



Intranodular and perinodular ultrasound radiomics distinguishes benign and malignant thyroid nodules: a multicenter study

Xuelin Zhu^{1,2#}, Jing Li^{3#}, Hao Li¹, Kaifeng Wang⁴, Jian Zhang⁵, Jian Meng⁶, Rong Wu⁷, Meilan Zhang^{3,8}, Hai Du⁸

¹The Faculty of Medicine, Qilu Institute of Technology, Jinan, China; ²Department of Ultrasound, Qingzhou People's Hospital, Qingzhou, China;

³Graduate School, Baotou Medical College, Baotou, China; ⁴The Second Clinical Medical College of Fujian Medical University, Fuzhou, China;

⁵Department of Imaging, The Affiliated Taizhou People's Hospital of Nanjing Medical University, Taizhou, China; ⁶Department of Ultrasound, North China University of Science and Technology Affiliated Hospital, Tangshan, China; ⁷Department of Ultrasound, Ordos Central Hospital, Ordos, China; ⁸Department of Radiology, Ordos Central Hospital, Ordos, China

Contributions: (I) Conception and design: X Zhu, H Du; (II) Administrative support: H Du; (III) Provision of study materials or patients: X Zhu, J Li, K Wang, J Zhang, J Meng, R Wu; (IV) Collection and assembly of data: X Zhu, J Li, H Li, K Wang, J Zhang, J Meng, R Wu, H Du; (V) Data analysis and interpretation: X Zhu, J Li, H Du; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work as co-first authors.

Correspondence to: Hai Du, MD. Department of Radiology, Ordos Central Hospital, No. 23 Ijinhuluo West Street, Dongsheng District, Ordos 017000, China. Email: nthaity@126.com.

Background: Ultrasound based radiomics prediction model can improve the differentiation ability of benign and malignant thyroid nodules to avoid overtreatment. This study evaluates the role of predictive models based on intranodular and perinodular ultrasound radiomics in distinguishing between benign and malignant thyroid nodules.

Methods: A total of 1,076 thyroid nodules were enrolled from three hospitals between 2016 and 2022, forming the training, validation and test cohorts. The clinical signature (Clinic_Sig) was developed based on clinical information and conventional morphological features of ultrasound. Expanding 1 pixel, 3 pixels, 5 pixels, 7 pixels, and 9 pixels outward from the thyroid nodule, six radiomics models were constructed using intranodular (intra) and combined radiomics (intranodular and perinodular: +p1,+p3,+p5,+p7,+p9) features. The model with the best area under the curve (AUC) was defined as radiomics signature (Rad_Sig). The combined model was constructed from Clinic_Sig and Rad_Sig. AUC and calibration curves were used to evaluate the predictive performance of the model. Decision curve analysis (DCA) was used to evaluate the clinical net benefit of the model.

Results: The intra+p1 radiomics model exhibited the highest efficacy (AUC =0.863) in the test cohort, which was combined with Clinic_Sig to construct the combined model. Compared with Clinic_Sig and Rad_Sig, the combined model showed the higher predictive performance, with AUCs of 0.942 (training), 0.894 (validation), and 0.933 (test). The calibration curve showed that the predicted probabilities of the combined model were in good agreement with the actual probabilities, and DCA indicated that it provided more net benefit than the treat-none or treat-all scheme.

Conclusions: The combined model based on clinical signatures, intranodular and perinodular ultrasound radiomics has the potential to effectively predict benign or malignant thyroid nodules.

Keywords: Intranodular; perinodular; ultrasound; radiomics; thyroid nodules

Submitted Sep 25, 2024. Accepted for publication Dec 10, 2024. Published online Dec 27, 2024.

doi: 10.21037/gs-24-416

View this article at: <https://dx.doi.org/10.21037/gs-24-416>

Introduction

The rate of thyroid malignancy is rising, suggesting that the discovery of incidental thyroid nodules will continue to be a relevant and increasingly salient topic within modern medicine (1). In the general population, detection of thyroid nodules occurs in up to 65% of cases (2). Nevertheless, most of these nodules are non-cancerous, with only around 7–15% proving to be malignant (3). The accurate differentiation between benign and malignant thyroid nodules is pivotal for clinical decision-making and management, in order to prevent unnecessary interventions, such as lifelong medication following thyroidectomy.

Preoperative ultrasound examination, as a crucial step in accurate disease assessment and surgical planning, emphasizes its undeniable significance (4). Thyroid ultrasound examination has become an indispensable component of thyroid nodule screening and diagnosis due to its convenience, non-invasiveness, and good cost-effectiveness. However, traditional ultrasound evaluation of benign and malignant thyroid nodules mainly relies on thyroid imaging reports and data systems (TI-RADS), which often involve a significant amount of subjective

judgment from doctors (5). To overcome this limitation, the research conducted by Wildman-Tobriner *et al.* delved into the potential of artificial intelligence (AI) in optimizing the American College of Radiology (ACR) TI-RADS. They emphasized that the integration of AI technology can improve specificity (6). Radiomics can deeply extract and quantify subtle features from traditional images that are imperceptible to the naked eye, providing a more objective and quantitative basis for clinical decision-making (7). This innovation not only reduces the influence of human factors, but also pushes the diagnosis of thyroid nodules towards a more accurate and effective track. Radiomics analysis is a quantitative and objective computer-based image analysis technique, which is widely used in the diagnosis, grading, staging, and prognosis prediction of organ diseases such as thyroid, breast, chest and lungs, liver, kidney, and gynecology (8). Several studies have established that incorporating radiomics modality can further improve the basic diagnostic performance when combined with clinical and ultrasound information. The prediction model established by Yoon *et al.* (9) using multivariate logistic regression analysis showed that the area under receiver operating characteristic (ROC) curve (AUC) of malignant thyroid nodules predicted by the model combining radiomics with clinical variables was significantly higher than that of the model consisting only of clinical variables (0.839 *vs.* 0.583). Liang *et al.* (10) compared the radiomics score with four TI-RADS scores, and found that the radiomics score model added more benefits than using any TI-RADS score model.

Some studies also indicated that the peritumoral region may provide supplementary information for tumor heterogeneity, so predictive models should not be limited to the intratumoral region alone (11,12). Hu *et al.* evaluated the pathological complete response of esophageal squamous cell carcinoma patients to neoadjuvant chemotherapy using intratumoral and peritumoral computed tomography (CT) radiomics, and concluded that the model combining intratumoral and peritumoral radiomics had the highest predictive performance (13). Thyroid malignant nodules are tumors with good prognosis, and there are currently no relevant research reports on whether the microenvironment around the thyroid nodules can provide valuable information for the differential diagnosis of benign and malignant. Hence, the aim of this study is to compare the predictive value of intranodular combining different pixel levels of peritumoral radiomics features for thyroid malignant nodules, and to explore the prediction

Highlight box

Key findings

- This study explores for the first time the role of ultrasound radiological features around nodules in distinguishing benign and malignant thyroid nodules, providing new directions for future research. The combined use of intranodular and perinodular ultrasound radiomics features has a higher predictive power for benign and malignant thyroid nodules than individual intranodular ultrasound features.

What is known and what is new?

- Previous research has shown that the combined clinical-radiomics model significantly outperforms either a standalone clinical model or a radiomics model in terms of predictive performance, offering clinicians a more solid and reliable foundation for their decision-making processes.
- Our study introduces the radiomics features of perinodular as an additional assessment dimension. Our aim is to achieve a more accurate portrayal of the biological characteristics and pathological status of nodules by integrating both the intranodular and perinodular comprehensive features.

What is the implication, and what should change now?

- This approach provides a potent technical backbone for the early detection of thyroid nodules, precise classification, and the tailored formulation of personalized treatment plans in clinical practice.

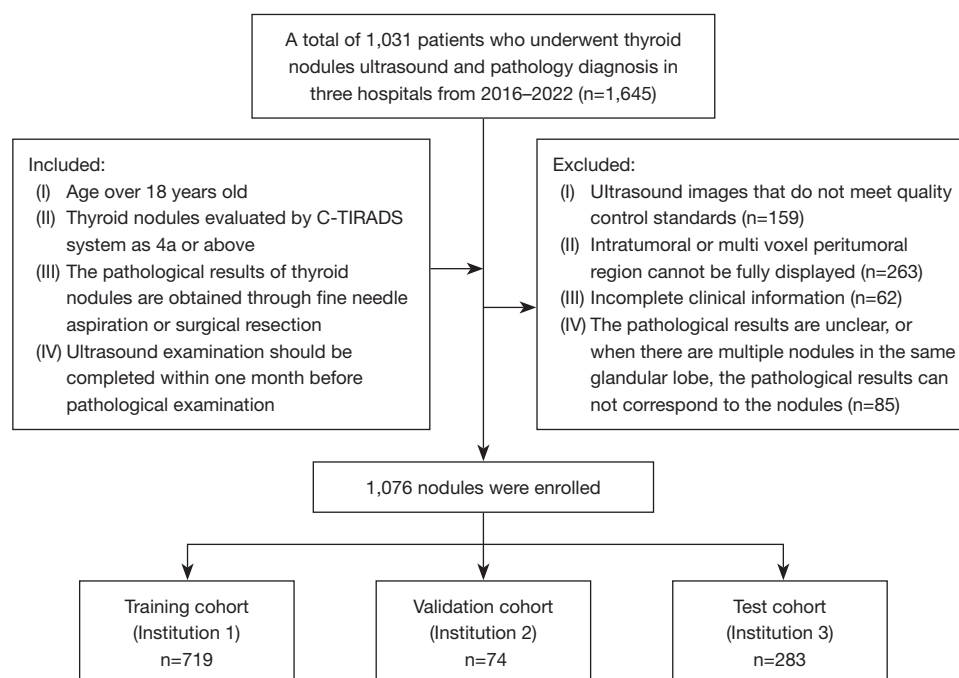


Figure 1 Flowchart revealing the study population based on exclusion criteria. C-TIRADS, Chinese-thyroid imaging, reporting, and data system; Institution 1, Erdos Central Hospital; Institution 2, Taizhou People's Hospital; Institution 3, North China University of Science and Technology Affiliated Hospital.

efficacy of combined models based on clinical features, intranodular and perinodular ultrasound radiomics features of thyroid nodules. We present this article in accordance with the TRIPOD reporting checklist (available at <https://gs.amegroups.com/article/view/10.21037/gs-24-416/rc>).

Methods

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the ethics committee of Ordos Central Hospital (No. 2024-100), and the same approval was recognized by the two other subcenters. Informed consent was waived for this retrospective study, and no personal information was disclosed.

Patients

The following inclusion criteria were applied: (I) age over 18 years old; (II) thyroid nodules evaluated by Chinese TI-RADS (C-TIRADS) system as 4A or above; (III) the pathological results of thyroid nodules are obtained through fine needle aspiration or surgical resection; (IV) the

interval between ultrasound and pathological examination results shall not exceed one month. Exclusion criteria encompassed: (I) the quality of ultrasound image does not meet the quality control standards; (II) intranodular or multi pixels perinodular region cannot be fully displayed; (III) the pathological results are unclear (according to the Bethesda System for Reporting Thyroid Cytopathology, under which the nodules were classified as Bethesda I, III, or IV, but there were no final pathological results), or when there are multiple nodules in the same glandular lobe, the pathological results can not correspond to the nodules; (IV) the patient's clinical information is incomplete. A flowchart showing the study population was presented as *Figure 1*. A total of 1,076 thyroid nodules from 817 patients at Erdos Central Hospital, Taizhou People's Hospital, and North China University of Science and Technology Affiliated Hospital between 2016 and 2022 were eventually included.

Gray-scale ultrasound image acquisition and clinical model construction

The parameters and inspection methods of the ultrasound instrument are shown in [Appendix 1](#). Clinical information

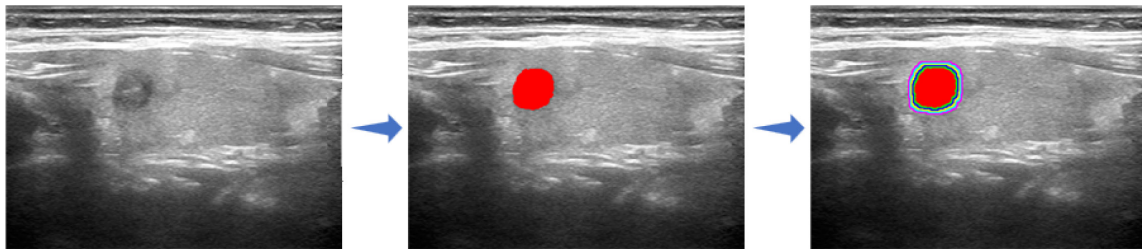


Figure 2 Examples of expanding mask at different distances from ultrasound image's original mask. The red area signifies the segmented nodal region, while the colored ring depicts the perinodal region.

of patients was obtained through the Hospital Information System, recording their gender and age. Two ultrasound physicians with ten years of experience in thyroid diagnosis evaluated the transverse and longitudinal dimension, direction of growth, composition, echogenicity, echogenic foci, and margin of the nodules. The disputed nodules were ultimately determined by an ultrasound physician with 23 years of diagnostic ultrasound experience. According to the C-TIRADS (14), solid composition, microcalcifications, markedly hypoechoic, ill-defined or irregular margins, and taller-than-wide were defined as malignant ultrasound features. Each malignant feature was scored 1 point, with 1 point classified as 4A, 2 points classified as 4B, 3–4 points classified as 4C, and 5 points classified as 5. Univariable logistic regression analyses were performed to select risk factors with $P < 0.05$. The selected factors were imported into backward stepwise multivariate regression for clinical signature (Clinic_Sig).

Nodule segmentation and data pre-processing

Two ultrasound physicians with a decade of thyroid diagnosis experience utilized open-source software ITK-SNAP (a tool for segmenting anatomical structures in medical images) for nodule segmentation (15). Inter- and intra-observer reproducibility of ultrasound radiomics features extraction were initially analyzed with the data of 30 randomly selected patients from each sequence in a double-blinded fashion by these two physicians.

Firstly, we resample all images into an isotropic dataset with a common voxel spacing of $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$. Secondly, to mitigate the influence of variations stemming from different ultrasound instruments and image batches, normalization preprocessing is applied to each group of radiomic features. Specifically, the Z-score normalization method is utilized to transform the pixel

values of the original images. This process converts data with significant differences in magnitude into standardized Z-scores, thereby ensuring the comparability of the data across different sources. For evaluating peritumoral region predictability, the mask padding toolkit on the OnekeyAI platform was utilized (16). This toolkit, implemented using the “SimpleITK” package in Python version 3.7, initiated the expansion process from 1 pixel outside the nodule and gradually incremented by 2 pixels along various radial directions (up to a maximum expansion of 9 pixels). This yielded five new masks representing diverse perinodular regions for each patient. Subsequently, nodules and the progressively expanding perinodular regions were employed for further analysis. *Figure 2* illustrated the expansion process of the nodule region.

Radiomics feature extraction and selection

All radiomics features were extracted from the region of interest (ROI) using the Pyradiomics package on the Python platform. These features can be grouped into three categories: (I) geometry; (II) intensity; and (III) texture. Geometric features describe the 3D shape attributes of the tumor. Intensity features portray the statistical distribution of pixel intensities within the tumor, focusing on the first order. Texture features capture the spatial distribution of patterns or intensities, considering both second and higher orders. Texture features were computed utilizing matrices including gray-level co-occurrence matrix (GLCM), gray-level run length matrix (GLRLM), gray-level size zone matrix (GLSZM), and neighborhood gray-level difference matrix (NGTDM). Overall, 1,648 features were extracted, encompassing 342 first-order features, 14 shape features, and 1,290 texture features.

Features with intraclass correlation coefficient (ICC) values greater than or equal to 0.85 were deemed robust,

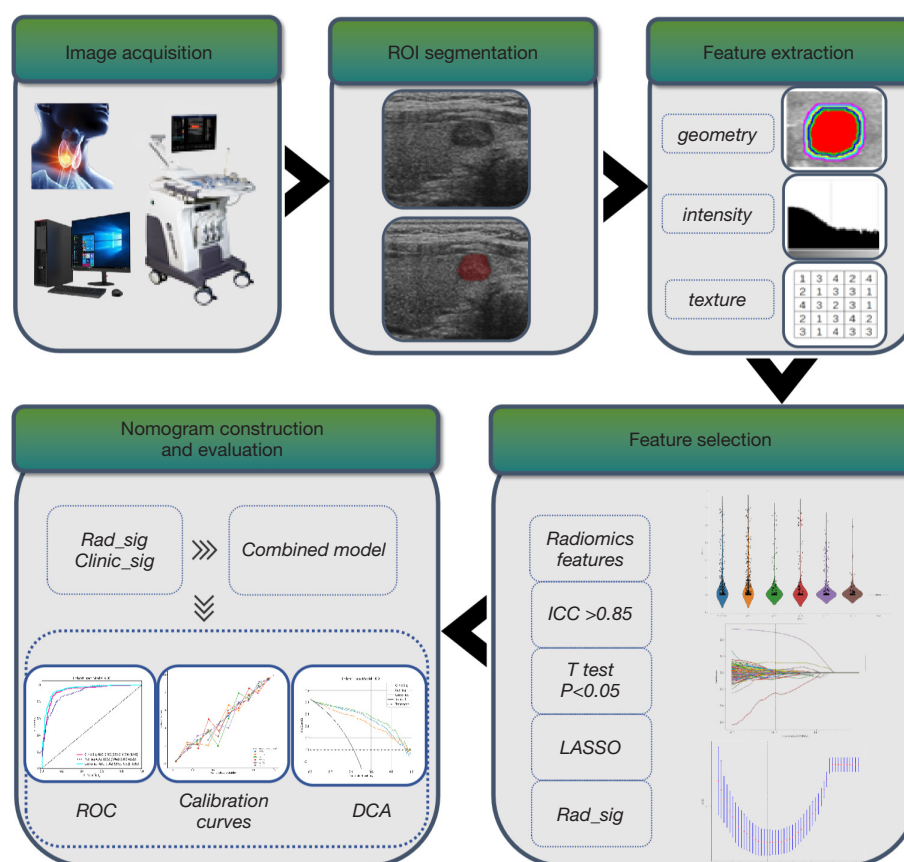


Figure 3 Flowchart of extraction of radiomics features, model establishment and performance evaluation. ROI, region of interest; ICC, intraclass correlation coefficient; LASSO, least absolute shrinkage and selection operator; Rad_Sig, radiomics signature; Clinic_Sig, clinical signature; ROC, receiver operating characteristic; DCA, decision curve analysis.

demonstrating resilience to inter- and intra-rater variability. Features with interobserver ICC values below 0.85 were excluded based on the 30-patient analysis results. Subsequently, a *t*-test was employed for feature screening, retaining radiomic features with *P* values below 0.05. For features exhibiting high repeatability, redundancy was reduced using the Spearman rank correlation coefficient. If the correlation coefficient between two features exceeded 0.9, one feature was retained, and the other was discarded. Finally, the least absolute shrinkage and selection operator (LASSO) with 10-fold cross-validation was employed to select the most predictive features. Features with non-zero coefficients were retained and used in machine learning models such as logistic regression (LR), support vector machine (SVM), random forest (RF), and XGBoost for model construction. The radiomics model with the best predictive performance was selected and defined as radiomics signature (Rad_Sig).

Construction and evaluation of the combined model

The combined model was established based on Clinic_Sig and Rad_Sig. Accuracy, specificity, sensitivity, positive predictive value (PPV), negative predictive value (NPV) and AUC were used to estimate the predictive performance of Clinic_Sig, Rad_Sig, combined model. The calibration curve was used to evaluate the consistency between the estimated probability and actual probability of the thyroid malignant nodules. Decision curve analysis (DCA) was used to assess the clinical usefulness by estimating the net benefit within threshold probabilities. A flowchart of extraction of radiomics features and model establishment was shown as Figure 3.

Statistical analysis

The independent sample *t*-test and the χ^2 test were

conducted to compare patients' clinical characteristics. The χ^2 test was used for discrete variables, and the *t*-test for continuous variables. All data analyses were performed using Python 3.7.12 on the OnekeyAI platform v3.1.8. Statistical analyses employed statsmodels v0.13.2, pyradiomics v3.0.1 for radiomics feature extraction, and scikit-learn 1.0.2 for machine learning algorithms like SVM.

Results

Clinical characteristics

A total of 1,076 thyroid nodules were evaluated as follows: C-TIRADS 4A: 386, C-TIRADS 4B: 314, C-TIRADS 4C: 273, C-TIRADS 5: 103. Among them, 594 were benign nodules and 482 were malignant nodules. In the training cohort, there were a total of 308 malignant nodules and 411 benign nodules. In the validation cohort, there were 66 malignant nodules and 8 benign nodules, while in the test cohort, there were 108 malignant nodules and 175 benign nodules. The pathological classification of thyroid nodules is shown in [Table S1](#). *Table 1* presented the clinical characteristics of the entire dataset. In the training cohort, significant distinctions were observed between benign and malignant nodules regarding age, transverse diameter, longitudinal diameter, direction of growth, echogenic foci, margin, echogenicity, and composition ($P < 0.001$). According to multivariate analysis, age, orientation, hyperechoic lesion, margin, echo, and composition were independent risk factors for predicting thyroid malignant nodules ($P < 0.05$), and these factors were included in Clinic_Sig.

Radiomics features selection and radiomics model construction

We have tried a total of 10 classifiers, in order to further optimize hyperparameters, we used LR, which is the most common machine learning model that can be used to compare different features as an aggregation model. *Table 2* and [Figure S1](#) displayed the predictive performance of intranodular, intranodular combined with perinodular radiomics models. The intra+p1 radiomics model exhibited AUC values of 0.902 (95% CI: 0.880–0.923) in the training cohort and 0.845 (95% CI: 0.720–0.969) in the validation cohort. Notably, in the test cohort, the model achieved an AUC of 0.863 (95% CI: 0.820–0.907), which was superior to all other perinodular models. Consequently,

the intra+p1 radiomics model was identified as having the best predictive performance among the peri-tumoral models, leading to its selection for further application. We conducted certain filtering and removal of features based on LASSO dimensionality reduction for avoiding overfitting caused by excessive features. The detailed removal process is shown in [Figure 4](#). Then, 28 intranodular features and 32 surrounding area features were extracted to form Rad_Sig, detailed features were shown in [Figure S2](#).

Construction and evaluation of the combined model

Table 3 compares the predictive abilities of Clinic_Sig, Rad_Sig, and the combined model. In the training, validation, and test cohorts, compared with other models, the AUCs of the combined model were 0.942 (95% CI: 0.923–0.959), 0.894 (95% CI: 0.776–1.000), and 0.933 (95% CI: 0.903–0.962), respectively, indicating superior performance in predicting benign and malignant thyroid nodules ([Figure 5](#)). The AUCs of the Clinic_Sig were 0.926 (95% CI: 0.906–0.946), 0.854 (95% CI: 0.695–1.000), 0.922 (95% CI: 0.889–0.954). While the AUCs of the Rad_Sig model were 0.902 (95% CI: 0.880–0.923), 0.845 (95% CI: 0.720–0.969), 0.863 (95% CI: 0.820–0.907). Calibration curves demonstrated substantial agreement between actual results and estimated probabilities for both feature models and the combined model, DCA curves revealed the potential clinical utility of each model, emphasizing the greater future applicability of the combined model (shown in [Figure 6](#)).

Discussion

Autoimmune thyroid disease (AITD) is the most common organic specific illness of the thyroid gland (17). In the pathological classification of thyroid nodules, the proportion of chronic lymphocytic thyroiditis (Hashimoto's thyroiditis) is relatively low, which might be attributed to our utilization of the C-TIRADS reporting system (14). Within this system, Hashimoto's thyroiditis is explicitly categorized as Class 1, indicating a low risk of malignancy. Given that our research focuses on evaluating nodules classified as 4A or higher, which possess a relatively higher potential for malignancy, and considering that in our ultrasound radiomics analysis of thyroid nodules and their surroundings, diffuse lesions such as Hashimoto's thyroiditis may not align with the segmentation of regions of interest, the inclusion criteria set for the study could directly impact the final research outcomes by excluding low-risk categories

Table 1 Clinical information and ultrasound features of the cohorts

Variables	Training cohort (n=719)			Validation cohort (n=74)			Test cohort (n=283)		
	Benign nodule (n=411)	Malignant nodule (n=308)	P	Benign nodule (n=8)	Malignant nodule (n=66)	P	Benign nodule (n=175)	Malignant nodule (n=108)	P
Age (years)	55.06±10.55	46.08±10.75	<0.001	53.50±12.25	53.09±12.09	0.93	57.41±9.90	51.20±9.86	<0.001
Gender			0.09			0.81			0.39
Female	315 (76.64)	253 (82.14)		6 (75.00)	42 (63.64)		120 (68.57)	80 (74.07)	
Male	96 (23.36)	55 (17.86)		2 (25.00)	24 (36.36)		55 (31.43)	28 (25.93)	
Transverse dimension	27.90±17.54	10.64±7.47	<0.001	24.34±9.74	9.74±9.72	<0.001	21.56±14.91	9.93±6.06	<0.001
Longitudinal dimension	21.84±14.42	9.39±5.52	<0.001	14.35±6.17	8.31±6.11	<0.001	12.55±8.68	7.62±3.97	<0.001
Direction of growth			<0.001			<0.001			<0.001
Wider-than-tall	392 (95.38)	195 (63.31)		8 (100.00)	36 (54.55)		172 (98.29)	89 (82.41)	
Taller-than-wide	19 (4.62)	113 (36.69)		0	30 (45.45)		3 (1.71)	19 (17.59)	
Composition			<0.001			0.03			<0.001
Mixed predominantly solid	199 (48.42)	15 (4.87)		3 (37.50)	2 (3.03)		46 (26.29)	0	
Mixed predominantly cystic	36 (8.76)	1 (0.32)		0	0		23 (13.14)	0	
Solid	176 (42.82)	292 (94.81)		5 (62.50)	64 (96.97)		106 (60.57)	108 (100.0)	
Echogenicity			<0.001			<0.001			<0.001
Mildly hypoechoic	81 (19.71)	256 (83.12)		3 (37.50)	58 (87.88)		77 (44.00)	104 (96.30)	
Isoechoic	36 (8.76)	1 (0.32)		0	0		25 (14.29)	0	
Anechoic	2 (0.49)	11 (3.57)		0	0		0	1 (0.93)	
Markedly hypoechoic	286 (69.59)	36 (11.69)		5 (62.50)	6 (9.09)		65 (37.14)	2 (1.85)	
Hyperechoic	6 (1.46)	4 (1.30)		0	2 (3.03)		8 (4.57)	1 (0.93)	
Echogenic foci			<0.001			0.31			<0.001
None	2 (0.49)	0		0	0		1 (0.57)	1 (0.93)	
Microcalcifications	357 (86.86)	160 (51.95)		6 (75.00)	33 (50.00)		150 (85.71)	49 (45.37)	
Macrocalcifications	24 (5.84)	117 (37.99)		2 (25.00)	22 (33.33)		12 (6.86)	43 (39.81)	
Peripheral calcifications	28 (6.81)	31 (10.06)		0	11 (16.67)		12 (6.86)	15 (13.89)	
Margin			<0.001			>0.99			<0.001
Smooth	388 (94.40)	180 (58.44)		0	0		159 (90.86)	40 (37.04)	
Irregular	1 (0.24)	1 (0.32)		0	0		0	1 (0.93)	
Ill-defined	22 (5.35)	127 (41.23)		8 (100.00)	66 (100.00)		16 (9.14)	67 (62.04)	

Data are presented as mean ± standard deviation or n (%).

Table 2 Predictive performance of thyroid nodules for benign and malignant based on ultrasound radiomics

Cohort	Model	AUC	95% CI	Accuracy	Sensitivity	Specificity	PPV	NPV
Training	intra	0.908	0.887–0.929	0.830	0.814	0.842	0.793	0.859
	in+p1	0.902	0.880–0.923	0.830	0.814	0.842	0.793	0.859
	in+p3	0.923	0.905–0.942	0.852	0.833	0.866	0.823	0.875
	in+p5	0.919	0.899–0.938	0.840	0.820	0.854	0.807	0.865
	in+p7	0.892	0.869–0.915	0.799	0.784	0.810	0.755	0.835
	in+p9	0.897	0.875–0.919	0.824	0.807	0.837	0.787	0.854
Validation	intra	0.871	0.770–0.973	0.541	0.485	1.000	1.000	0.190
	in+p1	0.845	0.720–0.969	0.554	0.500	1.000	1.000	0.195
	in+p3	0.775	0.543–1.000	0.527	0.500	0.750	0.943	0.154
	in+p5	0.858	0.743–0.973	0.514	0.455	1.000	1.000	0.455
	in+p7	0.905	0.827–0.983	0.500	0.439	1.000	1.000	0.178
	in+p9	0.854	0.744–0.964	0.500	0.439	1.000	1.000	0.178
Test	intra	0.849	0.803–0.894	0.766	0.766	0.766	0.667	0.843
	in+p1	0.863	0.820–0.907	0.794	0.776	0.806	0.709	0.855
	in+p3	0.837	0.790–0.884	0.794	0.785	0.800	0.706	0.859
	in+p5	0.785	0.729–0.840	0.730	0.692	0.754	0.632	0.692
	in+p7	0.846	0.799–0.893	0.770	0.757	0.777	0.675	0.840
	in+p9	0.837	0.790–0.883	0.770	0.748	0.783	0.678	0.835

AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; intra, intranodular; in+p1, intranodular and perinodular with 1 pixel; in+p3, intranodular and perinodular with 3 pixels; in+p5, intranodular and perinodular with 5 pixels; in+p7, intranodular and perinodular with 7 pixels; in+p9, intranodular and perinodular with 9 pixels.

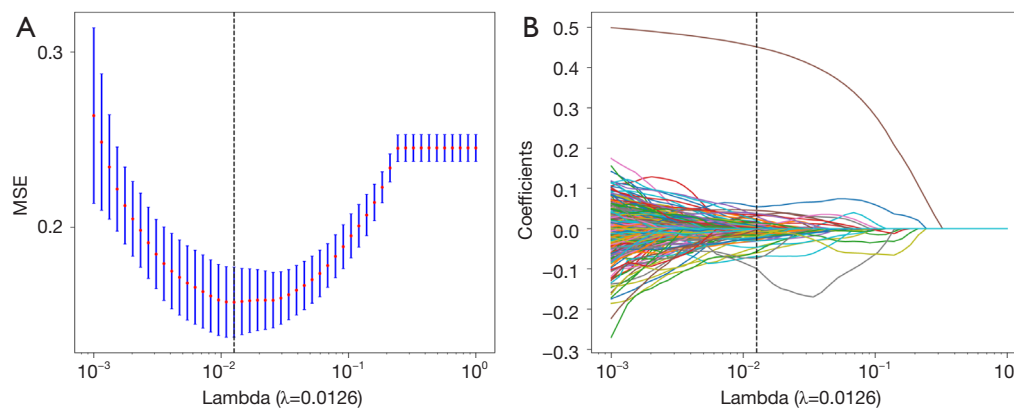


Figure 4 Selection of radiomics features via LASSO algorithm to establish Rad_Sig. (A) Tuning parameter selection by 10-fold cross-validation with minimum criteria. MSE (y-axis) was plotted as a function of Log(lambda) (x-axis). (B) LASSO coefficient profiles for the whole features. MSE, mean square error; LASSO, least absolute shrinkage and selection operator; Rad_Sig, radiomics signature.

Table 3 Comparison of all models in the three cohorts

Cohort	Model	AUC	95% CI	Accuracy	Sensitivity	Specificity	PPV	NPV
Training	Clinic_Sig	0.926	0.906–0.946	0.879	0.905	0.859	0.827	0.924
	Rad_Sig	0.902	0.880–0.923	0.830	0.859	0.808	0.769	0.885
	Combined	0.942	0.923–0.959	0.891	0.912	0.876	0.845	0.930
Validation	Clinic_Sig	0.854	0.695–1.000	0.865	0.879	0.750	0.967	0.429
	Rad_Sig	0.845	0.720–0.969	0.743	0.727	0.875	0.980	0.280
	Combined	0.894	0.776–1.000	0.878	0.879	0.875	0.983	0.467
Test	Clinic_Sig	0.922	0.889–0.954	0.851	0.907	0.817	0.752	0.935
	Rad_Sig	0.863	0.820–0.907	0.823	0.692	0.903	0.813	0.827
	Combined	0.933	0.903–0.962	0.869	0.869	0.869	0.802	0.916

AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; Rad_Sig, radiomics signature; Clinic_Sig, clinic signature.

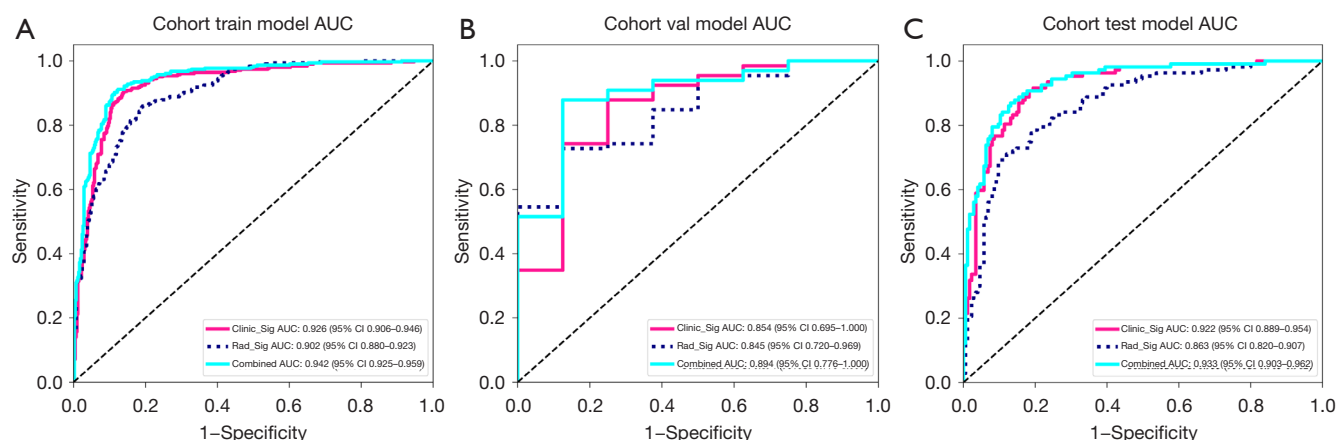


Figure 5 AUCs for the clinical, radiomics, and combined models for (A) training cohort, (B) validation cohort, and (C) test cohort. AUC, area under the curve; CI, confidence interval; Clinic_Sig, clinic signature; Rad_Sig, radiomics signature; Val, validation.

like typical Hashimoto's thyroiditis.

The tumor microenvironment, which refers to the internal environment in which tumor cells occur and exist, not only includes the tumor cells themselves, but also various surrounding cells such as fibroblasts, immune and inflammatory cells, and vascular cells. These cells organize in spatially structured arrangements, exhibiting microenvironmental niches, nutrient gradients and cell-cell interactions. In recent years, it has become increasingly appreciated that the tumor microenvironment plays a crucial role in mediating complex phenomena, such as tumor progression and response to treatment (18). Therefore, the imaging research of tumors has also expanded from intratumoral to peritumoral. Recently, multiple studies have

confirmed the diagnostic and predictive ability of radiomics in the peritumoral region in disease assessment (19–21). Currently, investigations into perinodular imaging have predominantly focused on CT and magnetic resonance imaging (MRI) (22,23). For example, Ding *et al.* (24) predicted sentinel lymph node metastasis of breast cancer by using the peritumoral MRI radiomics, concluding that the selection of tumor surrounding areas can affect the prediction efficiency. There are few reports on ultrasound based on peritumoral radiomics research, and the vast majority of studies on thyroid ultrasound imaging mainly focus on the intranodular regions (25,26). These studies did not consider the impact of perinodular regions, so we have developed radiomics models for different regions around

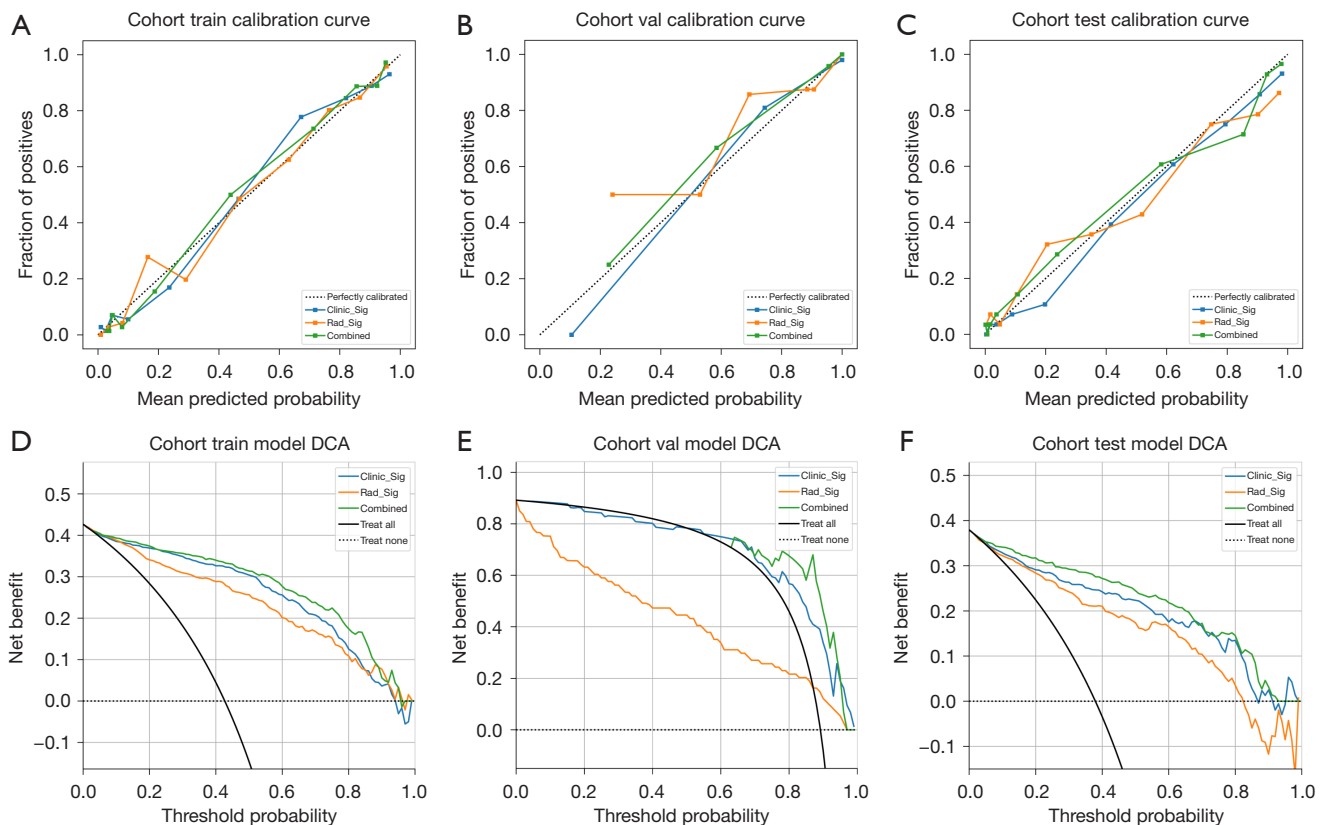


Figure 6 Calibration curves of (A) training cohort, (B) validation cohort and (C) test cohort. The closer fit of the diagonal curved line to the ideal straight line indicates the predictive accuracy of the combined model from the best model. DCA for three models in (D) training cohort, (E) validation cohort and (F) test cohort. Clinic_Sig, clinical signature; Rad_Sig, radiomics signature; DCA, decision curve analysis; Val, validation.

thyroid nodules. We expect that peritumoral radiomics can improve the prediction efficiency of benign and malignant thyroid nodules.

In this study, we first investigated the discriminative ability of intranodular, intranodular combined with five different perinodular regions radiomics features of thyroid for their malignancy. The results showed the intra+p1 model had the best predictive performance in the test cohort. Tang *et al.* (27) found that the closer the chosen peritumoral region was to the tumor, the richer its information content, leading to a higher number of features with meaningful distinctions that reflect variations in tumor characteristics. Zhou's team (28) compared the intratumoral and different peritumoral regions and also believed that using the smallest bounding box containing proximal peritumor tissue as input had higher accuracy compared to using tumor alone or larger boxes. Our conclusion was consistent with the above researches,

that is, 1 pixel extension of nodules showed the best predictive performance. We selected 28 intranodular and 32 perinodular features for the Rad_Sig, which showcased superior performance in the test cohort (AUC = 0.863, 95% CI: 0.820–0.907) when compared to intranodular model (AUC = 0.849, 95% CI: 0.803–0.894). It is worth noting that Rad_Sig exhibit lower performance than Clinic_Sig in predicting benign and malignant thyroid nodules, and this finding deserves further exploration. As an emerging technology, radiomics has shown great potential in medical image analysis, but its effectiveness in practical applications may still be constrained by various factors, including technological maturity, data quality, and the pathological and physiological characteristics of specific diseases (29). Sorrenti *et al.* also proposed a similar viewpoint in their research (30). Our original intention in developing this predictive model is to assist ultrasound doctors in improving their diagnostic capabilities through objective

means, thereby reducing unnecessary overtreatment and optimizing the allocation of medical resources. This vision is highly aligned with the current trend in the medical field to pursue precision medicine and reduce human error, so we adopt the method of subjective evaluation by senior doctors and introduce higher-level expert rulings to resolve disputes, which undoubtedly greatly enhances the reliability and consistency of the evaluation results. Although Rad_Sig performed slightly worse than Clinic_Sig in this study, it does not mean that radiomics has no value or potential in the diagnosis of thyroid nodules. On the contrary, this result may suggest that we need to further optimize radiomics algorithms, expand sample sizes, improve data quality, and explore fusion strategies between radiomics and clinical features in future research to fully leverage their complementary advantages in diagnosis. Our study was conducted across multiple centers, enhancing the credibility of our findings. However, compared to the intranodular model, the radiomics model of intranodular and perinodular with 1 pixel (intra+p1) did not significantly improve predictive performance. Thyroid tissue is recognized for its richness in blood vessels and lymphatic vessels, which can serve as routes for tumour invasion. While the mechanisms driving thyroid tumour metastasis are not fully comprehended, Pereira *et al.* (31) established that thyroid tumour invasion was not linked to the number of peritumoural lymphatic vessels. This factor might influence the presence of peri-tumoural lymphatic vessels in thyroid nodules, potentially impacting the specificity of radiomic features. Moreover, ultrasound imaging might grapple with reduced image clarity and limited valid information (32), further contributing to decreased specificity of extracted perineural features. The improvement of ultrasound image quality might help us discover more valuable image information.

Subsequently, a combined model centered on ultrasound images was established, combining clinical and radiomics features. Ultrasound radiomics, which analyzes the overall image of the lesion, can visualize tumor heterogeneity, and has yielded many achievements in the screening, diagnosis, and evaluation of thyroid tumors and can be used as an evaluation tool for physicians to assess the pathophysiologic status of various aspects of the thyroid (25). The combined model no longer only pays attention to the image features of patients, but also considers other clinical factors, which is expected to make the diagnosis and treatment of patients more personalized and comprehensive, and be used to solve clinical problems. In our previous study (33) on the

prediction model of pathological complete response in neoadjuvant chemotherapy for breast cancer, we found that the combined model (training: AUC =0.930, validation: AUC =0.895) is more effective than the single clinical model (training: AUC =0.869, validation: AUC =0.775, $P<0.05$). This study indicated that the combined model constructed by combining clinical features and radiomic features extracted from both intranodular and perinodular regions yielded superior predictive capability compared to utilizing Clinic_Sig and Rad_Sig individually. In the training cohort, the combined model achieved an AUC of 0.942, surpassing both Clinic_Sig (AUC =0.926) and Rad_Sig (AUC =0.902). Similarly, in the validation cohort, the combined model outperformed Clinic_Sig (AUC =0.854) and Rad_Sig (AUC =0.845) with an AUC of 0.894. This trend was also evident in the test cohort, where the combined model (AUC =0.933) outperformed Clinic_Sig (AUC =0.922) and Rad_sig (AUC =0.863). This further confirms the above research conclusions.

Nevertheless, there are three main limitations in this study. Firstly, our research is only based on the differentiation of benign and malignant thyroid nodules, and future research on thyroid nodules based on different pathological subtypes is an important direction. Secondly, although we have implemented a multicenter study, the overall sample size of our study is relatively small, and extracting a large number of image features from imaging data may require a larger sample size to establish a robust model. Finally, in our retrospective study, only grayscale two-dimensional ultrasound images were used. In the future, stricter quality control standards for ultrasound images and multimodal research combining elastography and contrast-enhanced ultrasound are the focus of our research.

Conclusions

In summary, our study delves into the potential of intranodular and perinodular ultrasound radiomics in predicting the benign and malignant thyroid nodules. In addition, we have designed a combined model that combines clinical and radiomics features, which is expected to serve as a valuable non-invasive predictive tool for clinical decision-making and tailored treatment strategies.

Acknowledgments

Funding: This work was supported by the Natural Science

Foundation of Inner Mongolia (No. 2023MS08031); Ordos Science and Technology Plan Project (No. Ordos 2019501); Weifang Youth Medical Talent Support Project and Qingdao University Medical Group Research Project (No. YLJT20242003).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://gs.amegroups.com/article/view/10.21037/gS-24-416/rc>

Data Sharing Statement: Available at <https://gs.amegroups.com/article/view/10.21037/gS-24-416/dss>

Peer Review File: Available at <https://gs.amegroups.com/article/view/10.21037/gS-24-416/prf>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://gs.amegroups.com/article/view/10.21037/gS-24-416/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the ethics committee of Ordos Central Hospital (No. 2024-100), and the same approval was recognized by the two other subcenters. Informed consent was waived for this retrospective study, and no personal information was disclosed.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Alexander EK, Doherty GM, Barletta JA. Management of thyroid nodules. *Lancet Diabetes Endocrinol* 2022;10:540-8.
- Alexander EK, Cibas ES. Diagnosis of thyroid nodules. *Lancet Diabetes Endocrinol* 2022;10:533-9.
- Wei P, Jiang N, Ding J, et al. The Diagnostic Role of Computed Tomography for ACR TI-RADS 4-5 Thyroid Nodules With Coarse Calcifications. *Front Oncol* 2020;10:911.
- Haddad RI, Bischoff L, Ball D, et al. Thyroid Carcinoma, Version 2.2022, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* 2022;20:925-51.
- Yucel S, Balci IG, Tomak L. Diagnostic Performance of Thyroid Nodule Risk Stratification Systems: Comparison of ACR-TIRADS, EU-TIRADS, K-TIRADS, and ATA Guidelines. *Ultrasound Q* 2023;39:206-11.
- Wildman-Tobriner B, Buda M, Hoang JK, et al. Using Artificial Intelligence to Revise ACR TI-RADS Risk Stratification of Thyroid Nodules: Diagnostic Accuracy and Utility. *Radiology* 2019;292:112-9.
- Summers RM. Artificial Intelligence of COVID-19 Imaging: A Hammer in Search of a Nail. *Radiology* 2021;298:E162-4.
- Guiot J, Vaidyanathan A, Deprez L, et al. A review in radiomics: Making personalized medicine a reality via routine imaging. *Med Res Rev* 2022;42:426-40.
- Yoon J, Lee E, Kang SW, et al. Implications of US radiomics signature for predicting malignancy in thyroid nodules with indeterminate cytology. *Eur Radiol* 2021;31:5059-67.
- Liang J, Huang X, Hu H, et al. Predicting Malignancy in Thyroid Nodules: Radiomics Score Versus 2017 American College of Radiology Thyroid Imaging, Reporting and Data System. *Thyroid* 2018;28:1024-33.
- Wang X, Zhao X, Li Q, et al. Can peritumoral radiomics increase the efficiency of the prediction for lymph node metastasis in clinical stage T1 lung adenocarcinoma on CT? *Eur Radiol* 2019;29:6049-58.
- Shi J, Dong Y, Jiang W, et al. MRI-based peritumoral radiomics analysis for preoperative prediction of lymph node metastasis in early-stage cervical cancer: A multi-center study. *Magn Reson Imaging* 2022;88:1-8.
- Hu Y, Xie C, Yang H, et al. Assessment of Intratumoral and Peritumoral Computed Tomography Radiomics for Predicting Pathological Complete Response to Neoadjuvant Chemoradiation in Patients With Esophageal Squamous Cell Carcinoma. *JAMA Netw Open* 2020;3:e2015927.
- Zhou J, Yin L, Wei X, et al. 2020 Chinese guidelines

- for ultrasound malignancy risk stratification of thyroid nodules: the C-TIRADS. *Endocrine* 2020;70:256-79.
15. Wang Y, Ding Y, Liu X, et al. Preoperative CT-based radiomics combined with tumour spread through air spaces can accurately predict early recurrence of stage I lung adenocarcinoma: a multicentre retrospective cohort study. *Cancer Imaging* 2023;23:83.
 16. Huang Y, Zhu T, Zhang X, et al. Longitudinal MRI-based fusion novel model predicts pathological complete response in breast cancer treated with neoadjuvant chemotherapy: a multicenter, retrospective study. *EClinicalMedicine* 2023;58:101899.
 17. Tywanek E, Michalak A, Świrski J, et al. Autoimmunity, New Potential Biomarkers and the Thyroid Gland-The Perspective of Hashimoto's Thyroiditis and Its Treatment. *Int J Mol Sci* 2024;25:4703.
 18. Elhanani O, Ben-Uri R, Keren L. Spatial profiling technologies illuminate the tumor microenvironment. *Cancer Cell* 2023;41:404-20.
 19. Pérez-Morales J, Tunali I, Stringfield O, et al. Peritumoral and intratumoral radiomic features predict survival outcomes among patients diagnosed in lung cancer screening. *Sci Rep* 2020;10:10528.
 20. Zhuo Y, Feng M, Yang S, et al. Radiomics nomograms of tumors and peritumoral regions for the preoperative prediction of spread through air spaces in lung adenocarcinoma. *Transl Oncol* 2020;13:100820.
 21. Sun R, Limkin EJ, Vakalopoulou M, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol* 2018;19:1180-91.
 22. Wu Y, Zhu M, Liu Y, et al. Peritumoral Imaging Manifestations on Gd-EOB-DTPA-Enhanced MRI for Preoperative Prediction of Microvascular Invasion in Hepatocellular Carcinoma: A Systematic Review and Meta-Analysis. *Front Oncol* 2022;12:907076.
 23. Xu H, Liu J, Chen Z, et al. Intratumoral and peritumoral radiomics based on dynamic contrast-enhanced MRI for preoperative prediction of intraductal component in invasive breast cancer. *Eur Radiol* 2022;32:4845-56.
 24. Ding J, Chen S, Serrano Sosa M, et al. Optimizing the Peritumoral Region Size in Radiomics Analysis for Sentinel Lymph Node Status Prediction in Breast Cancer. *Acad Radiol* 2022;29 Suppl 1:S223-8.
 25. Cao Y, Zhong X, Diao W, et al. Radiomics in Differentiated Thyroid Cancer and Nodules: Explorations, Application, and Limitations. *Cancers (Basel)* 2021;13:2436.
 26. Lu WW, Zhang D, Ni XJ. A Review of the Role of Ultrasound Radiomics and Its Application and Limitations in the Investigation of Thyroid Disease. *Med Sci Monit* 2022;28:e937738.
 27. Tang X, Huang H, Du P, et al. Intratumoral and peritumoral CT-based radiomics strategy reveals distinct subtypes of non-small-cell lung cancer. *J Cancer Res Clin Oncol* 2022;148:2247-60.
 28. Zhou J, Zhang Y, Chang KT, et al. Diagnosis of Benign and Malignant Breast Lesions on DCE-MRI by Using Radiomics and Deep Learning With Consideration of Peritumor Tissue. *J Magn Reson Imaging* 2020;51:798-809.
 29. Tessler FN, Thomas J. Artificial Intelligence for Evaluation of Thyroid Nodules: A Primer. *Thyroid* 2023;33:150-8.
 30. Sorrenti S, Dolcetti V, Radzina M, et al. Artificial Intelligence for Thyroid Nodule Characterization: Where Are We Standing? *Cancers (Basel)* 2022;14:3357.
 31. Pereira F, Pereira SS, Mesquita M, et al. Lymph Node Metastases in Papillary and Medullary Thyroid Carcinoma Are Independent of Intratumoral Lymphatic Vessel Density. *Eur Thyroid J* 2017;6:57-64.
 32. Nguyen DT, Kang JK, Pham TD, et al. Ultrasound Image-Based Diagnosis of Malignant Thyroid Nodule Using Artificial Intelligence. *Sensors (Basel)* 2020;20:1822.
 33. Zhu X, Shen J, Zhang H, et al. A Novel Combined Nomogram Model for Predicting the Pathological Complete Response to Neoadjuvant Chemotherapy in Invasive Breast Carcinoma of No Specific Type: Real-World Study. *Front Oncol* 2022;12:916526.

Cite this article as: Zhu X, Li J, Li H, Wang K, Zhang J, Meng J, Wu R, Zhang M, Du H. Intranodular and perinodular ultrasound radiomics distinguishes benign and malignant thyroid nodules: a multicenter study. *Gland Surg* 2024;13(12):2359-2371. doi: 10.21037/gs-24-416