probability of malignancy of PLG, followed by base, age, and symptom. Meanwhile, we can see that the correlation between “Portal CT value” and “Arterial CT value” was very high, and “Portal CT value” was also highly correlated with “Venous CT value.” This helps explain why only five features remained in the multivariate logistic regression analysis when using backward stepwise method

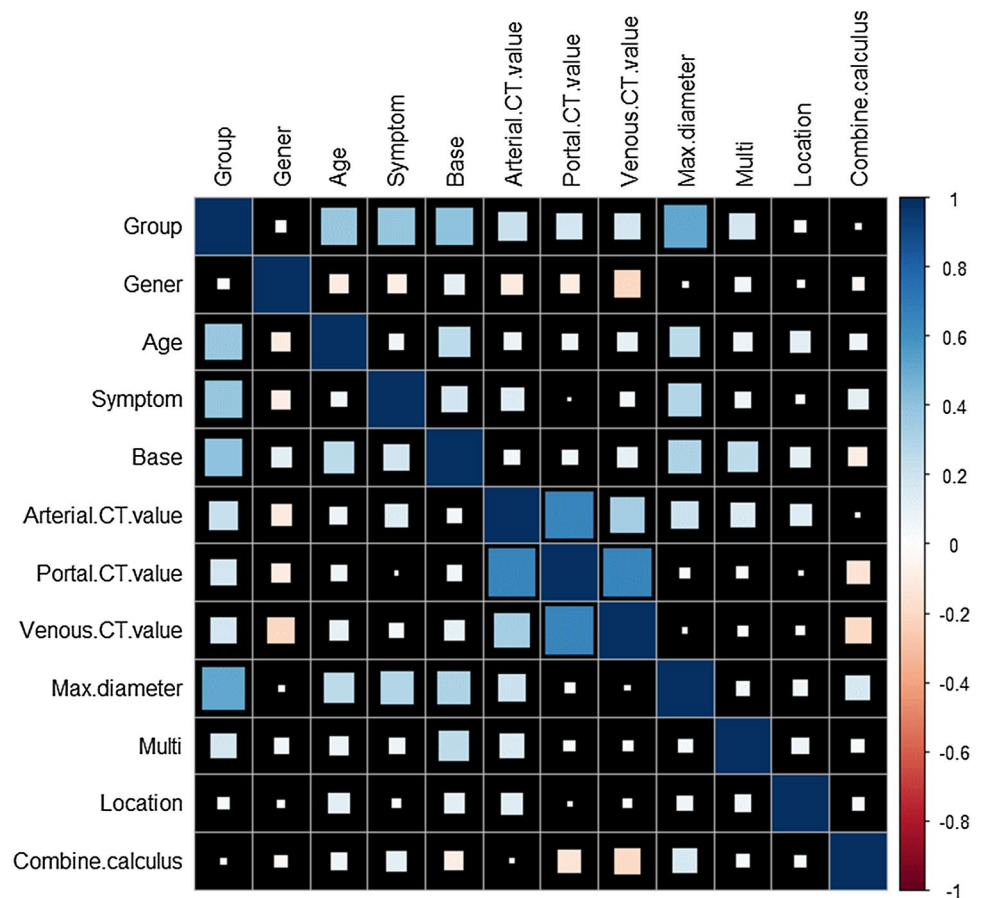


Table 3 Parameters of the multivariate logistic regression model

	Coefficient	Standard error	Z	P
Constant	-10.2585	2.2172	-4.63	0.0000
Age	0.0867	0.0286	3.03	0.0025
Symptom	1.3307	0.5965	2.23	0.0257
Base	1.5592	0.5899	2.64	0.0082
Arterial CT value	0.0230	0.0149	1.55	0.1216
Maximum diameter	1.0357	0.3943	2.63	0.0086

There was also a significant difference in combined model between benign and malignant PLG in the training group ($p < 0.001$), which was then confirmed in the testing group ($p < 0.001$). ROC curves were drawn in the training group and the testing group to evaluate the efficiency of the model. The AUC value was 0.994 in the training set and 0.931 in the testing set. The results are shown in Fig. 5. Although the AUC value of the above three models in the testing group showed a trend of $0.931 > 0.924 > 0.861$, the difference was not statistically significant after Delong test.

Discussion

In this study, we developed and validated radiomics signature-based, clinical feature-based, and combined diagnostic models for the preoperative prediction and differentiation of benign and malignant PLG. In addition, we calculated the AUC, sensitivity, and specificity of every model and compared the significant differences in the diagnostic efficiency of the three models. Our study suggested that combined diagnostic models based on the radiomics signature and clinical features can effectively differentiate benign and malignant gallbladder polyps ≥ 1 cm in size. However, after the Delong test, the differences between the AUC values of the combined model and the two sole models were statistically nonsignificant.

For the construction of the multivariate logical regression model based on clinical characteristics, 11 candidate clinical features were reduced to 5 features that influenced and determined the benignity and malignancy of the PLG and were

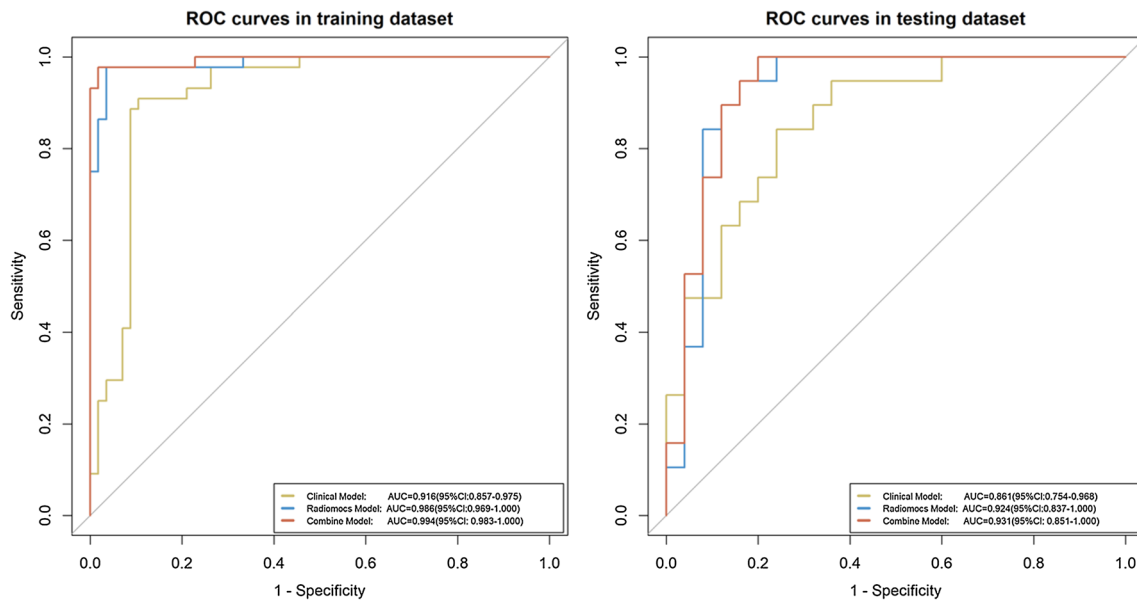


Fig. 5 ROC curves and AUC values for the clinical signature (yellow line) and radiomic signature (blue line) models and the ROC curve for the combination of the two signatures (red line). The left panel is in the training set, and the right panel is in the testing set

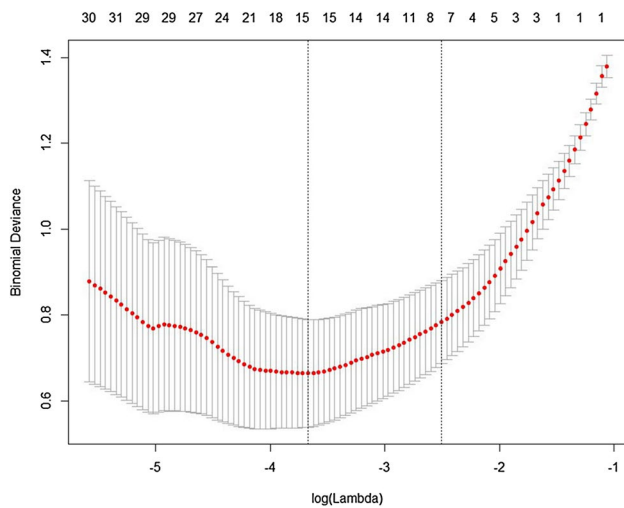


Fig. 6 Fig. 6: Feature selection using the least absolute shrinkage and selection operator (LASSO) binary regression model. The tenfold cross-validation was repeated 100 times to generate the optimal tuning parameter in the LASSO model. The value of lambda that gave the minimum binomial deviance was used to select features. Dotted vertical lines were drawn at the optimal values by using the minimum criteria and the 1 standard error of the minimum criteria. Seven features were remained in the final LASSO regression model

significantly different in the benign and malignant sets in the training group. Taking the above five factors as independent variables, a multivariate logistic regression model was built and validated in the testing group. The specificity and sensitivity were 0.680 and 0.895, respectively, in the testing group, and the AUC was 0.861. The results suggested that

the model had a certain level of accuracy. In addition, univariate correlation analysis showed that the feature of maximum diameter was most related to the benign and malignant nature of the PLG. Remarkably, in the testing group, there was no statistically significant difference in ages between those with benign and malignant PLG, possibly resulting from an unbalanced age distribution due to the small sample size.

For the construction of the radiomics signature, 396 candidate radiomics features were reduced to 15 potential predictors by shrinking the regression coefficients with the LASSO method, and the radiomics signature was finally constructed by multiple logistic regression with seven features selected, which influenced and determined the classification of the benignity and malignancy of the PLG. After validation in the testing group, the specificity and sensitivity were up to 0.880 and 0.895, respectively, and the AUC reached 0.924, which shows that the radiomics signature has high accuracy.

This study focused on PLG whose maximum diameter was not less than 10 mm because PLG smaller than 10 mm in diameter are generally managed by observation because of their low risk of malignancy [12], and it is difficult to obtain pathological diagnoses due to the lack of surgery. In clinical practice, the size of the PLG is one of the most important predictors for differentiating the nature of the lesions and an indication for surgery. However, a study showed that among the patients with PLG (≥ 10 mm) undergoing surgery, nearly 83% of the cases were benign [13], so it is not reliable to judge the nature of PLG only by their size [14]. In addition,

Table 4 Specific parameters of the radiomics model

	Coefficient	Standard error	Z	P
Constant	−12.3	3.27	−3.76	0.00017
Skewness	−9.55	3.38	−2.82	0.00478
Correlation_AllDirection_offset4_SD	1.45×10^8	6.42×10^7	2.25	0.02430
Correlation_angle45_offset7	15200	7.19×10^3	2.11	0.03518
HaralickCorrelation_angle135_offset7	6.38×10^{-9}	3.36×10^{-9}	1.90	0.05740
LongRunHighGreyLevelEmphasis_All-Direction_offset4_SD	-2.11×10^{-10}	1.58×10^{-10}	−1.34	0.18117
HighIntensitySmallAreaEmphasis	0.000401	2.62×10^{-4}	1.53	0.12685
ZonePercentage	0.781	0.291	2.69	0.00715

Table 5 Discriminant efficacy of Rad-score in the training and testing group

	AUC	95% CI	ACC	SPE	SEN	PPV	NPV	Cutoff value
Training set	0.986	0.969–1.000	0.970	0.965	0.977	0.956	0.982	0.370
Testing set	0.924	0.837–1.000	0.886	0.880	0.895	0.850	0.917	0.370

because of the heterogeneity in CT image acquisition, image preprocessing was applied before the extraction of radiomics features in this study.

Ultrasonography has been considered to be the preferred modality for the detection of PLG because it can not only discover polypoid gallbladder cavity lesions with high resolution but also visualize vessels in PLG with color Doppler imaging. In recent years, endoscopic ultrasonography (EUS) combined with contrast enhancement has been used in the diagnosis of malignant PLG [15]. However, it is often difficult to differentiate between benign and malignant PLG with an ultrasonography examination [16].

CT is also a common diagnostic modality for PLG in clinical practice, and its main advantage is to analyze the relationship between the lesion and surrounding tissues to provide help for the qualitative diagnosis of PLG combined with contrast-enhanced scanning. However, a study has shown that CT is even less effective than ultrasound in the differentiation of benign and malignant PLG [17]. Zhou et al. [5] found a seemingly good indicator based on the “delayed enhancement” effect of gallbladder cancer for the differentiation of benign and malignant PLG by measuring the CT values of lesions in arterial, portal and delayed phases. However, we also noted the limitation of this method that the cancerous area may not in the slices they selected and studied for locally cancerous polyps, which will lead to a wrong judgment because they were looking at benign areas, but the pathological diagnosis was actually malignant PLG. In this study, there were 6 cases of local adenocarcinomas in the background of adenoma, but the VOI was delineated in this study, which means that all slices of the lesion were included in the study and avoids that problem.

It is in recent years that MRI has been used as a diagnostic modality for the diagnosis of gallbladder cancer based on diffusion-weighted imaging (DWI); the main limitation of DWI is its limited sensitivity and specificity owing to the “T2 shine-through” effect [4].

The present study is the first time that radiomics based on CT images was used to differentiate the benignity and malignancy of PLG, which not only makes full use of available information from medical images without additional financial burden but also searches for invisible image features to help clinicians make objectively decisions.

This study has several limitations. First, the study was a retrospective analysis. Second, this was a single-center study, and further validation from multicenter and large sample studies is needed. Third, the fact that serological tumor marker characteristics were not included because of too many missing values.

Conclusion

In conclusion, this study presents a combined diagnostic model based on the radiomics signature and clinical features for the effective preoperative prediction and differentiation of benign and malignant PLG, which has the potential to provide complementary information for the decision-making of malignant risk of gallbladder polyps ≥ 1 cm in size before surgery.

Author contributions Xu and Niu contributed to the study design. Yang, Liu, and Chai carried out the collection and assembly of data and drafted the manuscript. Guo did the feature extraction statistical work. All authors reviewed the manuscript. All authors read and approved the final manuscript.

Funding No funding.

Data availability All authors have reviewed the final version of the paper and would like to take public responsibility for its content. The authors declare that they had full access to all of the data in this study and the authors take complete responsibility for the integrity of the data and the accuracy of the data analysis.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

Ethics approval Ethical approval was obtained for this retrospective analysis, and the informed consent requirement was waived.

Informed consent Not applicable.

References

- Park CH, Chung MJ, Oh TG et al. Differential diagnosis between gallbladder adenomas and cholesterol polyps on contrast-enhanced harmonic endoscopic ultrasonography. *Surgical endoscopy* 2013;27:1414-21.
- Lee KF, Wong J, Li JC, Lai PB. Polypoid lesions of the gallbladder. *American journal of surgery* 2004;188:186-90.
- Kim SW, Kim HC, Yang DM, Ryu JK, Won KY. Gallbladder carcinoma: causes of misdiagnosis at CT. *Clinical radiology* 2016;71:e96-109.
- Ogawa T, Horaguchi J, Fujita N et al. High b-value diffusion-weighted magnetic resonance imaging for gallbladder lesions: differentiation between benignity and malignancy. *Journal of gastroenterology* 2012;47:1352-60.
- Zhou W, Li G, Ren L. Triphasic Dynamic Contrast-Enhanced Computed Tomography in the Differentiation of Benign and Malignant Gallbladder Polypoid Lesions. *Journal of the American College of Surgeons* 2017;225:243-248.
- Zielinski MD, Atwell TD, Davis PW, Kendrick ML, Que FG. Comparison of surgically resected polypoid lesions of the gallbladder to their pre-operative ultrasound characteristics. *Journal of gastrointestinal surgery : official journal of the Society for Surgery of the Alimentary Tract* 2009;13:19-25.
- van Timmeren JE, Leijenaar RTH, van Elmpt W et al. Survival prediction of non-small cell lung cancer patients using radiomics analyses of cone-beam CT images. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 2017;123:363-369.
- Kim J, Choi SJ, Lee SH, Lee HY, Park H. Predicting Survival Using Pretreatment CT for Patients With Hepatocellular Carcinoma Treated With Transarterial Chemoembolization: Comparison of Models Using Radiomics. *AJR American journal of roentgenology* 2018;211:1026-1034.
- Yang L, Dong D, Fang M et al. Can CT-based radiomics signature predict KRAS/NRAS/BRAF mutations in colorectal cancer? *European radiology* 2018;28:2058-2067.
- Giraud P, Giraud P, Gasnier A et al. Radiomics and Machine Learning for Radiotherapy in Head and Neck Cancers. *Frontiers in oncology* 2019;9:174.
- Larue R, van Timmeren JE, de Jong EEC et al. Influence of gray level discretization on radiomic feature stability for different CT scanners, tube currents and slice thicknesses: a comprehensive phantom study. *Acta oncologica (Stockholm, Sweden)* 2017;56:1544-1553.
- Park JY, Hong SP, Kim YJ et al. Long-term follow up of gallbladder polyps. *Journal of gastroenterology and hepatology* 2009;24:219-22.
- Park JK, Yoon YB, Kim YT et al. Management strategies for gallbladder polyps: is it possible to predict malignant gallbladder polyps? *Gut and liver* 2008;2:88-94.
- Miwa H, Numata K, Sugimori K et al. Differential diagnosis of gallbladder polypoid lesions using contrast-enhanced ultrasound. *Abdominal radiology (New York)* 2019;44:1367-1378.
- Choi JH, Seo DW, Choi JH et al. Utility of contrast-enhanced harmonic EUS in the diagnosis of malignant gallbladder polyps (with videos). *Gastrointestinal endoscopy* 2013;78:484-93.
- Irie H, Kamochi N, Nojiri J, Egashira Y, Sasaguri K, Kudo S. High b-value diffusion-weighted MRI in differentiation between benign and malignant polypoid gallbladder lesions. *Acta radiologica (Stockholm, Sweden : 1987)* 2011;52:236-40.
- Jang JY, Kim SW, Lee SE et al. Differential diagnostic and staging accuracies of high resolution ultrasonography, endoscopic ultrasonography, and multidetector computed tomography for gallbladder polypoid lesions and gallbladder cancer. *Annals of surgery* 2009;250:943-9.

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