Am J Cancer Res 2024;14(9):4445-4458
www.ajcr.us /ISSN:2156-6976/ajcr0158774

Original Article

Preoperative CT-based intra- and peri-tumoral radiomic models for differentiating benign and malignant tumors of the parotid gland: a two-center study

Qian Shen[1,2*], Cong Xiang[3*], Kui Huang[4], Feng Xu[5], Fulin Zhao[5], Yongliang Han[1], Xiaojuan Liu[3], Yongmei Li[1]

[1]Department of Radiology, The First Affiliated Hospital of Chongqing Medical University, Chongqing 400016, China; [2]Department of Radiology, The Affiliated Stomatology Hospital of Southwest Medical University, Luzhou 646000, Sichuan, China; [3]School of Artificial Intelligence, Chongqing University of Technology, Chongqing 400016, China; [4]Department of Oral and Maxillofacial Surgery, The Affiliated Stomatology Hospital of Southwest Medical University, Luzhou 646000, Sichuan, China; [5]Department of Radiology, The Affiliated Hospital of Southwest Medical University, Luzhou 646000, Sichuan, China. [*]Equal contributors and co-first authors.

Received June 26, 2024; Accepted September 10, 2024; Epub September 15, 2024; Published September 30, 2024

Abstract: Objective: To investigate the ability of intra- and peritumoral radiomics based on three-phase computed tomography (CT) to distinguish between malignant and benign parotid tumors. Methods: We conducted a retrospective analysis of data from 374 patients with parotid gland tumors, all confirmed by histopathology. A total of 321 patients from Center 1 (January 2014 to January 2023) were randomly divided into the training set and internal testing set at a ratio of 7:3, whereas 53 patients from Center 2 (January 2020 to June 2022) constituted the external testing set. CT images of both the tumor and surrounding areas (2 mm and 5 mm areas surrounding the tumor) were reviewed, and their radiomic features were extracted for the construction of different radiomic models. In addition, a combined clinical-radiomic model was developed using multivariate logistic regression analysis. The model's predictive performance was evaluated using decision curve analysis (DCA) and receiver operating characteristic (ROC) curves. Results: Among the models evaluated, Tumor + External2 model demonstrated superior predictive performance. The areas under the curve (AUCs) of this model were 0.986 in the training set, 0.827 in the internal test set, and 0.749 in the external test set. For the clinical model, independent predictive factors included symptoms, boundaries, and lymph node swelling. The combined clinical-radiomic model achieved AUCs of 0.981, 0.842, and 0.749 in the three cohorts, outperforming both the Tumor model and the clinical model individually. Conclusion: The CT-based radiomic models incorporating intratumoral and peritumoral radiomic features can effectively distinguish malignant from benign parotid tumors, and the predictive accuracy is further improved by incorporating clinically independent predictors.

Keywords: Parotid neoplasms, radiomics, machine learning, intratumoral, peritumoral

Introduction

Among salivary gland tumors, parotid gland tumors have a higher incidence, with a benign-to-malignant ratio of approximately 4:1 [1]. While surgical treatment is commonly employed for patients with parotid gland tumors, it is important to note that different surgical approaches are selected based on whether the tumor is benign or malignant. For benign parotid tumors (BPTs), the preferred surgical option is superficial or local parotidectomy, whereas for malignant parotid tumors (MPTs), more invasive procedures are required, such as partial or total parotidectomy, postoperative chemoradiation, and facial nerve excision [2]. Thus, the selection of appropriate treatment heavily relies on accurate preoperative identification, which is crucial for patient prognosis.

Given that these tumors often lack distinct clinical manifestations, imaging analysis and fine needle aspiration biopsy (FNAB) are crucial for preoperative differentiation of benign and malignant parotid gland tumors. Although FNAB has been regarded as a routine clinical tech-

nique for the preoperative classification of parotid tumors, it is associated with severe surgical complications [3, 4]. Moreover, due to the overlapping radiological characteristics of malignant and benign parotid tumors, imaging results can sometimes be inconclusive, depending on the radiologist's expertise for interpretation [5]. Therefore, there is an urgent need to develop more efficient and non-invasive assessments to improve the preoperative discrimination of parotid tumors.

Radiomics offers a promising noninvasive approach for characterizing tumors and their adjacent microenvironments. This method converts traditional medical images into high-throughput quantitative imaging signals that beyond human vision scope, revealing inherent tumor heterogeneity and phenotypes [6, 7]. Prior research has underscored the robust differentiation capability of computed tomography (CT) radiomics in distinguishing lymphoma-associated malignant from benign ones in parotid glands and benign parotid tumors [8, 9]. In the context of parotid tumors, radiomics models offer non-invasive and objective assessments, aiding clinical decision-making. By extracting features from both the tumor and its surrounding tissue, radiomics models can capture the intricate interplay between the tumor and its microenvironment, which is crucial for understanding tumor biology and predicting clinical outcomes. Most previous radiomics studies have focused on distinguishing between specific subtypes of benign tumors or differentiating potential malignancies from benign tumors, predominantly focusing on the primary tumor area [10-12]. Nevertheless, recent investigations have revealed that the surrounding area may contain supplementary information regarding tumor heterogeneity across various cancer types [13]. Thus, the immediate peritumoral regions may present promising diagnostic value for extracting imaging biomarkers. Evidence suggests that radiomics features from the immediate surrounding areas adjacent to the tumor are valuable in distinguishing disease subtypes in breast cancer, lung cancer, and hepatocellular carcinoma [14, 15]. To our knowledge, no studies have utilized radiomics of the peritumoral regions to differentiate malignant from benign parotid tumors.

Therefore, we hypothesized that radiomics analysis of intratumoral regions would be helpful for clinical diagnosis. To test this, we constructed CT radiomics models based on the features extracted from different regions both around and within the tumor and evaluated the predictive performance of these models.

Materials and methods

Patients

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and was approved by the Institutional Review Board of the First Affiliated Hospital of Chongqing Medical University (No. K2023-414). The requirement for informed consent was waived in this study due to its retrospective nature.

We retrospectively collected data from 374 patients who were histopathologically diagnosed with benign parotid tumors (BPTs) or malignant parotid tumors (MPTs) at two centers (Figure S1). The inclusion criteria for this study: (1) patients had not received radiotherapy or chemotherapy and had no history of FNAB; (2) complete imaging and clinical data; and (3) all patients underwent dual-phase enhanced scanning and CT plain scan before surgery. The exclusion criteria: (1) CT images with obvious artifacts or noise; (2) tumor maximum diameter not exceeding 10 mm; and (3) presence of other types of tumors.

A total of 321 patients from Center 1 (January 2014 to January 2023) were randomly divided into a training set and an internal testing set at a ratio of 7:3, whereas 53 patients from Center 2 (January 2020 to June 2022) were regarded as the external testing set. The distribution of parotid tumors is presented in Table S1.

Image acquisition

Each patient underwent axial multi-phase scanning using multi-slice spiral CT equipment, including plain scanning, arterial phase, and venous phase imaging. Detailed information regarding the imaging protocols used in both centers is presented in Table S2.

Radiological and clinical data analysis

Patient imaging and clinical data were collected from the case management system, image archives, and other sources. A retrospective

analysis was conducted on clinical data, including variables such as smoking status and age. CT images were evaluated independently by two radiologists with extensive diagnostic experience, who were blinded to the pathological diagnoses of the parotid tumors. The two radiologists discussed and reached a consensus on different opinions. The radiological features assessed included tumor location, shape, number, maximum diameter, distribution, boundary, calcification, density, cystic areas, uniformity and degree of enhancement, peak enhancement phase, and presence of enlarged lymph nodes. Definitions for each of these radiological features are provided in Supplementary Appendix 1.

Image segmentation

Figure 1 provides an overview of the study workflow. The CT images of patients were stored in DICOM format using standard soft tissue settings. For evaluating histopathological results, two radiologists independently performed a blind evaluation, meaning they had no knowledge of the relevant tissue content. During the analysis, the radiologists manually segmented the region of interest (ROI). When analyzing axial multiphase CT images, they carefully delineated the edge of the tumor layer by layer, removing normal tissues, blood vessels, and other non-tumor areas. Inter-observer and intra-observer reproducibility were assessed using intragroup correlation coefficient (ICC). Forty samples were randomly selected, including 20 malignant and 20 benign tumor samples, and ROI segmentation was performed on these images by two radiologists. Both radiologists independently performed segmentation on the image, allowing for the evaluation of the consistency of radiological features identified by different observers. This segmentation process was repeated one month later, and the ICC values were found to be greater than 0.9, indicating excellent consistency.

After manual tumor segmentation, 2-mm and 5-mm peritumoral regions were automatically segmented using Python (version 3.7.12; http://www.python.org) (**Figure 2**). Next, the bone and air were filtered from the delineation by setting the maximum and minimum thresholds, and the final ROI borders (peritumoral regions) were manually adjusted to ensure accuracy [16, 17].

Radiomics feature extraction and selection

PyRadiomics in Python was used for feature extraction. Standardization and resampling techniques were applied to preprocess the images and data to ensure the consistency of the CT images across patients.

Features were extracted from five different ROIs-Tumor, External2, External5, Tumor + External2, and Tumor + External5- for each patient using the PyRadiomics Radiomics Feature Extractor toolbox. A total of 1688 radiological features were obtained for each segmentation, including 252 grayscale dependency matrices, 432 grayscale co-occurrence matrices, 90 adjacent grayscale difference matrices, 288 grayscale size region matrices, 288 grayscale run length matrices, among others. Subsequently, the extracted features were analyzed for both the tumor and its surrounding areas. The least absolute shrinkage and selection operator (LASSO) algorithm was used to select radiomic features. When optimizing the target, a loss function was used. To obtain accurate experimental results, the penalty term (λ) was added to the function; At the optimal λ, features with non-zero coefficients were retained, and their coefficients were ordered by magnitude. To avoid overfitting during the analysis process, each feature corresponded to 10 samples based on empirical rules, ultimately leading to the selection of the top 20 features. These algorithms for extracting radiomics features adhered to the Image Biomarker Standardization Initiative (IBSI) guidelines.

Before further analysis, the data of patients from Center 1 were randomized into the training and internal testing sets in a 7:3 ratio, whereas those from Center 2 were regarded as an independent external validation set. This cohort was collected from a different hospital to ensure that the model's performance could generalize to different populations. The extracted radiomics features were normalized using Z-scores to address the differences in the value scales of the feature. All feature selection processes were executed in the training dataset, and the intraclass correlation coefficients

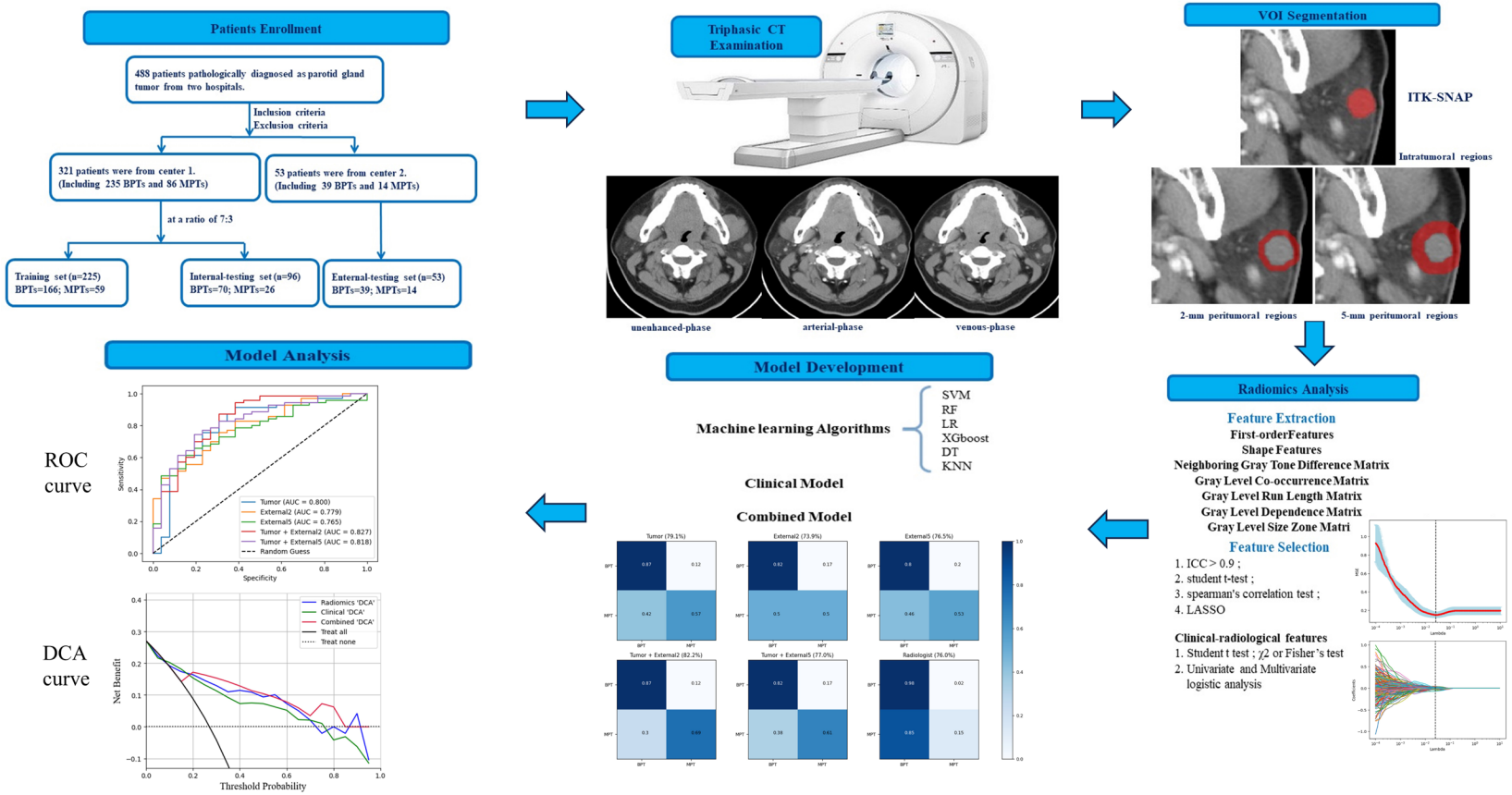

Figure 1. Workflow of this study.

Figure 2. Contrast-enhanced CT image from a pleomorphic adenoma patient, highlighted regions represent the primary tumor (A) and peritumoral region of 2 mm (B) and 5 mm (C).

(ICCs) between the features extracted by the two radiologists were calculated. Radiomics features with ICC value < 0.9 were eliminated. Among the remaining features, those representing significant differences between the BPT and MPT groups were identified using the student's t-test. And then, Spearman's correlation test was employed, and features with coefficients greater than 0.95 were removed to address redundancy. Finally, the top 20 discriminative radiomics features were selected using the LASSO algorithm.

Model construction and evaluation

The tumor model was constructed using radiomics features extracted from the tumor region [18]. The radiomics features were extracted from the tumor ROI across all three phases of CT scans: plain, arterial, and venous, using the PyRadiomics library. The LASSO algorithm was used to select the most discriminative features from the tumor ROI, resulting in a refined subset of features. These features were then used to construct the tumor model. Besides, a clinical model was developed based on independent clinical predictors identified through multivariate logistic regression analysis. These predictors were chosen based on their statistical significance in differentiating between benign and malignant parotid tumors. The clinical model was constructed using these predictors to assess their predictive power in the discrimination of tumor types.

Previous studies confirmed that CT image-based machine learning, support vector machine (SVM) classifiers, have high accuracy for discriminating MPTs from BPTs [18]. Therefore, we used the CT radiomics features to establish the SVM model as the baseline model.

In the training set, the synthetic minority oversampling technique (SMOTE) was used to balance the minority samples in a 1:1 ratio. Radiomics feature fusion was performed by combining the top 20 features from each region, resulting in a total of 40 features (Tumor + External2, and Tumor + External5). Subsequently, Principal Component Analysis (PCA) was used to determine the optimal number of retained features, aiming to preserve 95% of the total variance in the data. The PCA method is described in Supplementary Appendix 2. The receiver operating characteristic (ROC) curves of the radiomics models across the five cohorts were plotted, and the sensitivity and specificity as well as the area under the curve (AUC) was calculated. Calibration curves were plotted to evaluate the predicted probability of the models in the testing sets, with calibration performance quantified using Brier scores; a Brier score closer to zero indicates better calibration performance. The diagnostic confusion matrix of the radiologists was calculated and compared with that of the five radiomics signature models. Moreover, six different machine-learning algorithms were employed to identify the optimal classifier.

Machine learning is known to provide highly reliable, accurate, and objective models to assist in clinical decision-making [19]. The best radiomics signature model was selected to develop a combined model incorporating independent predictors. To evaluate the clinical efficiency of the model for tumor categorization, we quantified the net benefits of different

Table 1. Clinical and CT morphological characteristics of patients in the training and two validation cohorts

Variables	Training set			Internal-testing set			External-testing set		
	BPTs (n=166)	MPTs (n=59)	*P* value	BPTs (n=47)	MPTs (n=26)	*P* value	BPTs (n=39)	MPTs (n=14)	*P* value
Age[b] (years)	50.475 (15.518)	52.307 (14.539)	0.495	44.692 (17.211)	52.100 (14.539)	0.045*	55.857 (17.347)	58.231 (12.854)	0.694
Max-diameter[b] (cm)	2.429 (0.775)	2.201 (0.718)	0.039*	2.265 (1.064)	2.061 (0.746)	0.635	2.729 (1.013)	2.467 (0.810)	0.289
Sex[a] (F/M)	99/67	31/28	0.361	29/41	7/19	0.097	24/15	10/4	0.746
Smoking[a] (Yes/No)	75/91	19/40	0.092	75/91	19/40	0.092	17/22	4/10	0.362
Alcohol consumption[a] (Yes/No)	59/107	14/45	0.107	31/39	5/21	0.107	13/16	3/11	0.510
Number of nodules[a] (Single/Multiple)	178/12	59/0	0.075	64/6	26/0	0.186	31/8	11/3	1.000
Symptom[a] (With/Without)	15/151	20/39	< 0.001*	5/65	10/16	0.001*	10/29	6/8	0.311
Shape[a] (Round/Non-round)	129/37	49/10	0.458	51/19	24/2	0.052	30/9	12/2	0.706
Margin[a] (Clear/Unclear)	159/7	37/22	< 0.001*	65/5	12/14	< 0.001*	34/5	6/8	0.002*
Location[a] (Superficial/Deep/Both)	134/2/30	17/13/29	0.216	62/0/8	17/0/9	0.635	33/0/5	11/1/2	0.246
Distribution[a] (left/right/both)	86/70/8	28/29/1	0.425	32/36/2	14/12/0	0.574	13/21/5	8/6/0	0.173
Density[a] (Homogeneous/Heterogeneous)	66/100	29/30	0.223	21/49	13/13	0.093	17/22	5/9	0.755
Calcification[a] (With/Without)	4/162	19/40	0.654	2/68	1/25	1.000	0/39	0/14	1.000
Cystic areas (With/Without)	47/119	19/40	0.619	16/54	8/18	0.437	19/20	6/8	0.763
Enhanced-peak phase[a] (Arterial/Venous)	68/98	21/38	< 0.001*	38/32	16/10	0.645	27/12	11/3	0.732
Enhancement degree[a] (Slight/Moderate/Obvious)	13/44/109	8/19/32	0.230	5/18/9	3/9/14	0.005*	5/11/23	1/6/7	0.569
Enhanced uniformity[a] (Yes/No)	70/96	18/41	0.123	38/32	16/10	0.645	54/93	20/34	0.569
Enlarged lymph nodes[a] (With/Without)	1/165	15/44	< 0.001*	1/69	5/21	0.005*	7/32	4/10	0.001*

*Represents P < 0.05. [a]Categorical data are presented as numbers (n). [b]Quantitative data are presented as means (standard deviations) or medians (quartiles), *p* value was calculated using the independent samples t-test or Mann-Whitney U test. *p*-value was calculated with the χ^2 or Fisher's exact test. BPTs, benign parotid tumors; MPTs, malignant tumors; F, female; M, male.

Table 2. Univariable and multivariable logistic regression analysis of factors in the training cohort

Variable	Univariate analysis		Multivariate analysis	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Age (years)	1.000 (0.969-1.031)	0.975		
Max-diameter (cm)	1.487 (0.692-3.092)	0.309		
Sex	0.485 (0.132-1.779)	0.275		
Smoking	0.474 (0.116-1.934)	0.298		
Alcohol consumption	1.355 (0.372-4.934)	0.645		
Number of nodules	0.025 (0.001-0.736)	0.033*	0.144 (0.012-1.747)	0.128
Symptom	6.901 (2.271-20.974)	0.001*	5.707 (2.205-14.770)	< 0.001*
Shape	0.432 (0.136-1.377)	0.156		
Margin	30.316 (7.567-121.454)	< 0.001*	15.167 (5.116-44.968)	< 0.001*
Location	0.489 (0.228-1.052	0.067		
Distribution	2.000 (0.867-4.613)	0.104		
Density	0.750 (0.185-3.046)	0.688		
Calcification	0.089 (0.003-2.866)	0.172		
Cystic areas	1.347 (0.361-5.026)	0.658		
Enhanced peak phase	0.829 (0.300-2.294)	0.718		
Enhancement degree	0.728 (0.382-1.387)	0.334		
Enhanced uniformity	1.571 (0.485-5.086)	0.451		
Enlarged lymph nodes	806.789 (34.818-18694.732)	< 0.001*	131.615 (11.699-1480.670)	< 0.001*

*Represents P < 0.05. OR, odds ratio; CI, confidence interval.

threshold probabilities in the testing set using decision curve analysis (DCA).

Statistical analyses

All statistical analyses were performed using PyRadiomics in Python (version 3.7.12; http://www.python.org), R (version 3.6.3; https://www.r-project.org), and SPSS (version 26.0; IBM, Armonk, NY, USA) software. Categorical variables were presented as [n (%)] and analyzed using Chi-square test. Continuous variables were tested for normal distribution using the Shapiro-Wilk method. Normally distributed continuous variables were expressed as (Mean ± SD) and analyzed using the t-test with adjusted variance. Non-normally distributed continuous variables were presented as median (25th percentile, 75th percentile) and analyzed using the Wilcoxon rank-sum test. Two-sided P < 0.05 was deemed statistically significant for all statistical tests. Correlation analysis was performed using Spearman's correlation test. Multivariate logistic regression analysis was used to identify independent clinical predictors of malignancy. A clinical model was constructed using the identified predictors. The model was validated using the same internal and external validation strategies as the radiomics models. The "sklearn" and "RMDA" packages were used for plotting the curves of the ROC and the DCA, respectively.

Results

Population and radiological features of patients

Details of the patient's clinical and radiological features are presented in **Table 1**. There were no significant differences in the maximum diameter, age, sex, smoking, alcohol consumption, shape, location, distribution, calcification, enhancement degree, cystic areas, enhanced peak phase, density, and enhanced uniformity between malignant and benign parotid tumor cohorts (all *P* > 0.05) (**Table 2**). In addition, symptoms, boundaries, and enlarged lymph nodes were identified as independent predictors of malignant tumors through multivariate logistic regression analysis *(P* < 0.05). Based on these predictors, a clinical model was developed.

Radiomic signature models and performances

The selected features demonstrated high reproducibility, with ICCs exceeding 0.9, indicat-

ing that none were excluded during the reliability screening process. A total of 1688 radiomic features were extracted from each single ROI, resulting in 5064 features (Tumor, External2, External5) extracted from the images in the three scanning phases. Subsequently, the LASSO algorithm identified 20 discriminative radiomics characteristics for each intra- or peritumoral region (Figure S2).

The ROC curves of the five radiomics models in the training and testing sets are shown in **Figure 3**. In the internal and external testing sets, the Tumor + External2 model achieved AUCs of 0.827 and 0.745, which were higher than those of the Tumor (AUCs of 0.800 and 0.741), External2 (AUCs of 0.779 and 0.612), External5 (AUCs of 0.765 and 0.625), and Tumor + External5 (AUCs of 0.818 and 0.742) models. Furthermore, we comprehensively assessed the specificity, accuracy, sensitivity, negative predictive value (NPV), and positive predictive value (PPV) of the radiomics models, and the Tumor + External2 model displayed the best performance among the five models (**Table 3**).

The calibration curves, along with Brier scores, demonstrated good calibration for all five radiomics models in the testing sets (**Figure 4**), with the Tumor + External2 model showed the best calibration performance with the lowest Brier scores. The diagnostic confusion matrix for the radiologists was calculated and compared with those of the five radiomic signatures (Figure S3). This comparison revealed that the Tumor + External2 model surpassed both the other four models and the radiologists in terms of total diagnostic accuracy. Consequently, the Tumor + External2 model was selected as the best radiomic model for constructing a combined model.

Compared to other common machine learning algorithms, SVM algorithm achieved the highest AUC value based on the Tumor + External2 model. The results are presented in Table S3.

Combined model construction and validation

We established a combined model by incorporating the Tumor + External2 radiomics model with clinical predictors, including symptoms, enlarged lymph nodes, and borders. The combined clinical and radiomics model was constructed using SVM to evaluate and validate the diagnostic effectiveness of the different models. As shown in **Figure 5** and **Table 4**, the combined model had significantly better diagnostic performance compared to the other two models, with an ACU of 0.981. Analysis of the test set indicated that the combined model had better predictive performance than the tumor model alone. In addition, Figure S4 shows the DCA for the different models. The analysis reveals that, in the testing set, the combined model had a more significant net benefit in the classification of parotid gland tumors.

Discussion

Our study builds upon previous research in radiomics for differentiating between benign and malignant parotid tumors. Most prior studies have primarily focused on the use of intratumoral radiomics features [20], with limited attention given to the peritumoral region. For instance, Piludu et al. utilized radiomics features extracted from the tumor region to differentiate benign from malignant parotid tumors [21]. Similarly, Yu. et al. reported a radiomics model based solely on intratumoral features [22]. In this study, we evaluated the potential of intra-and peritumoral radiomics features to discriminate between malignant and benign parotid tumors. We found that the peritumoral parenchyma may contain some useful information for discriminating malignant from benign parotid tumors. The inclusion of intrinsic radiological features from the peritumoral region, irrespective of its size, resulted in more accurate indicators for discrimination. Specifically, the Tumor + External2 radiomics model demonstrated the best performance among the five models, both in the external (AUC=0.818) and internal (AUC=0.827) validation sets. Furthermore, compared with other machine learning algorithms, models based on SVM classifiers showed significantly better performance. Therefore, when clinical staff conducted preoperative evaluations, using a combination model of Tumor + External2 radiomics features and clinical predictive factors can achieve better evaluation results.

Previous radiomics research on parotid tumors mostly focused on the tumor parenchyma, without considering the tumor-surrounding tissue. However, surrounding tumor environment may

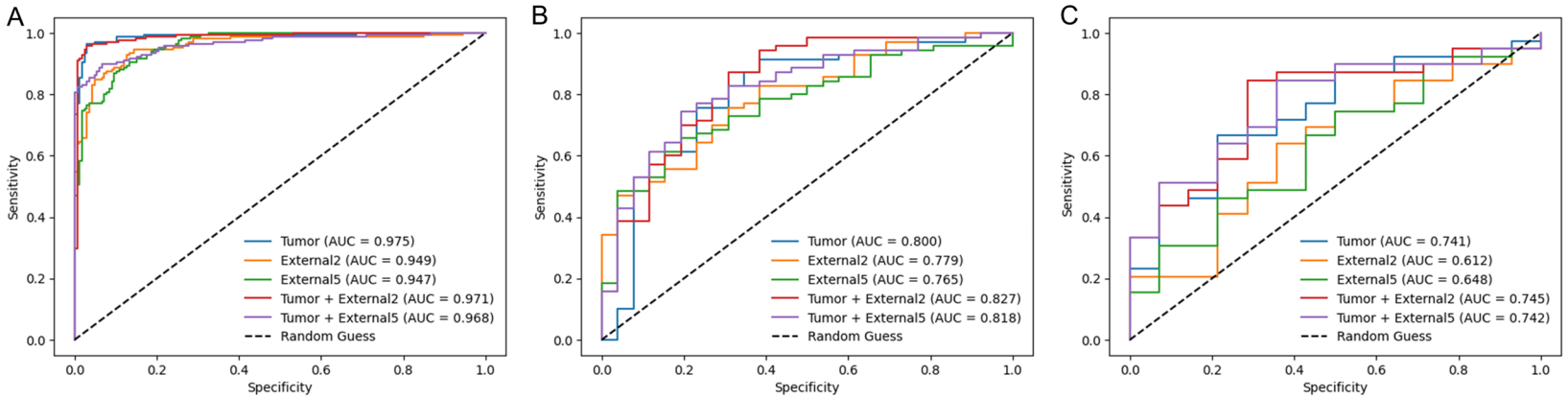

Figure 3. ROC curves of five radiomics signatures in the training (A), the internal-testing, (B) and the External-testing (C) sets.

Table 3. Prediction performance of five radiomics models (Tumor, External2, External5, Tumor + External2, and Tumor + External5) in three cohorts

Model	Cohort	AUC [95% CI]	Accuracy	Sensitivity	Specificity	PPV	NPV
Tumor	Training	0.975 [0.963-0.987]	0.963	0.957	0.969	0.969	0.958
	Internal-testing	0.800 [0.769-0.831]	0.791	0.871	0.576	0.847	0.625
	External-testing	0.741 [0.686-0.796]	0.716	0.743	0.642	0.852	0.473
External2	Training	0.949 [0.937-0.961]	0.897	0.873	0.921	0.917	0.879
	Internal-testing	0.779 [0.750-0.809]	0.739	0.828	0.500	0.816	0.520
	External-testing	0.612 [0.551-0.673]	0.679	0.794	0.357	0.775	0.384
External5	Training	0.947 [0.932-0.961]	0.879	0.909	0.849	0.857	0.903
	Internal-testing	0.765 [0.737-0.794]	0.729	0.800	0.538	0.823	0.500
	External-testing	0.625 [0.562-0.688]	0.679	0.743	0.500	0.805	0.411
Tumor + External2	Training	0.971 [0.962-0.981]	0.954	0.927	0.981	0.980	0.931
	Internal-testing	0.827 [0.799-0.855]	0.822	0.871	0.692	0.884	0.666
	External-testing	0.745 [0.699-0.791]	0.773	0.794	0.714	0.885	0.555
Tumor + External5	Training	0.968 [0.958-0.979]	0.936	0.903	0.969	0.967	0.909
	Internal-testing	0.818 [0.794-0.843]	0.770	0.828	0.615	0.852	0.571
	External-testing	0.742 [0.699-0.785]	0.773	0.820	0.642	0.864	0.562

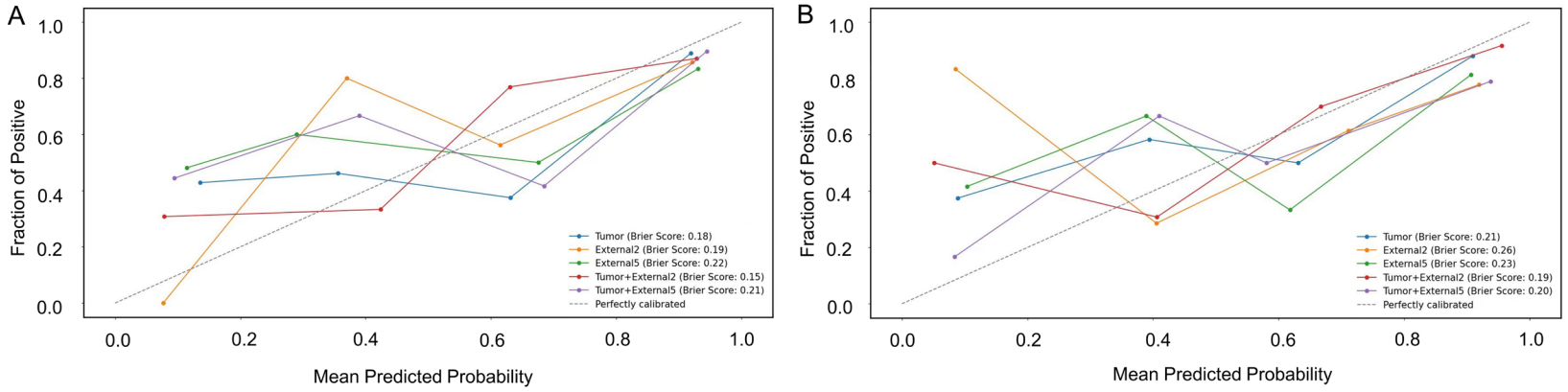

Figure 4. Calibration curves with Brier scores of five radiomics models in internal (A) and external testing (B) sets.

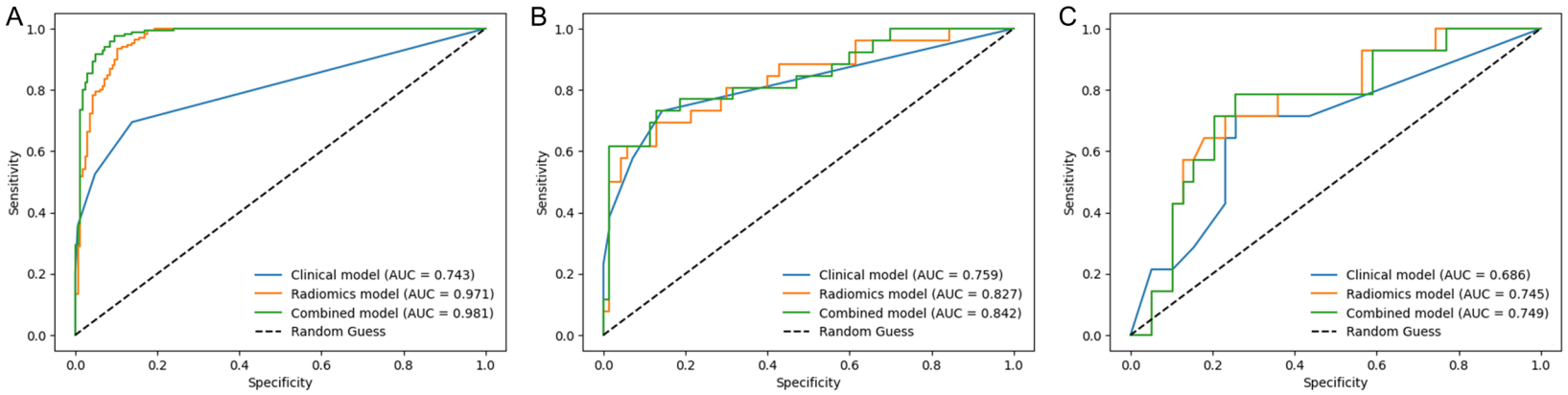

Figure 5. ROC curves for the radiomics, clinical, and combined models in predicting malignancy in the training (A), internal-testing (B), and External-testing (C) sets.

Table 4. Diagnostic performance of the clinical, radiomics, and combined models

Model	AUC [95% CI]	Accuracy	Sensitivity	Specificity
Training cohort				
Clinical model	0.743 (0.681-0.805)	0.840	0.951	0.525
Radiomics model	0.971 (0.962-0.981)	0.954	0.927	0.981
Combined model	0.981 (0.974-0.989)	0.960	0.927	0.981
Internal-testing cohort				
Clinical model	0.759 (0.733-0.743)	0.833	0.928	0.576
Radiomics model	0.827 (0.799-0.855)	0.822	0.871	0.692
Combined model	0.842 (0.817-0.866)	0.822	0.885	0.653
External-testing cohort				
Clinical model	0.686 (0.682-0.709)	0.735	0.743	0.714
Radiomics model	0.745 (0.699-0.791)	0.773	0.794	0.714
Combined model	0.749 (0.705-0.793)	0.773	0.794	0.714

AUC, area under curve; CI, confidence interval.

contain important biological information such as tumor invasion, tumor immune microenvironment, and neovascularization [23, 24]. Furthermore, recent studies have demonstrated that tumor-adjacent tissues may provide additional insights into the heterogeneity of tumors in various cancers. Beig et al. reported that radiomics features from the intranodular and perinodular regions can differentiate benign granulomas from non-small-cell lung cancer adenocarcinomas [16]. Braman et al. reported that the microenvironment surrounding breast cancer is related to its aggressiveness [25]. Chen et al. pointed out that combining radiomic characteristics from both the tumor and its surrounding area can enhance the evaluation of the immune core [26]. In this study, we constructed a model by extracting features from both the peritumoral and intratumoral regions to clarify whether the model can effectively discriminate benign and malignant parotid gland tumors.

Unlike previous studies on parotid tumors that mainly concentrated on intratumoral features and evaluated radiomic signatures alone, we established five radiomic models to compare their performance. The Tumor + External2 model exhibited the highest efficacy in differentiating malignant from benign parotid tumors. The tumor model (AUCs of 0.975, 95% CI, 0.963-0.987) and External2 model (AUCs of 0.949, 95% CI, 0.937-0.961) also showed good performance in the training set. Nevertheless, the performance on both the internal and external testing sets was unsatisfactory, and incorporating the two ROIs effectively enhanced the overall performance. Research has shown through numerous experiments that the radiomic characteristics of tumor surrounding tissues have certain value in clinical work. The features of different regions not only complement each other but also exhibit unique differences. By combining these features, the discrimination efficiency can be greatly improved [27]. Our findings indicate that wavelet features have the highest weight among the remaining features, aligning with the results of other studies [28, 29]. This consistency further proves the multi-scale spatial heterogeneity both around and within the tumor. Consequently, radiomic features could provide more valuable information on the tumor microenvironment as well as tumor biology, which are complementary to visual features.

In addition, we established a clinical-radiomic model that combined the independent predictors (symptom, boundary, and enlarged lymph nodes) and Tumor + External2 radiomic features. The results showed that the discrimination efficiencies of the radiomic models (AUCs of 0.827 and 0.745) in the internal and external testing sets were higher than those of the clinical models (AUCs of 0.759 and 0.686). The combined model showed excellent predictive performance for both the internal and external testing sets (AUCs of 0.842 and 0.749). These findings suggest that while radiomic models have superior predictive performance compared to clinical features, clinical features still play a significant role, which is consistent with previous studies [30]. Therefore, combining these clinical and radiomic features could more accurately discriminate between malignant and benign parotid tumors.

The differences in performance between our study and previous studies can be attributed to several factors. First, the inclusion of peritumoral features captures the biological interac-

tions between the tumor and its microenvironment, which are critical for understanding tumor behavior and predicting clinical outcomes. The peritumoral region provides valuable information about tumor invasion, immune infiltration, and angiogenesis, which are not captured by intratumoral features alone. Second, our study utilized data from two different centers, enhancing the generalizability of the findings. The diversity in patient populations and imaging protocols across centers contributes to the robustness of our radiomic models. Moreover, we employed advanced machine learning techniques, such as the LASSO algorithm for feature selection and support vector machines (SVM) for model construction, which optimized the models and enhanced their predictive performance. Our rigorous validation strategy, including internal and external testing sets, ensures the reliability and robustness of the radiomic models. This comprehensive validation approach is essential for clinical translation. Finally, the development of a combined model that integrates radiomic features with clinical predictors represents a novel approach to improving the accuracy of preoperative diagnosis. This combined model demonstrates improved performance compared to either radiomic or clinical models alone.

However, our study still has certain limitations. First, there may be some bias and interference in this study due to its retrospective nature and use of CT images obtained from various types of CT scanners. Second, although the reliability and reproducibility of radiomic feature extraction were satisfactory between the two observers, the segmentation of intratumoral regions was performed manually. Despite using an automatic technique for peritumoral region segmentation, a fully automatic segmentation method could improve stability and should be considered for future studies. Finally, due to the small sample sizes in this study, the prediction performance of the external testing set may be affected, and future validation with larger sample sizes are necessary.

Conclusion

In conclusion, CT-based radiomic models incorporating both intratumoral and peritumoral radiomic features can effectively distinguish malignant from benign parotid tumors. The predictive accuracy is further improved by combining clinically independent predictors, potentially providing an effective and non-invasive approach for clinical decision-making for patients with parotid tumors.

Acknowledgements

We would like to thank ShanXiong Chen, Ph.D., for constructive criticism of this manuscript. This study was supported by the Key Project of Technological Innovation and Application Development of Chongqing Science and Technology Bureau (No. CSTC2021 jscx-gksb-N0008), the Scientific and Technological Research Program of Chongqing Municipal Education Commission (No. KJQN202301162), and the Scientific Research Foundation of the Affiliated Stomatological Hospital of Southwest Medical University (No. 0802304002).

Disclosure of conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Address correspondence to: Yongmei Li, Department of Radiology, The First Affiliated Hospital of Chongqing Medical University, No. 1 Youyi Road, Yuanjiagang, Yuzhong District, Chongqing 400016, China. E-mail: lymzhang70@163.com; Xiaojuan Liu, School of Artificial Intelligence, Chongqing University of Technology, No. 69 Red Road, Banan District, Chongqing 400016, China. E-mail: liuxiaojuan0127@163.com

References

[1] Gandolfi MM and Slattery W 3rd. Parotid gland tumors and the facial nerve. Otolaryngol Clin North Am 2016; 49: 425-434.

[2] Moore MG, Yueh B, Lin DT, Bradford CR, Smith RV and Khariwala SS. Controversies in the workup and surgical management of parotid neoplasms. Otolaryngol Head Neck Surg 2021; 164: 27-36.

[3] Zbären P, Triantafyllou A, Devaney KO, Poorten VV, Hellquist H, Rinaldo A and Ferlito A. Preoperative diagnostic of parotid gland neoplasms: fine-needle aspiration cytology or core needle biopsy? Eur Arch Otorhinolaryngol 2018; 275: 2609-2613.

[4] Kato H, Kanematsu M, Watanabe H, Kajita K, Mizuta K, Aoki M and Okuaki T. Perfusion imag-

ing of parotid gland tumours: usefulness of arterial spin labeling for differentiating Warthin's tumours. Eur Radiol 2015; 25: 3247-3254.
[5] Vogl TJ, Albrecht MH, Nour-Eldin NA, Ackermann H, Maataoui A, Stöver T, Bickford MW and Stark-Paulsen T. Assessment of salivary gland tumors using MRI and CT: impact of experience on diagnostic accuracy. Radiol Med 2018; 123: 105-116.
[6] Elbuluk AM, Coxe FR, Schimizzi GV, Ranawat AS, Bostrom MP, Sierra RJ and Sculco PK. Abductor deficiency-induced recurrent instability after total hip arthroplasty. JBJS Rev 2020; 8: e0164.
[7] Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, Sanduleanu S, Larue RTHM, Even AJG, Jochems A, van Wijk Y, Woodruff H, van Soest J, Lustberg T, Roelofs E, van Elmpt W, Dekker A, Mottaghy FM, Wildberger JE and Walsh S. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017; 14: 749-762.
[8] Zheng YM, Xu WJ, Hao DP, Liu XJ, Gao CP, Tang GZ, Li J, Wang HX and Dong C. A CT-based radiomics nomogram for differentiation of lympho-associated benign and malignant lesions of the parotid gland. Eur Radiol 2021; 31: 2886-2895.
[9] Al Ajmi E, Forghani B, Reinhold C, Bayat M and Forghani R. Spectral multi-energy CT texture analysis with machine learning for tissue classification: an investigation using classification of benign parotid tumours as a testing paradigm. Eur Radiol 2018; 28: 2604-2611.
[10] Zheng YM, Chen J, Xu Q, Zhao WH, Wang XF, Yuan MG, Liu ZJ, Wu ZJ and Dong C. Development and validation of an MRI-based radiomics nomogram for distinguishing Warthin's tumour from pleomorphic adenomas of the parotid gland. Dentomaxillofac Radiol 2021; 50: 20210023.
[11] He Z, Mao Y, Lu S, Tan L, Xiao J, Tan P, Zhang H, Li G, Yan H, Tan J, Huang D, Qiu Y, Zhang X, Wang X and Liu Y. Machine learning-based radiomics for histological classification of parotid tumors using morphological MRI: a comparative study. Eur Radiol 2022; 32: 8099-8110.
[12] Yu Q, Ning Y, Wang A, Li S, Gu J, Li Q, Chen X, Lv F, Zhang X, Yue Q and Peng J. Deep learning-assisted diagnosis of benign and malignant parotid tumors based on contrast-enhanced CT: a multicenter study. Eur Radiol 2023; 33: 6054-6065.
[13] Alsahafi E, Begg K, Amelio I, Raulf N, Lucarelli P, Sauter T and Tavassoli M. Clinical update on head and neck cancer: molecular biology and ongoing challenges. Cell Death Dis 2019; 10: 540.
[14] Beig N, Khorrami M, Alilou M, Prasanna P, Braman N, Orooji M, Rakshit S, Bera K, Rajiah P, Ginsberg J, Donatelli C, Thawani R, Yang M, Jacono F, Tiwari P, Velcheti V, Gilkeson R, Linden P and Madabhushi A. Perinodular and intranodular radiomic features on lung CT images distinguish adenocarcinomas from granulomas. Radiology 2019; 290: 783-792.
[15] Braman N, Prasanna P, Whitney J, Singh S, Beig N, Etesami M, Bates DDB, Gallagher K, Bloch BN, Vulchi M, Turk P, Bera K, Abraham J, Sikov WM, Somlo G, Harris LN, Gilmore H, Plecha D, Varadan V and Madabhushi A. Association of peritumoral radiomics with tumor biology and pathologic response to preoperative targeted therapy for HER2 (ERBB2)-positive breast cancer. JAMA Netw Open 2019; 2: e192561.
[16] Xia TY, Zhou ZH, Meng XP, Zha JH, Yu Q, Wang WL, Song Y, Wang YC, Tang TY, Xu J, Zhang T, Long XY, Liang Y, Xiao WB and Ju SH. Predicting microvascular invasion in hepatocellular carcinoma using CT-based radiomics model. Radiology 2023; 307: e222729.
[17] Lin CH, Yan JL, Yap WK, Kang CJ, Chang YC, Tsai TY, Chang KP, Liao CT, Hsu CL, Chou WC, Wang HM, Huang PW, Fan KH, Huang BS, Tung-Chieh Chang J, Tu SJ and Lin CY. Prognostic value of interim CT-based peritumoral and intratumoral radiomics in laryngeal and hypopharyngeal cancer patients undergoing definitive radiotherapy. Radiother Oncol 2023; 189: 109938.
[18] Zheng Y, Zhou D, Liu H and Wen M. CT-based radiomics analysis of different machine learning models for differentiating benign and malignant parotid tumors. Eur Radiol 2022; 32: 6953-6964.
[19] Parmar C, Grossmann P, Bussink J, Lambin P and Aerts HJWL. Machine learning methods for quantitative radiomic biomarkers. Sci Rep 2015; 5: 13087.
[20] Wang Y, Zhu GQ, Yang R, Wang C, Qu WF, Chu TH, Tang Z, Yang C, Yang L, Zhou CW, Miao GY, Liu WR, Shi YH and Zeng MS. Deciphering intratumoral heterogeneity of hepatocellular carcinoma with microvascular invasion with radiogenomic analysis. J Transl Med 2023; 21: 734.
[21] Piludu F, Marzi S, Ravanelli M, Pellini R, Covello R, Terrenato I, Farina D, Campora R, Ferrazzoli V and Vidiri A. MRI-based radiomics to differentiate between benign and malignant parotid tumors with external validation. Front Oncol 2021; 11: 656918.
[22] Yu Q, Wang A, Gu J, Li Q, Ning Y, Peng J, Lv F and Zhang X. Multiphasic CT-based radiomics analysis for the differentiation of benign and malignant parotid tumors. Front Oncol 2022; 12: 913898.

[23] Chen Y, Xia Y, Tolat PP, Long L, Jiang Z, Huang Z and Tang Q. Comparison of conventional gadoxetate disodium-enhanced mri features and radiomics signatures with machine learning for diagnosing microvascular invasion. AJR Am J Roentgenol 2021; 216: 1510-1520.

[24] Saijo T, Ishii G, Ochiai A, Hasebe T, Yoshida J, Nishimura M and Nagai K. Evaluation of extratumoral lymphatic permeation in non-small cell lung cancer as a means of predicting outcome. Lung Cancer 2007; 55: 61-66.

[25] Braman NM, Etesami M, Prasanna P, Dubchuk C, Gilmore H, Tiwari P, Plecha D and Madabhushi A. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. Breast Cancer Res 2017; 19: 57.

[26] Chen S, Feng S, Wei J, Liu F, Li B, Li X, Hou Y, Gu D, Tang M, Xiao H, Jia Y, Peng S, Tian J and Kuang M. Pretreatment prediction of immunoscore in hepatocellular cancer: a radiomics-based clinical model based on Gd-EOB-DTPA-enhanced MRI imaging. Eur Radiol 2019; 29: 4177-4187.

[27] Chang R, Qi S, Zuo Y, Yue Y, Zhang X, Guan Y and Qian W. Predicting chemotherapy response in non-small-cell lung cancer via computed tomography radiomic features: Peritumoral, intratumoral, or combined? Front Oncol 2022; 12: 915835.

[28] Ji GW, Zhu FP, Xu Q, Wang K, Wu MY, Tang WW, Li XC and Wang XH. Radiomic features at contrast-enhanced CT predict recurrence in early stage hepatocellular carcinoma: a multi-institutional study. Radiology 2020; 294: 568-579.

[29] Liang W, Yang P, Huang R, Xu L, Wang J, Liu W, Zhang L, Wan D, Huang Q, Lu Y, Kuang Y and Niu T. A combined nomogram model to preoperatively predict histologic grade in pancreatic neuroendocrine tumors. Clin Cancer Res 2019; 25: 584-594.

[30] Xu Y, Shu Z, Song G, Liu Y, Pang P, Wen X and Gong X. The role of preoperative computed tomography radiomics in distinguishing benign and malignant tumors of the parotid gland. Front Oncol 2021; 11: 634452.

Supplementary Appendix 1. CT morphological characteristics

Max-diameter: The sizes of the tumors were measured by determining the maximal cross-sectional diameter.

Number: We observed the lesions of parotid tumor patients on the picture archiving and communication systems (PACS) of our hospital. If there was only one lesion, it was considered single, and if there were two or more lesions, it was multiple.

Symptoms (with or without): We assessed the symptoms based on the clinical record, including pain/tenderness or facial nerve palsy.

Location: superficial or deep lobe, defined by a dashed line delineated from the lateral edge of the mandible to the lateral border of the digastric muscle's posterior belly and retromandibular vein.

Density: The homogeneous or heterogeneous density of the lesion was assessed on the non-contrast CT.

Calcification: Calcification was defined as the CT value of the foci within the tumor is higher than 100Hu.

Cystic areas: cystic area was defined as having a CT scan attenuation of 20 HU or less.

Enhanced-peak phase: We measured CT values (in HU) on non-enhanced, arterial and venous CT scans by placing the largest possible circular region of interest within the solid portion of the lesion with caution to avoid the cystic area. The phase of the highest CT values was defined as enhanced-peak phase.

Enhancement degree: Obvious enhancement was defined as the CT value of tumor enhancement on postcontrast CT is 40Hu higher than it on non-enhanced CT scan. Slight enhancement was defined as the CT value of tumor enhancement on postcontrast CT below 20 Hu on the basis of non-enhanced CT value. Moderate enhanced CT values fell somewhere in between.

Enlarged lymph nodes (with or without): We evaluated ipsilateral lymph node metastases based on imaging features and intraoperative records. The maximal axial dimension criteria for metastatic lymph nodes on imaging were >15 mm for level I and II nodes, 8 mm for retropharyngeal nodes and 10 mm for all other node levels.

Supplementary Appendix 2. Principal Component Analysis (PCA) method

PCA is a statistical and data science technique designed to simplify the complexity inherent in high-dimensional data while retaining essential trends and patterns. The main idea behind PCA is to identify the directions in which the data varies the most. The first principal component accounts for the most variance in the data, the second principal component (uncorrelated with the first) accounts for the second most, and so on.

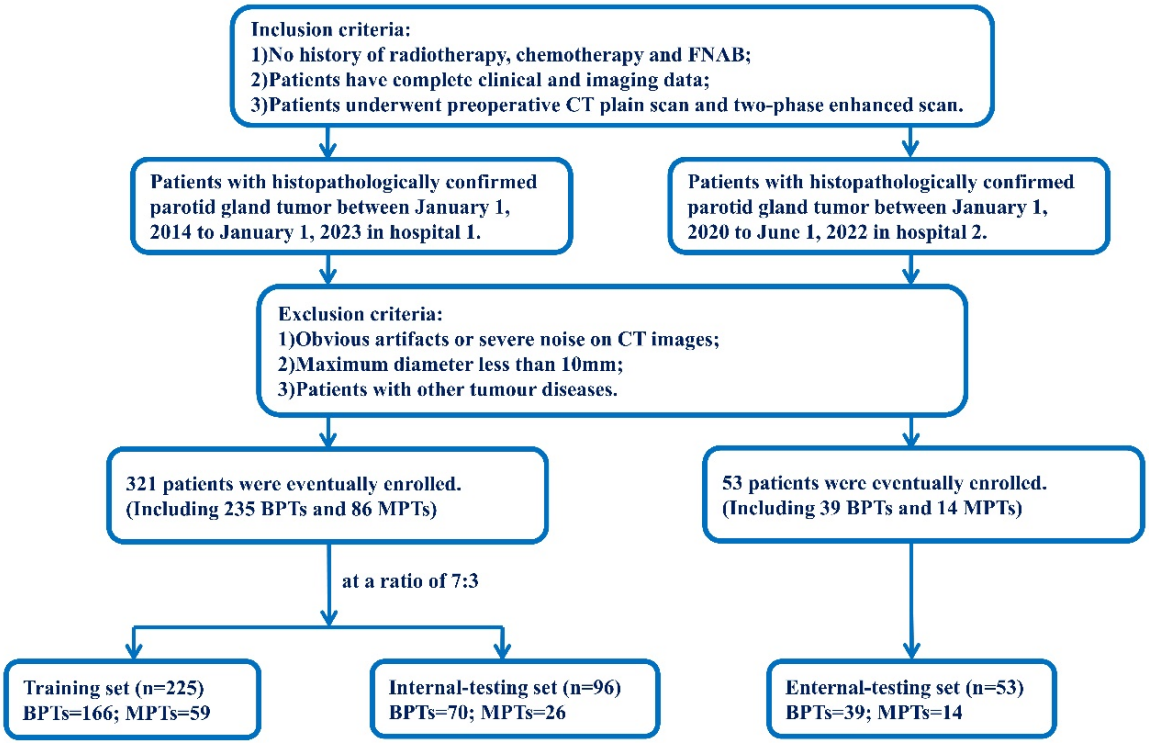

Figure S1. Flowchart for selecting the study population.

Table S1. Histopathological types and numbers of parotid tumors

BPT	Number (Center 1/Center 2)	MPT	Number (Center 1/Center 2)
Pleomorphic adenoma	128 (113/15)	Mucoepidermoid carcinoma	38 (30/8)
Warthin tumor	100 (82/18)	Adenoid cystic carcinoma	6 (4/2)
Basal cell adenoma	33 (28/5)	Acinic cell carcinoma	18 (15/3)
Myoepithelial tumor	3 (3/0)	Squamous cell carcinoma	5 (5/0)
Oncocytoma	4 (3/1)	Lymphoepithelial carcinoma	2 (2/0)
Ductal papillomas	2 (2/0)	Basal cell adenocarcinoma	2(2/0)
Lipoma	4 (4/0)	Myoepithelial carcinoma	7 (7/0)
		Salivary ductal carcinoma	9 (9/0)
		Lymphoma	6 (5/1)
		Secretory carcinoma	1 (1/0)
		Carcinoma in pleomorphic adenoma	1 (1/0)
		Undifferentiated carcinoma	3 (3/0)
		Eosinophilic cell carcinoma	2 (2/0)

Center 1, The First Affiliated Hospital of Chongqing Medical University; Center 2, The Affiliated Hospital of Southwest Medical University.

Table S2. CT protocols of the two centers

Parameters	Center 1			Center 2	
CT scanners	Discovery CT750 HD	SOMATOM Definition Flash	SOMATOM Definition Force	Philips iCT 256	United imaging uCT 550
Tube voltage	100-120 kV	100-120 kV	100 kV	100-120 kV	120 kV
Tube current	Automatic tube-current	Automatic tube-current	Automatic tube-current	Automatic tube-current	Automatic tube-current
Gantry rotation time	0.6 s	0.5 s	0.28 s	0.5 s	0.8 s
Detector collimation	64×0.625 mm	128×0.6 mm	128×0.6 mm	128×0.625 mm	64×0.6 mm
Section thickness	5 mm	5 mm	5 mm	5mm	5 mm
Section interval	5 mm	5 mm	5 mm	5 mm	5 mm
Image matrix	512×512	512×512	512×512	512×512	512×512
Contrast agent type	Omnipaque	Ioversol	Ioversol	Ioversol	Ioversol
Contrast agent concentration	300 mgI/mL	320 mgI/mL	320 mgI/mL	350 mgI/mL	350 mgI/mL
Contrast agent dosage	1.5 mL/kg	1.2 mL/kg	1.2 mL/kg	1.2 mL/kg	1.2 mL/kg
Contrast agent infused rate	3.0-4.0 mL/s	3.0-4.0 mL/s	3.0-4.0 mL/s	2.0-3.0 mL/s	2.0-3.0 mL/s
Arterial phase scan	30 s after the contrast injection	25 s after the contrast injection	25 s after the contrast injection	28 s after the contrast injection	25 s after the contrast injection
Venous phase scan	65 s after the contrast injection	60 s after the contrast injection	60 s after the contrast injection	65 s after the contrast injection	60 s after the contrast injection

Center 1, The First Affiliated Hospital of Chongqing Medical University; Center 2, The Affiliated Hospital of Southwest Medical University. CT: computed tomography.

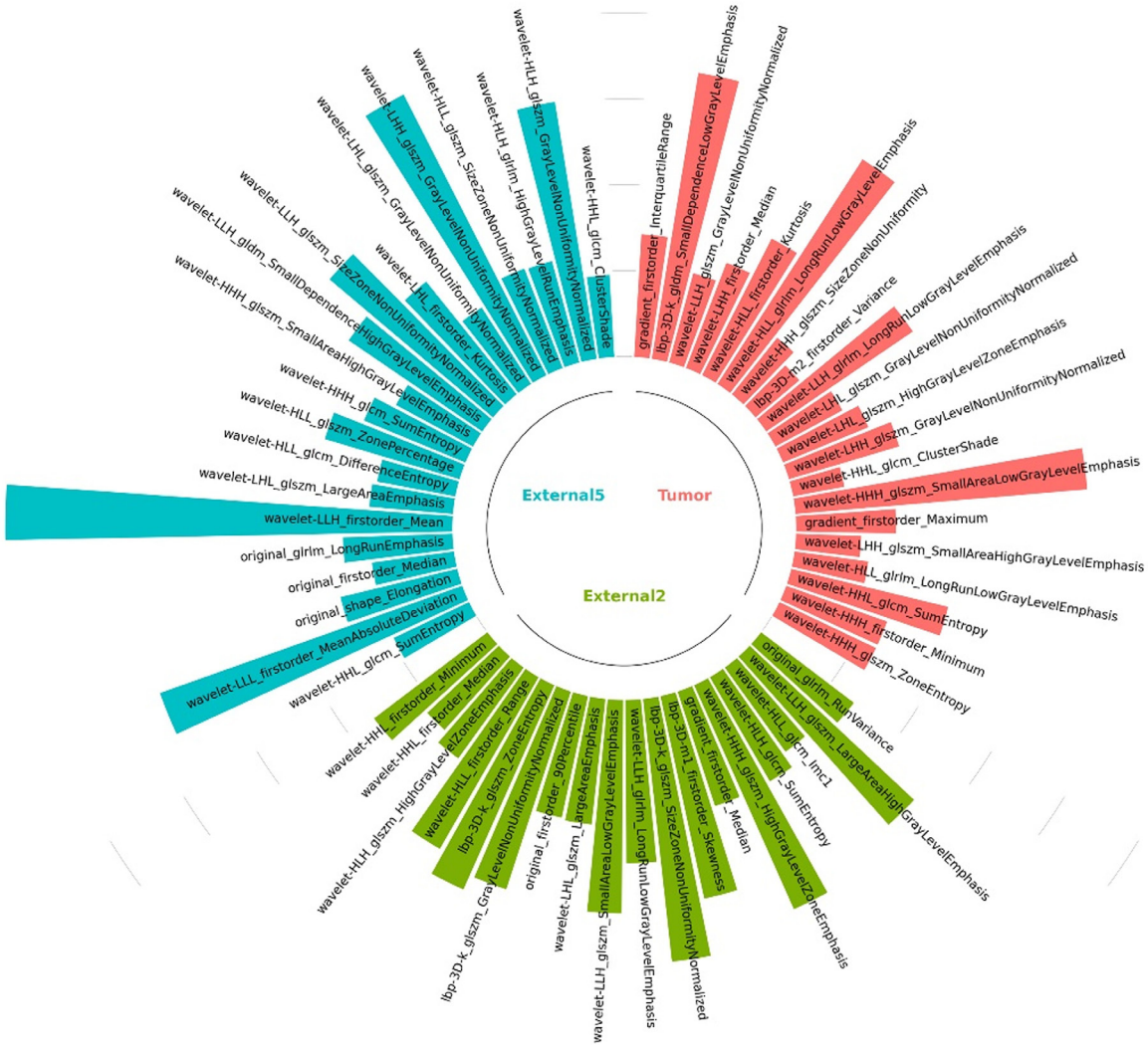

Figure S2. Radiomic feature selection results of Tumor, External2, External5, respectively.

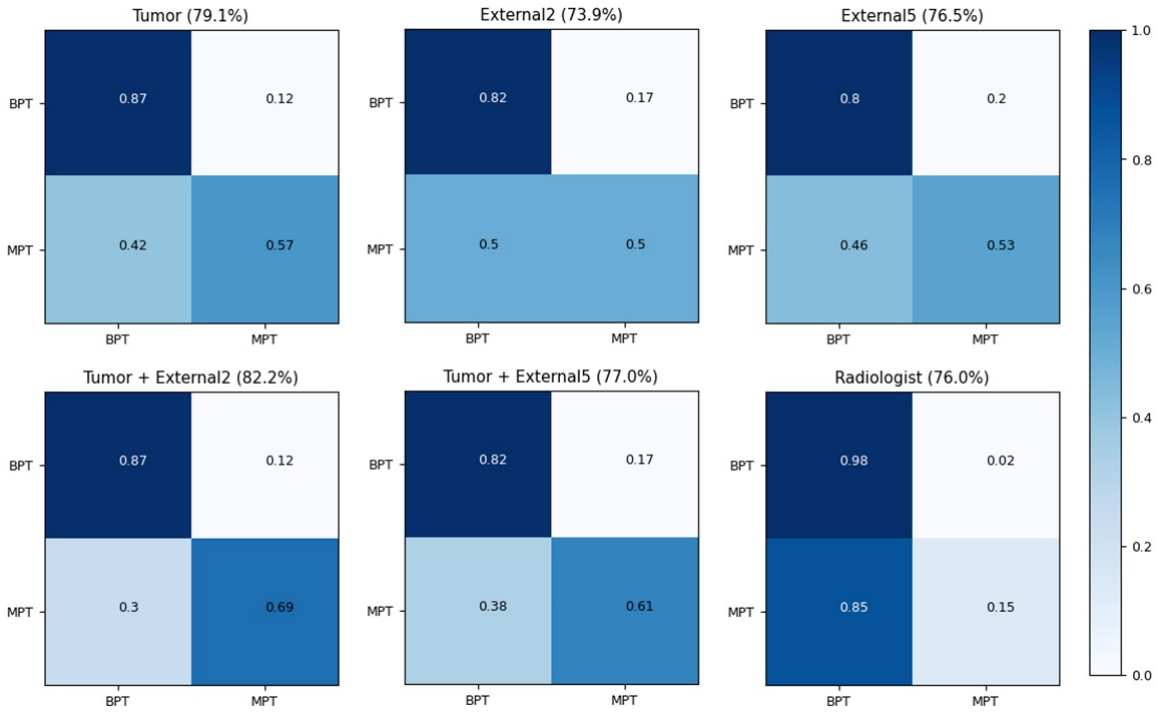

Figure S3. The 2×2 diagnostic confusion matrix analysis for different radiomic models, and the radiologist.

Table S3. The performance of Tumor + External2 radiomics using six machine learning methods

Model	AUC [95% CI]	Accuracy	Sensitivity	Specificity
Training set				
SVM	0.971 (0.962-0.981)	0.954	0.927	0.981
RF	0.965 (0.950-0.980)	0.930	0.921	0.939
LR	0.959 (0.952-0.966)	0.906	0.903	0.909
XGboost	0.988 (0.977-0.996)	1.000	1.000	1.000
DT	0.869 (0.826-0.912)	0.885	0.927	0.843
KNN	0.947 (0.927-0.967)	0.864	0.740	0.987
Internal-testing set				
SVM	0.827 (0.799-0.855)	0.822	0.871	0.692
RF	0.781 (0.747-0.815)	0.739	0.728	0.769
LR	0.801 (0.781-0.822)	0.822	0.857	0.730
XGboost	0.780 (0.748-0.813)	0.760	0.771	0.576
DT	0.728 (0.659-0.797)	0.750	0.771	0.692
KNN	0.761 (0.712-0.810)	0.677	0.642	0.769
External-testing set				
SVM	0.745 (0.701-0.785)	0.773	0.794	0.714
RF	0.721 (0.661-0.782)	0.622	0.666	0.500
LR	0.673 (0.613-0.732)	0.679	0.769	0.428
XGboost	0.684 (0.620-0.747)	0.716	0.794	0.5
DT	0.608 (0.467-0.750)	0.660	0.743	0.428
KNN	0.644 (0.554-0.733)	0.547	0.512	0.642

SVM: Support Vector Machine; RF: Random Forest; LR: Logistic Regression; DT: Decision Tree; KNN: k-Nearest Neighbor.

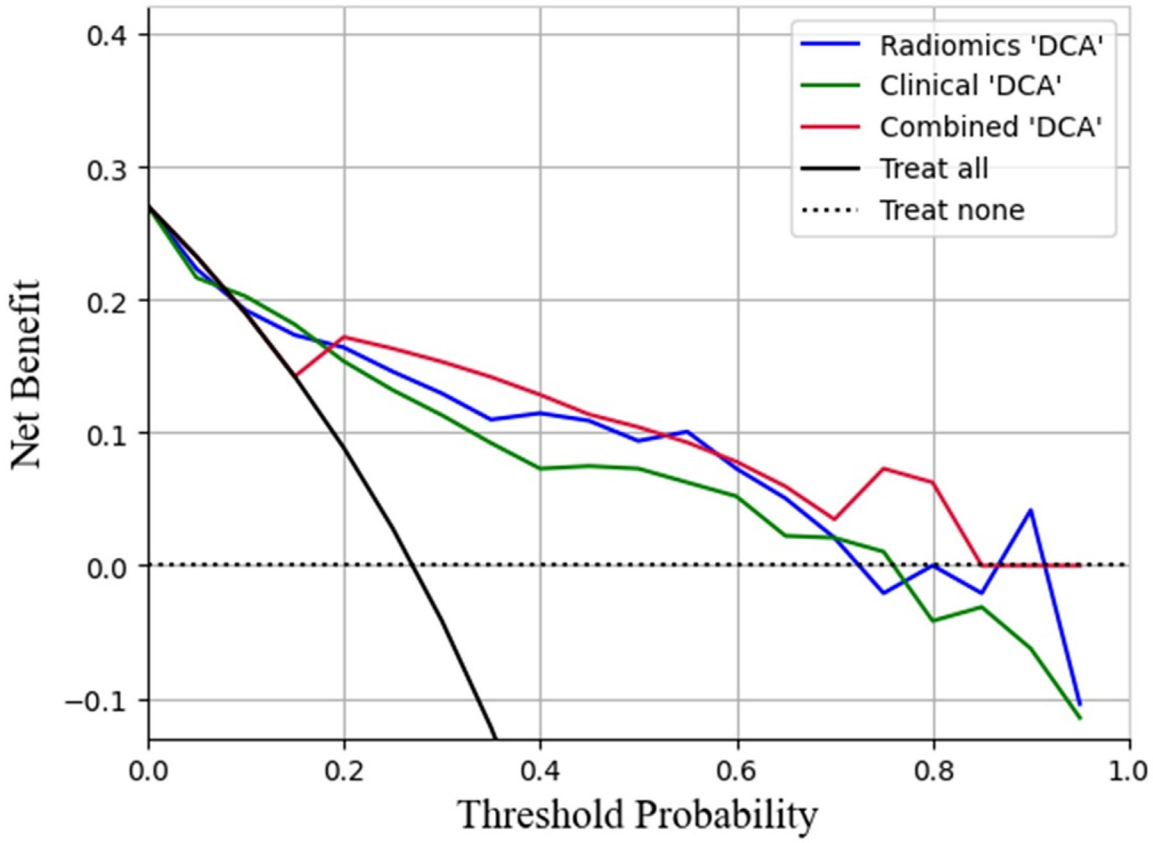

Figure S4. DCA to evaluate the clinical usefulness of the radiomic model, clinical model and combine model in classifying parotid gland tumors.