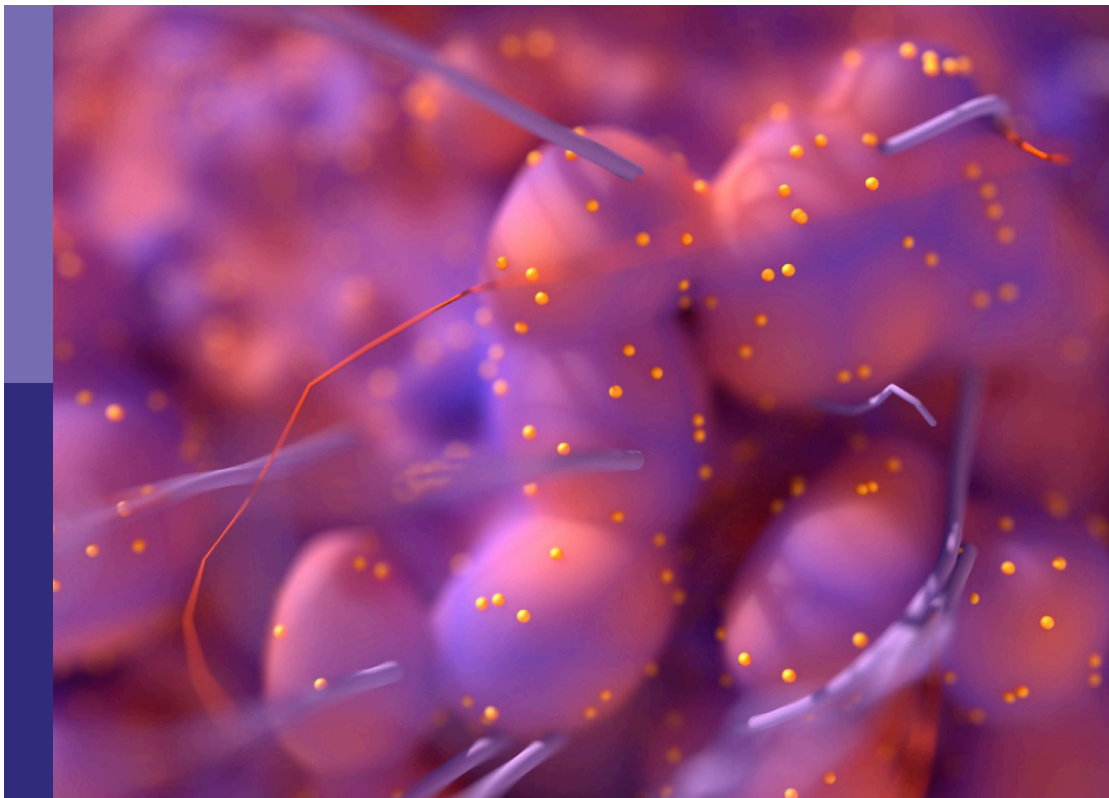
frontiers | Research Topics

# Cancer imaging techniques to distinguish benign and malignant tumors

**Edited by**
Damiano Caruso, Vikram D. Kodibagkar and Natalie Julie Serkova

**Published in**
Frontiers in Oncology



ISSN 1664-8714
ISBN 978-2-8325-2911-9
DOI 10.3389/978-2-8325-2911-9

## About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

# Cancer imaging techniques to distinguish benign and malignant tumors

**Topic editors**

Damiano Caruso — Sapienza University of Rome, Italy
Vikram D. Kodibagkar — Arizona State University, United States
Natalie Julie Serkova — University of Colorado Anschutz Medical Campus, United States

**Citation**

Caruso, D., Kodibagkar, V. D., Serkova, N. J., eds. (2023). *Cancer imaging techniques to distinguish benign and malignant tumors*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-2911-9

# Table of contents

frontiers in Oncology

ORIGINAL RESEARCH
published: 29 June 2021
doi: 10.3389/fonc.2021.592927

# Establishment of an Ultrasound Malignancy Risk Stratification Model for Thyroid Nodules Larger Than 4 cm

*Xuehua Xi[1†], Ying Wang[2†], Luying Gao[2†], Yuxin Jiang[2], Zhiyong Liang[3], Xinyu Ren[3], Qing Gao[2], Xingjian Lai[2], Xiao Yang[2], Shenling Zhu[2], Ruina Zhao[2], Xiaoyan Zhang[2] and Bo Zhang[1*]*

[1] Department of Ultrasound, China-Japan Friendship Hospital, Beijing, China, [2] Department of Ultrasound, Peking Union Medical College Hospital, Beijing, China, [3] Department of Pathology, Peking Union Medical College Hospital, Beijing, China

OPEN ACCESS

**Edited by:**
*Vikram D. Kodibagkar, Arizona State University, United States*

**Reviewed by:**
*Richa Vaish, Tata Memorial Hospital, India*
*Huan Liu, GE Healthcare, China*

***Correspondence:**
*Bo Zhang*
*thyroidus@163.com*

*[†]These authors have contributed equally to this work*

**Specialty section:**
*This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology*

**Received:** *08 August 2020*
**Accepted:** *14 June 2021*
**Published:** *29 June 2021*

**Citation:**
*Xi X, Wang Y, Gao L, Jiang Y, Liang Z, Ren X, Gao Q, Lai X, Yang X, Zhu S, Zhao R, Zhang X and Zhang B (2021) Establishment of an Ultrasound Malignancy Risk Stratification Model for Thyroid Nodules Larger Than 4 cm. Front. Oncol. 11:592927. doi: 10.3389/fonc.2021.592927*

**Background:** The incidence and mortality of thyroid cancer, including thyroid nodules > 4 cm, have been increasing in recent years. The current evaluation methods are based mostly on studies of patients with thyroid nodules < 4 cm. The aim of the current study was to establish a risk stratification model to predict risk of malignancy in thyroid nodules > 4 cm.

**Methods:** A total of 279 thyroid nodules > 4 cm in 267 patients were retrospectively analyzed. Nodules were randomly assigned to a training dataset (n = 140) and a validation dataset (n = 139). Multivariable logistic regression analysis was applied to establish a nomogram. The risk stratification of thyroid nodules > 4 cm was established according to the nomogram. The diagnostic performance of the model was evaluated and compared with the American College Radiology Thyroid Imaging Reporting and Data System (ACR TI-RADS), Kwak TI-RADS and 2015 ATA guidelines using the area under the receiver operating characteristic curve (AUC).

**Results:** The analysis included 279 nodules (267 patients, 50.6 ± 13.2 years): 229 were benign and 50 were malignant. Multivariate regression revealed microcalcification, solid mass, ill-defined border and hypoechogenicity as independent risk factors. Based on the four factors, a risk stratified clinical model was developed for evaluating nodules > 4 cm, which includes three categories: high risk (risk value = 0.8-0.9, with more than 3 factors), intermediate risk (risk value = 0.3-0.7, with 2 factors or microcalcification) and low risk (risk value = 0.1-0.2, with 1 factor except microcalcification). In the validation dataset, the malignancy rate of thyroid nodules > 4 cm that were classified as high risk was 88.9%; as intermediate risk, 35.7%; and as low risk, 6.9%. The new model showed greater AUC than ACR TI-RADS (0.897 *vs.* 0.855, p = 0.040), but similar sensitivity (61.9% *vs.* 57.1%, p = 0.480) and specificity (91.5% *vs.* 93.2%, p = 0.680).

**Conclusion:** Microcalcification, solid mass, ill-defined border and hypoechogenicity on ultrasound may be signs of malignancy in thyroid nodules > 4 cm. A risk stratification

model for nodules > 4 cm may show better diagnostic performance than ACR TI-RADS, which may lead to better preoperative decision-making.

**Keywords: thyroid nodules, thyroid cancer, ultrasound, risk stratification, size**

## INTRODUCTION

Thyroid nodules occur in up to 68% of people in the general population worldwide, and 5-15% of nodules are malignant (1, 2). Research shows that the incidence and mortality of thyroid cancer, including thyroid nodules > 4 cm, has been on the rise in recent years and warrants further research (3). Both the 2017 Thyroid Cancer Staging Manual of the American Joint Committee on Cancer (AJCC) and the 2015 Management Guidelines of the American Thyroid Association (ATA) (Referred to as ATA) for adult patients with thyroid nodules and differentiated thyroid cancer list thyroid nodules > 4 cm as an important factor for surgical decision-making, as integrated into the Tumor, Node, Metastasis (TNM) staging system (4, 5). Recent guidelines have suggested ultrasound risk stratification patterns to assess the malignant risk of thyroid nodules, including the American College Radiology Thyroid Imaging Reporting and Data System (ACR TI-RADS) (4, 6–9). However, these methods are based mostly on research of thyroid nodules < 4 cm. For example, fine-needle aspiration biopsy (FNAB) is considered to be the gold standard for preoperative diagnosis of thyroid cancer. However, FNAB shows lower sensitivity and higher rates of false negative results in the case of thyroid nodules > 4 cm (10–12). Decisions related to surgery and other treatments may be affected if pre-operative assessment of thyroid nodules is inaccurate or incomplete. Thus, distinguishing malignant from benign nodules pre-operatively would assist in diagnosis and decision-making. The aim of the current study was to identify factors that predict malignancy in thyroid nodules > 4 cm and construct an applicable risk stratification model.

## MATERIAL AND METHODS

Study protocols are shown in **Figure 1**.

### Patients

Consecutive patients with at least one thyroid nodule > 4 cm who underwent thyroidectomy at the Peking Union Medical College Hospital (Beijing, China) between 2010 and 2017 were reviewed retrospectively. The inclusion criteria were as follows :(1) the size of the nodule was > 4 cm in its longest diameter, as determined by ultrasonography; and (2) the nodule had not previously been treated surgically. The exclusion criteria were: (1) pathology results from surgical tissue were unavailable for the patient, or (2) ultrasound images were poor or incomplete. A total of 279 thyroid nodules in 267 patients were included for the study (**Figure 1**). The patients comprised 185 females aged 60.0 ± 13.2 yr and 82 men aged 49.9 ± 13.3 years (**Figure 1**).

The study was approved by the Institutional Review Board of Peking Union Medical College Hospital. All patients provided informed consent for their clinical data to be published anonymously for research purposes.

### Ultrasound Examination

Relevant clinical and ultrasound data of all cases were extracted from the central hospital database. Ultrasound examinations were performed with Phillips IU 22, GE Logiq 9 or GE Logiq 7 devices equipped with a linear array probe of 8-15 MHz A convex array probe of 5-12 MHz was used for larger thyroid nodules. Ultrasound images were retrospectively reviewed by two radiologists who had more than 5 years' experience analyzing thyroid ultrasound, and who were blinded to patients' clinical and pathological results. The age and sex of the patients were recorded, as were ultrasound features of each nodule, including size, composition, echogenicity, margin, shape, border, calcification and halo. The two radiologists resolved any inconsistencies in their reviews through discussion.

All thyroid nodules were also evaluated using ACR TI-RADS, Kwak TI-RADS and ATA (4, 6, 9). According to the ACR TI-RADS, points were given for all ultrasound features in a nodule. Features suggesting malignancy were awarded additional points. The total points determined the nodule's ACR TI-RADS level, which ranged from TR1 (benign) to TR5 (high probability of malignancy) (**Table 1**).

### Statistical Analysis

Data analysis was performed using SPSS 19.0 (IBM, Chicago, IL, USA) and $p < 0.05$ as the definition of statistical significance. Nodules were randomly assigned to a training dataset or validation dataset (13). Continuous data were reported as means ± SD, and inter-group differences were assessed for significance using Student's *t*-test. Differences in categorical data were assessed using the $\chi^2$-test or Fisher's exact test as appropriate. Categorical variables were classified based on clinical and ultrasound findings. The continuous variable age was transformed into a categorical variable (≥55 or < 55 years) based on a previous report (5).

The variables that were identified as statistically significant prognostic factors were assessed in multivariate logistic regression analysis. A nomogram was constructed based on the results of multivariate analysis and validated using the validation dataset, using the *rms* package in R 3.6.0. The diagnostic performance of the nomogram was evaluated using the concordance index

**Abbreviations:** ACR TI-RADS, American College Radiology Thyroid Imaging Reporting and Data System; AUC, area under the receiver operating characteristic curve; AJCC, American Joint Committee on Cancer; ATA, American Thyroid Association; TNM, Tumor, Node, Metastasis; FNAB, fine-needle aspiration biopsy; C-index, concordance index; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval.

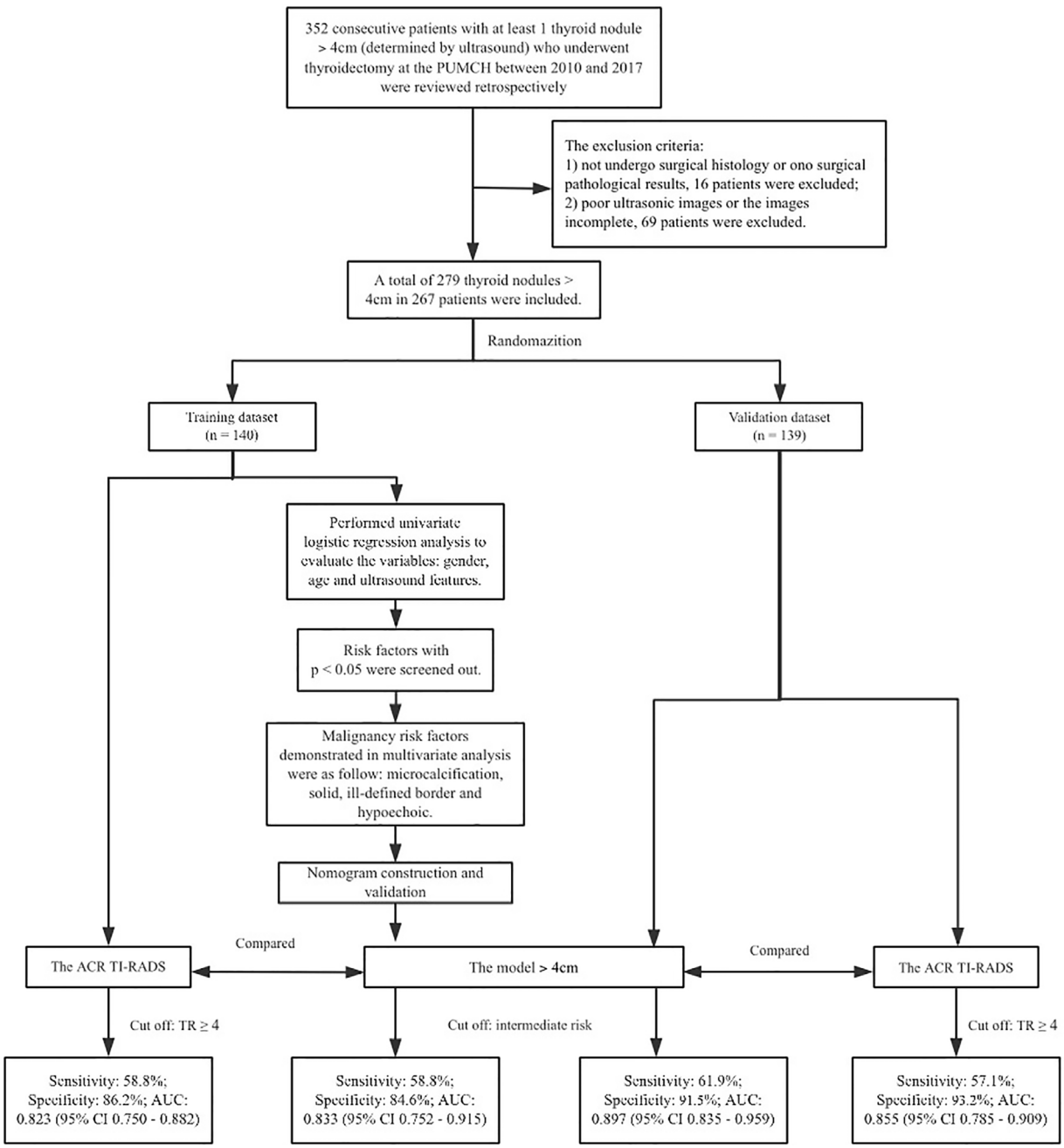



**FIGURE 1** | The study protocol.

(C-index) and area under the receiver operating characteristic curve (AUC). Bootstrapping validation (1,000 bootstrap resamples) was used to calculate a relative corrected C-index (14). A calibration curve (1,000 bootstrap resamples) was generated to verify the calibration of the prediction nomogram.

A model for risk stratification of thyroid nodules > 4 cm was established according to the nomogram. Nodules were classified as high, intermediate, or low risk. The cut-off values for the three-level risk stratification were determined according to AUC, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Accuracy was calculated according to the cut-off value.

Similarly, the diagnostic performance of the ACR TI-RADS was evaluated in terms of AUC, sensitivity, specificity, PPV, NPV and accuracy. The results were compared between this reference standard and proposed model.

**TABLE 1** | Point system in ACR TI-RADS (6).

| Points | Categories | | | | |
|---|---|---|---|---|---|
| | Composition | Echogenicity | Shape | Margin | Echogenic foci |
| 0 | cystic or almost completely cystic; spongiform | anechoic | wider-than-tall | smooth;ill-defined | none;large comet-tail artifacts |
| 1 | mixed cystic and solid | Hyperechoic;isoechoic | | | macrocalcifications |
| 2 | solid or almost completely solid | hypoechoic | | lobulated or irregular | peripheral (rim) calcifications |
| 3 | | very hypoechoic | taller-than-wide | extra-thyroidal extension | punctate echogenic foci |

# RESULTS

## Clinical Characteristics of the Sample

The study comprised 279 nodules in 267 patients, including 50 (17.9%) malignant and 229 (82.1%) benign nodules (**Table 2**). Of the 267 patients with a mean age of 50.6 ± 13.2 years (range, 17-80 years), 185 were women (60.0 ± 13.2 years) and 82 were men (49.9 ± 13.3 years).

Nodules were randomly assigned to a training dataset (n=140) or validation dataset (n=139) (**Table 2**). There was no significant difference in the ultrasound features between the two datasets. In this study, 264 (94.6%) thyroid nodules were performed with Phillips IU 22, 12 (4.3%, 1 malignant and 11 benign) were performed with GE Logiq 9, and 3 (1.1%, 3 benign) were performed with GE Logiq 7. The interclass correlation coefficient (ICC) of the two radiologists was 0.86 (95% CI, 0.76-0.92).

## Model Construction and Validation

In the training dataset, 29 nodules (19.3%) were malignant (**Table 2**). We performed univariate logistic regression analysis using age, sex, composition, echogenicity, border, margin, shape, calcification, and halo. All variables except age and shape were identified as statistically significant risk factors (**Table 2**). These risk factors were then included in the multivariable analysis.

**TABLE 2** | Ultrasound features and ACR TI-RADS levels for 279 thyroid nodules > 4 cm.

| Feature or level | All nodules (n = 279) | Training dataset (n = 140) | | | Validation dataset (n = 139) | | |
|---|---|---|---|---|---|---|---|
| | | Pathological diagnosis | | p-value | Pathological diagnosis | | p-value |
| | | Malignant (n = 29) | Benign (n = 111) | | Malignant (n = 21) | Benign (n = 118) | |
| Ultrasound features | | | | | | | |
| Size (cm) | 5.24 ± 1.1 | 5.7 ± 1.5 | 5.1 ± 1.0 | 0.349 | 5.0 ± 0.5 | 5.3 ± 1.2 | |
| Composition | | | | <0.001 | | | <0.001 |
| Solid | 64 (22.9) | 15 (51.7) | 13 (11.7) | | 14 (66.7) | 22 (18.6) | |
| Mixed cystic and solid/Cystic | 215 (77.1) | 14 (48.3) | 98 (88.3) | | 7 (33.3) | 96 (81.4) | |
| Echogenicity | | | | 0.001 | | | <0.001 |
| Hypoechoic | 122 (43.7) | 22 (75.9) | 44 (39.6) | | 17 (81.0) | 39 (33.1) | |
| Other* | 157 (56.3) | 7 (24.1) | 76 (60.4) | | 4 (19.0) | 79 (66.9) | |
| Border | | | | <0.001 | | | <0.001 |
| Ill-defined | 28 (10) | 9 (31.0) | 6 (5.4) | | 8 (4.2) | 5 (4.2) | |
| Defined | 251 (90) | 20 (69.0) | 105 (94.6) | | 13 (61.9) | 113 (95.8) | |
| Margin | | | | <0.001 | | | <0.001 |
| Irregular/lobulated | 42 (15.1) | 10 (34.5) | 7 (6.3) | | 10 (47.6) | 15 (12.7) | |
| Smooth | 237 (84.9) | 19 (65.5) | 104 (93.7) | | 11 (52.4) | 103 (87.3) | |
| Shape | | | | – | | | – |
| Taller-than-wide >1 | 0 (-) | 0 (-) | 0 (-) | | 0 (-) | 0 (-) | |
| ≤1 | 279 (100) | 29 (100) | 111 (-) | | 29 (100) | 111 (-) | |
| Calcification | | | | <0.001 | | | <0.001 |
| Microcalcification | 21 (7.5) | 9 (31) | 3 (2.7) | | 8 (38.1) | 1 (0.8) | |
| Others# | 258 (92.5) | 20 (69) | 108 (97.3) | | 13 (61.9) | 117 (99.2) | |
| Halo | | | | 0.014 | | | 0.185 |
| Absent | 127 (45.5) | 17 (58.6) | 43 (38.7) | | 12 (57.1) | 55 (46.6) | |
| Irregular | 50 (17.9) | 8 (27.6) | 20 (18.0) | | 5 (23.8) | 17 (14.4) | |
| Regular and thin | 102 (36.5) | 4 (13.8) | 48 (43.2) | | 4 (19.0) | 46 (39.0) | |
| ACR TI-RADS | | | | <0.001 | | | <0.001 |
| TR5 | 16 (5.7) | 8 (27.6) | 1 (0.9) | | 6 (28.6) | 1 (0.8) | |
| TR4 | 35 (12.5) | 10 (34.5) | 7 (6.3) | | 7 (33.3) | 11 (9.3) | |
| TR3 | 150 (53.8) | 10 (34.5) | 68 (61.3) | | 8 (38.1) | 64 (54.2) | |
| TR2 | 73 (26.2) | 1 (3.4) | 34 (30.6) | | 0 (-) | 38 (32.2) | |
| TR1 | 5 (1.8) | 0 (-) | 1 (0.9) | | 0 (-) | 4 (3.4) | |

*Values are mean ± SD or n (%), unless otherwise noted. *Including isoechoic, hyperechoic, and anechoic. #Including macrocalcification, large comet-tail artifacts and none. ACR, American College Radiology; TI-RADS, Thyroid Imaging Reporting and Data System.*

Malignancy risk factors in multivariate analysis were as follows (**Table 3**): microcalcification [odds ratio (OR) 8.37, 95% confidence interval (CI) 1.641-42.724, p = 0.011], solid (OR 1.49, 95% CI 0.391-2.566, p = 0.008), ill-defined border (OR 4.40, 95% CI 1.074-18.031, p = 0.039) and hypoechogenicity (OR 2.94, 95% CI 1.031-8.389, p = 0.044).

A nomogram that integrated all four significant independent factors was constructed (**Figure 2**). The model showed a C-index of 0.833 (95% CI 0.752-0.915) for predicting malignancy in the training dataset (**Figure 3A**) and 0.897 (95% CI 0.835-0.9591) for predicting malignancy in the validation dataset (**Figure 3C**). Calibration curves for the probability of malignancy showed a good correlation between the nomogram-predicted and observed values (**Figures 3B, D**).

The risk value of each factor was calculated using the nomogram. The risk value was 0.34 for microcalcification (100 points), 0.21 for solid and ill-defined border (70 points), and 0.14 for hypoechogenicity (50 points). Using the model, all nodules were assigned to one of three risk categories (**Table 4**): nodules with more than 3 factors were classified as high risk (0.8-0.9); nodules with 2 factors or microcalcification, as intermediate risk (0.3-0.7); and nodules with 1 factor except microcalcification, as low risk (0.1-0.2).

In the training dataset, the malignancy rate of thyroid nodules > 4 cm that were classified as high risk was 90.0%; intermediate risk, 42.1%; and low risk, 10.8%. The corresponding malignancy rates in the validation dataset were 88.9%, 35.7%, and 6.9% (**Table 4**). The risk stratification of the model was compared with the ACR TI-RADS (p < 0.001, **Table 5**).

## Diagnostic Efficiency of the Model

Receiver operating characteristic curves demonstrated that the best cut-off value of the model was intermediate risk. In the training dataset, the model had a sensitivity of 58.8%, specificity of 84.6%, NPP of 93.7%, PPV of 34.5%, accuracy of 81.4% and AUC of 0.833 (95% CI 0.752-0.915) (**Table 6**). The AUC of the nomogram was higher than that of ACR TI-RADS (0.823, 95% CI 0.750-0.882, p = 0.011). However, the model was similar to ACR TI-RADS in sensitivity (58.8% *vs.* 58.8%, p = 0.181) and

**TABLE 3** | Multivariate binary logistic regression in the training dataset.

| Factor | β | B. E | Wals | P value | Exp (B) | 95% CI |
|---|---|---|---|---|---|---|
| Microcalcification | 2.125 | 0.831 | 6.533 | 0.011 | 8.374 | 1.641-42.724 |
| Solid | 1.479 | 0.555 | 7.105 | 0.008 | 4.388 | 1.479-13.016 |
| Ill-defined border | 1.482 | 0.719 | 4.242 | 0.039 | 4.401 | 1.074-18.031 |
| Hypoechogenicity | 1.079 | 0.535 | 4.066 | 0.044 | 2.940 | 1.031-8.3888 |
| Intercept | -2.8219 | 0.459 | -6.15 | <0.0001 | | |

*Those factors with statistically significance (P<0.01) in the univariate analysis were added to the logistic analysis. CI, confidence interval.*

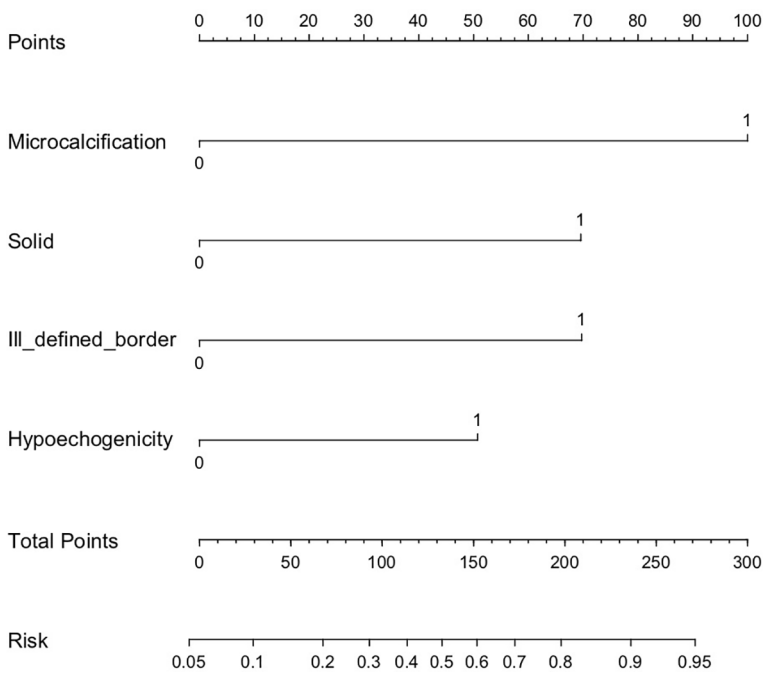



**FIGURE 2** | Risk stratification nomogram to predict the probability of malignancy of thyroid nodules > 4 cm. For example, a solid hypoechogenic thyroid nodule > 4 cm that shows the ultrasound feature of microcalcification, but not ill-defined border receives 221 points and a malignancy risk of 0.88 (high risk).

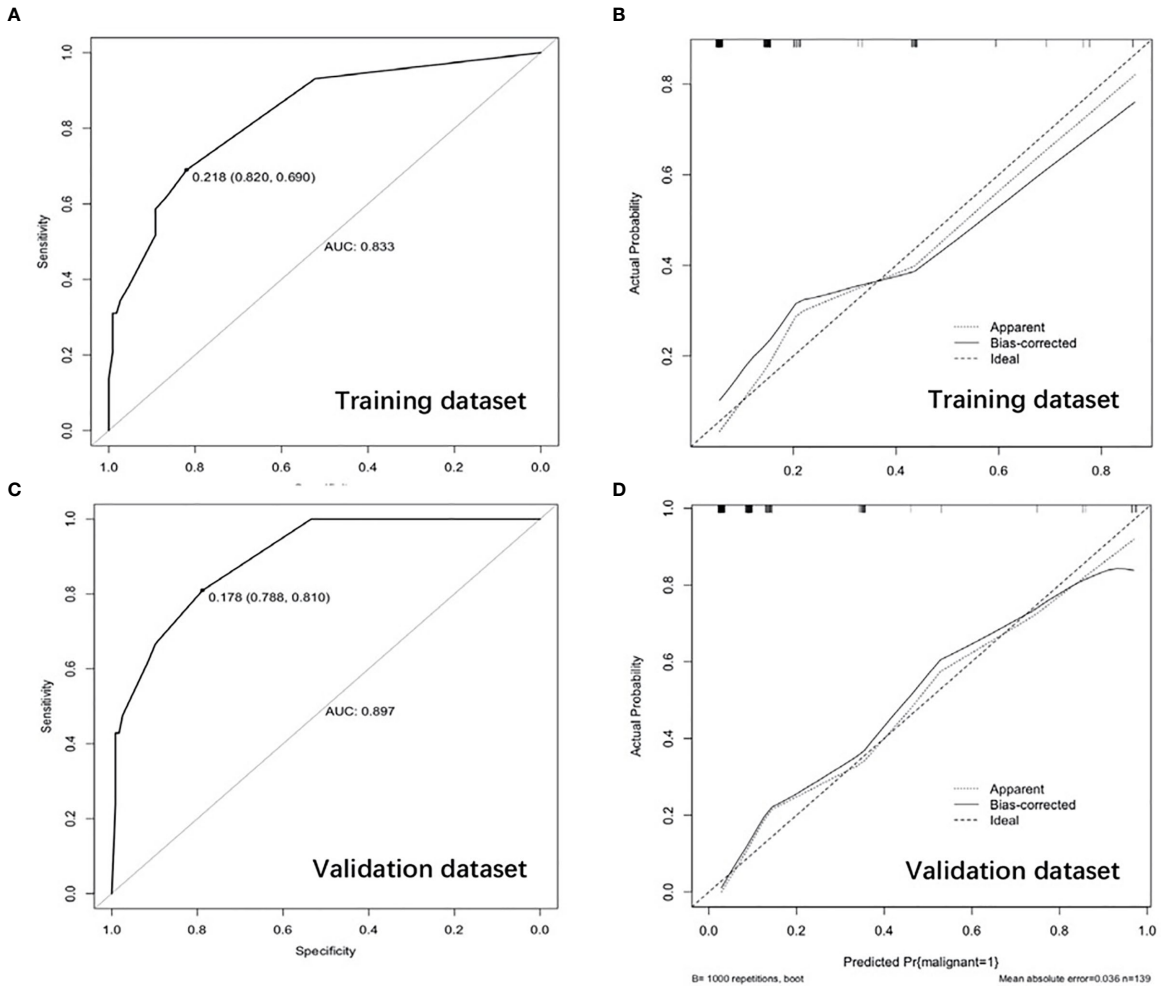



**FIGURE 3** | Receiver operating characteristic curve in the training dataset **(A)** and validation dataset **(C)**. Calibration curve showing nomogram-predicted malignancy compared with the actual malignancy in the training dataset **(B)** and validation dataset **(D)**.

specificity (84.6% *vs.* 86.2%, p = 0.424) (**Table 6**). The AUC of the nomogram was higher than that of Kwak TI-RADS (0.817, 95% CI 0.741-0.871, p = 0.214) and ATA (0.812, 95% CI 0.735-0.874, p = 0.158). There was no significant statistical difference in the AUC value among the ACR TI-RADS, Kwak TI-RADS and ATA (p > 0.05).

For the validation dataset, the model showed a sensitivity of 61.9%, specificity of 91.5%, NPV of 93.1%, PPV of 56.5%,

**TABLE 4** | Risk stratification model for thyroid nodules > 4 cm.

| Patterns | Points | Risk value | No. malignant nodules (malignancy rate, %) | |
|---|---|---|---|---|
| | | | training dataset (n = 29) | Validation dataset (n = 21) |
| High risk | | | 9 (90.0) | 8 (88.9) |
| More than 3 factors | 190-290 | 0.8-0.9 | | |
| Intermediate risk | | | 8 (42.1) | 5 (35.7) |
| 2 factors | 120-170 | 0.4-0.7 | | |
| Microcalcification | 100 | 0.34 | | |
| Low risk | | | 12 (10.8) | 8 (6.9) |
| Solid | 70 | 0.21 | | |
| Ill-defined border | 70 | 0.21 | | |
| Hypoechoic | 50 | 0.14 | | |

*The four significant factors were microcalcification, solid, ill-defined border, and hypoechoic.*

**TABLE 5 |** Comparison of malignancy risk stratification between the proposed model and ACR TI-RADS.

| ACR TI-RADS | Proposed model* | | | | | |
|---|---|---|---|---|---|---|
| | Training dataset (n = 140) | | | Validation dataset (n = 139) | | |
| | High risk | Intermediate risk | Low risk | High risk | Intermediate risk | Low risk |
| TR ≥4 | 8 (100) | 9 (66.7) | 9 (44.4) | 9 (88.9) | 11 (36.4) | 5 (20.0) |
| TR = 3 | 2 (50.0) | 10 (20.0) | 66 (10.6) | 0 (-) | 3 (33.3) | 69 (10.1) |
| TR ≤ 2 | 0 (-) | 0 (-) | 36 (2.8) | 0 (-) | 0 (-) | 42 (0) |

*Values are n (%) for malignant tumors, based on pathology of surgical samples. *The risk stratification of the model > 4 cm was compared with ACR TI-RADS in training dataset and validation dataset (p < 0.001). ACR, American College Radiology; TI-RADS, Thyroid Imaging Reporting and Data System.*

accuracy of 87.1% and AUC of 0.897 (95% CI 0.835-0.959, **Table 6**). The AUC of the nomogram was higher than that of ACR TI-RADS (0.855, 95% CI 0.785-0.909, p = 0.040, **Table 6**). However, the model was similar to ACR TI-RADS in sensitivity (61.9% *vs.* 57.1%, p = 0.480) and specificity (91.5% *vs.* 93.2%, p = 0.680) (**Table 6**). The AUC of the nomogram was higher than that of Kwak TI-RADS (0.878, 95% CI 0.811-0.927, p = 0.445) and ATA (0.831, 95% CI 0.758-0.889, p = 0.205). There was no significant statistical difference in the AUC value among the ACR TI-RADS, Kwak TI-RADS and ATA (p > 0.05).

## DISCUSSION

In the present study, we established a model to predict the risk of malignancy for thyroid nodules > 4 cm. This model incorporated four factors relatively easy to determine from conventional ultrasound imaging of nodules: composition, echogenicity, border, and calcification. We observed that the model achieved satisfactory diagnostic performance in both the training and validation datasets. Furthermore, the proposed model predicted malignancy better than ACR TI-RADS in both datasets, although it showed similar specificity and sensitivity as the reference standard.

The thyroid nodules > 4cm have its unique ultrasonic risk stratification. In our study, multivariate regression revealed the following independent risk factors for thyroid cancer: microcalcification (OR 8.37, 95% CI 1.641-42.724), solid (OR 1.49, 95% CI 0.391-2.566), ill-defined border (OR 4.40, 95% CI 1.074-18.031) and hypoechogenicity (OR 2.94, 95% CI 1.031-8.389). Our findings were consistent with the three suspicious malignant signs (microcalcification, solid and hypoechogenicity) in ACR TI-RADS and Kwak TI-RADS, which confirmed the effectiveness of the existing guidelines. These factors were incorporated together to develop a nomogram. This nomogram could be a useful and convenient tool in clinical practice to evaluate the malignancy risk of thyroid nodules > 4cm. The model showed a C-index of 0.833 (95% CI 0.752-0.915) for predicting malignancy in the training dataset and 0.897 (95% CI 0.835-0.9591) for predicting malignancy in the validation dataset. Calibration curve plotting demonstrated its significant predictive and discriminatory capacity in the validation cohort. Using the model, all nodules were assigned to one of three risk categories): high risk (0.8-0.9), intermediate risk (0.3-0.7) and low risk (0.1-0.2). In the training dataset, the malignancy rate of thyroid nodules > 4 cm that were classified as high risk was 90.0%; intermediate risk, 42.1%; and low risk, 10.8%. The corresponding malignancy rates in the validation dataset were 88.9%, 35.7%, and 6.9%.

The risk stratification of the model > 4cm was different with the ACR TI-RADS (p < 0.001), which indicating that the thyroid nodules > 4cm have unique characteristics of ultrasonic risk stratification.

The proposed model may be convenient to implement in the clinic. Of various factors linked to nodule malignancy, including solid, hypoechogenicity, microcalcification, taller-than-wide shape and irregular/lobulated margin (4, 6, 9), our model identified only four independent risk factors: solid, hypoechogenicity, microcalcification, and ill-defined border. This suggests that fewer ultrasound features can still provide reliable predictions of malignancy. Taller-than-wide shape is considered an insensitive but highly specific indicator of malignancy (9, 15–18), especially in sub-centimeter thyroid nodules (15, 16), and it is assigned more points (3 points) in ACR TI-RADS to reflect an association with malignancy (6). However, none of the thyroid nodules in either the training or validation datasets had taller-than-wide shape, which may indicate that shape has no diagnostic value for thyroid

**TABLE 6 |** Diagnostic performance between the proposed model and ACR TI-RADS.

| Model | Se | Sp | PPV | NPV | Ac | AUC | 95% CI |
|---|---|---|---|---|---|---|---|
| Training | | | | | | | |
| model | 58.8% | 84.6% | 34.5% | 93.7% | 81.4% | 0.833 | 0.752-0.915 |
| ACR TI-RADS | 58.8% | 86.2% | 37.0% | 93.8% | 82.9% | 0.823 | 0.750-0.882 |
| *P*-value | 0.181 | 0.424 | | | | 0.011 | |
| Validation | | | | | | | |
| model | 61.9% | 91.5% | 56.5% | 93.1% | 87.1% | 0.897 | 0.835-0.959 |
| ACR-TI-RASA | 57.1% | 93.2% | 60.0% | 92.4% | 87.8% | 0.855 | 0.785-0.909 |
| *P*-value | 0.480 | 0.680 | | | | 0.040 | |

*ACR, American College Radiology; AUC, area under the curve; CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value; Se, sensitivity; Sp, specificity; TI-RADS, Thyroid Imaging Reporting and Data System.*

nodules > 4 cm. This may reflect that ultrasonography is less accurate at assessing shape, the larger the thyroid nodule is. Conversely, our study identified ill-defined border as a predictor of malignancy, similar to a previous study (19), but this factor is assigned 0 point in ACR TI-RADS. Our model and associated nomogram may be clinically easier to use than ACR TI-RADS and more accurate for thyroid nodules > 4 cm.

The model > 4cm provides a better diagnostic efficiency than the ACR TI-RADS, kwak TI-RADS and ATA.

According to Kim's meta-analysis (20), the overall diagnostic performance of the three risk stratification systems (the ACR TI-RADS, kwak TI-RADS and ATA) of the representative society guidelines were comparable. In this study, there was no significant statistical difference in the AUC value among the ACR TI-RADS, Kwak TI-RADS and ATA in two datasets ($p > 0.05$). The AUC value were also no significant statistical difference ($p > 0.05$) among the model > 4cm, Kwak TI-RADS and ATA, similar to Shen's study (21).

The ATA guidelines cannot cover all nodules. For example, in this study, there were two hyperechoic thyroid nodules with microcalcification which were not belong to any risk stratification of ATA guidelines. The ACR convened committees developed a set of standard terms (lexicon) for ultrasound reporting and proposed a TI-RADS based on the lexicon. All nodules can be scored in ACR TI-RADS. And ACR-TIRADS showed the lowest rate of unnecessary FNAB and highest rate of malignancy in FNAB (22–24). So, the model > 4cm was mainly compared to ACR TI-RADS. The AUC of the model > 4cm was higher than of the ACR TI-RADS, whether in training dataset or in validation dataset, and the difference has statistically significant ($p < 0.05$). And the model was similar to ACR TI-RADS in sensitivity (61.9% *vs.* 57.1%, $p = 0.480$) and specificity (91.5% *vs.* 93.2%, $p = 0.680$), which were higher than Ha's study (25). This predictive model can use fewer indicators to diagnose the risk of malignant thyroid nodules larger than 4cm, and its diagnostic efficiency was consistent with that of ACR TI-RADS.

One benign thyroid nodule predicted by our model to be malignant and classified as TR5 in ACR TI-RADS illustrates the shortfalls of both systems. Ultrasonography showed that the nodule was solid, hypoechoic, and microcalcified, and that it had irregular margins and an ill-defined border. The nodule received 290 points, and malignant risk was > 0.95 according to the model. Pathology analysis of surgical samples indicated Riedel's thyroiditis, a rare inflammatory process involving thyroid and surrounding cervical tissues that is associated with systemic fibrosis (26). The nodules associated with this condition show nonspecific ultrasound features and so are often misdiagnosed (27, 28).

This study had limitations. First, it was a retrospective study, so confounding factors could not be controlled. Second, all patients underwent thyroidectomy, which may have led to selection bias. The clinical utility of this proposed model as a preoperative decision-making tool should be explored in prospective studies.

## CONCLUSION

Thyroid nodules > 4 cm merit a unique ultrasonic risk stratification, and the model proposed here may outperform ACR TI-RADS. The model should be tested in large prospective studies for its ability to guide preoperative decisions.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of Peking Union Medical College Hospital. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

## AUTHOR CONTRIBUTIONS

BZ conceived and designed the study. All the other authors collected the data. XY and SZ performed the analysis. XX prepared all the figures and tables. XX, YW, and LG were major contributors in writing the manuscript. BZ edited the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by a grant from Spatial-Temporal Mapping Analysis on Chinese Cancer Burden (2018-12M-3-003) and a grant from the National Natural Science Foundation of China (81971627).

## REFERENCES

1. Guth S, Theune U, Aberle J, Galach A, Bamberger CM. Very High Prevalence of Thyroid Nodules Detected by High Frequency (13 MHz) Ultrasound Examination. *Eur J Clin Invest* (2019) 39(8):699–706. doi: 10.1111/j.1365-2362.2009.02162.x
2. Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, et al. Revised American Thyroid Association Management Guidelines for Patients With Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* (2009) 19 (11):1167–214. doi: 10.1089/thy.2009.0110
3. Lim H, Devesa SS, Sosa JA, Check D, Kitahara CM. Trends in Thyroid Cancer Incidence and Mortality in the United States, 1974-2013. *JAMA* (2017) 317 (13):1338–48. doi: 10.1001/jama.2017.2719
4. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association Management Guidelines for Adult

Patients With Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* (2016) 26(1):1–133. doi: 10.1089/thy.2015.0020
5. Perrier ND, Brierley JD, Tuttle RM. Differentiated and Anaplastic Thyroid Carcinoma: Major Changes in the American Joint Committee on Cancer Eighth Edition Cancer Staging Manual. *CA Cancer J Clin* (2018) 68(1):55–63. doi: 10.3322/caac.21439
6. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, et al. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. *J Am Coll Radiol* (2017) 14(5):587–95. doi: 10.1016/j.jacr.2017.01.046
7. Horvath E, Majlis S, Rossi R, Franco C, Niedmann JP, Castro A, et al. An Ultrasonogram Reporting System for Thyroid Nodules Stratifying Cancer Risk for Clinical Management. *J Clin Endocrinol Metab* (2009) 94(5):1748–51. doi: 10.1210/jc.2008-1724
8. Park JY, Lee HJ, Jang HW, Kim HK, Yi JH, Lee W, et al. A Proposal for a Thyroid Imaging Reporting and Data System for Ultrasound Features of Thyroid Carcinoma. *Thyroid* (2009) 19(11):1257–64. doi: 10.1089/thy.2008.0021
9. Kwak JY, Han KH, Yoon JH, Moon HJ, Son EJ, Park SH, et al. Thyroid Imaging Reporting and Data System for US Features of Nodules: A Step in Establishing Better Stratification of Cancer Risk. *Radiology* (2011) 260 (3):892–9. doi: 10.1148/radiol.11110206
10. McCoy KL, Jabbour N, Ogilvie JB, Ohori NP, Carty SE, Yim JH. The Incidence of Cancer and Rate of False-Negative Cytology in Thyroid Nodules Greater Than or Equal to 4 cm in Size. *Surgery* (2007) 142: (6):837–44. doi: 10.1016/j.surg.2007.08.012
11. Pinchot SN, Al-Wagih H, Schaefer S, Sippel R, Chen H. Accuracy of Fine-Needle Aspiration Biopsy for Predicting Neoplasm or Carcinoma in Thyroid Nodules 4 cm or Larger. *Arch Surg* (2009) 144(7):649–55. doi: 10.1001/archsurg.2009.116
12. Godazandeh G, Kashi Z, Zargarnataj S, Fazli M, Ebadi R, Kerdabadi EH. Evaluation the Relationship Between Thyroid Nodule Size With Malignancy and Accuracy of Fine Needle Aspiration Biopsy (FNAB). *Acta Inform Med* (2016) 24(5):347–50. doi: 10.5455/aim.2016.24.347-350
13. Schrag A, Anastasiou Z, Ambler G, Noyce A, Walters K. Predicting Diagnosis of Parkinson's Disease: A Risk Algorithm Based on Primary Care Presentations. *Mov Disord* (2019) 34(4):480–6. doi: 10.1002/mds.27616
14. Pencina MJ, D'Agostino RB. Overall C as a Measure of Discrimination in Survival Analysis: Model Specific Population Value and Confidence Interval Estimation. *Stat Med* (2004) 23(13):2109–23. doi: 10.1002/sim.1802
15. Kwak JY, Kim EK, Kim MJ, Son EJ. Significance of Sonographic Characterization for Managing Subcentimeter Thyroid Nodules. *Acta Radiol* (2009) 50(8):917–23. doi: 10.1080/02841850903062724
16. Sharma A, Gabriel H, Nemcek AA, Nayar R, Du H, Nikolaidis P. Subcentimeter Thyroid Nodules: Utility of Sonographic Characterization and Ultrasound-Guided Needle Biopsy. *AJR Am J Roentgenol* (2011) 197: W1123–1128. doi: 10.2214/AJR.10.5684
17. Na DG, Baek JH, Sung JY, Kim J-H, Kim JK, Choi YJ, et al. Thyroid Imaging Reporting and Data System Risk Stratification of Thyroid Nodules: Categorization Based on Solidity and Echogenicity. *Thyroid* (2016) 26:562–72. doi: 10.1089/thy.2015.0460
18. Moon WJ, Jung SL, Lee JH, Na DG, Baek JH, Lee YH, et al. Benign and Malignant Thyroid Nodules: US Differentiation—Multicenter Retrospective Study. *Radiology* (2008) 247:762–70. doi: 10.1148/radiol.2473070944
19. Chen L, Zhang J, Meng L, Lai Y, Huang W. A New Ultrasound Nomogram for Differentiating Benign and Malignant Thyroid Nodules. *Clin Endocrinol* (2019) 90(2):351–9. doi: 10.1111/cen.13898
20. Kim PH, Suh CH, Baek JH, Chung SR, Choi YJ, Lee JH. Unnecessary Thyroid Nodule Biopsy Rates Under Four Ultrasound Risk Stratification Systems: A Systematic Review and Meta-Analysis. *Eur Radiol* (2020) 31(5):2877–85. doi: 10.1007/s00330-020-07384-6
21. Shen Y, Liu M, He J, Wu S, Chen M, Wan Y, et al. Comparison of Different Risk-Stratification Systems for the Diagnosis of Benign and Malignant Thyroid Nodules. *Front Oncol* (2019) 9:378. doi: 10.3389/fonc.2019.00378
22. Xu T, Wu Y, Wu RX, Zhang YZ, Gu JY, Ye XH, et al. Validation and Comparison of Three Newly-Released Thyroid Imaging Reporting and Data Systems for Cancer Risk Determination. *Endocrine* (2019) 64(2):299–307. doi: 10.1007/s12020-018-1817-8
23. Ha EJ, Na DG, Baek JH, Sung JY, Kim JH, Kang SY. US Fine-Needle Aspiration Biopsy for Thyroid Malignancy: Diagnostic Performance of Seven Society Guidelines Applied to 2000 Thyroid Nodules. *Radiology* (2018) 287(3):893–900. doi: 10.1148/radiol.2018171074
24. Castellana M, Castellana C, Treglia G, Giorgino F, Giovanella L, Russ G, et al. Performance of Five Ultrasound Risk Stratification Systems in Selecting Thyroid Nodules for FNA. *J Clin Endocrinol Metab* (2020) 105(5):1659–69. doi: 10.1210/clinem/dgz170
25. Ha SM, Baek JH, Na DG, Suh CH, Chung SR, Choi YJ, et al. Diagnostic Performance of Practice Guidelines for Thyroid Nodules: Thyroid Nodule Size Versus Biopsy Rates. *Radiology* (2019) 291(1):92–9. doi: 10.1148/radiol.2019181723
26. Hennessey JV. Clinical Review: Riedel's Thyroiditis: A Clinical Review. *J Clin Endocrinol Metab* (2011) 96:3031–41. doi: 10.1210/jc.2011-0617
27. Walsh J, Griffin TP, Ryan CB, Fitzgibbon J, Sheahan P, Murphy MS. A Case Report Demonstrating How the Clinical Presentation of the Diffuse Sclerosing Variant of Papillary Thyroid Carcinoma Can Mimic Benign Riedel's Thyroiditis. *Case Rep Endocrinol* (2015) 2015:686085. doi: 10.1155/2015/686085
28. Fatourechi MM, Hay ID, McIver B, Sebo TJ, Fatourechi V. Invasive Fibrous Thyroiditis (Riedel Thyroiditis): The Mayo Clinic Experience, 1976-2008. *Thyroid* (2011) 21(7):765–72. doi: 10.1089/thy.2010.0453

**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.

frontiers in Oncology

ORIGINAL RESEARCH
published: 18 October 2021
doi: 10.3389/fonc.2021.552634

# MRI-Based Machine Learning in Differentiation Between Benign and Malignant Breast Lesions

*Yanjie Zhao[1‡], Rong Chen[2‡], Ting Zhang[1], Chaoyue Chen[3], Muhetaer Muhelisa[1], Jingting Huang[1], Yan Xu[4*] and Xuelei Ma[1*†]*

*[1] Department of Biotherapy, West China Hospital and State Key Laboratory of Biotherapy, Sichuan University, Chengdu, China, [2] Department of Radiology, Guiqian International General Hospital, Guiyang, China, [3] Department of Neurosurgery, West China Hospital, Sichuan University, Chengdu, China, [4] Department of Breast and Thyroid Surgery, Daping Hospital, Army Military Medical University, Chongqing, China*

**Background:** Differential diagnosis between benign and malignant breast lesions is of crucial importance relating to follow-up treatment. Recent development in texture analysis and machine learning may lead to a new solution to this problem.

**Method:** This current study enrolled a total number of 265 patients (benign breast lesions: malignant breast lesions = 71:194) diagnosed in our hospital and received magnetic resonance imaging between January 2014 and August 2017. Patients were randomly divided into the training group and validation group (4:1), and two radiologists extracted their texture features from the contrast-enhanced T1-weighted images. We performed five different feature selection methods including Distance correlation, Gradient Boosting Decision Tree (GBDT), least absolute shrinkage and selection operator (LASSO), random forest (RF), eXtreme gradient boosting (Xgboost) and five independent classification models were built based on Linear discriminant analysis (LDA) algorithm.

**Results:** All five models showed promising results to discriminate malignant breast lesions from benign breast lesions, and the areas under the curve (AUCs) of receiver operating characteristic (ROC) were all above 0.830 in both training and validation groups. The model with a better discriminating ability was the combination of LDA + gradient boosting decision tree (GBDT). The sensitivity, specificity, AUC, and accuracy in the training group were 0.814, 0.883, 0.922, and 0.868, respectively; LDA + random forest (RF) also suggests promising results with the AUC of 0.906 in the training group.

**Conclusion:** The evidence of this study, while preliminary, suggested that a combination of MRI texture analysis and LDA algorithm could discriminate benign breast lesions from malignant breast lesions. Further multicenter researches in this field would be of great help in the validation of the result.

**Keywords: linear discriminant analysis, differential diagnosis, machine learning, breast lesion, texture analysis, MRI**

OPEN ACCESS

***Edited by:***
*Natalie Julie Serkova, University of Colorado, United States*

***Reviewed by:***
*Mario Sansone, University of Naples Federico II, Italy*
*Ashis Kumer Biswas, University of Colorado Denver, United States*

****Correspondence:***
*Xuelei Ma*
*drmaxuelei@gmail.com*
*Yan Xu*
*xy931@163.com*

***†ORCID:***
*Xuelei Ma*
*orcid.org/0000-0002-9148-5001*

*‡These authors have contributed equally to this work*

***Specialty section:***
*This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology*

***Received:*** *16 April 2020*
***Accepted:*** *24 September 2021*
***Published:*** *18 October 2021*

***Citation:***
*Zhao Y, Chen R, Zhang T, Chen C, Muhelisa M, Huang J, Xu Y and Ma X (2021) MRI-Based Machine Learning in Differentiation Between Benign and Malignant Breast Lesions. Front. Oncol. 11:552634. doi: 10.3389/fonc.2021.552634*

# INTRODUCTION

Breast cancer is increasingly acknowledged as a serious, worldwide public concern in women (1, 2). Several researchers have reported the incidence of breast cancer increases with age (3). This malignant and complex lesion, with a spectrum of its different subtypes, has resulted in various treatment modality, followed with heterogeneous responses and clinical outcomes (4). Early detection, diagnosis, and treatment are of crucial importance in improving the prognosis of the patients. In clinical practice, magnetic resonance imaging (MRI) is strongly suggested as the primary examination method of breast cancer for its non-ionizing radiation damage, high soft tissue resolution, and advantages in identifying the location and size of the lesions (5–8). However, a considerable problem with this kind of application is that there is hardly any competent method to separate the MRI patterns of benign breast lesions from the patterns of malignant breast lesions due to modest specificity, which usually leads to over- or undertreatment and unnecessary biopsy (9, 10). Although some studies have demonstrated that a mass size and non-mass enhancement with segmental or regional distribution indicate a breast papilloma with malignant lesions, in most cases, the differences were rather imperceptible (11). One major reason for this dilemma is the overlap between morphologic and kinetic characteristics between benign and malignant lesions (12, 13). Therefore, the importance has been raised to establish an efficient method to distinguish malignant breast lesions from benign breast lesions.

Recent development in texture analysis (TA), also known as radiomics, has led to a renewed solution to this demanding problem. TA, a mathematical method to quantify the heterogeneity in images by calculating the voxel intensity, has been applied to medical imaging (including computed tomography and MRI) and received satisfying results in the diagnosis of various lesions (14–16). Since TA can acquire the additional quantified information from the images that are not discernible to the human eye, more studies have been long established in the advantages of TA in facilitating differential diagnosis (17, 18). However, to date, limited researches have emerged to apply TA in differential diagnosis between benign and malignant breast lesions. The aim of this research has therefore been to adopt and evaluate machine learning algorithm combined with MRI TA in the discrimination of benign and malignant breast lesions.

# MATERIALS AND METHODS

## Patient Selection

We retrospectively searched for patients diagnosed with benign or malignant breast lesions from January 1, 2014, to August 3, 2017, in the institution's database. Eligibility criteria required patients to have 1) histopathological report of biopsy, 2) detailed electronic medical records, and 3) diagnostic MR scanning records before chemotherapy or surgical resection. Patients were excluded from the study if they had 1) existence of motion artifact on MR images and 2) received certain treatments (including surgical resection, chemotherapy, or radiotherapy) before MR scanning.

## MR Imaging Sequence Selection

For all the patients enrolled in this study, after laying the patient in a prone position, contrast-enhanced T1-weighted sequence was available, and the imaging was performed using a 1.5-T MR scanner with a bilateral, dedicated, 16-channel phased-array breast coil (Magnetom Aera, Siemens Medical Solutions, Germany). Dynamic series consisted of seven individual dynamic images with axial fat-suppressed T1-weighted imaging (T1WI), and the parameters were as follows: repetition time/echo time (TR/TE) = 4.62/1.75 ms, slice thickness = 1.5 mm, and space = 0 mm. Of these dynamic images, one was obtained before the intravenous injection, and six were obtained after the intravenous injection. Gadolinium-DTPA (Magnevist, Berlin, Germany) was injected as the contrast agent (0.15 mmol/kg bodyweight) at a rate of 2.0 ml/s, followed by a 15-ml saline flush.

## Image Processing and Lesion Segmentation

In this study, after a preliminary assessment on images, we selected contrast-enhanced T1-weighted (TIC) images for further analysis. Image series were imported from radiomics platform as Digital Imaging and Communications in Medicine (DICOM) files. We performed TA on LIFEx software (version 5.10, French Alternative Energies and Atomic Energy Commission) (19). We followed the guidelines of the image biomarker standardization initiative (IBSI) and manually segmented the two-dimensional region of interest (ROI) of the benign and malignant breast lesions depending on the imaging characteristic differences between the lesions and normal tissue (20). Two radiologists, blinded to the patients' electronic medical record and histopathological diagnosis, used LIFEx software to draw the ROI with the assistance of a senior radiologist. For the purpose of higher validity as well as reproducibility, disagreements were addressed to the senior radiologist and received further discussion. The process was performed under the software protocols, and the ROI was carefully drawn layer by layer on the axial plane along the boundary of the lesions.

## Texture Feature Extraction

No specific preprocessing was conducted in the present study. The image gray-level intensity was normalized to a scale of 1 to 64 (19). Image processing steps including interpolation, re-segmentation, and discretization were automatically performed by the radiomic-specific LIFEx software. For each patients' image series, 45 texture features recommended by the IBSI were extracted from the delineated ROIs by first order or higher order. We extracted histogram-based indices as first-order statistics. In the higher-order statistics, texture features were calculated from six matrices: gray-level co-occurrence matrix (GLCM), gray-level run length matrix (GLRLM), gray-level zone length matrix (GLZLM) (also known as gray-level size zone matrix (GLSZM)), histogram-based matrix (HISTO), neighborhood gray-level dependence matrix (NGLDM), and Shape. The main texture features for all the included patients are recorded in **Supplementary Table 1**.

## Texture Feature Selection

In **Supplementary Table 2**, we listed the definition and description of every kind of our extracted texture features. In this study, the capacity of the machine learning algorithm was limited; therefore, we could not take all the features into the analysis. Feature selection was conducted to determine the most related texture features and, additionally, avoid overfitting. Moreover, in order to find out optimal texture features, five independent selection methods were adopted, including distance correlation, gradient boosting decision tree (GBDT), least absolute shrinkage and selection operator (LASSO), random forest (RF), and extreme gradient boosting (XGBoost). These selection methods created five subsets and formed five different datasets.

## Classification

Linear discriminant analysis (LDA) is a supervised pattern recognition technique that can separate groups by searching for one or several linear combination or discriminant of predictors that maximize the ratio of between-class variance and minimize the ratio of within-class variance. Related packages were downloaded from the scikit-learn, and the models were constructed by default (21). This case study established five independent classification models on the basis of LDA algorithm: distance correlation + LDA, RF + LDA, LASSO + LDA, XGBoost + LDA, and GBDT + LDA. We randomly divided the texture features of both benign and malignant breast lesions into the training groups and validation groups (training group:validation group = 4:1). The ratio of benign and malignant breast lesions in both groups was proportional to the ratio of total benign and malignant breast lesions enrolled in our study. After training, the models were later applied with the data from the validation group, and the performance was evaluated. For each model, the randomized procedure was repeated over 100 times to appraise the robustness of the machine learning algorithm. The sensitivity, specificity, areas under the curve (AUCs), and accuracy, which represented the ability of the model to distinguish the two lesions, were later calculated in both groups. The comparison between the five different models (distance correlation + LDA, RF + LDA, LASSO + LDA, XGBoost + LDA, and GBDT + LDA) was conducted to choose a relatively suitable model with the optimal discriminative ability of benign and malignant breast lesions. The flowchart of the MRI classification of benign and malignant breast lesions in different datasets is summarized in **Figure 1**.

## Ethics Approval

Studies involving human participants were reviewed and approved by the medical ethics committee of Daping Hospital. The patients/participants' legal guardian provided written informed consent to participate in this study.

# RESULTS

## Patient Characteristics

A total number of 314 patients were primarily selected from the database of the institution after a retrospective review of their MR images and electronic medical record. Among them, 51 patients were excluded according to the exclusion criteria. In the remaining patients, 71 were histopathology-proven benign breast lesions, while 194 were histopathology-proven malignant breast lesions. **Table 1** displays the overview of baseline characteristics of the included patients. The median age and the age range for benign and malignant breast lesion groups were 31.9 (19–45) and 51.9 years (27–83), with a standard deviation of 6.0 and 10.3 years. In the benign lesion group, the majority comprised plasma cell mastitis (90.1%), while seven (9.9%) other patients were diagnosed with granulomatous mastitis. The numbers of malignant lesions for non-invasive carcinoma, invasive carcinoma, and others were 11 (5.7%), 194 (92.2%), and 4 (2.1%), respectively. All the patients underwent diagnostic MRI examination between January 2014 and August 2017. **Figure 2** illustrates the contrast-enhanced T1-weighted MR images of two examples (A: malignant breast lesions; B: benign breast lesions).

## Benign *Versus* Malignant Breast Lesions

In the present study, we conducted five different texture feature selection methods on the statistics, including distance correlation, GBDT, LASSO, RF, and XGBoost. All the selected texture features are recorded and summarized in **Supplementary Table 3**. We constructed five different models based on the LDA algorithm and five different datasets of texture features. All these models had achieved high discriminant performance, and all the AUCs in these training groups are above 0.850. The detailed performance of these models (sensitivity, specificity, AUC, and accuracy) is listed in **Table 2**. The result suggested that the GBDT + LDA model achieved a statistically higher discriminative ability among the others and received the highest AUC in the training groups as well as validation groups. The sensitivity, specificity, AUC, and accuracy of the training groups' model were 0.814, 0.779, 0.922, and 0.868, respectively; as for the validation group, the results were 0.883, 0.892, 0.911, and 0.868, respectively. The RF + LDA model also showed optimal discriminant ability with the AUC > 0.9 in the training group.

**Figure 3** directly illustrates the discriminative function of the GBDT + LDA models. Note that there was little overlap between the distribution of benign breast lesion groups (triangles) and malignant breast lesion groups (circles) and the distribution of group centroids (squares). The promising results indicated that GBDT + LDA models provided a qualitative separation between benign and malignant breast lesions. **Figure 4** reveals an example of 100 independent training cycles in the GBDT + LDA models, in which the findings are shown in the distribution of the direct LDA function (A: benign breast lesions; B: malignant breast lesions).

# DISCUSSION

Prior studies had stressed the importance of distinguishing benign breast lesions from malignant breast lesions in view of redundant invasive examination and the significant differences in

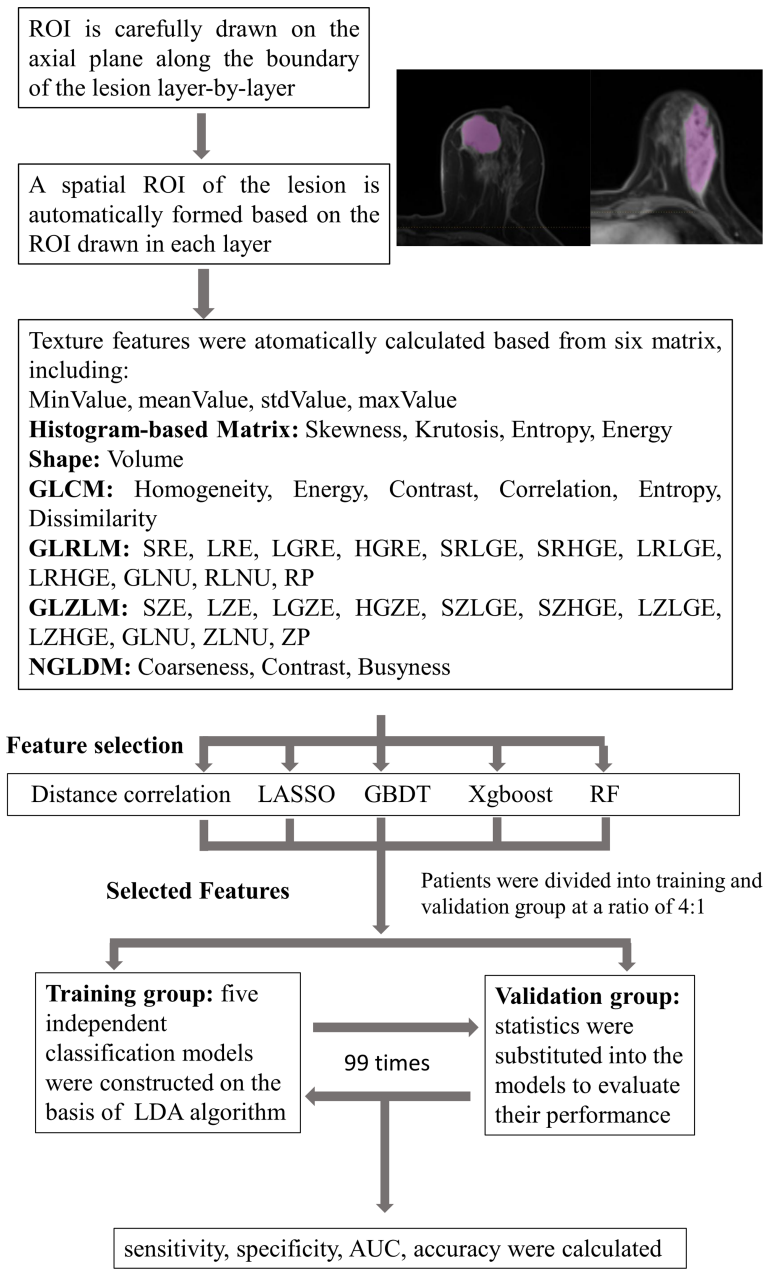



**FIGURE 1** | Flowchart of the MRI classification process by different selection methods. ROI, region of interest; GLCM, gray-level co-occurrence matrix; GLRLM, gray-level run length matrix; GLZLM, gray-level zone length matrix; NGLDM, neighborhood gray-level dependence matrix; LASSO, least absolute shrinkage and selection operator; GBDT, gradient boosting decision tree; RF, random forest; LDA, linear discriminant analysis; AUC, area under the receiver operating characteristic curve.

their treatment strategies and prognostic results (22). Traditionally, MR scanning was recommended as a sensitive modality to detect and diagnose breast lesions, together with mammography, ultrasound, and image-guided needle biopsy. However, many uncertainties still existed about the accuracy of this modality for the characteristics of benign and malignant breast lesions resemblance on the conventional MR images (12). In the current study, we examined the discriminative ability of MRI-based TA by combining five different extracted MRI-based texture feature datasets with a supervised pattern recognition technique to establish five LDA-based models. The results indicated that all these models presented good discriminant ability. Moreover, a combination of GBDT selection method for TA and LDA algorithm for classification exhibited a better performance by statistics among the others. These findings highlight the potential usefulness of machine learning to complete the separation of benign and malignant breast lesions in clinical practice.

**TABLE 1** | Baseline characteristics of the 93 patients included in the analysis.

| Characteristics | Benign lesions n = 71 (%) | Malignant lesions n = 194 (%) |
|---|---|---|
| Mean age (years; SD) | 31.9 ± 6 | 51.9 ± 10.3 |
| Location | | |
| Left | 23 (32.4%) | 89 (45.9%) |
| Right | 43 (60.6%) | 105 (54.1%) |
| Bilateral | 5 (7.0%) | 0 (0.0%) |
| Pathology | | |
| PCM | 64 (90.1%) | |
| GM | 7 (9.9%) | |
| Non-invasive carcinoma | | 11 (5.7%) |
| Invasive carcinoma | | 194 (92.2%) |
| Others[†] | | 4 (2.1%) |

*PCM, plasma cell mastitis; GM, granulomatous mastitis.*

[†]*"Others" refers to carcinoma with medullary features, tubular carcinoma, invasive cribriform carcinoma, and invasive papillary carcinoma each in the present study.*

TA was a statistical method focused on the analytic techniques and the description of image texture, which was formerly defined as the repeating patterns of local variations in gray-level intensities (23). Advances in TA had enabled recent researches to visualize spatial histologic heterogeneity, capturing image patterns that were unrecognizable to human eyes. A considerable amount of literature has been published applying TA to discriminate breast lesions. A previous study had tried to adopt a combination of texture features (GLCM entropy, GLCM Sum Average, and GLCM Homogeneity) and morphology features to diagnose benign and malignant lesions from a 2D slice of 3D images (24). This idea was further extended by researchers, and a study was conducted to investigate the utility of 3D breast lesion characterization (GLCM) by Student's t-test in distinguishing benign and malignant lesions (25). Compared with these studies, we enrolled a larger group of texture features generated from different matrixes. Moreover, we combined machine learning to select significant texture features, and the results of our models showed a reasonably high sensitivity, specificity, AUC, and accuracy. In recent years, other studies had tried to relate breast lesions TA of contrast-enhanced MR images to the underlying lesion subtypes and received satisfying results (26).

The past decade had seen the rapid development of machine learning applied to MR images in different fields (27–29). After evaluation of the performance of the selection methods and classification algorithms, predictive models were created for tumor grading, diagnosis of interest, and clinical outcome. The association between molecular expression (Ki67 and HER2) and contrast-enhanced MRI features were also observed in some studies (30, 31). Other researchers investigated machine learning in different radiological techniques such as mammograms and ultrasound (32, 33). Previous research had indicated that integration of 10-fold cross-validation method and machine learning into the interpretation of MR images can help to make decisive rules to manage suspicious breast lesions (34). Another study also achieved promising diagnostic results in discriminating breast lesions with deep learning method using ResNet50 (35). In our study, we built five different classification models based on LDA algorithm, and a relatively optimal machine learning model with a combination of LDA + GBDT demonstrated a non-inferior accuracy of 0.868 and 0.892 in the training and validation groups, respectively. The current findings to evaluate the performance of five different feature selection methods combined with LDA algorithms add to a growing body of literature on computer-aided diagnosis of breast lesions.

Many studies in the medical image field of machine learning algorithm had noticed the influence on the diagnostic performance caused by the adoption of different texture features. Recently, the IBSI had standardized the extraction of image biomarkers from imaging

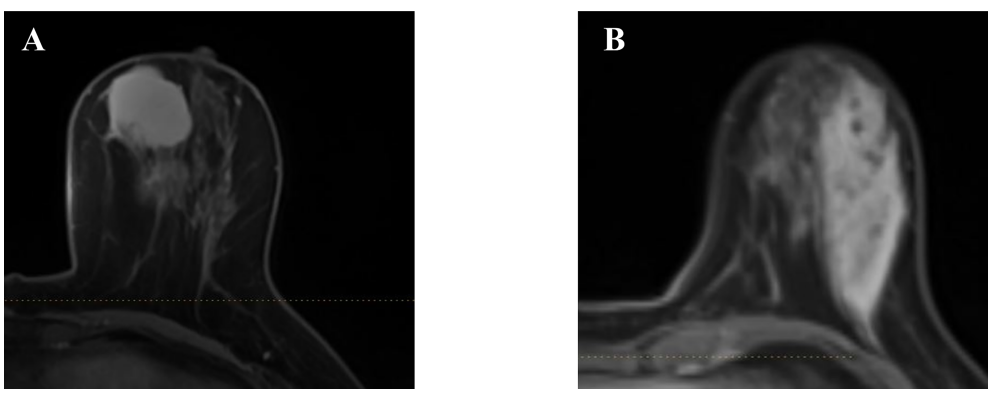



**FIGURE 2** | Two examples of the axial plane of contrast-enhanced T1-weighted MR images. **(A)** Malignant breast lesions. **(B)** Benign breast lesions.

**TABLE 2** | The performance of five different models.

| | Training | | | | Validation | | | |
|---|---|---|---|---|---|---|---|---|
| | **Sensitivity** | **Specificity** | **Accuracy** | **AUC** | **Sensitivity** | **Specificity** | **Accuracy** | **AUC** |
| Distance Correlation | 0.655 | 0.801 | 0.777 | 0.859 | 0.635 | 0.811 | 0.787 | 0.835 |
| RF | 0.753 | 0.863 | 0.839 | 0.906 | 0.675 | 0.865 | 0.825 | 0.881 |
| LASSO | 0.733 | 0.839 | 0.818 | 0.863 | 0.745 | 0.856 | 0.836 | 0.869 |
| XGBoost | 0.702 | 0.813 | 0.795 | 0.899 | 0.701 | 0.837 | 0.815 | 0.899 |
| **GBDT** | **0.814** | **0.883** | **0.868** | **0.922** | **0.779** | **0.892** | **0.868** | **0.911** |

*LASSO, least absolute shrinkage and selection operator; GBDT, gradient boosting decision tree; RF, random forest; LDA, linear discriminant analysis; AUC, area under the receiver operating characteristic curve.*
*We highlighted a relatively better performed model in bold values.*

to present a high-throughput quantitative image analysis. All the texture features included in this study were recommended in the IBSI feature reference values and added to the quality of our research (20). Previous researches had discussed various kinds of selection methods including Student's t-test, Mann–Whitney U test, ReliefF algorithm (36, 37). Based on the results, the interference of the selection method cannot be ruled out. Compared with previous studies, we extracted a relatively large number of parameters from different matrices, which increased the possibility to select the optimal features. Moreover, in order to establish the optimal classifier, we chose five different feature selection methods (distance correlation, RF, LASSO, XGBoost, and GBDT) with which the performances were later evaluated. Overall, the result of this study indicates that all the models showed good performance and that the GBDT + LDA model achieved a better performance for benign breast lesions from malignant breast lesions with the highest AUC of 0.922 in the training group. GBDT was proposed as a tree-based algorithm based on a greedy strategy (called gradient boosting) that evaluates the importance of a texture feature through the time it used as branching point for the tree. However, the results of this study must be interpreted with caution because no significant differences were observed between the performance of all the models, and the variance in AUC may result from the statistical group. Therefore, the results can only be suggested as a hypothesis generation and required verification from future, larger studies.

Several limitations to this pilot study need to be acknowledged. First, this study is single-centered. The selection bias for patients is unavoidable and may have influenced the analysis. Second, the sample size was relatively small, and a greater size of the sample is expected for further study to validate the results. Third, the results obtained in this study did not receive external validation in other datasets, and the diagnostic ability of this model may be influenced due to different MR scanners and image processing procedure. Fourth, only texture features from contrast-enhanced T1-weighted

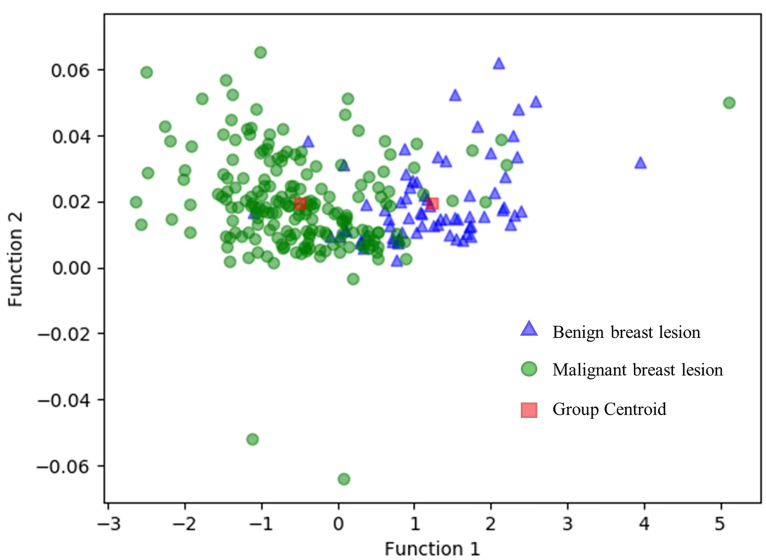



**FIGURE 3** | Discriminative function of the GBDT + LDA models. Distribution of the benign and malignant breast lesions that originated from multiple dimensions were reduced and reflected to a two-dimension plane. Little overlap was observed between the distribution of benign breast lesion groups (triangles) and malignant breast lesion groups (circles) and the distribution of group centroids (squares). It suggests a qualitative separation between benign and malignant breast lesions. GBDT, gradient boosting decision tree; LDA, linear discriminant analysis.

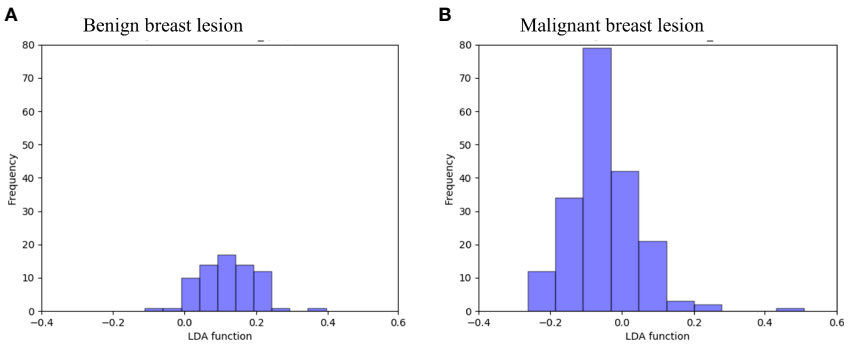



**FIGURE 4** | Examples of the 100 data analysis training cycles. Distribution of the direct LDA function determining benign and malignant breast lesions illustrates a promising performance of the GBDT + LDA models. GBDT, gradient boosting decision tree; LDA, linear discriminant analysis. **(A)** Distribution of the LDA function determined for the benign breast lesions for one cycle; **(B)** distribution of the LDA function determined for the malignant breast lesions for one cycle.

(TIC) images were introduced in this study. Further studies are required to explore the classifier adapted with texture features from other sequences.

## CONCLUSION

The study was undertaken to design optimal classification model using texture features combined with machine learning algorithm and evaluate its sensitivity, specificity, AUC, and accuracy. We used five selection methods and established five discriminative models, and their performances were evaluated. In general, therefore, it seems that texture features have potential to be utilized in discriminating benign breast lesions from malignant breast lesions. More broadly, future multicenter researches with more patients in this field would be of great help to validate this preliminary result.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in https://scikit-learn.org/stable/supervised_learning.html#supervised-learning.

## ETHICS STATEMENT

Studies involving human participants were reviewed and approved by medical ethics committee of Daping Hospital. The patients/participants' legal guardian provided written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

YZ extracted the texture features from MR images and wrote the manuscript. RC contributed equally to this paper and extracted the texture features from MR images. CC, MM, TZ, and JH conducted all the analysis in this paper. XM is the first corresponding author of this paper and proposed the topic and methods of this paper. YX is the second corresponding author of this paper and provides the detailed information as well as the MR images of the patients. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the National Natural Science Foundation of China under Grant No. 81472482; and the Clinical Technology Innovation and Cultivation Project of Army Military Medical University of China under Grant No. CX2019LC120.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.552634/full#supplementary-material

## REFERENCES

1. Torre LA, Siegel RL, Ward EM, Jemal A. Global Cancer Incidence and Mortality Rates and Trends–an Update. *Cancer Epidemiology Biomarkers Prevention: Publ Am Assoc Cancer Research Cosponsored by Am Soc Prev Oncol* (2016) 25(1):16–27. doi: 10.1158/1055-9965.Epi-15-0578
2. DeSantis CE, Ma J, Goding Sauer A, Newman LA, Jemal A. Breast Cancer Statistics, 2017, Racial Disparity in Mortality by State. *CA: Cancer J Clin* (2017) 67(6):439–48. doi: 10.3322/caac.21412
3. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2018. *CA: Cancer J Clin* (2018) 68(1):7–30. doi: 10.3322/caac.21442
4. Diaby V, Tawk R, Sanogo V, Xiao H, Montero AJ. A Review of Systematic Reviews of the Cost-Effectiveness of Hormone Therapy, Chemotherapy, and

Targeted Therapy for Breast Cancer. *Breast Cancer Res Treat* (2015) 151 (1):27–40. doi: 10.1007/s10549-015-3383-6
5. Orel SG, Schnall MD. MR Imaging of the Breast for the Detection, Diagnosis, and Staging of Breast Cancer. *Radiology* (2001) 220(1):13–30. doi: 10.1148/radiology.220.1.r01jl3113
6. Marino MA, Helbich T, Baltzer P, Pinker-Domenig K. Multiparametric MRI of the Breast: A Review. *J Magnetic Resonance Imaging: JMRI* (2018) 47 (2):301–15. doi: 10.1002/jmri.25790
7. Kuhl CK, Strobel K, Bieling H, Leutner C, Schild HH, Schrading S. Supplemental Breast MR Imaging Screening of Women With Average Risk of Breast Cancer. *Radiology* (2017) 283(2):361–70. doi: 10.1148/radiol.2016161444
8. Malur S, Wurdinger S, Moritz A, Michels W, Schneider A. Comparison of Written Reports of Mammography, Sonography and Magnetic Resonance Mammography for Preoperative Evaluation of Breast Lesions, With Special Emphasis on Magnetic Resonance Mammography. *Breast Cancer research: BCR* (2001) 3(1):55–60. doi: 10.1186/bcr271
9. Huang W, Tudorica LA, Li X, Thakur SB, Chen Y, Morris EA, et al. Discrimination of Benign and Malignant Breast Lesions by Using Shutter-Speed Dynamic Contrast-Enhanced MR Imaging. *Radiology* (2011) 261 (2):394–403. doi: 10.1148/radiol.11102413
10. Furman-Haran E, Schechtman E, Kelcz F, Kirshenbaum K, Degani H. Magnetic Resonance Imaging Reveals Functional Diversity of the Vasculature in Benign and Malignant Breast Lesions. *Cancer* (2005) 104 (4):708–18. doi: 10.1002/cncr.21225
11. Wang LJ, Wu P, Li XX, Luo R, Wang DB, Guan WB. Magnetic Resonance Imaging Features for Differentiating Breast Papilloma With High-Risk or Malignant Lesions From Benign Papilloma: A Retrospective Study on 158 Patients. *World J Surg Oncol* (2018) 16(1):234. doi: 10.1186/s12957-018-1537-9
12. Heller SL, Moy L. Imaging Features and Management of High-Risk Lesions on Contrast-Enhanced Dynamic Breast MRI. *AJR Am J Roentgenology* (2012) 198 (2):249–55. doi: 10.2214/ajr.11.7610
13. Kuhl CK, Klaschik S, Mielcarek P, Gieseke J, Wardelmann E, Schild HH. Do T2-Weighted Pulse Sequences Help With the Differential Diagnosis of Enhancing Lesions in Dynamic Breast MRI? *J Magnetic Resonance Imaging: JMRI* (1999) 9(2):187–96. doi: 10.1002/(sici)1522-2586(199902)9:2<187::aid-jmri6>3.0.co;2-2
14. Kim JH, Ko ES, Lim Y, Lee KS, Han BK, Ko EY, et al. Breast Cancer Heterogeneity: MR Imaging Texture Analysis and Survival Outcomes. *Radiology* (2017) 282(3):665–75. doi: 10.1148/radiol.2016160261
15. Nakajo M, Jinguji M, Nakabeppu Y, Nakajo M, Higashi R, Fukukura Y, et al. Texture Analysis of (18)F-FDG PET/CT to Predict Tumour Response and Prognosis of Patients With Esophageal Cancer Treated by Chemoradiotherapy. *Eur J Nucl Med Mol Imaging* (2017) 44(2):206–14. doi: 10.1007/s00259-016-3506-2
16. Han X, Yang J, Luo J, Chen P, Zhang Z, Alu A, et al. Application of CT-Based Radiomics in Discriminating Pancreatic Cystadenomas From Pancreatic Neuroendocrine Tumors Using Machine Learning Methods. *Front Oncol* (2021) 11:606677. doi: 10.3389/fonc.2021.606677
17. Guo C, Zhuge X, Wang Q, Xiao W, Wang Z, Wang Z, et al. The Differentiation of Pancreatic Neuroendocrine Carcinoma From Pancreatic Ductal Adenocarcinoma: The Values of CT Imaging Features and Texture Analysis. *Cancer Imaging: Off Publ Int Cancer Imaging Soc* (2018) 18(1):37. doi: 10.1186/s40644-018-0170-8
18. Xu H, Deng L, Tian R, Ma X. Editorial: Novel Methods for Oncologic Imaging Analysis: Radiomics, Machine Learning, and Artificial Intelligence. *Front Oncol* (2021) 11:628310. doi: 10.3389/fonc.2021.628310
19. Nioche C, Orlhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, et al. Lifex: A Freeware for Radiomic Feature Calculation in Multimodality Imaging to Accelerate Advances in the Characterization of Tumor Heterogeneity. *Cancer Res* (2018) 78(16):4786–9. doi: 10.1158/0008-5472.Can-18-0125
20. Zwanenburg A, Vallières M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-Based Phenotyping. *Radiology* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145
21. *Scikit-Learn*. Available at: https://scikit-learn.org/stable/.
22. Medeiros LR, Duarte CS, Rosa DD, Edelweiss MI, Edelweiss M, Silva FR, et al. Accuracy of Magnetic Resonance in Suspicious Breast Lesions: A Systematic Quantitative Review and Meta-Analysis. *Breast Cancer Res Treat* (2011) 126 (2):273–85. doi: 10.1007/s10549-010-1326-9
23. Castellano G, Bonilha L, Li LM, Cendes F. Texture Analysis of Medical Images. *Clin Radiol* (2004) 59:1061–9. doi: 10.1016/j.crad.2004.07.008
24. Nie K, Chen JH, Yu HJ, Chu Y, Nalcioglu O, Su MY. Quantitative Analysis of Lesion Morphology and Texture Features for Diagnostic Prediction in Breast MRI. *Acad Radiol* (2008) 15(12):1513–25. doi: 10.1016/j.acra.2008.06.005
25. Chen W, Giger ML, Li H, Bick U, Newstead GM. Volumetric Texture Analysis of Breast Lesions on Contrast-Enhanced Magnetic Resonance Images. *Magnetic Resonance Med* (2007) 58(3):562–71. doi: 10.1002/mrm.21347
26. Waugh SA, Purdie CA, Jordan LB, Vinnicombe S, Lerski RA, Martin P, et al. Magnetic Resonance Imaging Texture Analysis Classification of Primary Breast Cancer. *Eur Radiol* (2016) 26(2):322–30. doi: 10.1007/s00330-015-3845-6
27. Kocak B, Durmaz ES, Kadioglu P, Polat Korkmaz O, Comunoglu N, Tanriover N, et al. Predicting Response to Somatostatin Analogues in Acromegaly: Machine Learning-Based High-Dimensional Quantitative Texture Analysis on T2-Weighted MRI. *Eur Radiol* (2019) 29(6):2731–9. doi: 10.1007/s00330-018-5876-2
28. Liao X, Cai B, Tian B, Luo Y, Song W, Li Y. Machine-Learning Based Radiogenomics Analysis of MRI Features and Metagenes in Glioblastoma Multiforme Patients With Different Survival Time. *J Cell Mol Med* (2019) 23 (6):4375–85. doi: 10.1111/jcmm.14328
29. Ramkumar S, Ranjbar S, Ning S, Lal D, Zwart CM, Wood CP, et al. MRI-Based Texture Analysis to Differentiate Sinonasal Squamous Cell Carcinoma From Inverted Papilloma. *AJNR Am J Neuroradiology* (2017) 38(5):1019–25. doi: 10.3174/ajnr.A5106
30. Ma W, Ji Y, Qi L, Guo X, Jian X, Liu P. Breast Cancer Ki67 Expression Prediction by DCE-MRI Radiomics Features. *Clin Radiol* (2018) 73 (10):909.e901–909.e905. doi: 10.1016/j.crad.2018.05.027
31. Bitencourt AGV, Gibbs P, Rossi Saccarelli C, Daimiel I, Lo Gullo R, Fox MJ, et al. MRI-Based Machine Learning Radiomics can Predict HER2 Expression Level and Pathologic Response After Neoadjuvant Therapy in HER2 Overexpressing Breast Cancer. *EBioMedicine* (2020) 61:103042. doi: 10.1016/j.ebiom.2020.103042
32. Ribli D, Horváth A, Unger Z, Pollner P, Csabai I. Detecting and Classifying Lesions in Mammograms With Deep Learning. *Sci Rep* (2018) 8(1):4165. doi: 10.1038/s41598-018-22437-z
33. Tagliafico AS, Valdora F, Mariscotti G, Durando M, Nori J, La Forgia D, et al. An Exploratory Radiomics Analysis on Digital Breast Tomosynthesis in Women With Mammographically Negative Dense Breasts. *Breast (Edinburgh Scotland)* (2018) 40:92–6. doi: 10.1016/j.breast.2018.04.016
34. Ellmann S, Wenkel E, Dietzel M, Bielowski C, Vesal S, Maier A, et al. Implementation of Machine Learning Into Clinical Breast MRI: Potential for Objective and Accurate Decision-Making in Suspicious Breast Masses. *PloS One* (2020) 15(1):e0228446. doi: 10.1371/journal.pone.0228446
35. Zhou J, Zhang Y, Chang KT, Lee KE, Wang O, Li J, et al. Diagnosis of Benign and Malignant Breast Lesions on DCE-MRI by Using Radiomics and Deep Learning With Consideration of Peritumor Tissue. *J Magnetic Resonance Imaging: JMRI* (2020) 51(3):798–809. doi: 10.1002/jmri.26981
36. Jiang Z, Song L, Lu H, Yin J. The Potential Use of DCE-MRI Texture Analysis to Predict HER2 2+ Status. *Front Oncol* (2019) 9:242. doi: 10.3389/fonc.2019.00242
37. Kocak B, Durmaz ES, Ates E, Sel I, Turgut Gunes S, Kaya OK, et al. Radiogenomics of Lower-Grade Gliomas: Machine Learning-Based MRI Texture Analysis for Predicting 1p/19q Codeletion Status. *Eur Radiol* (2020) 30(2):877–86. doi: 10.1007/s00330-019-06492-2

**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.

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

ORIGINAL RESEARCH
published: 21 January 2022
doi: 10.3389/fonc.2021.681489

# A Computed Tomography Nomogram for Assessing the Malignancy Risk of Focal Liver Lesions in Patients With Cirrhosis: A Preliminary Study

*Hongzhen Wu[1,2], Zihua Wang[3], Yingying Liang[2], Caihong Tan[2], Xinhua Wei[2], Wanli Zhang[2], Ruimeng Yang[2], Lei Mo[2] and Xinqing Jiang[1,2]**

[1] Department of Radiology, The First Affiliated Hospital of Jinan University, Guangzhou, China, [2] Department of Radiology, Guangzhou First People's Hospital, School of Medicine, South China University of Technology, Guangzhou, China, [3] Department of Radiology, Foshan Hospital of Traditional Chinese Medicine, Foshan, China

OPEN ACCESS

**Edited by:**
Damiano Caruso,
Sapienza University of Rome, Italy

**Reviewed by:**
Lian-Ming Wu,
Shanghai JiaoTong University, China
Zhongxiang Ding,
Zhejiang University, China

***Correspondence:**
Xinqing Jiang
gzcmmcjxq@163.com;
eyjiangxq@scut.edu.cn

**Specialty section:**
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

**Received:** 16 March 2021
**Accepted:** 27 December 2021
**Published:** 21 January 2022

**Citation:**
Wu H, Wang Z, Liang Y, Tan C, Wei X, Zhang W, Yang R, Mo L and Jiang X (2022) A Computed Tomography Nomogram for Assessing the Malignancy Risk of Focal Liver Lesions in Patients With Cirrhosis: A Preliminary Study. Front. Oncol. 11:681489. doi: 10.3389/fonc.2021.681489

**Purpose:** The detection and characterization of focal liver lesions (FLLs) in patients with cirrhosis is challenging. Accurate information about FLLs is key to their management, which can range from conservative methods to surgical excision. We sought to develop a nomogram that incorporates clinical risk factors, blood indicators, and enhanced computed tomography (CT) imaging findings to predict the nature of FLLs in cirrhotic livers.

**Method:** A total of 348 surgically confirmed FLLs were included. CT findings and clinical data were assessed. All factors with $P < 0.05$ in univariate analysis were included in multivariate analysis. ROC analysis was performed, and a nomogram was constructed based on the multivariate logistic regression analysis results.

**Results:** The FLLs were either benign (n = 79) or malignant (n = 269). Logistic regression evaluated independent factors that positively affected malignancy. AFP (OR = 10.547), arterial phase hyperenhancement (APHE) (OR = 740.876), washout (OR = 0.028), satellite lesions (OR = 15.164), ascites (OR = 156.241), and nodule-in-nodule architecture (OR =27.401) were independent predictors of malignancy. The combined predictors had excellent performance in differentiating benign and malignant lesions, with an AUC of 0.959, a sensitivity of 95.24%, and a specificity of 87.5% in the training cohort and AUC of 0.981, sensitivity of 94.74%, and specificity of 93.33% in the test cohort. The C-index was 96.80%, and calibration curves showed good agreement between the nomogram predictions and the actual data.

**Conclusions:** The nomogram showed excellent discrimination and calibration for malignancy risk prediction, and it may aid in making FLLs treatment decisions.

**Keywords: focal liver lesion, liver, computed tomography, nomogram, neoplasm**

## INTRODUCTION

It is challenging for abdominal radiologists to detect and characterize focal liver lesions (FLLs) in patients with cirrhosis (1–3). Cirrhosis is a major risk factor for hepatocellular carcinoma (HCC) (4). The diagnosis of benign and malignant FLLs in cirrhotic individuals is important. However, in cirrhotic livers, these lesions may lack typical imaging features (1). Therefore, the final diagnosis may need to be verified by tissue sampling (1, 5). Accurate descriptions of FLLs guide their management, which ranges from conservative treatment to surgical excision (6). Proper identification can prevent unnecessary biopsy and allow the appropriate treatment to be selected (5, 7).

Computed tomography (CT) is commonly used to diagnose and manage patients with chronic liver disease (8–10). Dynamic CT has high sensitivity and specificity for diagnosing FLLs (11). The wash-in and washout of contrast agents can assist distinguishing HCC from other FLLs (12). Clinical and laboratorial risk factors, such as age and sex, are also helpful. However, many cases require further imaging or histopathological examination to confirm the diagnosis (5, 10). Therefore, it is helpful to develop a scoring system combining clinical information and imaging results to evaluate the malignancy of FLLs.

A nomogram is a graphical statistical tool that combines variables into a continuous scoring system to calculate precise risk probabilities for specific individual outcomes (13–15). This instrument is an important tool in modern medical decision-making in specialties of oncology (16–21), such as differentiating focal nodular hyperplasia (15) or hepatocellular adenoma (16) from HCC in noncirrhotic patients, prediction of microvascular invasion and liver failure after hepatectomy in patients with in HCC (22–25) and efficacy evaluation of intrahepatic cholangiocarcinoma after hepatectomy (26). CT-based nomograms may be used to predict the nature of FLLs, aiding clinicians in selecting the best management plan. We sought to develop a nomogram that incorporates clinical risk factors, blood indicators, and enhanced CT imaging findings to predict the nature of FLLs in cirrhotic livers.

## MATERIAL AND METHODS

### Patients

Our institutional review board approved this retrospective single-center study with a waiver for the requirement to obtain informed consent. The subjects had to have liver cirrhosis of any etiology, surgically confirmed after CT scans between January 2017 and December 2020.

The inclusion criteria were as follows: (a) serological markers, such as serum total bilirubin, plasma albumin, prothrombin time, blood platelet count and α-fetoprotein (AFP), measured simultaneously before surgery; (b) confirmation by pathology; (c) no history of preoperative anticancer therapy, including transcatheter arterial chemoembolization (TACE), percutaneous radiofrequency ablation (PRFA), or percutaneous ethanol injection (PEI); and (d) multidetector CT imaging. The exclusion criteria were as follows: (a) no pathological confirmation; (b) undergoing treatment before imaging or surgery; (c) incomplete serological markers before surgery; and (d) poor image quality. For patients with multiple lesions, we analyzed the largest lesion. A total of 348 surgically confirmed FLLs in 348 patients were included in the study. A total of 295 patients diagnosed from 2017 to June 2020 were included in a training cohort, and 53 patients diagnosed from July 2020 to December 2020 were included in a test cohort.

Age, sex, and basic patient information were collected. Routine examinations included serum total bilirubin, total plasma protein, prothrombin time, blood platelet counts, tumor marker AFP levels, and hepatitis (B and C) results. The Child-Pugh classification was determined for each patient based on the above variables. AFP > 20.0 ng/ml was the threshold for positivity (27).

### CT Technique

The area from the diaphragm to the pubic symphysis was examined on a multidetector CT scanner (Aquilion 16, Toshiba Medical Systems Corporation; Tochigi ken, Japan; Brilliance 64, Philips, Netherlands; Aquilion ONE TSX-301A, Toshiba Medical Systems Corporation; Tochigi ken, Japan) with plain and dynamic contrast-enhanced scans as follows: tube voltage of 120 kVp, tube current of 200 mA, slice thickness of 5 mm, and rotation time of 0.5 seconds. The helical pitch was 0.9, the field of view was 35 to 40 cm, the matrix was 512 × 512, and a standard reconstruction algorithm was used. After plain CT scan, the patients received 80-100 mL of contrast agent (iodipamide, 370 mg I/mL, Bracco) at a rate of 3.5-4.0 mL/s, followed by 20 mL of saline solution through the elbow vein using a power injector. Scans in the arterial phase (AP, 35 seconds), portal venous phase (PVP, 70 seconds), and equilibrium phase (EP, 3 minutes) were obtained.

### CT Imaging Analysis

Two radiologists(7 and 13 years of abdominal diagnostic experience) independently evaluated the CT images from the Picture Archiving and Communication System. They knew the purpose and design of the study, but they were unaware of the patient demographics, clinical history, clinical reports, and reference criteria. For each lesion, the readers assessed the presence of imaging features mainly based on Liver Imaging Reporting and Data System (LI-RADS) version 2018 (28, 29), some features reported in the literature (30, 31) or commonly used in reports. The included features were as follows: tumor size (maximum (Dmax) cross-section diameter), non-rim arterial phase hyperenhancement (APHE), non-peripheral washout in the PVP (washout), enhancing capsule in either the portal venous or delay phases (appearance detected as enhancing rim), blood product in mass (bleeding within or around the lesion without surgery, trauma or intervention), fat (excess fat in the whole or part of the mass relative to the background liver), necrosis (areas within the tumor without obvious enhancement), infiltrative appearance (invasion), mural nodules (peripheral nodules within the lesion attached to the tumor wall), satellite lesions (nodules in the surrounding parenchyma resemble the main lesion), halo enhancement(solar

enhancement in parenchyma around the lesion), peritumoral enhance(rim arterial phase hyperenhancement), vein tumor thrombus (VTT, definite enhancement of soft tissue in the portal vein), delayed enhancement(progressive enhancement in the center of the lesion), internal artery (small vessels in the arterial stage), non-enhancing "capsule" (capsule appearance not detected as enhancing rim), mosaic architecture(random distribution of internal nodules or compartments, often with different radiographic features), nodule-in-nodule architecture (the internal nodules were small and larger than the external nodules, with different imaging features), corona enhancement, lymph node enlargement (short diameter >10mm) and ascites (perihepatic water density).

## Statistical Analysis

Statistical analysis was performed using IBM SPSS Statistics version 25 (SPSS Inc., Chicago, IL, USA) and R version 3.3.4 (www.R-project.org). The Mann-Whitney U test was used for continuous variables, and the $\chi^2$ or Fisher exact test was used for categorical variables. To test the consistency of the two readers, a Kendall correlation coefficient was used to measure the index, and a kappa consistency test was used for the counting index. The k values were considered poor for a k of 0.01 to 0.20; fair for a k of 0.21 to 0.40; moderate for a k of 0.41 to 0.60; good for a k of 0.61 to 0.80; and excellent for a k of 0.81 to 1.00. Univariate analysis was used to compare the differences in clinical factors (all independent clinical risk factors, blood markers, and CT findings) between the two cohorts, and multivariate logistic regression analysis was used to establish a clinical factor model, with significant variables in the univariate analysis as the input. The odds ratio (OR) was used as a relative risk estimate for each risk factor and is presented with its corresponding 95% confidence interval (CI). After establishing the combined predictor, receiver operating characteristic (ROC) analysis was performed to calculate the area under the curve (AUC), sensitivity and specificity. A nomogram was developed by scaling the regression coefficients into a multiple logistic regression of 0-100 points. Important malignancy factors in the multivariate analysis were included in the nomograms. The total score is the sum of the points for each independent variable and is converted to the prediction probability. The nomogram's performance was measured by the consistency index (C-index) and calibrated with 1,000 bootstrap samples to reduce overfitting bias (32). A calibration curve was plotted to evaluate the actual observations *vs* the nomogram predictions of the benignity or malignancy of lesions. Decision curve analysis (DCA) showed that the net clinical benefit was correlated with the diagnostic procedure including the established nomogram (33). All statistical tests were two-sided, and a P-value < 0.05 was considered statistically significant.

# RESULTS

## Histopathologic Results

A total of 348 FLLs were identified in the patients, including HCC (n = 196), cholangiocarcinoma (n = 43), metastasis (n = 24), neuroendocrine tumor (n = 2), undifferentiated sarcoma (n = 1), perivascular epithelioid cell tumor (n = 1), papillary neoplasm (n = 1), combined HCC- cholangiocarcinoma (n = 1), cyst (n = 32), hemangioma (n = 14), abscess (n = 10), focal nodular hyperplasia (n = 10), regenerative nodules (n = 4), adenoma (n = 4), parasitization (n = 2), extramedullary hemopoiesis (n = 2), and tuberculosis (n = 1). These FLLs were either benign (n = 79) or malignant (n = 269), as shown in **Figure 1**, and the baseline characteristics of the patients are shown in **Table 1** and **Table 2**.

## Interobserver Agreement

The indexes of lesion size (Dmax) had good consistency among observers (P > 0.05). The consistency value of the counting indexes among observers was greater than 0.75, indicating good consistency between observers.

## Univariate Analysis of Independent Predictors of Malignancy

For categorical variables, sex ($\chi^2$ = 19.836, P < 0.001), pathogenesis ($\chi^2$ = 66.657, P < 0.001), AFP ($\chi^2$ = 61.136, P < 0.001), APHE ($\chi^2$ = 87.075, P < 0.001), washout ($\chi^2$ = 26.807, P < 0.001), enhancing capsule ($\chi^2$ = 5.798, P = 0.016), blood product in mass ($\chi$2 = 9.882, P = 0.002), necrosis ($\chi^2$ = 27.228, P < 0.001), infiltrative appearance ($\chi^2$ = 9.529, P=0.002), mural nodules ($\chi^2$ = 5.882, P = 0.015), satellite lesions ($\chi^2$ = 9.529, P = 0.002), VTT ($\chi^2$ = 6.264, P = 0.012), internal artery ($\chi^2$ = 64.521, P < 0.001), ascites ($\chi^2$ = 5.663, P=0.017), non-enhancing capsule ($\chi^2$ = 44.611, P < 0.001), mosaic architecture ($\chi^2$ = 57.635, P < 0.001), and nodule-in-nodule architecture ($\chi^2$ = 62.087, P < 0.001) were significantly different between the two cohorts (**Table 1**). There

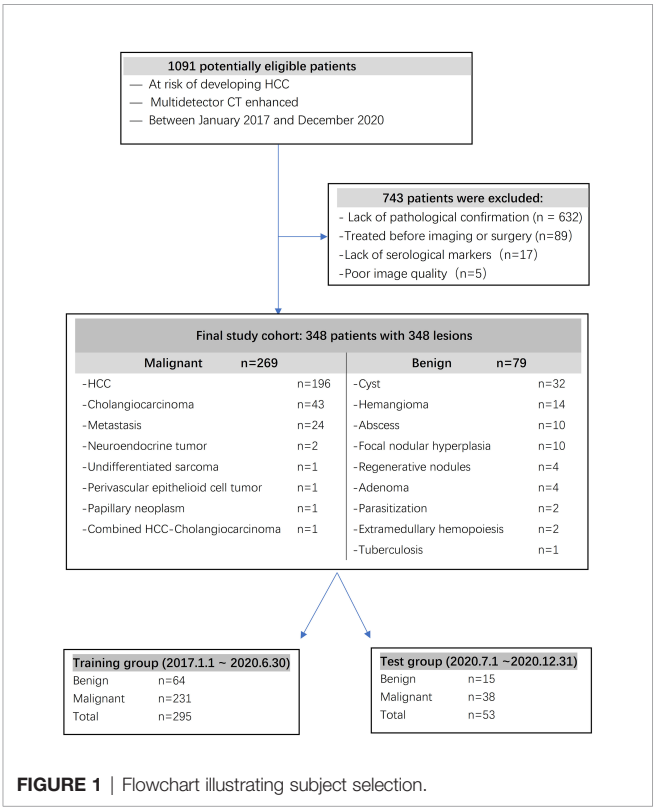



**FIGURE 1** | Flowchart illustrating subject selection.

**TABLE 1 |** Characteristics of patients in the training and test cohorts (1).

| Parameters | | Training (n = 295) | | | | | Test (n = 53) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Benign | Malignant | Total | $\chi^2$ | P | Benign | Malignant | Total | $\chi^2$ | P |
| **Sex** | Male | 34 | 186 | 220 | 19.836 | <0.001 | 11 | 33 | 44 | 1.39 | 0.238 |
| | Female | 30 | 45 | 75 | | | 4 | 5 | 9 | | |
| **CGOLF** | A | 52 | 156 | 208 | 4.764 | 0.092 | 12 | 29 | 41 | 0.08 | 0.773 |
| | B | 12 | 73 | 85 | | | 3 | 9 | 12 | | |
| | C | 0 | 2 | 2 | | | 0 | 0 | 0 | | |
| **Pathogenesis** | Alcohol | 39 | 29 | 68 | 66.657 | <0.001 | 10 | 5 | 15 | 15.18 | <0.001 |
| | HBV | 25 | 195 | 220 | | | 5 | 33 | 38 | | |
| | HCV | 0 | 6 | 6 | | | 0 | 0 | 0 | | |
| | PBC | 0 | 1 | 1 | | | 0 | 0 | 0 | | |
| **AFP** | - | 60 | 89 | 149 | 61.136 | <0.001 | 15 | 14 | 29 | 17.31 | <0.001 |
| | + | 4 | 142 | 146 | | | 0 | 24 | 24 | | |
| **APHE** | - | 59 | 63 | 122 | 87.075 | <0.001 | 14 | 7 | 21 | 25.23 | <0.001 |
| | + | 5 | 168 | 173 | | | 1 | 31 | 32 | | |
| **Washout** | - | 50 | 96 | 146 | 26.807 | <0.001 | 12 | 15 | 27 | 7.07 | 0.008 |
| | + | 14 | 135 | 149 | | | 3 | 23 | 26 | | |
| **Capsule** | - | 30 | 71 | 101 | 5.798 | 0.016 | 9 | 11 | 20 | 4.41 | 0.036 |
| | + | 34 | 160 | 194 | | | 6 | 27 | 33 | | |
| **Blood** | - | 56 | 156 | 212 | 9.882 | 0.002 | 13 | 22 | 35 | 3.97 | 0.046 |
| | + | 8 | 75 | 83 | | | 2 | 16 | 18 | | |
| **Necrosis** | - | 26 | 28 | 54 | 27.228 | <0.001 | 4 | 3 | 7 | 3.31 | 0.069 |
| | + | 38 | 203 | 241 | | | 11 | 35 | 46 | | |
| **Infiltrative** | - | 56 | 157 | 213 | 9.529 | 0.002 | 13 | 24 | 37 | 2.82 | 0.093 |
| | + | 8 | 74 | 82 | | | 2 | 14 | 16 | | |
| **Mural** | - | 31 | 74 | 105 | 5.882 | 0.015 | 6 | 13 | 19 | 0.16 | 0.692 |
| | + | 33 | 157 | 190 | | | 9 | 25 | 34 | | |
| **Satellite** | - | 56 | 157 | 213 | 9.529 | 0.002 | 13 | 23 | 36 | 3.37 | 0.066 |
| | + | 8 | 74 | 82 | | | 2 | 15 | 17 | | |
| **Halo** | - | 58 | 201 | 259 | 0.610 | 0.435 | 14 | 32 | 46 | 0.78 | 0.377 |
| | + | 6 | 30 | 36 | | | 1 | 6 | 7 | | |
| **PE** | - | 51 | 183 | 234 | 0.007 | 0.935 | 12 | 33 | 45 | 0.39 | 0.531 |
| | + | 13 | 48 | 61 | | | 3 | 5 | 8 | | |
| **VTT** | - | 64 | 210 | 274 | 6.264 | 0.012 | 15 | 34 | 49 | 1.71 | 0.191 |
| | + | 0 | 21 | 21 | | | 0 | 4 | 4 | | |
| **DE** | - | 58 | 213 | 271 | 0.168 | 0.682 | 13 | 36 | 49 | 1.00 | 0.316 |
| | + | 6 | 18 | 24 | | | 2 | 2 | 4 | | |
| **Internal** | - | 42 | 36 | 78 | 64.521 | <0.001 | 10 | 9 | 19 | 8.64 | 0.003 |
| | + | 22 | 195 | 217 | | | 5 | 29 | 34 | | |
| **LN** | - | 63 | 220 | 283 | 1.315 | 0.252 | 15 | 35 | 50 | 1.26 | 0.263 |
| | + | 1 | 11 | 12 | | | 0 | 3 | 3 | | |
| **Ascites** | - | 63 | 205 | 268 | 5.663 | 0.017 | 15 | 35 | 50 | 1.26 | 0.263 |
| | + | 1 | 26 | 27 | | | 0 | 3 | 3 | | |
| **Fat** | - | 63 | 227 | 290 | 0.009 | 0.926 | – | – | – | – | – |
| | + | 1 | 4 | 5 | | | – | – | – | | |
| **NC** | - | 34 | 207 | 241 | 44.611 | <0.001 | 11 | 36 | 47 | 4.91 | 0.027 |
| | + | 30 | 24 | 54 | | | 4 | 2 | 6 | | |
| **Mosaic** | - | 35 | 26 | 61 | 57.635 | <0.001 | 9 | 8 | 17 | 7.49 | 0.006 |
| | + | 29 | 205 | 234 | | | 6 | 30 | 36 | | |
| **Nodule** | - | 58 | 81 | 139 | 62.087 | <0.001 | 15 | 17 | 32 | 13.73 | <0.001 |
| | + | 6 | 150 | 156 | | | 0 | 21 | 21 | | |
| **Corona** | - | 64 | 221 | 285 | 2.868 | 0.090 | 15 | 36 | 51 | 0.82 | 0.365 |
| | + | 0 | 10 | 10 | | | 0 | 2 | 2 | | |

*CGOLF, Child grading of liver function; HBV, Hepatitis C virus; HCV, Hepatitis C virus; PBC, Primary biliary cirrhosis; APHE, non-rim Arterial phase hyperenhancement; Blood, Blood product in mass; Infiltrative, infiltrative appearance; Mural, Mural nodules; Satellite, Satellite lesions; Halo, Halo enhancement; PE, Peritumoral enhancement; VTT, vein tumor thrombus; DE, Delayed enhancement; Internal, Internal artery; LN, Lymph node; NC, Non-enhancing "capsule"; Mosaic, Mosaic architecture; Nodule, Nodule-in-nodule architecture; Corona, Corona enhancement.*

**TABLE 2** | Characteristics of patients in the training and test cohorts (2).

| Parameters | Training | | | | | | | | Test | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Benign | | | Malignant | | | Z | P | Benign | | | Malignant | | | Z | P |
| | 25th | Median | 75th | 25th | Median | 75th | | | 25th | Median | 75th | 25th | Median | 75th | | |
| **Age(year)** | 49.3 | 56.0 | 64.0 | 47.0 | 56.0 | 63.0 | -0.697 | 0.486 | 46.0 | 58.0 | 65.0 | 50.5 | 57.0 | 63.0 | -0.761 | 0.446 |
| **STB (umol/L)** | 12.1 | 21.4 | 43.4 | 15.4 | 25.0 | 59.5 | 2.029 | 0.042 | 12.4 | 23.1 | 69.4 | 13.7 | 20.8 | 34.8 | -0.227 | 0.820 |
| **PA(g/L)** | 14.4 | 63.9 | 73.9 | 18.9 | 63.0 | 68.4 | -0.419 | 0.675 | 11.2 | 19.0 | 71.8 | 21.8 | 65.0 | 74.0 | 1.343 | 0.179 |
| **PT(s)** | 12.6 | 13.6 | 14.3 | 12.9 | 13.7 | 14.3 | 1.260 | 0.208 | 13.0 | 13.9 | 14.6 | 12.8 | 14.1 | 15.9 | 0.445 | 0.657 |
| **BP (X10$^9$/L)** | 173.5 | 234.5 | 289.3 | 138.0 | 189.0 | 251.0 | -3.073 | 0.002 | 173.0 | 206.0 | 284.0 | 158.8 | 187.5 | 289.0 | -0.632 | 0.527 |
| **Dmax (cm)** | 4.8 | 6.5 | 8.7 | 3.7 | 5.4 | 8.7 | -1.848 | 0.065 | 4.4 | 6.5 | 8.0 | 3.4 | 4.9 | 8.8 | -0.622 | 0.534 |

*STB, Serum total bilirubin; PA, Plasma albumin; BP, Blood platelet; PT, Prothrombin time; Dmax, Maximum cross-section diameter.*

were significant differences in blood platelet counts (P =0.002) and serum total bilirubin (P =0.042) between the benign and malignant cohorts, but there were no significant differences in the other indexes (P > 0.05) (**Table 2**).

## Multivariable Factors Associated With Malignancy

Logistic regression evaluated the independent factors affecting malignancy. AFP (OR = 10.547, 95% CI 2.083- 53.401; P = 0.004), APHE (OR = 740.876, 95% CI 56.527- 9710.303; P < 0.001), washout (OR = 0.028, 95% CI 0.002-0.348; P = 0.005), satellite lesions (OR = 15.164, 95% CI 2.199-104.579; P = 0.006), ascites (OR = 156.241, 95% CI 1.822-13394.835; P = 0.026), and nodule-in-nodule architecture (OR = 27.401, 95% CI 4.982-150.700; P < 0.001) were independent predictors of malignancy (**Table 3**). The independent factors are used to build a model as combined predictors. The results of ROC analysis of the combined predictors for predicting malignant lesions are shown in (**Figures 2A, B**). In the training cohort, the AUC was 0.959, the sensitivity was 95.24%, and the specificity was 87.50%, while in the test cohort, the AUC was 0.981, the sensitivity was 94.74%, and the specificity was 93.33%. The order of AUCs was as follows: test cohort > training cohort > APHE > nodule-in-nodule architecture > AFP > washout > satellite lesions > ascites (**Table 4**).

## Malignancy Risk and the Prediction Nomogram

Based on the independent factors, we established a nomogram of the corresponding scoring system using RMS package in R as shown in **Figure 3**. Total points = 51 (AFP) + 100 (APHE) – 41 (washout) + 30 (satellite lesions) + 68 (ascites) + 54 (nodule-in-nodule architecture). The probability of malignant FLLs was approximately 80% for a patient with a score of 104, and with a cutoff point of 50%, the lesion

**TABLE 3** | The results of multivariate logistic regression analysis.

| Parameters | B | S.E. | Wald | P | OR | 95% C.I. for OR | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower | Upper |
| **Sex** | -0.477 | 0.673 | 0.503 | 0.478 | 0.621 | 0.166 | 2.320 |
| **Pathogenesis** | – | – | 0.402 | 0.940 | – | – | – |
| **Pathogenesis (1)** | -14.731 | 40193.1 | 0.000 | 1.000 | 0 | 0 | – |
| **Pathogenesis (2)** | -14.257 | 40193.1 | 0.000 | 1.000 | 0 | 0 | – |
| **Pathogenesis (3)** | 1.931 | 42542.7 | 0.000 | 1.000 | 6.895 | 0 | – |
| **AFP** | 2.356 | 0.828 | 8.104 | 0.004 | 10.547 | 2.083 | 53.401 |
| **APHE** | 6.608 | 1.313 | 25.334 | 0.000 | 740.876 | 56.527 | 9710.303 |
| **Washout** | -3.564 | 1.279 | 7.762 | 0.005 | 0.028 | 0.002 | 0.348 |
| **Capsule** | 0.993 | 0.726 | 1.867 | 0.172 | 2.698 | 0.650 | 11.207 |
| **Blood** | 1.451 | 0.962 | 2.272 | 0.132 | 4.266 | 0.647 | 28.140 |
| **Necrosis** | -1.186 | 0.926 | 1.641 | 0.200 | 0.306 | 0.050 | 1.875 |
| **Infiltrative** | -0.002 | 0.888 | 0.000 | 0.998 | 0.998 | 0.175 | 5.688 |
| **Mural nodules** | -0.445 | 0.721 | 0.381 | 0.537 | 0.641 | 0.156 | 2.632 |
| **Satellite lesions** | 2.719 | 0.985 | 7.616 | 0.006 | 15.164 | 2.199 | 104.579 |
| **DE** | 18.150 | 6298.88 | 0.000 | 0.998 | 7.6E+07 | 0 | – |
| **Internal** | 0.797 | 0.884 | 0.813 | 0.367 | 2.219 | 0.392 | 12.555 |
| **Ascites** | 5.051 | 2.271 | 4.947 | 0.026 | 156.241 | 1.822 | 13394.835 |
| **NC** | -1.410 | 1.033 | 1.863 | 0.172 | 0.244 | 0.032 | 1.849 |
| **Mosaic** | 0.035 | 0.945 | 0.001 | 0.971 | 1.035 | 0.162 | 6.596 |
| **Nodule** | 3.311 | 0.870 | 14.488 | 0.000 | 27.401 | 4.982 | 150.700 |
| **STB** | 0.021 | 0.014 | 2.161 | 0.142 | 1.021 | 0.993 | 1.050 |
| **Blood platelet** | 0.001 | 0.003 | 0.243 | 0.622 | 1.001 | 0.996 | 1.007 |
| **Constant** | 11.839 | 40193.1 | 0.000 | 1.000 | 1 | 138612.5 | – |

*APHE, non-rim Arterial phase hyperenhancement; Blood, Blood products in mass; DE, Delayed enhancement; Internal, Internal artery; NC, Non-enhancing "capsule"; Mosaic, Mosaic architecture; Nodule, Nodule-in-nodule architecture; STB, Serum total bilirubin.*

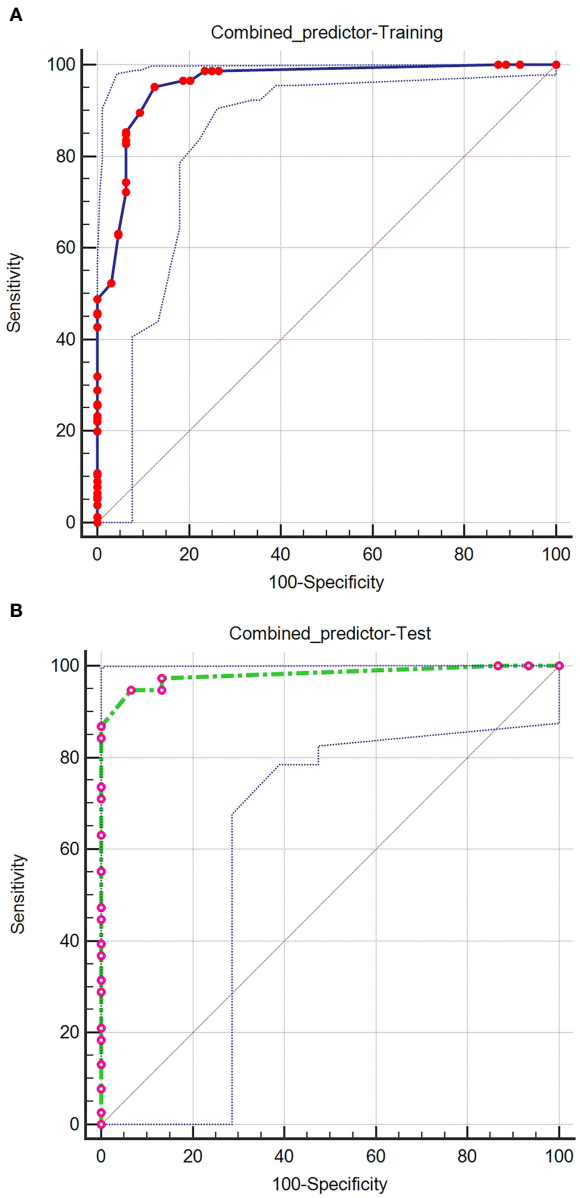



**FIGURE 2** | ROC curves of the combined predictors for predicting malignant lesions in the training **(A)** and test **(B)** cohorts.

was considered malignant. With a 50% cutoff point, a score of more than 80 points indicated malignant FLLs with a C-index of 96.80%. Moreover, calibration curves showed good agreement between the nomogram predictions and the actual data (**Figures 4A, B**). The DCA results are shown in **Figures 5A, B**.

## DISCUSSION

In this study, we established a precise nomogram based on CT imaging findings for predicting the malignancy of FLLs. The results indicate a good identification effect.

**TABLE 4** | Results of receiver operating characteristic (ROC) analysis.

| Group | AUC | 95%CI | Youden index | Sensitivity (%) | Specificity (%) | P |
|---|---|---|---|---|---|---|
| **Training#** | 0.959 | 0.929,0.978 | 0.827 | 95.24 | 87.50 | <0.001 |
| **Test#** | 0.981 | 0.899,0.999 | 0.881 | 94.74 | 93.33 | <0.001 |
| **AFP** | 0.776 | 0.724,0.822 | 0.552 | 61.47 | 93.75 | <0.001 |
| **APHE** | 0.825 | 0.776,0.866 | 0.649 | 72.73 | 92.19 | <0.001 |
| **Washout** | 0.683 | 0.626,0.736 | 0.366 | 58.44 | 78.12 | <0.001 |
| **Satellite lesions** | 0.598 | 0.539,0.654 | 0.195 | 32.03 | 87.50 | <0.001 |
| **Ascites** | 0.548 | 0.490,0.606 | 0.097 | 11.26 | 98.44 | <0.001 |
| **Nodule** | 0.778 | 0.726,0.824 | 0.556 | 64.94 | 90.62 | <0.001 |

*# as combined predictors.*

*APHE, non-rim Arterial phase hyperenhancement; Nodule, Nodule-in-nodule architecture.*

Triphasic CT scans are effective tools for differentiating benign and malignant FLLs, as the sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy were 100%, 80%, 94.5%, 100% and 95.5%, respectively (34). Diffusion weighted imaging (DWI) may provide additional information for the differentiation of HCC from nodules with abnormal hyperplasia by detecting the movement of freely diffusible water molecules. DWI model has a reference value for describing FLLs, the distributed diffusion coefficient, which shows good diagnostic performance (35).

In the univariate analysis of independent predictors of malignancy, most factors showed significant differences between the benign and malignant cohorts. Male patients with higher serum total bilirubin, higher AFP levels and lower blood platelet counts may be prone to malignant lesions. In CT, the presence of the following features indicates a high possibility of malignant lesions: APHE, washout, enhancing capsule, blood product in mass, necrosis, infiltrative appearance, mural nodules, satellite lesions, VTT, internal artery, ascites, non-enhancing capsule, mosaic architecture, and nodule-in-nodule architecture. Multivariable factors including AFP, APHE, washout, satellite lesions, ascites, and nodule-in-nodule architecture, were independent predictors of malignancy. A previous study reported that capsular enhancement is an important imaging biomarker for predicting high-grade HCC, and non-enhancing capsule is not significantly associated with high-grade HCC (36). However, our study showed that in our cohort, capsular enhancement was not an independent predictor of malignancy. The reason may be that some of the benign lesions such as abscesses, also showed capsular enhancement. APHE and washout are the main features of HCC, while nodule-in-nodule architecture is an auxiliary feature of HCC (37, 38). Nodule-in-nodule architecture, defined as a small nodule within the lesion, is an independent predictor of microvascular invasion (MVI) of HCC (36) and is proven to be an independent predictor of malignancy in our study. HCC accounts for the majority of malignant lesions, and these three features are independent predictors for the identification of benign and malignant FLLs in this study. A diagnostic model including AFP, sex, age and prothrombin time (ASAP model) has been shown to accurately predict the development of HCC in patients at high risk of hepatitis B virus. The ASAP model performed well in both the

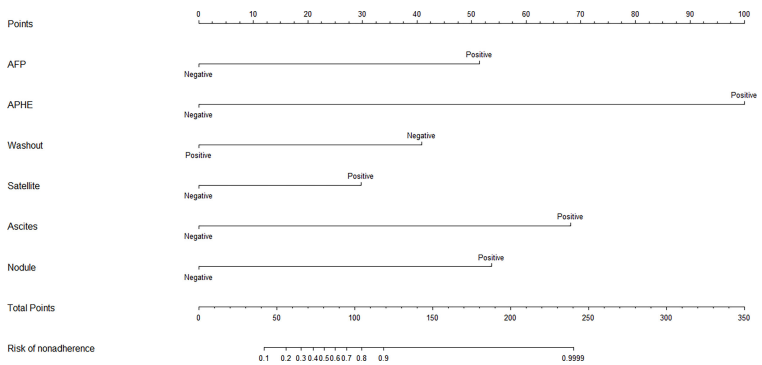



**FIGURE 3** | Based on the independent factors, the nomogram analysis method was used to establish a prediction scoring system. To use the nomogram, find the score for each variable on the corresponding axis and the total scores for all of the variables, and draw a line from the total score axis to the malignant risk axis to determine the malignancy risk (APHE, nonrim Arterial phase hyperenhancement; Satellite,Satellite lesions; Nodule,Nodule-in-nodule architecture).

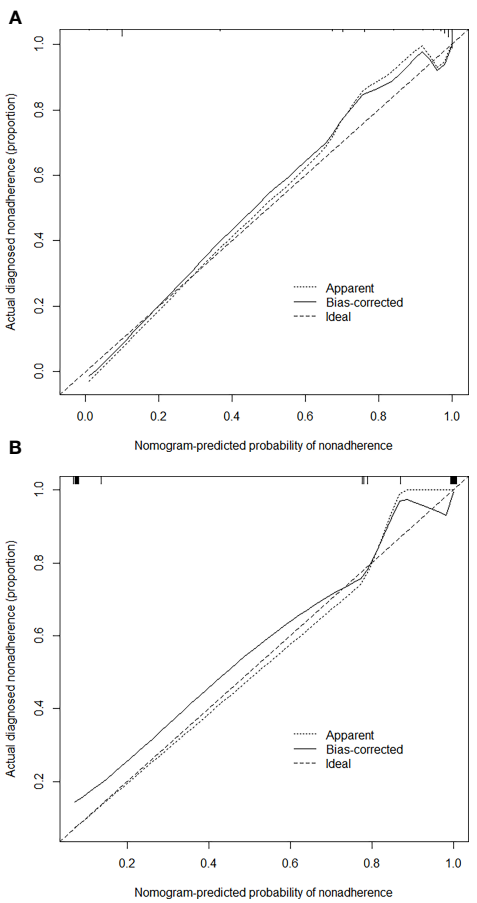



**FIGURE 4** | Calibration curves of the nomogram for estimating the malignancy risk in the training **(A)** and test **(B)** cohorts. On the calibration curve, the x-axis is the nomogram-predicted probability of malignancy, and the y-axis is the actual probability. The dotted line represents the ideal curve; the small, dotted line is the nomogram curve; and the straight line is the bias-corrected curve.

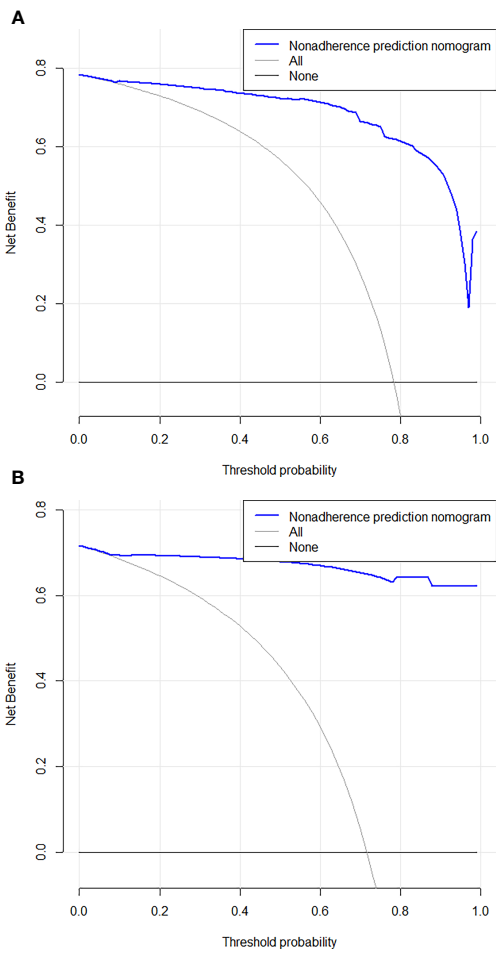



**FIGURE 5** | Decision curve analysis for the combined nomogram in the training **(A)** and test **(B)** cohorts. The Y axis represents the net benefit, and the X axis represents the threshold probability. The solid blue line shows the expected net benefit per patient based on Nomogram.

test and validation groups (39). In our study, with the exception of AFP, sex, serum total bilirubin, blood platelet counts, and pathogenesis were not independent predictors of malignancy, although there were statistically significant differences in these factors between the benign and malignant cohorts. The order of AUC values suggested that the combined predictors were the strongest in the diagnosis of malignant lesions (sensitivity = 95.24%, specificity = 87.50%). This model also had high diagnostic sensitivity (94.74%) and specificity (93.33%) in the test group.

We established a nomogram with a corresponding scoring system. With a cutoff point of 50%, it can accurately determine if an FLL is malignant. An example of image scoring is illustrated (**Figures 6A–C**).This lesion demonstrates no arterial phase hyperenhancement (APHE,0 score points, **Figure 6A**), presence of washout (minus 41 score points, **Figure 6B**), without satellite lesions (0 score points), ascites (68 score points, **Figure 6C**) and nodule-in-nodule architecture (54 score points, **Figure 6C**), with AFP>20.0 ng/ml (51 score points). The nomogram equation therefore would be as follows: 51 (AFP) + 100 (APHE) -41 (washout) + 30 (satellite lesions) + 68 (ascites) + 54 (nodule-in-nodule architecture) =132, indication malignancy according to the nomogram score > 80 points. Washout is one of the main features in the diagnosis of HCC, but in our study, it was negatively associated with benign and malignant tumors. One of the possible reasons is that some benign lesions, such as adenomas, can show washout, while some HCC lesions lack washout. In addition, washout is a purely visual criterion, which may result in observer-dependent bias. According to the literature reports, quantitative washout assessment in LI-RADS has the opposite effect in HCC diagnosis and needs to be redefined (40).

In this study, the C-index (96.80%) and calibration curve demonstrated that our nomogram was accurate in predicting the malignancy of lesions. However, there are some limitations to our model. First, this was a small, single-center, retrospective study that lacked a external validation group, which may alter the scoring system's efficiency. Therefore, multicenter and large-scale studies are necessary to improve the scoring system, and a prospective study is needed to confirm its reliability. Second, cirrhosis is likely to increase the possibility of HCC, which may cause selective bias. Third, rare lesions of liver were really small in our study. Assessing more data can make the model more generalizable.

## CONCLUSION

Based on AFP and CT findings including APHE, washout, satellite lesions, ascites, and nodule-in-nodule architecture, we developed an objective scoring system to predict the risk of malignancy. This model may aid in making informed treatment decisions for FLLs. A large-scale, prospective validation study is needed to assess the broad applicability of the nomogram.

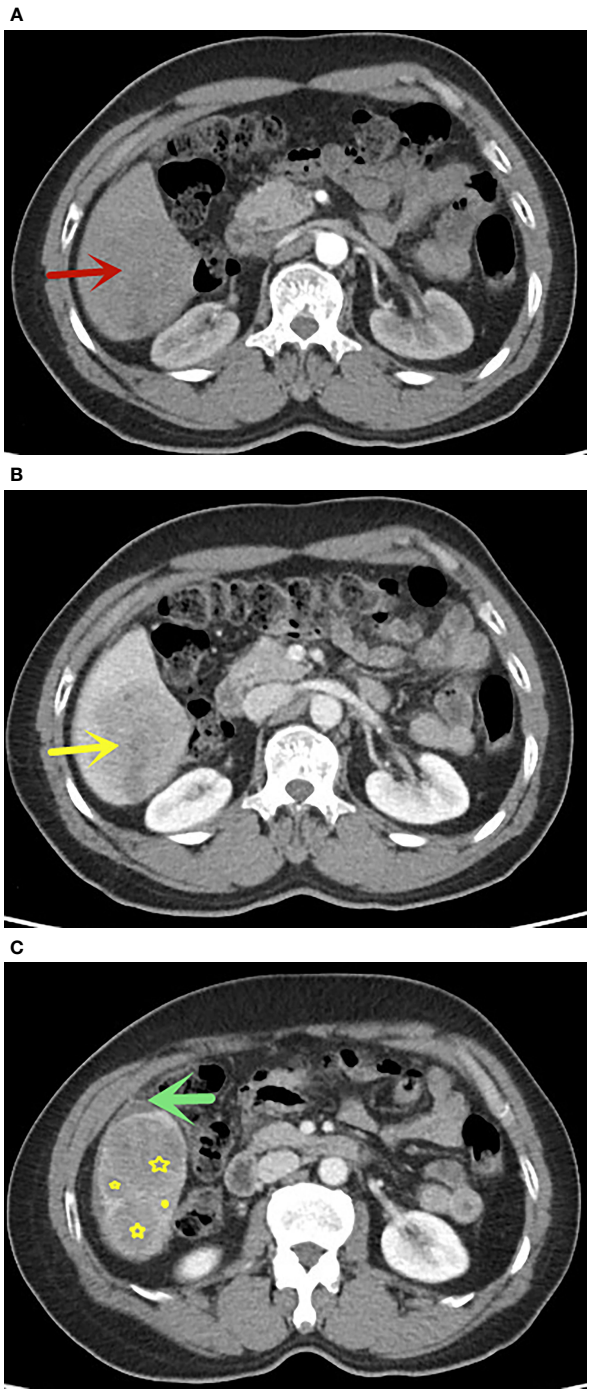



**FIGURE 6** | A 58 years old man with FLLs in segment VI, AFP>20.0 ng/ml, **(A)** no APHE in arterial phase (red arrow), **(B)** washout in portal venous phase, **(C)** presence of nodule-in-nodule architecture(yellow star) and ascites (green arrow), without satellite lesions. This lesion indicates malignancy according to the nomogram score greater than 80 points. The pathological result was hepatocellular carcinoma.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

This retrospective study was approved by the institutional review board of the hospital (K-2019-079-02) and based on the principles of the Helsinki declaration.

## AUTHOR CONTRIBUTIONS

HW, ZW, and CT designed the study, analyzed the data and drafted manuscript. XW, WZ, LM, and XJ contributed significantly to revise the manuscript. YL and RY polished the language. All authors contributed to the article and approved the submitted version.

## FUNDING

This work has received funding from the Science Foundation of Guangzhou First People's Hospital (M2019013), Guangzhou Health Science and Technology Project (20201A011013, 20211A011013), the Traditional Chinese Medicine Scientific Research Project of Guangdong Traditional Chinese Medicine Bureau (20201290), the Science and Technology Project of Guangzhou (201904010422, 202102010025), the Natural Science Foundation of Guangdong Province (2021A1515011350) and the Special Fund for the Construction of High-level Key Clinical Specialty (Medical Imaging) in Guangzhou Guangzhou Key Laboratory of Molecular Imaging and Clinical Translational Medicine.

## REFERENCES

1. Vernuccio F, Cannella R, Porrello G, Calandra A, Midiri M, Furlan A, et al. Uncommon Imaging Evolutions of Focal Liver Lesions in Cirrhosis. *Abdom Radiol (NY)* (2019) 44(9):3069–77. doi: 10.1007/s00261-019-02101-4
2. Faruque J, Rubin DL, Beaulieu CF, Napel S. Modeling Perceptual Similarity Measures in CT Images of Focal Liver Lesions. *J Digit Imaging* (2013) 26 (4):714–20. doi: 10.1007/s10278-012-9557-4
3. Hunia-Pajak K, Gadziala E, Pluta M, Blotnicki M, Partyka M, Gutkowski K. A Pitfall in Diagnosing Focal Lesions of the Liver. *Pol Arch Intern Med* (2019) 129(1):59–60. doi: 10.20452/pamw.4375
4. Nomair AM, Madkour MA, Shamseya MM, Elsheredy HG, Shokr A. Profiling of Plasma Metabolomics in Patients With Hepatitis C-Related Liver Cirrhosis and Hepatocellular Carcinoma. *Clin Exp Hepatol* (2019) 5(4):317–26. doi: 10.5114/ceh.2019.89478
5. Bottcher J, Hansch A, Pfeil A, Schmidt P, Malich A, Schneeweiss A, et al. Detection and Classification of Different Liver Lesions: Comparison of Gd-EOB-DTPA-Enhanced MRI Versus Multiphasic Spiral CT in a Clinical Single Centre Investigation. *Eur J Radiol* (2013) 82(11):1860–9. doi: 10.1016/j.ejrad.2013.06.013
6. Ronot M, Di Renzo S, Gregoli B, Duran R, Castera L, Van Beers BE, et al. Characterization of Fortuitously Discovered Focal Liver Lesions: Additional Information Provided by Shearwave Elastography. *Eur Radiol* (2015) 25 (2):346–58. doi: 10.1007/s00330-014-3370-z
7. Jansen MJA, Kuijf HJ, Veldhuis WB, Wessels FJ, Viergever MA, Pluim JPW. Automatic Classification of Focal Liver Lesions Based on MRI and Risk Factors. *PloS One* (2019) 14(5):e0217053. doi: 10.1371/journal.pone.0217053
8. Luersen GF, Bhosale P, Szklaruk J. State-Of-the-Art Cross-Sectional Liver Imaging: Beyond Lesion Detection and Characterization. *J Hepatocell Carcinoma* (2015) 2:101–17. doi: 10.2147/JHC.S85201
9. Thuring J, Rippel O, Haarburger C, Merhof D, Schad P, Bruners P, et al. Multiphase CT-Based Prediction of Child-Pugh Classification: A Machine Learning Approach. *Eur Radiol Exp* (2020) 4(1):20. doi: 10.1186/s41747-020-00148-3
10. Wang Q, Shi G, Qi X, Fan X, Wang L. Quantitative Analysis of the Dual-Energy CT Virtual Spectral Curve for Focal Liver Lesions Characterization. *Eur J Radiol* (2014) 83(10):1759–64. doi: 10.1016/j.ejrad.2014.07.009
11. Ominde ST, Mutala TM. Multicentre Study on Dynamic Contrast Computed Tomography Findings of Focal Liver Lesions With Clinical and Histological Correlation. *SA J Radiol* (2019) 23(1):1667. doi: 10.4102/sajr.v23i1.1667
12. Alhasan A, Cerny M, Olivie D, Billiard JS, Bergeron C, Brown K, et al. LI-RADS for CT Diagnosis of Hepatocellular Carcinoma: Performance of Major and Ancillary Features. *Abdom Radiol (NY)* (2019) 44(2):517–28. doi: 10.1007/s00261-018-1762-2
13. Zhang X, Yuan K, Wang H, Gong P, Jiang T, Xie Y, et al. Nomogram to Predict Mortality of Endovascular Thrombectomy for Ischemic Stroke Despite Successful Recanalization. *J Am Heart Assoc* (2020) 9(3):e014899. doi: 10.1161/JAHA.119.014899
14. Berardi G, Morise Z, Sposito C, Igarashi K, Panetta V, Simonelli I, et al. Development of a Nomogram to Predict Outcome After Liver Resection for Hepatocellular Carcinoma in Child-Pugh B Cirrhosis. *J Hepatol* (2020) 72 (1):75–84. doi: 10.1016/j.jhep.2019.08.032
15. Nie P, Yang G, Guo J, Chen J, Li X, Ji Q, et al. A CT-Based Radiomics Nomogram for Differentiation of Focal Nodular Hyperplasia From Hepatocellular Carcinoma in the Non-Cirrhotic Liver. *Cancer Imaging* (2020) 20(1):20. doi: 10.1186/s40644-020-00297-z
16. Nie P, Wang N, Pang J, Yang G, Duan S, Chen J, et al. CT-Based Radiomics Nomogram: A Potential Tool for Differentiating Hepatocellular Adenoma From Hepatocellular Carcinoma in the Noncirrhotic Liver. *Acad Radiol* (2021) 28(6):799–807. doi: 10.1016/j.acra.2020.04.027
17. Wang L, Dong T, Xin B, Xu C, Guo M, Zhang H, et al. Integrative Nomogram of CT Imaging, Clinical, and Hematological Features for Survival Prediction of Patients With Locally Advanced Non-Small Cell Lung Cancer. *Eur Radiol* (2019) 29(6):2958–67. doi: 10.1007/s00330-018-5949-2
18. Nie P, Yang G, Wang Z, Yan L, Miao W, Hao D, et al. A CT-Based Radiomics Nomogram for Differentiation of Renal Angiomyolipoma Without Visible Fat From Homogeneous Clear Cell Renal Cell Carcinoma. *Eur Radiol* (2020) 30 (2):1274–84. doi: 10.1007/s00330-019-06427-x
19. Li M, Li X, Guo Y, Miao Z, Liu X, Guo S, et al. Development and Assessment of an Individualized Nomogram to Predict Colorectal Cancer Liver Metastases. *Quant Imaging Med Surg* (2020) 10(2):397–414. doi: 10.21037/qims.2019.12.16
20. Matulewicz RS, Rademaker A, Meeks JJ. A Simplified Nomogram to Assess Risk of Bladder Cancer in Patients With a New Diagnosis of Microscopic Hematuria. *Urol Oncol* (2020) 38(4):240–46. doi: 10.1016/j.urolonc.2019.12.010
21. Zhang QW, Gao YJ, Zhang RY, Zhou XX, Chen SL, Zhang Y, et al. Personalized CT-Based Radiomics Nomogram Preoperative Predicting Ki-67 Expression in Gastrointestinal Stromal Tumors: A Multicenter Development and Validation Cohort. *Clin Transl Med* (2020) 9(1):12. doi: 10.1186/s40169-020-0263-4
22. Chong H, Zhou P, Yang C, Zeng M. An Excellent Nomogram Predicts Microvascular Invasion That Cannot Independently Stratify Outcomes of Small Hepatocellular Carcinoma. *Ann Transl Med* (2021) 9(9):757. doi: 10.21037/atm-20-7952

23. Ma X, Wei J, Gu D, Zhu Y, Feng B, Liang M, et al. Preoperative Radiomics Nomogram for Microvascular Invasion Prediction in Hepatocellular Carcinoma Using Contrast-Enhanced CT. *Eur Radiol* (2019) 29(7):3595–605. doi: 10.1007/s00330-018-5985-y
24. Gao SX, Liao R, Wang HQ, Liu D, Luo F. A Nomogram Predicting Microvascular Invasion Risk in BCLC 0/A Hepatocellular Carcinoma After Curative Resection. *BioMed Res Int* (2019) 2019:9264137. doi: 10.1155/2019/9264137
25. Cai W, He B, Hu M, Zhang W, Xiao D, Yu H, et al. A Radiomics-Based Nomogram for the Preoperative Prediction of Posthepatectomy Liver Failure in Patients With Hepatocellular Carcinoma. *Surg Oncol* (2019) 28:78–85. doi: 10.1016/j.suronc.2018.11.013
26. Cai Y, Zhang B, Li J, Li H, Liu H, Xie K, et al. A Novel Nomogram Based on Hepatic and Coagulation Function for Evaluating Outcomes of Intrahepatic Cholangiocarcinoma After Curative Hepatectomy: A Multi-Center Study of 653 Patients. *Front Oncol* (2021) 11:711061(711061). doi: 10.3389/fonc.2021.711061
27. Samoylova ML, Mehta N, Roberts JP, Yao FY. Predictors of Ultrasound Failure to Detect Hepatocellular Carcinoma. *Liver Transpl* (2018) 24(9):1171–77. doi: 10.1002/lt.25202
28. Chernyak V, Fowler KJ, Kamaya A, Kielar AZ, Elsayes KM, Bashir MR, et al. Liver Imaging Reporting and Data System (LI-RADS) Version 2018: Imaging of Hepatocellular Carcinoma in At-Risk Patients. *Radiology* (2018) 289 (3):816–30. doi: 10.1148/radiol.2018181494
29. Cerny M, Bergeron C, Billiard J-S, Murphy-Lavallée J, Olivié D, Bérubé J, et al. LI-RADS for MR Imaging Diagnosis of Hepatocellular Carcinoma: Performance of Major and Ancillary Features. *Radiology* (2018) 288(1):118–28. doi: 10.1148/radiol.2018171678
30. Kohno S, Arizono S, Isoda H, Yoshizawa A, Togashi K. Imaging Findings of Hemorrhagic Hepatic Cysts With Enhancing Mural Nodules. *Abdom Radiol (NY)* (2019) 44(4):1205–12. doi: 10.1007/s00261-019-01898-4
31. Martel G, Alsharif J, Aubin JM, Marginean C, Mimeault R, Fairfull-Smith RJ, et al. The Management of Hepatobiliary Cystadenomas: Lessons Learned. *HPB (Oxford)* (2013) 15(8):617–22. doi: 10.1111/hpb.12026
32. Gao S-X, Liao R, Wang H-Q, Liu D, Luo F. A Nomogram Predicting Microvascular Invasion Risk in BCLC 0/A Hepatocellular Carcinoma After Curative Resection. *BioMed Res Int* (2019) 2019:1–8. doi: 10.1155/2019/9264137
33. Ni FB, Lin Z, Fan XH, Shi KQ, Ao JY, Wang XD, et al. A Novel Genomic-Clinicopathologic Nomogram to Improve Prognosis Prediction of Hepatocellular Carcinoma. *Clin Chim Acta* (2020) 504:88–97. doi: 10.1016/j.cca.2020.02.001
34. Hafeez S, Alam MS, Sajjad Z, Khan ZA, Akhter W, Mubarak F. Triphasic Computed Tomography (CT) Scan in Focal Tumoral Liver Lesions. *J Pak Med Assoc* (2011) 61(6):571–5.
35. Kim HC, Seo N, Chung YE, Park MS, Choi JY, Kim MJ. Characterization of Focal Liver Lesions Using the Stretched Exponential Model: Comparison With Monoexponential and Biexponential Diffusion-Weighted Magnetic Resonance Imaging. *Eur Radiol* (2019) 29(9):5111–20. doi: 10.1007/s00330-019-06048-4
36. Wei H, Jiang H, Liu X, Qin Y, Zheng T, Liu S, et al. Can LI-RADS Imaging Features at Gadoxetic Acid-Enhanced MRI Predict Aggressive Features on Pathology of Single Hepatocellular Carcinoma? *Eur J Radiol* (2020) 132:109312. doi: 10.1016/j.ejrad.2020.109312
37. Chernyak V, Flusberg M, Law A, Kobi M, Paroder V, Rozenblit AM. Liver Imaging Reporting and Data System: Discordance Between Computed Tomography and Gadoxetate-Enhanced Magnetic Resonance Imaging for Detection of Hepatocellular Carcinoma Major Features. *J Comput Assist Tomogr* (2018) 42(1):155–61. doi: 10.1097/RCT.0000000000000642
38. Cerny M, Chernyak V, Olivie D, Billiard JS, Murphy-Lavallee J, Kielar AZ, et al. LI-RADS Version 2018 Ancillary Features at MRI. *Radiographics* (2018) 38(7):1973–2001. doi: 10.1148/rg.2018180052
39. Yang T, Xing H, Wang G, Wang N, Liu M, Yan C, et al. A Novel Online Calculator Based on Serum Biomarkers to Detect Hepatocellular Carcinoma Among Patients With Hepatitis B. *Clin Chem* (2019) 65(12):1543–53. doi: 10.1373/clinchem.2019.308965
40. Stocker D, Becker AS, Barth BK, Skawran S, Kaniewska M, Fischer MA, et al. Does Quantitative Assessment of Arterial Phase Hyperenhancement and Washout Improve LI-RADS V2018-Based Classification of Liver Lesions? *Eur Radiol* (2020) 30(5):2922–33. doi: 10.1007/s00330-019-06596-9


**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.

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 25 October 2022
DOI 10.3389/fonc.2022.1028577



OPEN ACCESS

EDITED BY
Margarida Julià-Sapé,
Universitat Autònoma de Barcelona, Spain

REVIEWED BY
Ran Zhang,
Huiying Medical Technology Co., Ltd., China
Octavian Sabin Tataru,
George Emil Palade University of Medicine, Pharmacy, Sciences and Technology of Târgu Mureș, Romania

*CORRESPONDENCE
Ming-Zhao Xiao
xmz.2004@163.com
Qing Jiang
300899@hospital.cqmu.edu.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 26 August 2022
ACCEPTED 07 October 2022
PUBLISHED 25 October 2022

CITATION
He Q-H, Tan H, Liao F-T, Zheng Y-N, Lv F-J, Jiang Q and Xiao M-Z (2022) Stratification of malignant renal neoplasms from cystic renal lesions using deep learning and radiomics features based on a stacking ensemble CT machine learning algorithm. *Front. Oncol.* 12:1028577. doi: 10.3389/fonc.2022.1028577



# Stratification of malignant renal neoplasms from cystic renal lesions using deep learning and radiomics features based on a stacking ensemble CT machine learning algorithm

Quan-Hao He[1†], Hao Tan[1†], Fang-Tong Liao[1], Yi-Neng Zheng[2], Fa-Jin Lv[2], Qing Jiang[3*] and Ming-Zhao Xiao[1*]

[1]Department of Urology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China, [2]Department of Radiology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China, [3]Department of Urology, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China

Using nephrographic phase CT images combined with pathology diagnosis, we aim to develop and validate a fusion feature-based stacking ensemble machine learning model to distinguish malignant renal neoplasms from cystic renal lesions (CRLs). This retrospective research includes 166 individuals with CRLs for model training and 47 individuals with CRLs in another institution for model testing. Histopathology results are adopted as diagnosis criterion. Nephrographic phase CT scans are selected to build the fusion feature-based machine learning algorithms. The pretrained 3D-ResNet50 CNN model and radiomics methods are selected to extract deep features and radiomics features, respectively. Fivefold cross-validated least absolute shrinkage and selection operator (LASSO) regression methods are adopted to identify the most discriminative candidate features in the development cohort. Intraclass correlation coefficients and interclass correlation coefficients are employed to evaluate feature's reproducibility. Pearson correlation coefficients for normal distribution features and Spearman's rank correlation coefficients for non-normal distribution features are used to eliminate redundant features. After that, stacking ensemble machine learning models are developed in the training cohort. The area under the receiver operator characteristic curve (ROC), calibration curve, and decision curve analysis (DCA) are adopted in the testing cohort to evaluate the performance of each model. The stacking ensemble machine learning algorithm reached excellent diagnostic performance in the testing dataset. The calibration plot shows good stability when using the stacking ensemble model. Net benefits presented by DCA are higher than the Bosniak 2019 version classification when employing any machine learning algorithm. The fusion feature-based machine learning algorithm accurately distinguishes malignant renal neoplasms from CRLs,

which outperformed the Bosniak 2019 version classification, and proves to be more applicable for clinical decision-making.

KEYWORDS

renal neoplasms, machine learning, Bosniak-2019 classification, cystic renal lesions, radiomics analysis

## Introduction

The detection rate of cystic renal lesions (CRLs) is increasing rapidly due to the growing use of computed tomography (CT). However, a considerable number of CRLs are simple renal cysts or benign cystic renal neoplasms according to histopathologic results after surgery; only a proportion of CRLs are malignant renal neoplasms. Cystic renal neoplasms refer to a diverse category of kidney tumors according to the WHO kidney tumor classification, which have a wide range of biological profiles, and the necessity of early surgical treatment for malignant CRL cannot be overstated (1). Since the components of CRL must be identified correctly in order to determine the appropriate treatment strategies, CT imaging is routinely used to distinguish CRLs. Meanwhile, CRLs are difficult to diagnose and manage especially in the early stage as they show a complex pattern in CT images including thickness of septation, enhancement of the mural nodule, and calcifications (2). In an attempt to identify malignant CRL at early stage, standardize the terminology explaining complicated renal cysts, and provide classification standards for stratifying surgically required renal lesions, the Bosniak classification system was created (3). In the 2019 version of the Bosniak classification system, more discriminative and quantitative criteria were introduced to improve the specificity in identifying higher-risk CRL categories and explicit detailed meanings about key terms to promote agreement and consistency among different readers. Based on the updated Bosniak classification, one or more enhancing nodules in CRL with obtuse margins (more than 4 mm) or with acute margins represent malignant renal neoplasms. Thickened wall or septa with enhancement in CRL also suggests the possibility of malignancy (4). However, according to the Bosniak classification, these high-risk CRL (IIF, III, and IV) could still be benign renal cysts rather than malignant neoplasms. Inaccurate treatment and related diagnostic errors caused by the misapplication of Bosniak classification may increase excessive medical care following undesirable outcomes including renal function impairment, re-operation, and neoplastic transplantation (5, 6). It has been shown that the diagnostic performances of the Bosniak 2019 classification criteria do not improve considerably compared with the previous version. A large number of previous Class III lesions will be reclassified as IIF according to the 2019 Bosniak version, resulting in decreased sensitivity (7, 8). The majority of Bosniak I and II lesions are benign renal cysts and Bosniak IIF, III, and IV lesions are more likely to be renal neoplasms. Approximately 10%–20% of Bosniak IIF lesions, 50% of Bosniak III lesions, and 90% of Bosniak IV lesions are considered to be renal neoplasms according to the latest research (9). To increase the diagnosis sensitivity and overcome the limits of biased visual image assessment, quantitative image analysis approaches using machine learning techniques also known as radiomics have gained popularity in recent years (10). In this study, we aim to develop and validate a stacking ensemble-based machine learning model using deep learning and radiomics features to stratify malignant cystic renal neoplasm more precisely.

## Materials and methods

### Participant enrollment and pathology assessment

In this retrospective study, data originated from abdominal CT scans or dedicated CT urography (CTU) scans in two separate institutions comprising unenhanced phases, corticomedullary phases, and nephrographic phases (Vue PACS, Carestream Health Inc & General Electric Advantage Workstation). Ethics committees in both institutions approved this retrospective investigation. Candidate participants included those with renal cysts larger than 1 cm, those with no surgery history (renal needle biopsy, nephrolithotomy, nephrectomy, or partial nephrectomy), those without conditions linked to multiple renal cysts (polycystic disease, Von Hippel–Lindau syndrome, or autosomal dominant polycystic kidney disease), and those with less than 25% solid portion in cystic lesions. Each individual in this study could only include verified cystic renal masses based on the final pathology findings, ensuring a realistic and reliable model's presentation. The detailed selection process and the pathological results of two cohorts are displayed in Figure 1.

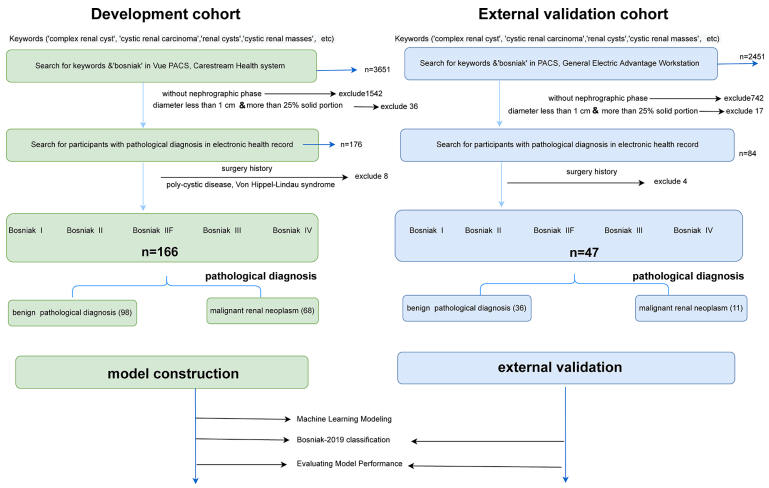



FIGURE 1
Flowchart illustrating how the cystic renal masses were enrolled and the distribution of pathology results in the final datasets. Detailed inclusion and exclusion criteria are also displayed in the flowchart. CRLs were classified as benign or malignant CRL based on pathological findings. After that, training datasets were applied to build machine learning models, and testing datasets were adopted to reclassify cystic renal masses based on the Bosniak 2019 version and assess model performance.

## Radiomics feature extraction

For extracting handcrafted radiomics features, radiomics feature class can be divided into three groups (1): first-order statistics, (2) shape features, and (3) second-order features. The image type of radiomics features can be divided into three groups: (1) original, (2) log, and (3) wavelet. Using the standard sample parameters setting provided in the official Pyradiomics YAML file, we extracted 1,231 handcrafted radiomics features in each individual.

## 3DResnet50 feature extraction

For extracting deep learning features, we defined a 3D-cropbox containing CRL area. The 3D-cropbox's width and length match the maximum cross-section area of the CRL, while its height matches the dimensions comprising the CRL area. In the 3D-cropbox, the area outside the ROI will be filled with black. After CRL regions have been segmented, the informative slices (the consecutive axial slices containing the full CRL area) will be cropped and resized to 14 mm * 128 mm * 128 mm (14-layer CT scans, default Voxel spacing is 1mm). The preprocessed images will be selected as the input of the convolutional neural network (CNN) model with pretrained weights. Deep learning features in each individual originated from the preprocessed CT images in the ROI area with 14 consecutive slices. Figure 2 displays the detailed workflow of the 3D-cropbox. By removing the last layer of the pretrained model, disabling gradient updates, and adding a 3D maximum pooling layer, we extracted 2,048 deep learning features in each individual. The detailed 3DResnet50 structure can be found in the Supplementary Table 1.

## Radiomics feature harmonization

CT acquisition and reconstruction parameters have a direct impact on handcrafted radiomics features. However, it is impractical to standardize platforms and parameters in advance across different institutions (11). ComBat harmonization, one of the strategies designed to deal with the batch effect, has been widely used in genomics-related research. In this study, ComBat harmonization methods are used to address the difference in extracted radiomics features originating from different image acquisition procedures.

## Correlation coefficient test

The intraclass correlation coefficients and interclass correlation coefficients are utilized to determine selected features with high reproducibility and robustness in image processing setting parameters. The interclass correlation coefficient results are extracted from two independent readers who randomly selected 25% participants in the enrolled datasets. The intraclass correlation coefficient results are calculated from one reader who randomly outlined the same participants in the enrolled datasets at different times (1 month interval) (12).

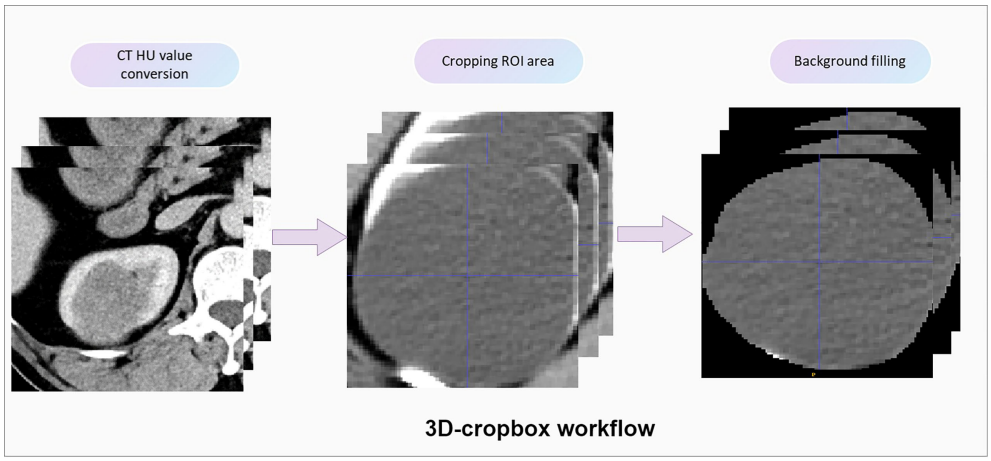



FIGURE 2
Detailed workflow of the 3D-cropbox. To assure that deep learning features retrieved from the 3D-cropbox are entirely tumor characteristics, the area outside the ROI will be filled with black.

## Quality control methods

The quality control approach for radiomics analysis and deep learning feature extraction consists of five steps: (1) image quality control, (2) region-of-interest (ROI) quality control, (3) feature extraction quality control, (4) feature selection quality control, and (5) machine learning algorithm quality control. We follow the recommendations from the Image Biomarker Standardization Initiative (IBSI) (13). Radiomics quality scores (RQS) are adopted to assess the reliability in this research (14). Detailed quality control methods and RQS calculation results are introduced in the Supplementary File.

## Statistical analysis

All ROIs are achieved through ITK-SNAP (version 3.6.0), and radiomics features extraction are executed using Pyradiomics package (version 3.0.1). The pretrained 3DResnet50 model weights come from 23 medical datasets (including brain MR images and lung CT images), which has been an open source published in Tencent's Medicalnet project (https://github.com/Tencent/MedicalNet). Deep learning features are extracted after adding a 3D maximum pooling layer in the 3DResnet50 model. After feature extraction, the least absolute shrinkage and selection operator (LASSO) approach is selected to choose the most recognized feature subsets in the training datasets (15, 16). We use the fivefold cross-validation method to select candidate variables. Pearson correlation coefficients for normal distribution and Spearman's rank correlation coefficients for non-normal distribution are used to identify whether redundant features existed in the primary selected radiomic features. Meanwhile, the intraclass correlation coefficients and interclass correlation coefficients are used to assess the final selected feature's reproducibility. Fivefold cross-validation Grid Search methods are adopted for parameter tuning in the stacking model construction. Figure 3 depicts the detailed workflow in model construction. Receiver operating characteristic (ROC) curve analysis and accuracy score (ACC) are used to evaluate each model's performance. DeLong tests are used to evaluate whether the statistical significance of area under the ROC curve (AUC) value exists in four models compared with the Bosniak 2019 version. Calibration curve is adopted to evaluate consistency performances of four models in the testing dataset. Decision curve analysis (DCA) is adopted to assess the clinical practicality. Two-sided $p$-value less than 0.05 is considered statistically significant. All four machine learning models are implemented using the scikit-learn package, and all statistical analysis and plot drawings are implemented using python (3.9 version) and R software (4.0.5 version).

# Results

## Participants

This retrospective study was approved by both hospital ethics committees, and private information of all patients had been de-identified. Among those participants included in development cohort (The First Affiliated Hospital of

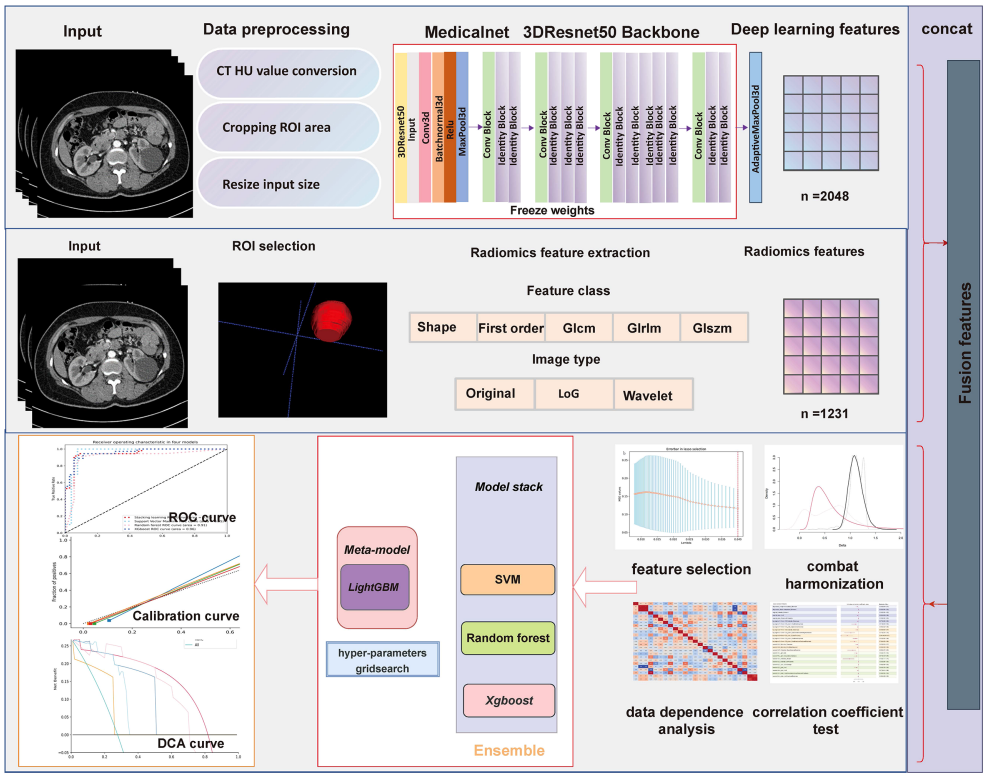



FIGURE 3
Flowchart presents the detailed procedure in fusion feature-based model construction and assessment methodology.

Chongqing Medical University), 98 individuals were diagnosed as benign CRL and 68 individuals were diagnosed as malignant CRL. In the pathological results of testing cohort (The Second Affiliated Hospital of Chongqing Medical University), 11 individuals have malignant CRL diagnosed and 36 individuals have benign CRL diagnosed. Detailed characteristic distributions are displayed in Table 1.

## Machine learning algorithm performance in CRL classification

Figures 4 and 5 show the detailed performance of base models and stacking ensemble model, respectively. The AUC values in the base models (Random Forest, Support Vector Machine, Xgboost) and the final model all demonstrate good discriminative ability (Figure 4). The $p$-value in the DeLong test shows that the results of the stacking ensemble radiomics model are statistically significant when compared to the Bosniak classification ($p < 0.05$). Detailed performance of four machine learning models compared with the Bosniak 2019 version classification in the training and testing cohorts is displayed in Table 2. In Figure 5, all four machine learning models illustrate good calibration performance.

## Clinical impact of the machine learning model compared with the Bosniak 2019 classification

DCA for four machine learning models in the testing dataset reveal that all four machine learning models deliver a higher net benefit than "none" and "all" treatment methods under most threshold probabilities (Figure 6). Meanwhile, according to the histopathologic results in the testing dataset, all four machine learning models provide a higher net benefit than the management guideline based on the Bosniak classification in terms of correctly stratifying cyst lesions into malignant renal neoplasms and benign renal masses, implying that using machine learning algorithm will provide better clinical decision support. Four instances of the final model's performance in the testing cohort are presented in Figure 7. Detailed confusion matrixes for four machine learning models and the Bosniak classification are displayed in Supplementary Material.

10.3389/fonc.2022.1028577

TABLE 1 Detailed distribution of the Bosniak 2019 classification and pathology results in the training cohort and testing cohort.

| Bosniak 2019 version | | Bosniak I | Bosniak II | Bosniak IIF | Bosniak III | Bosniak IV |
|---|---|---|---|---|---|---|
| | **Pathology analysis** | 59 | 17 | 23 | 20 | 47 |
| | **Benign results** | | | ***n* = 98** | | |
| | Simple renal cysts (89) | 59 | 15 | 10 | 5 | 0 |
| | Cystic nephroma (4) | 0 | 0 | 1 | 1 | 2 |
| | Renal angiomyolipoma (5) | 0 | 0 | 2 | 1 | 2 |
| **Training** | **Malignance results** | | | ***n* = 68** | | |
| **cohort** | Unclassified renal cell carcinoma (5) | 0 | 0 | 0 | 0 | 5 |
| | Clear cell renal cell carcinoma (38) | 0 | 0 | 3 | 8 | 27 |
| | Papillary renal cell carcinoma (10) | 0 | 1 | 1 | 1 | 7 |
| | Chromophobe renal cell carcinoma (3) | 0 | 0 | 0 | 1 | 2 |
| | Multilocular cystic renal neoplasm of low malignant potential (8) | 0 | 0 | 4 | 2 | 2 |
| | Tubulocystic renal cell carcinoma (4) | 0 | 1 | 2 | 1 | 0 |
| | **Pathology analysis** | 13 | 12 | 9 | 3 | 10 |
| | **Benign results** | | | ***n* = 36** | | |
| | Simple renal cysts (34) | 13 | 12 | 7 | 2 | 0 |
| **Testing** | Cystic nephroma (1) | 0 | 0 | 0 | 0 | 1 |
| **cohort** | Mixed epithelial and stromal tumor (1) | 0 | 0 | 1 | 0 | 0 |
| | **Malignance results** | | | ***n* = 11** | | |
| | Clear cell renal cell carcinoma (9) | 0 | 0 | 0 | 1 | 8 |
| | Multilocular cystic renal neoplasm of low malignant potential (2) | 0 | 0 | 1 | 0 | 1 |

## Discussion

Although the updated Bosniak classification system has a close correlation with cystic renal tumors, it has limits when analyzing the pathological results of Bosniak IIF, III, and IV classified CRLs, leading to inappropriate surgical operations and excessive follow-up costs. Identifying low-risk Bosniak classified CRLs can help prevent overtreatment and an increase in healthcare costs (17, 18). Previous findings suggested that the progression of Bosniak IIF cystic renal

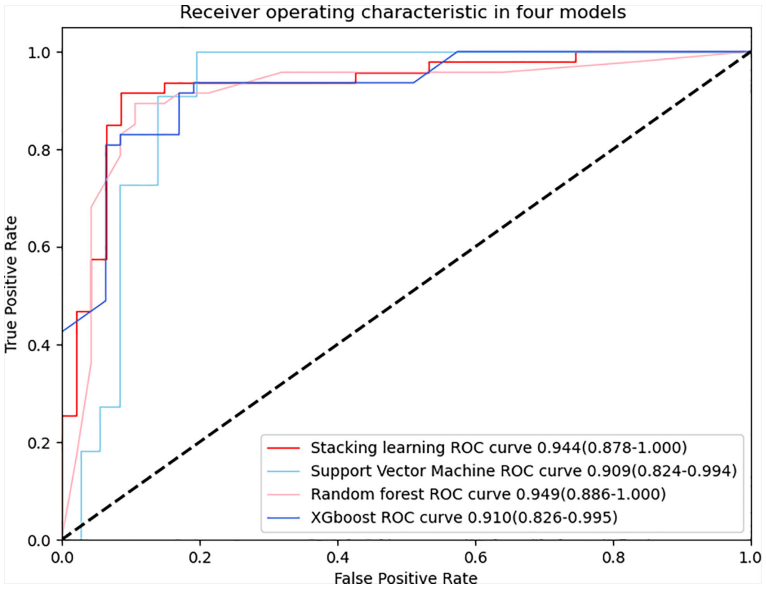



FIGURE 4
ROC curve is adopted to evaluate the diagnostic efficacy in four models. All four models have similar and excellent performance and detailed AUC values are respectively displayed in the graph, indicating the robust performance of quantitative features after LASSO selection.

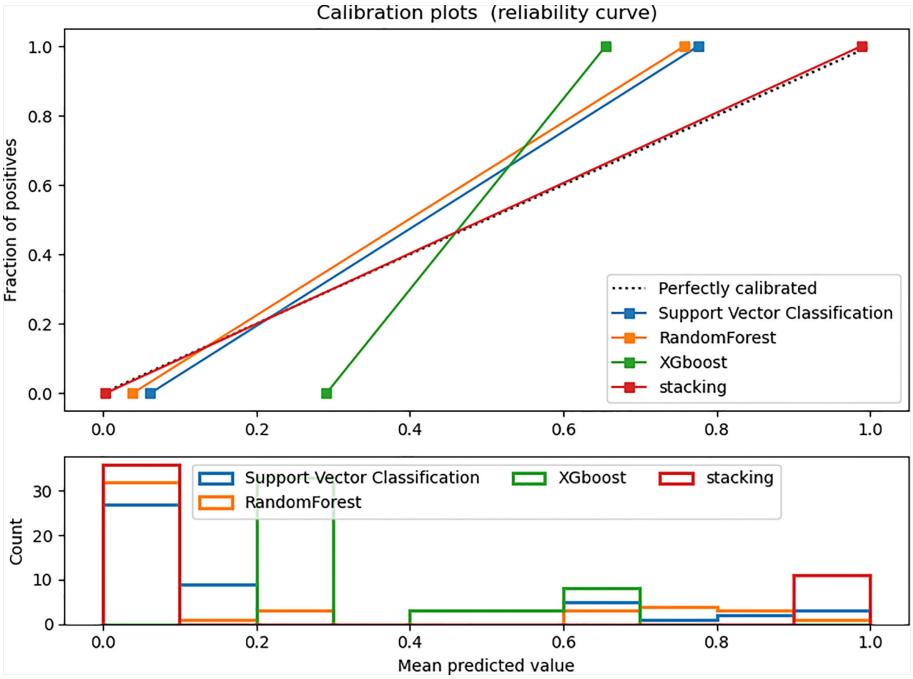



FIGURE 5
Calibration curve in the testing dataset. The ideal prediction curve is represented by the black dashed line. The model is more accurate as the actual prediction curve gets closer to the dashed line. The bar chart below depicts the average distribution of four models' predicted probabilities.

masses was 4 years (19). The high-risk Bosniak CRL had a quick progression that required radical nephrectomy rather than inappropriate surgical procedures like renal cyst decortication (20, 21). In this retrospective study, using stacking ensemble machine learning methods, we achieved excellent diagnostic performance in discriminating benign CRLs from malignant CRLs, which outperformed the Bosniak classification system. In the final model, we included 19 reproducible and discriminating radiomics features and 5 deep learning features, which displayed robust and consistent performance between cross-validation datasets and testing datasets.

TABLE 2 The performance of predictive models and the Bosniak 2019 version classification in the testing cohort and training cohort.

| | Model | AUC (95% CI) | ACC (95% CI) | Sensitivity | Specificity | *P*-value in the DeLong test |
|---|---|---|---|---|---|---|
| **Train cohort (5-fold cross-validation)** | Stacking | 0.948 (0.912–0.984) | 0.916 (0.915–0.917) | 0.912 (0.884–0.979) | 0.918 (0.864–0.973) | $p < 0.001$ |
| | Xgboost | 0.918 (0.867–0.986) | 0.916 (0.915–0.917) | 0.882 (0.806–0.959) | 0.939 (0.891–0.986) | $p = 0.051$ |
| | Random forest | 0.955 (0.925–0.985) | 0.904 (0.903–0.905) | 0.926 (0.864–0.989) | 0.888 (0.825–0.950) | $p < 0.001$ |
| | Support Vector Machine | 0.941 (0.902–0.980) | 0.886 (0.884–0.887) | 0.956 (0.907–1.000) | 0.837 (0.764–0.910) | $p = 0.002$ |
| | Bosniak 2019 classification | 0.863 (0.816–0.910) | 0.843 (0.842–0.845) | 0.971 (0.930–1.000) | 0.755 (0.670–0.840) | **ref** |
| **Test cohort** | Stacking | 0.944 (0.878–1.000) | 0.936 (0.934–0.939) | 1.000 (1.000–1.000) | 0.917 (0.826–1.000) | $p = 0.014$ |
| | Xgboost | 0.910 (0.826–0.995) | 0.851 (0.846–0.856) | 1.000 (1.000–1.000) | 0.806 (0.676–0.935) | $p = 0.168$ |
| | Random forest | 0.949 (0.886–1.000) | 0.915 (0.912–0.918) | 1.000 (1.000–1.000) | 0.889 (0.786–0.992) | $p = 0.010$ |
| | Support Vector Machine | 0.909 (0.824–0.994) | 0.851 (0.846–0.856) | 1.000 (1.000–1.000) | 0.806 (0.676–0.935) | $p = 0.128$ |
| | Bosniak 2019 classification | 0.847 (0.771–0.924) | 0.766 (0.758–0.773) | 1.000 (1.000–1.000) | 0.694 (0.544–0.845) | **ref** |

AUC, area under the receiver operating characteristic curve; ACC, accuracy score; ref, reference in AUC DeLong test, 95%CI; 95% confidence interval.

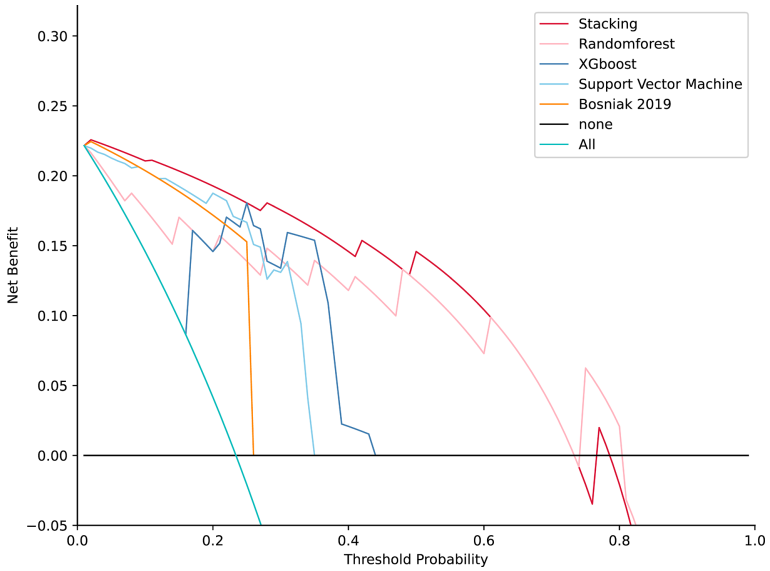



FIGURE 6
Decision curve analysis for four models compared with the Bosniak 2019 version in testing datasets. The net benefit is represented on the *y*-axis and the corresponding threshold probability is represented on the *x*-axis. The stacking model is represented by the red line. The Bosniak 2019 version is represented by the yellow line. All machine learning models present better performance and offered more net benefit than the Bosniak model.

The following elements contribute to the stacking ensemble model's reliability (1): The research procedure adheres to IBSI guidelines. (2) By using histopathologic examinations as diagnostic criterion, the diagnostic accuracy in this study is confirmed. (3) A stacking ensemble machine learning algorithm prevents overfitting in the training dataset. (4) In the testing step, the stacking model shows good diagnostic performance. (5) The RQS analysis results of this study was 16, which indicated that the quality of this radiomics study was reliable and reproducible. The points were accumulated by adhering to the following criteria: image protocol quality compliance (+1), feature reduction or adjustment for multiple testing compliance (+3), discrimination method with resampling method compliance (+2), calibration statistics method compliance (+1), validation from another institute compliance (+3), comparison to "gold standard" compliance (+2), potential clinical utility (+2), open-sourced code (+1), and open-sourced radiomics features compliance (+1). The updated 2019 version of the Bosniak classification intends to address inter-reader variability and improve diagnostic performance in predicting malignancy CRL. However, the proposed classification ability has yet to be confirmed (22). In this study, two different well-trained readers still disagreed on some CRL issues. In contrast, the stacking decision algorithm demonstrated strong and consistent performance without requiring subjective judgment across the testing datasets.

Many earlier studies have shown that machine learning approaches can be used to stratify CRL (23, 24). However, only a few studies rely on pathology as the diagnostic criteria (25). Miskin et al. used a CT texture-based machine learning method to distinguish CRLs as benign cysts and potentially malignant cysts based on cystic renal mass reclassification using the Bosniak 2019 version (26). However, the Bosniak classification does not have a precise pathological standard, and benign lesions may still be present in these potentially malignant groups, which limited the clinical value. Recently, Reinhold et al. used a CT-based radiomics model with a clinical decision algorithm to distinguish malignant CRLs from CRLs (27). However, they defined benign CRLs as non-imaging changes over 4 years' follow-up rather than pathological diagnostic criteria that could lead to actual biases, and the ability to distinguish benign from malignant CRLs remains debatable since benign CRLs were not defined by a pathological standard. In this study, all included CRLs have postoperative pathological results, which ensured that the model performance was reliable. Stacking algorithms demonstrated high specificity and sensibility, which may potentially impact clinical practice when radiologists or urologists try to evaluate and determine the right surgical procedure for CRLs.

Although the final machine learning model effectively predicted the outcome of CRL pathology, several limitations should also be mentioned. First, limited by the clinical sample size of CRLs, the

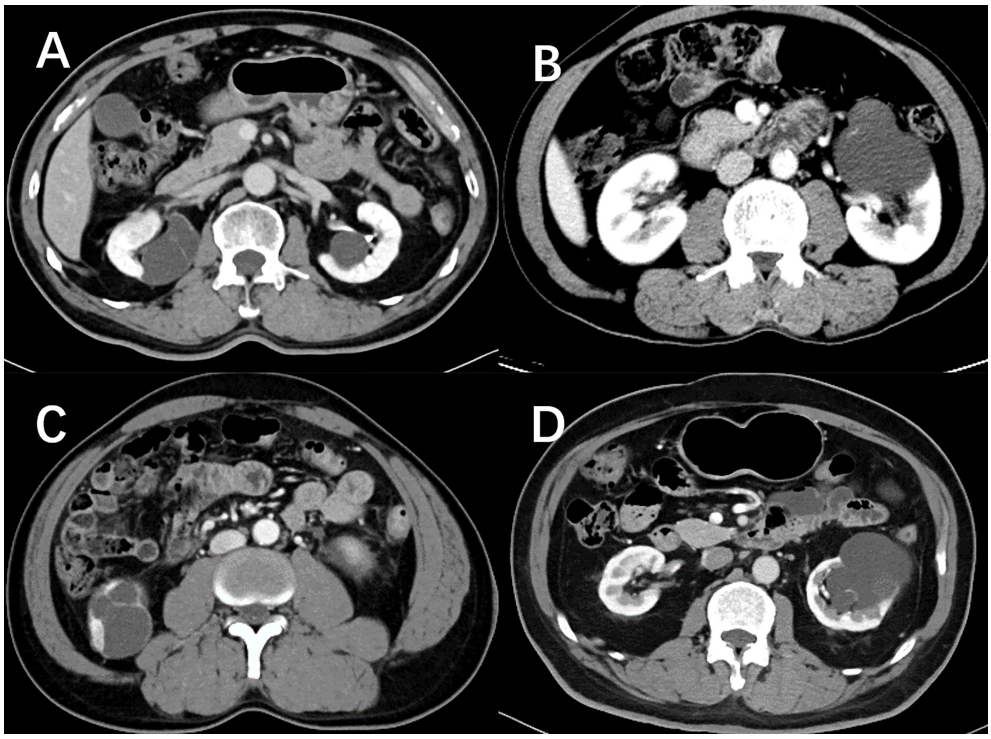



FIGURE 7
Nephrographic phase images for four cystic renal lesions in the testing datasets. Cystic renal lesion in **(A)** is benign 2019 Bosniak II lesion and cystic renal lesion in **(C)** is identified as clear cell renal cell carcinoma according to the histopathologic results after surgery, which stacking model properly predicted as cystic renal neoplasms. Cystic renal lesion in **(B, D)** are classified as 2019 Bosniak IIF, which are accurately identified as benign CRL by the stacking machine learning model.

diagnostic performance of our machine learning model in large samples still needs to be validated while testing datasets were used in this work. Second, we adopted nephrographic phase CT images to build the machine learning algorithm rather than using triple-phase CT images. In the Bosniak 2019 version, MRI standard features were formally added, whereas there is very little research focusing on renal cysts' textural features in MR images (28, 29). Previous studies have shown that using renal mass protocol MRI (with subtraction images) can help identify whether there are underlying enhancing features related to malignancy (2, 30). Nephrographic phase CT images and corticomedullary phase CT images for the radiomics model all demonstrated good discriminatory capability when compared with the Bosniak 2019 version categories (31). Future studies could attempt to integrate the triple-phase CT images and MR images by sequence-to-sequence models like the recurrent neural network (RNN) and vision transformer (VIT) (32, 33). Meanwhile, cystic nephroma is typically prevalent in female patients aged 50 to 60 years, which indicates that clinical variables such as age and gender may be potential predictors, and a mixture model that combines radiomics features with clinical features may further improve diagnostic performance (34).

## Conclusion

In summary, a stacking fusion feature-based machine learning meta model demonstrates good discrimination capability in stratifying malignant cystic renal neoplasms in CRLs across the testing datasets, which will be beneficial in diagnosing malignant CRLs at a curable stage, reducing overdiagnosis and overtreatment in CRLs.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

Q-HH and HT contributed equally to this work and share the first authorship. Q-HH, HT, F-JL, and Y-NZ designed the study; Q-HH, HT, and F-TL performed the experiments and collected the data. Q-HH analyzed the data and wrote the manuscript. M-ZX and QJ reviewed and edited the manuscript. All authors read and approved the final manuscript.

## Funding

This study was supported by the National Key Research and Development Project, No. 2020YFC2005900.

## Acknowledgments

We appreciate all radiologists and related staff in the First Affiliated Hospital of Chongqing Medical University for their assistance in data collection and ROI sketching.

## 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.

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.1028577/full#supplementary-material

## References

1. Moch H, Cubilla AL, Humphrey PA, Reuter VE, Ulbright TM. The 2016 who classification of tumours of the urinary system and Male genital organs-part a: Renal, penile, and testicular tumours. *Eur Urol* (2016) 70(1):93–105. doi: 10.1016/j.eururo.2016.02.029

2. Hu EM, Zhang A, Silverman SG, Pedrosa I, Wang ZJ, Smith AD, et al. Multi-institutional analysis of ct and mri reports evaluating indeterminate renal masses: Comparison to a national survey investigating desired report elements. *Abdom Radiol* (2018) 43(12):3493–502. doi: 10.1007/s00261-018-1609-x

3. Smith AD, Allen BC, Sanyal R, Carson JD, Zhang H, Williams JH, et al. Outcomes and complications related to the management of bosniak cystic renal lesions. *AJR Am J roentgenol* (2015) 204(5):W550–6. doi: 10.2214/ajr.14.13149

4. Dana J, Gauvin S, Zhang M, Lotero J, Cassim C, Artho G, et al. Ct-based bosniak classification of cystic renal lesions: Is version 2019 an improvement on version 2005? *Eur Radiol* (2022). doi: 10.1007/s00330-022-09082-x

5. Yang B, Qiu C, Wan S, Liu J, Li Q, Mai Z, et al. Long-term follow-up study of the malignant transformation potential of the simple renal cysts. *Trans Androl Urol* (2020) 9(2):684–9. doi: 10.21037/tau.2020.03.29

6. Soputro NA, Kapoor J, Zargar H, Dias BH. Malignant ascites following radical nephrectomy for cystic renal cell carcinoma. *BMJ Case Rep* (2021) 14(7): e243103. doi: 10.1136/bcr-2021-243103

7. Schoots IG, Zaccai K, Hunink MG, Verhagen P. Bosniak classification for complex renal cysts reevaluated: A systematic review. *J Urol* (2017) 198(1):12–21. doi: 10.1016/j.juro.2016.09.160

8. Yan JH, Chan J, Osman H, Munir J, Alrasheed S, Flood TA, et al. Bosniak classification version 2019: Validation and comparison to original classification in pathologically confirmed cystic masses. *Eur Radiol* (2021) 31(12):9579–87. doi: 10.1007/s00330-021-08006-5

9. Spiesecke P, Reinhold T, Wehrenberg Y, Werner S, Maxeiner A, Busch J, et al. Cost-effectiveness analysis of multiple imaging modalities in diagnosis and follow-up of intermediate complex cystic renal lesions. *BJU Int* (2021) 128(5):575–85. doi: 10.1111/bju.15353

10. Corrias G, Micheletti G, Barberini L, Suri JS, Saba L. Texture analysis imaging "What a clinical radiologist needs to know". *Eur J Radiol* (2022) 146:110055. doi: 10.1016/j.ejrad.2021.110055

11. Hu Y, Xie C, Yang H, Ho JWK, Wen J, Han L, et al. Computed tomography-based deep-learning prediction of neoadjuvant chemoradiotherapy treatment response in esophageal squamous cell carcinoma. *Radiother Oncol* (2021) 154:6–13. doi: 10.1016/j.radonc.2020.09.014

12. Pleil JD, Wallace MAG, Stiegel MA, Funk WE. Human biomarker interpretation: The importance of intra-class correlation coefficients (Icc) and their calculations based on mixed models, anova, and variance estimates. *J Toxicol Environ Health Part B Crit Rev* (2018) 21(3):161–80. doi: 10.1080/10937404.2018.1490128

13. Zwanenburg A, Vallières M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* (2020) 295 (2):328–38. doi: 10.1148/radiol.2020191145

14. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14(12):749–62. doi: 10.1038/nrclinonc.2017.141

15. Perez-Ortiz M, Gutierrez PA, Tino P, Hervas-Martinez C. Oversampling the minority class in the feature space. *IEEE Trans Neural Networks Learn Syst* (2016) 27(9):1947–61. doi: 10.1109/tnnls.2015.2461436

16. Vasquez MM, Hu C, Roe DJ, Chen Z, Halonen M, Guerra S. Least absolute shrinkage and selection operator type methods for the identification of serum biomarkers of overweight and obesity: Simulation and application. *BMC Med Res Method* (2016) 16(1):154. doi: 10.1186/s12874-016-0254-8

17. Campbell SC, Clark PE, Chang SS, Karam JA, Souter L, Uzzo RG. Renal mass and localized renal cancer: Evaluation, management, and follow-up: Aua guideline: Part I. *J Urol* (2021) 206(2):199–208. doi: 10.1097/ju.0000000000001911

18. Campbell SC, Uzzo RG, Karam JA, Chang SS, Clark PE, Souter L. Renal mass and localized renal cancer: Evaluation, management, and follow-up: Aua guideline: Part ii. *J Urol* (2021) 206(2):209–18. doi: 10.1097/ju.0000000000001912

19. Boissier R, Ouzaid I, Nouhaud FX, Khene Z, Dariane C, Chkir S, et al. Long-term oncological outcomes of cystic renal cell carcinoma according to the bosniak classification. *Int Urol Nephrol* (2019) 51(6):951–8. doi: 10.1007/s11255-019-02085-6

20. Huang Z, Wang H, Ji Z. Giant polycystic papillary renal cell carcinoma: A case report and literature review. *Front Oncol* (2022) 12:876217. doi: 10.3389/fonc.2022.876217

21. Xv Y, Lv F, Guo H, Liu Z, Luo D, Liu J, et al. A ct-based radiomics nomogram integrated with clinic-radiological features for preoperatively predicting

Who/Isup grade of clear cell renal cell carcinoma. *Front Oncol* (2021) 11:712554. doi: 10.3389/fonc.2021.712554

22. Pacheco EO, Torres US, Alves AMA, Bekhor D, D'Ippolito G. Bosniak classification of cystic renal masses version 2019 does not increase the interobserver agreement or the proportion of masses categorized into lower bosniak classes for non-subspecialized readers on ct or Mr. *Eur J Radiol* (2020) 131:109270. doi: 10.1016/j.ejrad.2020.109270

23. Gillingham N, Chandarana H, Kamath A, Shaish H, Hindman N. Bosniak iif and iii renal cysts: Can apparent diffusion coefficient-derived texture features discriminate between malignant and benign iif and iii cysts? *J Comput assist tomogr* (2019) 43(3):485–92. doi: 10.1097/rct.0000000000000851

24. Lee Y, Kim N, Cho KS, Kang SH, Kim DY, Jung YY, et al. Bayesian Classifier for predicting malignant renal cysts on mdct: Early clinical experience. *AJR Am J roentgenol* (2009) 193(2):W106–11. doi: 10.2214/ajr.08.1858

25. Li Y, Dai C, Bian T, Zhou J, Xiang Z, He M, et al. Development and prospective validation of a novel weighted quantitative scoring system aimed at predicting the pathological features of cystic renal masses. *Eur Radiol* (2019) 29 (4):1809–19. doi: 10.1007/s00330-018-5722-6

26. Miskin N, Qin L, Matalon SA, Tirumani SH, Alessandrino F, Silverman SG, et al. Stratification of cystic renal masses into benign and potentially malignant: Applying machine learning to the bosniak classification. *Abdom Radiol (New York)* (2021) 46(1):311–8. doi: 10.1007/s00261-020-02629-w

27. Dana J, Lefebvre TL, Savadjiev P, Bodard S, Gauvin S, Bhatnagar SR, et al. Malignancy risk stratification of cystic renal lesions based on a contrast-enhanced ct-based machine learning model and a clinical decision algorithm. *Eur Radiol* (2022) 32(6):4116–27. doi: 10.1007/s00330-021-08449-w

28. Ferreira AM, Reis RB, Kajiwara PP, Silva GEB, Elias J, Muglia VF. Mri evaluation of complex renal cysts using the bosniak classification: A comparison to ct. *Abdom Radiol* (2016) 41(10):2011–9. doi: 10.1007/s00261-016-0797-5

29. Krishna S, Schieda N, Pedrosa I, Hindman N, Baroni RH, Silverman SG, et al. Update on mri of cystic renal masses including bosniak version 2019. *J magn reson Imaging JMRI* (2021) 54(2):341–56. doi: 10.1002/jmri.27364

30. Davenport MS, Hu EM, Smith AD, Chandarana H, Hafez K, Palapattu GS, et al. Reporting standards for the imaging-based diagnosis of renal masses on ct and mri: A national survey of academic abdominal radiologists and urologists. *Abdom Radiol* (2017) 42(4):1229–40. doi: 10.1007/s00261-016-0962-x

31. Zhang Y, Zhao Y, Lv Y, Gu X. Value of quantitative ctta in differentiating malignant from benign bosniak iii renal lesions on ct images. *J Comput assist tomogr* (2021) 45(4):528–36. doi: 10.1097/rct.0000000000001181

32. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* (2017) 77(21):e104–e7. doi: 10.1158/0008-5472.Can-17-0339

33. Han K, Wang Y, Chen H, Chen X, Guo J, Liu Z, et al. A survey on vision transformer. *IEEE Trans Pattern Anal Mach Intell* (2022). doi: 10.1109/tpami.2022.3152247

34. Terada N, Arai Y, Kinukawa N, Yoshimura K, Terai A. Risk factors for renal cysts. *BJU Int* (2004) 93(9):1300–2. doi: 10.1111/j.1464-410X.2004.04844.x

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 27 October 2022
DOI 10.3389/fonc.2022.978164



OPEN ACCESS

EDITED BY
Damiano Caruso,
Sapienza University of Rome, Italy

REVIEWED BY
Natalia D Gladkova,
Privolzhsky Research Medical University (PIMU), Russia
Fajin Dong,
Jinan University, China

*CORRESPONDENCE
Hua-Rong Ye
yehuarong@hotmail.com
Xin-Wu Cui
cuixinwu@live.cn
Zhen-Zhong Guo
zhongbujueqi@hotmail.com

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 25 June 2022
ACCEPTED 12 October 2022
PUBLISHED 27 October 2022

CITATION
Zhang G, Yu J, Lei Y-M, Hu J-R, Hu H-M, Harput S, Guo Z-Z, Cui X-W and Ye H-R (2022) Ultrasound super-resolution imaging for the differential diagnosis of thyroid nodules: A pilot study. *Front. Oncol.* 12:978164. doi: 10.3389/fonc.2022.978164



# Ultrasound super-resolution imaging for the differential diagnosis of thyroid nodules: A pilot study

Ge Zhang[1,2†], Jing Yu[1†], Yu-Meng Lei[1], Jun-Rui Hu[3], Hai-Man Hu[4], Sevan Harput[5], Zhen-Zhong Guo[2*], Xin-Wu Cui[6*] and Hua-Rong Ye[1*]

[1]Department of Medical Ultrasound, China Resources & Wisco General Hospital, Wuhan University of Science and Technology, Wuhan, China, [2]Hubei Province Key Laboratory of Occupational Hazard Identification and Control, Wuhan University of science and technology, Wuhan, China, [3]Department of Chemistry and Chemical Engineering, Queen's University Belfast, Belfast, United Kingdom, [4]Department of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan, China, [5]Department of Electrical and Electronic Engineering, London South Bank University, London, United Kingdom, [6]Department of Medical Ultrasound, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

**Objective:** Ultrasound imaging provides a fast and safe examination of thyroid nodules. Recently, the introduction of super-resolution imaging technique shows the capability of breaking the Ultrasound diffraction limit in imaging the micro-vessels. The aim of this study was to evaluate its feasibility and value for the differentiation of thyroid nodules.

**Methods:** In this study, B-mode, contrast-enhanced ultrasound, and color Doppler flow imaging examinations were performed on thyroid nodules in 24 patients. Super-resolution imaging was performed to visualize the microvasculature with finer details. Microvascular flow rate (MFR) and micro-vessel density (MVD) within thyroid nodules were computed. The MFR and MVD were used to differentiate the benign and malignant thyroid nodules with pathological results as a gold standard.

**Results:** Super-resolution imaging (SRI) technique can be successfully applied on human thyroid nodules to visualize the microvasculature with finer details and obtain the useful clinical information MVD and MFR to help differential diagnosis. The results suggested that the mean value of the MFR within benign thyroid nodule was 16.76 $\pm$ 6.82 mm/s whereas that within malignant thyroid was 9.86 $\pm$ 4.54 mm/s. The mean value of the MVD within benign thyroid was 0.78 while the value for malignant thyroid region was 0.59. MFR and MVD within the benign thyroid nodules were significantly higher than those within the malignant thyroid nodules respectively ($p < 0.01$).

**Conclusions:** This study demonstrates the feasibility of ultrasound super-resolution imaging to show micro-vessels of human thyroid nodules *via* a clinical ultrasound platform. The important imaging markers, such as MVD and MFR, can be derived from SRI to provide more useful clinical information. It has the potential to be a new tool for aiding differential diagnosis of thyroid nodules.

KEYWORDS

thyroid nodule, super-resolution, ultrasound imaging, microbubble, contrast-enhanced ultrasound

## Introduction

Thyroid nodules are very common all over the world (1, 2). The main objective of the thyroid nodules-diagnosis is to distinguish malignant nodules from benign ones. Ultrasound (US) is a primary medical imaging tool in determining the risk stratification of thyroid nodules, which is critical for clinical management of thyroid nodules, which offers guidance for the application of fine-needle aspiration (FNA), as well as treatment decisions in patients with thyroid nodules (3, 4). Malignant US imaging patterns of thyroid nodules including solid composition, ill-defined margins, hypoechogenicity, microcalcifications, taller than wide and a lack of "halo" have been widely defined in previous studies (5).

Thyroid Imaging Reporting and Data System (TI-RADS), as the most used ultrasound-based malignancy risk stratification systems of thyroid nodules, is not enough to exclude or confirm thyroid malignancy sensitively and specifically (6, 7). TI-RADS classification is largely affected by the inter-observer variability, which results in suboptimal sensitivity and specificity (8). Furthermore, atypical benign and malignant nodules, especially TI-RADS 3 and 4 nodules, have a certain overlap in patterns during the routine US examination. These overlaps can easily lead to misdiagnosis of diseases and thus challenges for clinicians. Therefore, it is particularly important to explore a new way to accurately diagnose benign and malignant nodules.

Vascular distribution and flow characteristics within the nodule are broadly believed to play an important role in determining tumor characteristics (9). A number of scholars believe that angiogenesis has a great importance for distinguishing benign and malignant thyroid nodules. Micro-vessel density (MVD) has been regarded as the gold standard for the evaluation of tumor angiogenesis (10). Furthermore, some studies have indicated that counting micro-vessels may reflect local extrathyroidal and vascular invasion (11). Jiang performed immunohistochemical staining for CD31 and CD34 to obtain MVD on 122 thyroid nodules with different pathological types. The results showed the MVD in benign thyroid nodules was significantly higher than that in malignant thyroid nodules (12). Considering the invasiveness of acquiring MVD in this way, a new method to obtain micro-vessel density in thyroid nodules non-invasively still requires further exploration.

Color Doppler flow imaging (CDFI) is a widely used US blood flow imaging since the motion of red blood cells acting as scatterers. However, only relatively fast flow (larger than 1 cm/s) in the large vessels can be detected (13). Thus, CDFI is not an ideal imaging approach to observe the microvasculature and also the relatively slow flows in these micro-vessels. In the last several decades, the application of contrast-enhanced ultrasound (CEUS) on thyroid-related diseases has greatly improved. CEUS can significantly enhance the US contrast echo of the bloodstream, nevertheless, the imaging resolution of CEUS is still limited by the diffraction limit of US applied (14). Therefore, CEUS is still insufficient to observe the microvasculature at the capillary level and also the microvascular flow rate (MFR) within the micro-vessels. Besides, no guidelines recommend CDFI and CEUS as a routine method for US risk stratification of thyroid nodules (15–17).

In the last few years, inspired by optical super-resolution imaging (SRI), US SRI technique can bypass the compromise between penetration and resolution as the conventional US imaging is limited in image resolution by the diffraction to the scale of wavelength. After the injection of a low concentration of microbubble contrast agents, spatially isolated microbubbles can be localized, and their displacement can be tracked within a subwavelength resolution (18, 19). Therefore, US SRI and super-resolved velocity map (SRVM) can be generated at such a submicron scale, providing possibility to bridge the gap between multiple imaging techniques and histopathology (20, 21). After the generation of US SRI and SRVM, a number of useful clinical parameters could be derived from them, such as MVD and MFR. These clinical parameters could effectively help clinicians for clinical decision-making and also the disease managements. US SRI was firstly demonstrated *in vitro* using a single micro-channel in 2011 to show that US SRI could remarkably break the inherent diffraction limit (22). US SRI technique was then implemented in the mouse ear in 2015. This study firstly introduced the concept of SRVM and showed that

the magnitude and direction of flow velocity within the microvasculature can be derived (23). Moreover, US SRI was demonstrated in the rat brain (18) and a cancer model (24) respectively to show its potential to be a useful tool for a more comprehensive understanding and diagnostics of various disease progressions that alter the microvascular blood flow.

To our best knowledge, there have been no studies on the performance of SRI in discriminating between benign and malignant thyroid nodules. This study intended to explore whether SRI technique could image micro-vessels in thyroid nodules. Based on this, the CEUS images obtained by a low concentration of microbubble injection were used to perform US SRI processing. MFR and MVD within the thyroid area were further calculated to probe a new technical method for distinguishing benign and malignant thyroid nodules.

# Materials and methods

## Clinical data acquisition

From November 2021 to December 2021, 24 patients with thyroid nodules were enrolled (comprising 24 thyroid nodules that were classified as TI-RADS category 3 or 4). Patients aged ≥ 18 years with solid thyroid nodules that were > 5 mm in maximum diameter can be included in the study. Potential patients with any of the following conditions: large calcifications of the nodule with posterior acoustic shadow; previous medication, radiation, or surgery of the thyroid; hyperthyroidism, thyroiditis; or currently pregnant were excluded. All 24 thyroid nodules in 24 patients were diagnosed based on histopathological results from surgical excision or FNA. This study was approved by the Institutional Review Board of China Resources & Wisco General Hospital (approval number: HRWGZYY20210021). Written informed consent was obtained from each patient who underwent CEUS examinations.

All the ultrasound examinations were conducted by one senior radiologist who has received standardized training of thyroid nodules before collecting materials in this study. All patients underwent B-mode, CDFI and CEUS imaging respectively with an US platform (Resona 9s, Mindray Bio-Medical Electronics Co. Ltd., Shenzhen, China) and an L14-5WU linear array transducer (9.0 MHz center frequency, bandwidth 4 MHz - 14 MHz). During the scanning, the patients were asked to be supine and to expose the neck fully. The thyroid nodules were scanned in transverse and longitudinal sections. The patients were asked to avoid swallowing and hold their breath briefly during the scan to avoid the motion. According to the TI-RADS classification, the location, size, composition, echogenicity, margins, calcifications, and shape of thyroid nodules were observed and recorded, and then the nodules were assigned into different TI-RADS categories according to their total scores. Meanwhile, CDFI was also conducted to observe the blood flow and vascular morphology in and around the nodules. All the image datasets were acquired respectively at the same plane with the most abundant blood supply in each nodule.

For the SRI data acquisition, a low concentration of SonoVue (Bracco, Milan, Italy) microbubble solution was injected intravenously as a bolus of 0.1 mL as this technique requires to localize spatially isolated individual microbubble signals. Real-time dual-mode images (B-mode and CEUS) were utilized to guide the image plane and observe the microbubble signals after the injection. For each dataset, more than 1,500 images were acquired at a frame rate of 80 Hz. A mechanical index (MI) of 0.08 was used to avoid the microbubble destruction during the CEUS examinations. The other imaging parameters were set on the US system as follows: gain 10 dB, dynamic range 100. A stack of several thousands of US images was obtained for thyroid nodules for further US SRI processing. For each dataset, the pathological results received from biopsy and quantifications achieved from US SRI were compared.

For the routine CEUS examination, 2.4 mL of SonoVue was rapidly injected intravenously followed by a 5 mL of physiological saline solution to flush and promote the perfusion of the contrast agent in the blood vessels. The other imaging parameters remained the same as the previous SR examination. The data acquisition pipeline is shown in Figure 1.

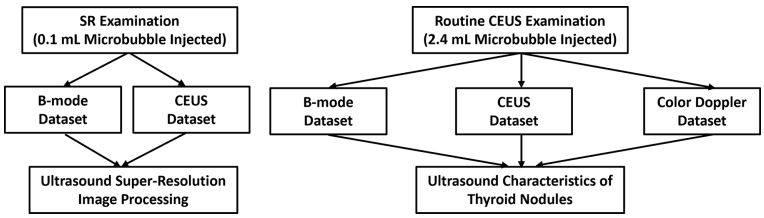



FIGURE 1
Data acquisition pipeline. A low concentration of microbubbles (0.1 mL) was injected to acquire B-mode and CEUS datasets for further US SRI processing. Then a routine concentration of microbubbles (2.4 mL) was injected to acquire B-mode, CEUS and CDFI datasets for US characteristics of thyroid nodules.

## Ultrasound imaging processing

The image data was obtained *via* the clinical US platform. The super-localization processing was performed offline using MATLAB (version R2021b; MathWorks Inc., Natick, MA, USA). A previous established two-stage motion correction (25–27) was applied to correct the tissue motion during the scanning.

For each dataset, a singular value decomposition (SVD) processing technique was used to filter out the tissue signal and retain the microbubble signals. Super-Localization processing was performed on each frame after setting an image pixel value threshold to reject the noise and detect potential microbubble signals. Each observed point spread function (PSF) was compared with a calibration PSF according to their area (A), intensity (I), and shape/eccentricity (E). These parameters were used to discard potential non-microbubble signals and noises. All the observed PSFs with the corresponding three attributes were summarized into three matrices. All the values were normalized in each matrix. The location of each spatially isolated microbubbles was calculated by the "centroid" method. The centroid of each localized microbubble was computed by its intensity-weighted center of mass. All the localizations from all the images were assembled into the final SRI (28).

The MFR was calculated based on the region of interest (ROI) which was manually drawn on MATLAB referring to the contours of benign and malignant thyroid nodules on both the B-mode image and the corresponding SRVM. To compute the MFR, the tracking method computes the best correlated bubble signals within a search window between neighboring images. Briefly, each microbubble detected in the image H and each of the microbubbles in the image H+1 was recorded within a search window. Since the frame rated of 80 Hz was used, 800 micrometers was set as the maximum search window so that flow rate up to 15 mm/s can be tracked. For each signal in the frame H, a paired signal in the image H+1 was identified if they have the maximum normalized cross-correlation value above a determined threshold of 0.8. The distribution of microvascular flow rate was visualized by the histogram of microbubble velocities.

The MVD was defined as tracked microbubble area divided by the ROI area. The ROI was also manually drawn on MATLAB referring to the contours of benign and malignant thyroid nodules on both the B-mode image and the corresponding SRI. The post-processing flow chart used in this study is shown in Figure 2.

## Statistical analysis

Results were expressed as the mean ± standard deviation. Statistical significance was analyzed using the two-sample t-test. All the continuous variables conformed to a normal distribution. A chi-square test was performed for independence to observe if two variables are related. A p value smaller than 0.05 was regarded to suggest statistical significance; a p-value smaller than 0.01, strong significance; and a p value smaller than 0.001, high significance.

# Results

## Ultrasound images of thyroid nodules

As shown in Table 1, clinical information, and grayscale US characteristics of 24 thyroid nodules who underwent thyroid

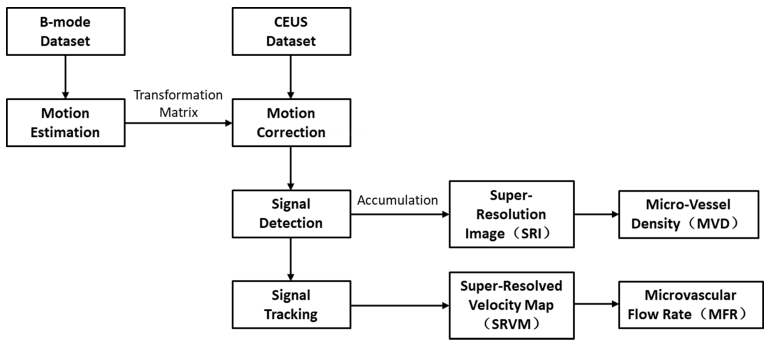



FIGURE 2
Post-processing flow chart to compute US SRI and the corresponding parameters. Motion estimation is performed on the original B-mode dataset, the generated transformation matrix is applied on the corresponding CEUS dataset to compensate for the movements. Then the bubble signals are detected and tracked over frames to compute the final SR images.

10.3389/fonc.2022.978164

TABLE 1 Patient clinical information and grayscale US characteristics of thyroid nodules for 24 patients.

| Characteristic | Benign | Malignant |
|---|---|---|
| Number of nodules | 12 | 12 |
| Age# | 50.9 (31-82) | 44.7 (29-53) |
| Gender | | |
| Female n (%) | 8 (66.7) | 9 (75.0) |
| Male | 4 (33.3) | 3 (25.0) |
| Size | | |
| ≤1cm | 5 (41.7) | 6 (50.0) |
| 1-2cm | 5 (41.7) | 5 (41.7) |
| >2cm | 2 (16.7) | 1 (8.3) |
| Composition | | |
| Cystic | 0 (0.0) | 0 (0.0) |
| Sponge-like | 0 (0.0) | 0 (0.0) |
| Mixed | 0 (0.0) | 0 (0.0) |
| Solid | 12 (100) | 12 (100) |
| Echogenicity | | |
| Anechogenicity | 0 (0.0) | 0 (0.0) |
| Iso- or hyperechogenicity | 4 (33.3) | 0 (0.0) |
| Hypoechogenicity | 8 (66.7) | 4 (33.3) |
| Marked hypoechogenicity | 0 (0.0) | 8 (66.7) |
| Shape | | |
| Wider than tall | 12 (100) | 7 (58.3) |
| Taller than wide | 0 (0.0) | 5 (41.7) |
| Microcalcifcation | | |
| Microcalcifcation | 0 (0.0) | 9 (75.0) |
| No microcalcifcation | 12 (100) | 3 (25.0) |
| Margin | | |
| Well defined | 9 (75.0) | 0 (0.0) |
| Poorly defined | 1 (8.3) | 7 (58.3) |
| Irregularity or lobuling | 1 (8.3) | 2 (16.7) |
| Extracapsular spread | 1 (8.3) | 3 (25.0) |
| TI-RADS classification | | |
| TI-RADS 3 | 6 (50.0) | 1 (8.3) |
| TI-RADS 4 | 6 (50.0) | 11 (91.7) |

unless otherwise specified, data in parentheses are percentages. #Numbers in parentheses are the range.

imaging from the China Resources & WISCO General Hospital medical records database (12 cases in each group of benign and malignant groups) were obtained. 24 TI-RADS category 3 and 4 thyroid nodules were identified as 12 malignant thyroid nodules and 12 benign nodules *via* the pathological results; all malignant nodules were papillary carcinomas (n = 12). Benign nodules included: thyroid adenoma (n = 6), nodular goiter (n = 5), and 1 other (subacute thyroiditis, n = 1). The gender, age, nodule size and composition were not found to differ significantly between malignant and benign nodules ($p > 0.05$). Among the sonographic features, hypoechogenicity or marked hypoechogenicity and microcalcification and aspect ratio had statistic differences between malignant and benign nodules ($p < 0.05$).

Various imaging techniques were performed to depict the microvasculature of benign and malignant thyroid nodules on human. SRI was also performed to visualize the microvasculature. Each moving isolated microbubble in the representative benign and malignant thyroid nodules was tracked locally to generate SRI and super-resolved velocity map (Videos 1, 2). Figure 3 above showed the B-mode and the corresponding CDFI, CEUS and SRI and SRVM images of the representative benign (A-E) and malignant (F-J) thyroid nodule respectively. Figure 4 showed two zoomed-in sections of the benign thyroid nodule in Figures 3A–E as the white and yellow boxes indicated in Figures 3A–E respectively. Figure 5 showed two zoomed-in sections of the malignant thyroid nodule in Figures 3F–J as the white and yellow boxes indicated in Figures 3F-J respectively.

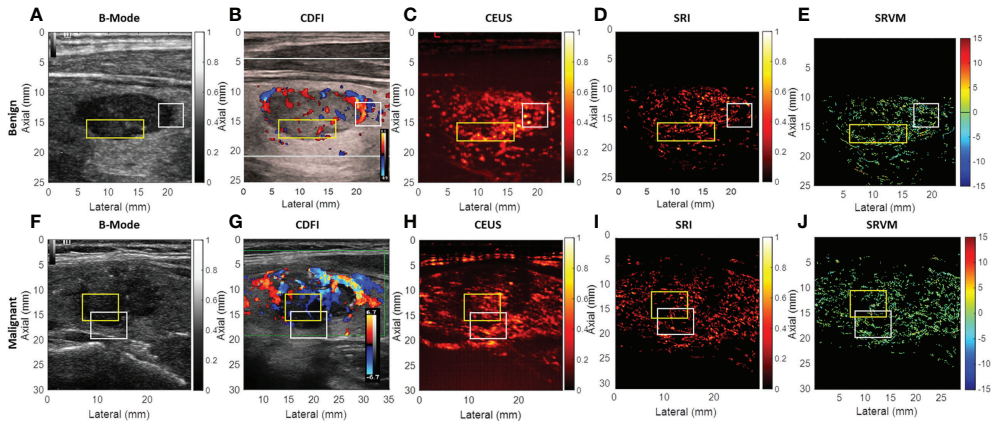



FIGURE 3
US images of the representative benign and malignant thyroid nodules. **(A)** B-mode image of the benign thyroid nodule. **(B)** CDFI images show the flow within the benign thyroid nodules. **(C)** CEUS image shows the accumulation of microbubbles along the frames within the benign thyroid nodule. **(D)** SRI shows the microvasculature within the benign thyroid nodule. **(E)** SRVM shows the flow velocity within the benign thyroid nodule. **(F)** B-mode image of the malignant thyroid nodule. **(G)** CDFI images show the flow within the malignant thyroid nodules. **(H)** CEUS image shows the accumulation of microbubbles along the frames within the malignant thyroid nodule. **(I)** SRI shows the microvasculature within the malignant thyroid nodule. **(J)** SRVM shows the flow velocity within the malignant thyroid nodule.

For CDFI, various flow patterns can be observed in both benign and malignant thyroid nodules, which include spotty, short lines, branching, and tortuous. Moreover, only large vessels with relatively fast blood flows can be observed as CDFI technique was not sensitive enough to visualize the microvascular flow. However, SRVM can visualize the microvascular flow with finer details and reveal the MFR within the microvasculature clearly whereas CDFI cannot as demonstrated in Figures 4B, E, G, J and 5B, E, G, J. Similar to CDFI, in the SRVM, the red color represented the relatively high flow rate and blue color represented the relatively low flow rate. It should be noted that a number of microvascular flow patterns shown in SRVM cannot be observed in CDFI.

For CEUS, after the injection of microbubble solution, the CEUS images showed an iso-enhancement pattern in the benign thyroid nodule and a heterogeneous hypo-enhancement pattern in malignant thyroid nodule. Microvasculature cannot be clearly seen on CEUS since the image resolution is inherently limited by the US transmission frequency. After super-localization processing, SRI can reveal the microvasculature within the thyroid nodule with much greater detail. Two adjacent micro-vessels cannot be observed on CEUS whereas they can be clearly seen in SRI as demonstrated in Figures 4C, D and 5C, D. Additionally, the tortuosity of individual micro-vessels was clearly reflected in SRI as SRI offered a better image resolution which could distinguish adjacent micro-vessels. However, it is challenging to observe tortuosity of individual micro-vessels in CEUS as the contrast signal of two adjacent micro-vessels may overlap together into a large vessel as shown in (Figures 4C, D, H, I and 5C, D, H, I).

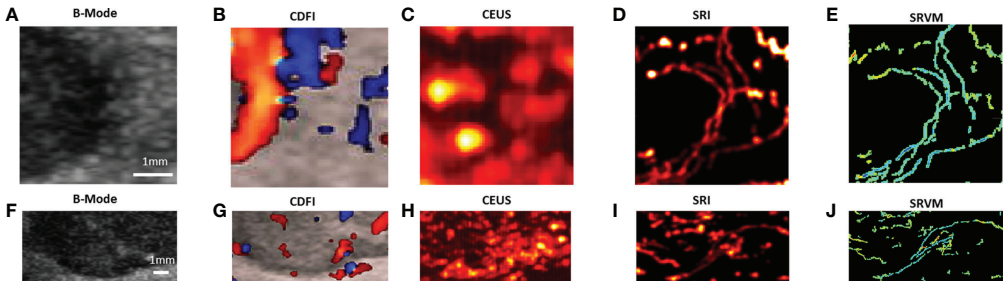



FIGURE 4
Two zoomed-in sections showing the detailed comparisons of benign thyroid nodule between **(A, F)** B-mode, **(B, G)** CDFI, **(C, H)** CEUS, **(D, I)** SRI and **(E, J)** SRVM as the white and yellow boxes indicated in Figures 3A–E.

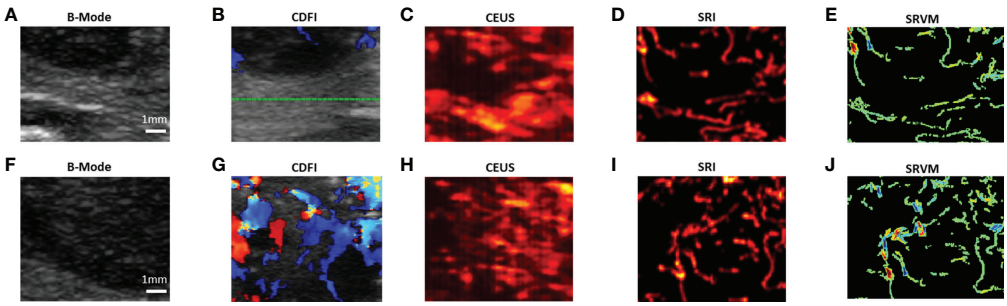



FIGURE 5
Two zoomed-in sections showing the detailed comparisons of malignant thyroid nodule between **(A, F)** B-mode, **(B, G)** CDFI, **(C, H)** CEUS, **(D, I)** SRI and **(E, J)** SRVM as the white and yellow boxes indicated in Figures 3F–J.

## Quantification of ultrasound super-resolution imaging

The MFR and MVD in both benign and malignant thyroid nodules were quantified respectively as shown in Figure 6 and Table 2. The mean value of the MFR in benign thyroid nodule region was 16.76 ± 6.82 mm/s whereas that in malignant thyroid region was 9.86 ± 4.54 mm/s. The mean value of the MVD in benign thyroid region was 0.78 while the value for malignant thyroid region was 0.59. A chi-square test was performed for independence to observe if two variables are related. A p value of $< 0.001$ suggests that two variables are related. The results showed that MFR and MVD within the benign thyroid nodules were 41.2% and 24.4% significantly higher than those within the malignant thyroid nodules respectively.

# Discussion

In this study, we utilized a novel US SRI technique to visualize the microvasculature and the corresponding microvascular flow rate within thyroid nodules. Compared with conventional CDFI and CEUS modalities, SRI could reveal more blood flow information within thyroid nodules; our results demonstrate the feasibility of using SRI to visualize microvasculature in human thyroid nodules. In addition, US SRI is capable of generating maps of microvascular morphology and blood flow velocity with a capillary-level image resolution. These features allow for a detailed analysis of local hemodynamics within the thyroid region, showing that the MFR and MVD within benign thyroid nodules are significantly higher than those within malignant thyroid nodules. This work may provide new important imaging markers

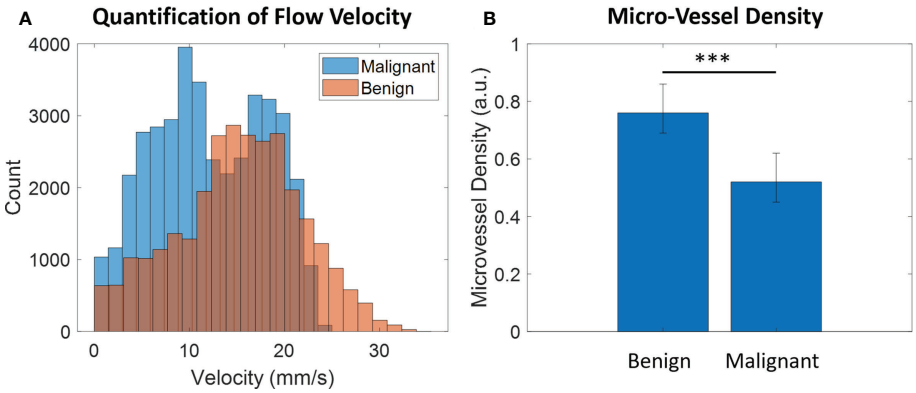



FIGURE 6
Quantification of **(A)** MFR and **(B)** MVD in both benign and malignant thyroid nodules. *** represents that there is a highly significant difference between two groups (p-value < 0.001).

10.3389/fonc.2022.978164

TABLE 2 Chi-square test for MVD and MFR.

| Parameters | Mean (Benign) | Std. (Benign) | Mean (Malignant) | Std. (Malignant) | Chi-Square | *p* |
|---|---|---|---|---|---|---|
| **MVD** | 0.78 | 0.12 | 0.59 | 0.10 | 56.87 | <0.001 |
| **MFR** | 16.76 | 6.82 | 9.86 | 4.54 | 98.55 | <0.001 |

for clinicians to differentiate benign and malignant thyroid nodules based on the microhemodynamics of thyroid nodules. For the first time as far as we are aware, US SRI technique was applied to help differentiate the benign and malignant thyroid nodules in patients with a clinically available US system and scanner. Image features, such as MFR and MVD, computed from the US SRI were evaluated to differentiate the benign and malignant thyroid nodules.

Previous studies examining vascularity within thyroid nodules by CDFI to differentiate benign from malignant thyroid nodules have shown controversial results. By analyzing 698 thyroid nodules of multiple studies, the results showed that 55.56% of the studies showed increased vascularity in malignant thyroid nodules, while the remaining studies showed no difference or even a decrease (29). The possible reasons for this discrepancy may be the differences in nodule size and pathological type. These different manifestations of vascularity in the benign and malignant thyroid nodules limit the application of CDFI in the differentiation of benign and malignant nodules. Our results demonstrated no significant difference was showed in blood flow patterns between benign and malignant thyroid nodules as showed in Figure 3. This diagnostic variability may be since CDFI cannot overcome the diffraction limit of ultrasound to truly reveal intra-nodular microvascular information (14).

CEUS makes up for deficiencies of conventional B-mode ultrasound and CDFI for imaging various diseases (30). The main adding value of CEUS is the visualization of sequence and intensity of vascular perfusion and hemodynamics. However, to date, no specific contrast enhancement pattern can be used to diagnose benign and malignant thyroid nodules alone (31). Widely accepted CEUS enhancement patterns for diagnosing malignant thyroid nodules include the following criterion: hypo-enhancement, heterogeneous and slow wash-in and wash-out curve lower than in normal thyroid tissue. The pathological basis may be related to the malignant nodules' complex neovascularization. Once the tumor grows beyond the newly formed vasculature, necrosis and embolism may occur in the tumor, resulting in an uneven enhancement or low enhancement pattern (5, 12, 32). Combining CEUS and quantitative analysis, our results also showed that malignant thyroid nodules tend to have low enhancement or uneven enhancement but without statistical difference. This may be due to the small sample size. Despite the high sensitivity and specificity of CEUS in diagnosing nodules, there is still a 12.5% missed diagnosis rate and a 13.67% misdiagnosis rate (5). Thus, a new method to obtain vascular distribution and flow characteristics in thyroid nodules non-invasively still requires further exploration.

US SRI has shown the capability to noninvasively visualize and quantify microvascular structures and blood flow dynamics with resolution below the wave diffraction limit (33–35). A large number of previous studies have demonstrated the capability of applying US SRI technique *in-vitro* and in various animals (23, 28, 34, 36). Recently, several studies have shown that, US SRI technique can be successfully applied on human brain and kidney to further help clinicians for medical diagnosis (37–39). In 2018, first-in-human applications of the US SRI techniques using clinical scanners were demonstrated in human breast cancer (40). Opacic et al. showed that US SRI could provide a number of clinical parameters such as velocities and directions of movement in individual vessels which can differentiate tumors with different vascular phenotypes, and it may provide opportunities for functional characterization of tumors and the assessment of therapy response. To assess any early changes in breast cancer following vascular disrupting agent (VDA) treatment, a study using 3-dimensional (3-D) super-resolution ultrasound (SR-US) imaging to observe MVD in breast lesions showed that the MVD of each tumor volume was significantly reduced at 24 hours after VDA treatment (41), indicating the feasibility of super-resolution technique for disease diagnosis *via* calculating micro-vessel density and microvascular flow rate in masses. However, there is no previous study reported the application of US SRI on human thyroid nodule to differentiate the benign and malignancy. In this study, our results confirmed that the super-resolution technique can not only clearly display the micro-vessels in thyroid nodules, but also quantify the MFR and MVD in the nodules. In addition, the MFR and MVD within benign thyroid nodules were significantly higher than those within malignant thyroid nodules, which showed the consistency with the previous study. The possible explanations could be that, in malignant thyroid region, the micro-vessels are more irregular and tortuous. Previous study has also reported that, some malignant thyroid nodules are completely avascular (42).

There are a few challenges existing in the general SRI technique. A previous study has reported a successful two-stage motion correction technique on the super-resolution imaging of human lower limb (28). This study has shown that two-stage motion correction could significantly correct the in-plane motion for the super-resolution imaging. For the out-of-plane motion, it is still a challenge for the 2D US imaging. Therefore, a 3D super-resolution

imaging technique equipped with a 2D array probe is expected to overcome this problem. It should be noted that, the CDFI shown in Figure 3 were not acquired at the same time as the CEUS image acquisition. Therefore, it may contribute to the reason that the CDFI images were not well corresponded to the final SR image. Another challenge exists in this study is that the acquisition of CEUS images for SR image processing requires the patients and the image transducer to remain stationary during the image acquisition. This is another general challenge for all the SR image processing (43). This is because it is difficult to ask patients to hold their breath for more than 10 s to acquire the image data. Moreover, heart beating is another inevitable factor that contributes to the tissue motion. All these factors would affect the SRI reconstruction of microvasculature.

There are some other limitations existing in this study. First, the sample size is relatively small in this study. Nevertheless, multiple studies have demonstrated the feasibility and reliability of super-resolution techniques to provide a number of clinical parameters of micro-vessels at the micron scale. A larger sample size is desired to draw a more solid conclusion to differentiate the benign and malignant thyroid nodules using US SRI. SR image generation is quite time-consuming and requires offline processing. The graphics processing unit (GPU) parallel processing is also desired to be integrated into the SR processing in the future to further accelerate the SR image generation. Second, the MVD in benign and malignant nodules could not be obtained by immunohistochemical staining in this study, however the difference in MVD obtained by SRI in benign and malignant thyroid nodules showed consistency with previous study (12), manifesting that the MVD within the benign thyroid nodules were significantly higher than those within the malignant thyroid nodules.

In conclusion, US SRI has been successfully demonstrated on human thyroid nodule to show finer microvascular details at a submicron scale image resolution and obtain important imaging markers, such as MFR and MVD within the thyroid nodule region which cannot be obtained from conventional B-mode, CEUS and CDFI techniques. Conventional multimodal US imaging techniques together with SRI technique could effectively improve the diagnostic accuracy in differentiating benign and malignant thyroid nodules as more useful clinical information (such as MVD and MFR parameters) can be provided for radiologists to differentiate benign and malignant thyroid nodules. The gap between conventional imaging techniques and histopathology can be bridged *via* US SRI to provide a new approach for clinicians to manage thyroid nodules and avoid unnecessary punctures.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving human participants were reviewed and approved by The Institutional Review Board of China Resources & Wisco General Hospital. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

GZ, JY, Y-ML, J-RH, M-HY, HC, NL, and H-MH contributed to the collection of relevant literature. Sevan Harput contributed the technical support of SRI technique. GZ and JY contributed to the literature analysis and manuscript preparation. GZ and JY sorted out the literature and wrote the manuscript. Z-ZG, X-WC and H-RY were responsible for design of this article and provided data acquisition, analysis, and interpretation. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the Key Research and Development Project of Hubei Province (No.2020BCB022), the National Natural Science Foundation of China (No. 82071953), the joint fund project of the Hubei Provincial Health and Family Planning Commission (No.WJ2019H197) and Hubei Province Key Laboratory of Occupational Hazard Identification and Control, Wuhan University of science and technology (No: 0HIC2019G08).

## Acknowledgments

The authors would like to thank the help from Prof. Mengxing Tang at Imperial College London in the completion of this article and the reviewers for reviewing this article. Further, we also would like to thank the help from M-H Y, HC and NL for the collection of relevant literature.

## 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.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.978164/full#supplementary-material

## References

1. Zhou J, Yin L, Wei X, Zhang S, Song Y, Luo B, et al. 2020 Chinese Guidelines for ultrasound malignancy risk stratification of thyroid nodules: the c-TIRADS. *Endocrine* (2020) 70(2):256–79. doi: 10.1007/s12020-020-02441-y

2. Hoang JK, Asadollahi S, Durante C, Hegedus L, Papini E, Tessler FN. An international survey on utilization of five thyroid nodule risk stratification systems: A needs assessment with future implications. *Thyroid* (2022) 32(6):675–81. doi: 10.1089/thy.2021.0558

3. Peng JY, Pan FS, Wang W, Wang Z, Shan QY, Lin JH, et al. Malignancy risk stratification and FNA recommendations for thyroid nodules: A comparison of ACR TI-RADS, AACE/ACE/AME and ATA guidelines. *Am J Otolaryngol* (2020) 41(6):102625. doi: 10.1016/j.amjoto.2020.102625

4. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American thyroid association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid* (2016) 26(1):1–133. doi: 10.1089/thy.2015.0020

5. Radzina M, Ratniece M, Putrins DS, Saule L, Cantisani V. Performance of contrast-enhanced ultrasound in thyroid nodules: Review of current state and future perspectives. *Cancers (Basel)* (2021) 13(21):5469. doi: 10.3390/cancers13215469

6. Sorrenti S, Dolcetti V, Radzina M, Bellini MI, Frezza F, Munir K, et al. Artificial intelligence for thyroid nodule characterization: Where are we standing? *Cancers (Basel)* (2022) 14(14):3357. doi: 10.3390/cancers14143357

7. Lin Y, Lai S, Wang P, Li J, Chen Z, Wang L, et al. Performance of current ultrasound-based malignancy risk stratification systems for thyroid nodules in patients with follicular neoplasms. *Eur Radiol* (2022) 32(6):3617–30. doi: 10.1007/s00330-021-08450-3

8. Watkins L, O'Neill G, Young D, McArthur C. Comparison of British thyroid association, American college of radiology TIRADS and artificial intelligence TIRADS with histological correlation: diagnostic performance for predicting thyroid malignancy and unnecessary fine needle aspiration rate. *Br J Radiol* (2021) 94(1123):20201444. doi: 10.1259/bjr.20201444

9. Zhang L, Gu J, Zhao Y, Zhu M, Wei J, Zhang B. The role of multimodal ultrasonic flow imaging in thyroid imaging reporting and data system (TI-RADS) 4 nodules. *Gland Surg* (2020) 9(5):1469–77. doi: 10.21037/gs-20-641

10. Rajabi M, Mousa SA. The role of angiogenesis in cancer treatment. *Biomedicines* (2017) 5(2):34. doi: 10.3390/biomedicines5020034

11. Skuletic V, Radosavljevic GD, Pantic J, Markovic BS, Jovanovic I, Jankovic N, et al. Angiogenic and lymphangiogenic profiles in histological variants of papillary thyroid carcinoma. *Pol Arch Intern Med* (2017) 127(6):429–37. doi: 10.20452/pamw.3999

12. Jiang J, Shang X, Zhang H, Ma W, Xu Y, Zhou Q, et al. Correlation between maximum intensity and microvessel density for differentiation of malignant from benign thyroid nodules on contrast-enhanced sonography. *J Ultrasound Med* (2014) 33(7):1257–63. doi: 10.7863/ultra.33.7.1257

13. Huang C, Zhang W, Gong P, Lok UW, Tang S, Yin T, et al. Super-resolution ultrasound localization microscopy based on a high frame-rate clinical ultrasound scanner: an in-human feasibility study. *Phys Med Biol* (2021) 66(8):10.1088/1361-6560/abef45. doi: 10.1088/1361-6560/abef45

14. Qiu L, Zhang J, Yang Y, Zhang H, Lee FF, He Q, et al. *In vivo* assessment of hypertensive nephrosclerosis using ultrasound localization microscopy. *Med Phys* (2022) 49(4):2295–308. doi: 10.1002/mp.15583

15. Lu R, Meng Y, Zhang Y, Zhao W, Wang X, Jin M, et al. Superb microvascular imaging (SMI) compared with conventional ultrasound for evaluating thyroid nodules. *BMC Med Imaging* (2017) 17(1):65. doi: 10.1186/s12880-017-0241-5

16. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, et al. ACR thyroid imaging, reporting and data system (TI-RADS): White paper of the ACR TI-RADS committee. *J Am Coll Radiol* (2017) 14(5):587–95. doi: 10.1016/j.jacr.2017.01.046

17. Russ G, Bonnema SJ, Erdogan MF, Durante C, Ngu R, Leenhardt L. European Thyroid association guidelines for ultrasound malignancy risk stratification of thyroid nodules in adults: The EU-TIRADS. *Eur Thyroid J* (2017) 6(5):225–37. doi: 10.1159/000478927

18. Errico C, Pierre J, Pezet S, Desailly Y, Lenkei Z, Couture O, et al. Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging. *Nature* (2015) 527(7579):499–502. doi: 10.1038/nature16066

19. Demene C, Robin J, Dizeux A, Heiles B, Pernot M, Tanter M, et al. Transcranial ultrafast ultrasound localization microscopy of brain vasculature in patients. *Nat BioMed Eng* (2021) 5(3):219–28. doi: 10.1038/s41551-021-00697-x

20. Zhu J, Rowland EM, Harput S, Riemer K, Leow CH, Clark B, et al. 3D super-resolution US imaging of rabbit lymph node vasculature in vivo by using microbubbles. *Radiology* (2019) 291(3):642–50. doi: 10.1148/radiol.2019182593

21. Harput S, Christensen-Jeffries K, Ramalli A, Brown J, Zhu J, Zhang G, et al. 3-d super-resolution ultrasound imaging with a 2-d sparse array. *IEEE Trans Ultrasonics Ferroelectrics Frequency Control* (2020) 67(2):269–77. doi: 10.1109/TUFFC.2019.2943646

22. Couture O, Besson B, Montaldo G, Fink M, Tanter M eds. Microbubble ultrasound super-localization imaging (MUSLI). In: *2011 IEEE International Ultrasonics Symposium* (2011):1285–87. doi: 10.1109/ULTSYM.2011.6293576

23. Christensen-Jeffries K, Browning RJ, Tang MX, Dunsby C, Eckersley RJ. *In vivo* acoustic super-resolution and super-resolved velocity mapping using microbubbles. *IEEE Trans Med Imaging* (2015) 34(2):433–40. doi: 10.1109/TMI.2014.2359650

24. Lin F, Shelton SE, Espindola D, Rojas JD, Pinton G, Dayton PA. 3-d ultrasound localization microscopy for identifying microvascular morphology features of tumor angiogenesis at a resolution beyond the diffraction limit of conventional ultrasound. *Theranostics* (2017) 7(1):196–204. doi: 10.7150/thno.16899

25. Zhang G, Harput S, Hu H, Christensen-Jeffries K, Zhu J, Brown J, et al. Fast acoustic wave sparsely activated localization microscopy: Ultrasound super-resolution using plane-wave activation of nanodroplets. *IEEE Trans Ultrasonics Ferroelectrics Frequency Control* (2019) 66(6):1039–46. doi: 10.1109/TUFFC.2019.2906496

26. Zhang G, Harput S, Toulemonde M, Broughton-Venner J, Zhu J, Riemer K, et al. Acoustic wave sparsely activated localization microscopy (AWSALM): In vivo fast ultrasound super-resolution ultrasound imaging using. *2019 IEEE International Ultrasonics Symposium (IUS)* (2019) 1930–33. doi: 10.1109/ULTSYM.2019.8926069

27. Harput S, Zhang G, Toulemonde M, Zhu J, Christensen-Jeffries K, Brown J, et al. Activation and 3D imaging of phase-change nanodroplet contrast agents with a 2D ultrasound probe. In: *2019 IEEE International Ultrasonics Symposium (IUS)* (2019) 2275–8. doi: 10.1109/ULTSYM.2019.8925892

28. Harput S, Christensen-Jeffries K, Brown J, Li Y, Williams KJ, Davies AH, et al. Two-stage motion correction for super-resolution ultrasound imaging in human lower limb. *IEEE Trans Ultrason Ferroelectr Freq Control* (2018) 65(5):803–14. doi: 10.1109/TUFFC.2018.2824846

29. Yang GCH, Fried KO. Most thyroid cancers detected by sonography lack intranodular vascularity on color Doppler imaging: Review of the literature and sonographic-pathologic correlations for 698 thyroid neoplasms. *J Ultrasound Med* (2017) 36(1):89–94. doi: 10.7863/ultra.16.03043

30. Kloth C, Kratzer W, Schmidberger J, Beer M, Clevert DA, Graeter T. Ultrasound 2020 - diagnostics & therapy: On the way to multimodal ultrasound: Contrast-enhanced ultrasound (CEUS), microvascular Doppler techniques, fusion imaging, sonoelastography, interventional sonography. *Rofo* (2021) 193(1):23–32. doi: 10.1055/a-1217-7400

31. Zhang J, Zhang X, Meng Y, Chen Y. Contrast-enhanced ultrasound for the differential diagnosis of thyroid nodules: An updated meta-analysis with comprehensive heterogeneity analysis. *PLoS One* (2020) 15(4):e0231775. doi: 10.1371/journal.pone.0231775

32. Wu Y, Zhou C, Shi B, Zeng Z, Wu X, Liu J. Systematic review and meta-analysis: diagnostic value of different ultrasound for benign and malignant thyroid nodules. *Gland Surg* (2022) 11(6):1067–77. doi: 10.21037/gs-22-254

33. Yang Z, Samanta S, Yan W, Yu B, Qu J. Super-resolution microscopy for biological imaging. *Adv Exp Med Biol* (2021) 3233:23–43. doi: 10.1007/978-981-15-7627-0_2

34. Christensen-Jeffries K, Brown J, Aljabar P, Tang M, Dunsby C, Eckersley RJ. 3-d *In vitro* acoustic super-resolution and super-resolved velocity mapping using microbubbles. *IEEE Trans Ultrason Ferroelectr Freq Control* (2017) 64(10):1478–86. doi: 10.1109/TUFFC.2017.2731664

35. Chen Q, Yu J, Rush BM, Stocker SD, Tan RJ, Kim K. Ultrasound super-resolution imaging provides a noninvasive assessment of renal microvasculature changes during mouse acute kidney injury. *Kidney Int* (2020) 98(2):355–65. doi: 10.1016/j.kint.2020.02.011

36. Brown J, Christensen-Jeffries K, Harput S, Zhang G, Zhu J, Dunsby C, et al. Investigation of microbubble detection methods for super-resolution imaging of microvasculature. *IEEE Trans Ultrason Ferroelectr Freq Control* (2019) 66(4):676–91. doi: 10.1109/TUFFC.2019.2894755

37. Lindsey BD, Light ED, Nicoletto HA, Bennett ER, Laskowitz DT, Smith SW. The ultrasound brain helmet: new transducers and volume registration for *in vivo* simultaneous multi-transducer 3-d transcranial imaging. *IEEE Trans Ultrason Ferroelectr Freq Control* (2011) 58(6):1189–202. doi: 10.1109/TUFFC.2011.1929

38. Soulioti DE, Espindola D, Dayton PA, Pinton GF. Super-resolution imaging through the human skull. *IEEE Trans Ultrason Ferroelectr Freq Control* (2020) 67 (1):25–36. doi: 10.1109/TUFFC.2019.2937733

39. Dekkers JF, Alieva M, Wellens LM, Ariese HCR, Jamieson PR, Vonk AM, et al. High-resolution 3D imaging of fixed and cleared organoids. *Nat Protoc* (2019) 14(6):1756–71. doi: 10.1038/s41596-019-0160-8

40. Opacic T, Dencks S, Theek B, Piepenbrock M, Ackermann D, Rix A, et al. Motion model ultrasound localization microscopy for preclinical and clinical multiparametric tumor characterization. *Nat Commun* (2018) 9(1):1527. doi: 10.1038/s41467-018-03973-8

41. Oezdemir I, Li J, Song J, Hoyt K eds. 3-d super-resolution ultrasound imaging for monitoring early changes in breast cancer after treatment with a vascular-disrupting agent. *2021 IEEE International Ultrasonics Symposium (IUS)* (2021):1–4. doi: 10.1109/IUS52206.2021.9593426

42. Yongfeng Z, Ping Z, Hong P, Wengang L, Yan Z. Superb microvascular imaging compared with contrast-enhanced ultrasound to assess microvessels in thyroid nodules. *J Med Ultrason (2001)* (2020) 47(2):287–97. doi: 10.1007/s10396-020-01011-z

43. Dencks S, Piepenbrock M, Opacic T, Krauspe B, Stickeler E, Kiessling F, et al. Clinical pilot application of super-resolution US imaging in breast cancer. *IEEE Trans Ultrason Ferroelectr Freq Control* (2019) 66(3):517–26. doi: 10.1109/TUFFC.2018.2872067

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 10 November 2022
DOI 10.3389/fonc.2022.990603



OPEN ACCESS

EDITED BY
Vikram D. Kodibagkar,
Arizona State University, United States

REVIEWED BY
Pietro Giorgio Calò,
University of Cagliari, Italy
Liu Yiqiang,
Beijing Cancer Hospital, China

*CORRESPONDENCE
Jun Li
1287424798@qq.com

†These authors have contributed equally to this article

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 10 July 2022
ACCEPTED 05 October 2022
PUBLISHED 10 November 2022

CITATION
Wang S-R, Li Q-L, Tian F, Li J, Li W-X, Chen M, Sang T, Cao C-L and Shi L-N (2022) Diagnostic value of multiple diagnostic methods for lymph node metastases of papillary thyroid carcinoma: A systematic review and meta-analysis.
*Front. Oncol.* 12:990603.
doi: 10.3389/fonc.2022.990603



# Diagnostic value of multiple diagnostic methods for lymph node metastases of papillary thyroid carcinoma: A systematic review and meta-analysis

Si-Rui Wang[1,2†], Qiao-Li Li[1,2†], Feng Tian[3], Jun Li[1,2*], Wen-Xiao Li[1,2], Ming Chen[1], Tian Sang[1,2], Chun-Li Cao[1] and Li-Nan Shi[1,2]

[1]Department of Ultrasound, the First Affiliated Hospital of Medical College, Shihezi University, Shihezi, Xinjiang, China, [2]NHC Key Laboratory of Prevention and Treatment of Central Asia High Incidence Diseases (First Affiliated Hospital, School of Medicine, Shihezi University), Shihezi, Xinjiang, China, [3]Department of Neurology, the First Affiliated Hospital of Medical College, Shihezi University, Shihezi, Xinjiang, China

**Objective:** This study compared the diagnostic value of various diagnostic methods for lymph node metastasis (LNM) of papillary thyroid carcinoma (PTC) through network meta-analysis.

**Methods:** In this experiment, databases such as CNKI, Wanfang, PubMed, and Web of Science were retrieved according to the Cochrane database, Prisma, and NMAP command manual. A meta-analysis was performed using STATA 15.0, and the value of the surface under the cumulative ranking curve (SUCRA) was used to determine the most effective diagnostic method. Quality assessments were performed using the Cochrane Collaboration's risk of bias tool, and publication bias was assessed using Deeks' funnel plot.

**Results:** A total of 38 articles with a total of 6285 patients were included. A total of 12 diagnostic methods were used to study patients with LNM of PTC. The results showed that 12 studies were direct comparisons and 8 studies were indirect comparisons. According to the comprehensive analysis of the area of SUCRA, US+CT(86.8) had the highest sensitivity, FNAC had the highest specificity (92.4) and true positive predictive value (89.4), and FNAC+FNA-Tg had higher negative predictive value (99.4) and accuracy (86.8). In the non-invasive method, US+CT had the highest sensitivity, and the sensitivity (SEN) was [OR=0.59, 95% confidence interval (CI): (0.30, 0.89]. Among the invasive methods, the combined application of FNAC+FNA-Tg had higher diagnostic performance. The sensitivity was [OR=0.62, 95% CI: (0.26, 0.98)], the specificity (SPE) was [OR=1.12, 95% CI: (0.59, 1.64)], the positive predictive value was [OR=0.98, 95% CI: (0.59, 1.37)], the negative predictive value was [OR=0.64, 95% CI (0.38, 0.90)], and the accuracy was [OR=0.71, 95% CI: (0.31, 1.12)].

10.3389/fonc.2022.990603

**Conclusion:** In the non-invasive method, the combined application of US+CT had good diagnostic performance, and in the invasive method, the combined application of FNAC+FNA-Tg had high diagnostic performance, and the above two methods were recommended.

KEYWORDS

lymph nodes metastasis, diagnostic value, network meta-analysis, multiple diagnostic methods, papillary thyroid carcinoma

# 1 Introduction

Thyroid cancer (TC) is one of the most common endocrine tumors worldwide, with an incidence rate of 3.1% and a mortality rate of 0.4% (1). The increase in papillary thyroid carcinoma (PTC) is the main reason for the increased incidence of adenocarcinoma (2). Although PTC has a good prognosis, the probability of distant cervical lymph node metastasis (LNM) in PTC patients reaches 90% (3), so the status of LN is also an important basis for judging recurrence and LN dissection (4). At the same time, for patients with LNM of PTC, the operation caused by persistent LNs recurrence will increase the risk of postoperative complications such as dyspnea, asphyxia, hypoparathyroidism, etc. (5, 6). Therefore, preoperative diagnosis is an important means for PTC patients to avoid persistent LNs recurrence and reduce complications.

At present, the commonly used methods for diagnosing the metastasis of LNM of PTC include CT, magnetic resonance imaging (MRI), IWBS, US, strain elastography (SE-US), Contrast-Enhanced Ultrasonography (CEUS), FNAC and other detection methods. But each of these methods has its advantages and disadvantages; In a single diagnostic method, ultrasound (US) is the main basis for clinical diagnosis of PTC (7), and the accuracy of the ultrasound results in the neck is high (8). However, the identification of LNM of PTC often needs to be judged by doctors' experience, which is often regarded as an inaccurate method (9). For patients with lymph node infiltration or distant metastasis of DTC after surgical resection, ultrasound-guided fine-needle aspiration cytology (FNAC) can be performed, and the sensitivity can reach 70–80% (10), while the sensitivity of computed tomography (CT) can reach 94.5% (11). In recent years, combined diagnostic methods such as US +CT (12) and US+ Contrast-Enhanced Ultrasonography (CEUS) (13) have also been commonly used to diagnose the LNM of PTC. Among these studies, the preoperative diagnostic performance of different diagnostic methods for LNM of PTC has been evaluated and analyzed or only one or two diagnostic methods have been compared, but a thorough evaluation has not been performed. Since different diagnostic methods have different diagnostic efficacy, and the diagnostic efficacy of the same diagnostic method is different in different studies, we conducted a meta-analysis to evaluate the diagnostic performance of different diagnostic methods for LNM of PTC so as to obtain the optimal diagnostic protocol.

In this study, we summarized the available evidence, investigated and compared a variety of different diagnostic techniques through network meta-analysis and the use of two or more published direct comparative studies, and concluded the optimal diagnostic protocol for LNM of PTC through comprehensive analysis. The results of this study will provide more evidence-based data for the development of guidelines and will guide patients with LNM of PTC to use appropriate diagnostic methods for preoperative or postoperative evaluation.

**Abbreviations:** PTC, papillary thyroid carcinoma; PTMC, papillary thyroid microcarcinoma; FTC, Follicular thyroid carcinoma; MTC, Medullary thyroid cancer; SE-US,strain elastography; FNAC, ultrasound-guided fine-needle aspiration cytology; FNA-Tg, fine-needle aspiration thyroglobulin; IWBS,131I whole-body; 18F-FDG PET/CT, F-18 fluorodeoxy glucose positron emission tomography/computed tomography; CT, Computed tomography; US, Ultrasound; CEUS, Contrast-Enhanced Ultrasonography; MRI,Magnetic Resonancediffusion Tension Imaging; SEN, Sensitivity; SPE, Specificity; PPV, positive predictive value; NPV, negative predictive value.

# 2 Methods

## 2.1 Retrieval strategy

We used keywords such as "Thyroid Neoplasms", "lymph node", "Neoplasm Staging", "Lymphatic Metastasis", and"Elasticity Imaging Techniques" in PubMed, Embase, Web of Science, CNKI, and Wanfang databases for retrieval. In order to obtain more sufficient data, we also screened the references of the retrieved articles (Table 1).

### 2.1.1 Inclusion and exclusion criteria

Inclusion criteria: ① Research subjects: PTC patients diagnosed with LNM. ② Study type: A randomized controlled

TABLE 1 Search strategy.

| No. | Retrieval type |
|---|---|
| #1 | lymph node【Mesh】 |
| #2 | Neoplasm Staging【Mesh】 |
| #3 | Staging, Neoplasm【Title/Abstract】 |
| #4 | Tumor Staging【Title/Abstract】 |
| #5 | TNM Staging System【Title/Abstract】 |
| #6 | TNM Classifications【Title/Abstract】 |
| #7 | preoperative staging【Title/Abstract】 |
| #8 | Lymphatic Metastasis【Mesh】 |
| #9 | Lymphatic Metastases【Title/Abstract】 |
| #10 | Lymph Node Metastasis【Title/Abstract】 |
| #11 | Lymph Node Metastasis【Title/Abstract】 |
| #12 | Metastasis, Lymph Node【Title/Abstract】 |
| #13 | #1OR #2 OR #3 OR#4OR#5OR#6 OR #7 OR#8 OR#9OR #10 OR#11OR#12 |
| #14 | Thyroid Neoplasms【Mesh】 |
| #15 | Neoplasm, Thyroid【Title/Abstract】 |
| #16 | thyroid carcinoma【Title/Abstract】 |
| #17 | thyroid cancer【Title/Abstract】 |
| #18 | thyroid neoplasm【Title/Abstract】 |
| #19 | Cancer of the Thyroid【Title/Abstract】 |
| #20 | Thyroid Cancers【Title/Abstract】 |
| #21 | #14OR#15OR#16 OR #17 OR#18 OR#19OR #20 |
| #22 | ultrasound contrast【Title/Abstract】 |
| #23 | Elasticity Imaging Techniques【Mesh】 |
| #24 | Elastography【Title/Abstract】 |
| #25 | Elastograms【Title/Abstract】 |
| #26 | B-mode【Title/Abstract】 |
| #27 | Ultrasonography【Mesh】 |
| #28 | Diagnostic Ultrasound【Title/Abstract】 |
| #29 | Ultrasound Imaging【Title/Abstract】 |
| #30 | Ultrasonic Imaging【Title/Abstract】 |
| #31 | Ultrasonic Diagnosis【Title/Abstract】 |
| #32 | Ultrasonic Diagnosis【Title/Abstract】 |
| #33 | Ultrasound-Guided Fine-Needle aspiration【Title/Abstract】 |
| #34 | Magnetic Resonance Imaging【Mesh】 |
| #35 | MRI Scan【Title/Abstract】 |
| #36 | Positron Emission Tomography Computed Tomography【Mesh】 |
| #37 | PET-CT Scan【Title/Abstract】 |
| #38 | 18F-FDG PET/CT【Title/Abstract】 |
| #39 | 131I【Title/Abstract】 |
| #40 | IWBS【Title/Abstract】 |
| #41 | 131I scintigraphy【Title/Abstract】 |
| #42 | #23 OR#24OR#25OR#26 OR #27 OR#28 OR#29OR #30 OR#31OR#32#33 OR#34OR#35OR#36 OR #37 OR#38 OR#39OR #40 OR#41 |
| #43 | #13 AND#21AND#42 |

study was conducted and two or more functional or non-invasive diagnostic methods should be included. ③ Gold standard: postoperative histopathology reports. ④Outcome indicators: It can reflect the sensitivity (SEN), specificity (SPE), positive predictive value (PPV), negative predictive value (NPV), and accuracy of various diagnostic methods for LNM in patients with PTC.

Exclusion criteria: ① There were less than 2 diagnostic methods in the study; ② There were no clear inclusion and exclusion criteria in the study; ③Reviews and lecture-type literature. ④ Literature for which the full text cannot be obtained.

## 2.2 Literature screening

Two researchers (Qiao-Li Li, Si-Rui Wang) read the titles and abstracts of the retrieved literature respectively. According to the inclusion and exclusion criteria established in this study, the literatures that did not meet the inclusion criteria were excluded, and the literatures that might meet the inclusion criteria and other related literatures were obtained and the relevant literatures were further intensively read. Articles with disagreements that were difficult to determine whether to be included were determined through discussion or consultation with a third party (Jun Li).

## 2.3 Quality evaluation of literature

For the RCT study, according to the method provided by the Cochrane Handbook, the research group adopted the risk of bias assessment tool of the Cochrane Collaboration (14) (RevMan v.5.3.5, Cochrane Collaboration, Oxford, UK), evaluated the methodological quality of the included studies from 6 aspects, and made the judgment of "yes" (low bias), "no" (high bias) and "unclear" (lack of relevant information or uncertainty of bias).

## 2.4 Data extraction

The data extracted in this study mainly included: (1) Characteristics of studies in the literature (author, publication time, country, study type, gold standard, etc.) (2) Subject characteristics (sample size, gender of patients, mean age or age range) (3) Effect indicators.

## 2.5 Statistical analysis

We grouped them according to different diagnostic methods and performed a network meta-analysis using the extracted diagnostic tools for the diagnostic performance of lymph node metastases.

We used Stata software (version-15.1) to perform the aggregation and analysis of NMA using Markov Chain Monte Carlo Subset Simulation in a Bayesian-based framework

according to the instruction manual for Prisma NMA (15). We used the nodal method to quantify and demonstrate the agreement between direct and indirect comparisons. Through the calculation of the instructions in the Stata software and whether the P value was greater than 0.05, it was judged whether the consistency check was passed.

Network diagrams of different exercise interventions were presented and described using Stata software. In the resulting network diagram, each node represented a different diagnostic approach, and the lines connecting the nodes represented direct head-to-head comparisons between interventions. The size of each node and the width of the connecting lines were proportional to the number of studies.

The diagnostic performance of each diagnostic method was analyzed by the area under SUCRA, and the certainty that one method was superior to the other was measured. Although SUCRA could effectively express the percentage of diagnostic performance of each diagnostic method, there was a possibility that there was no actual clinical significance between the diagnostic methods. In order to detect whether there was publication bias in some studies, funnel plots were generated for each diagnostic efficacy, and symmetry criterion were used to check.

# 3 Results

## 3.1 Selection and characteristics of literature

A total of 4490 articles were extracted through literature search of the database and reference extraction, of which we excluded 1247 duplicate articles. Of the remaining 3243 articles, there were 2829 articles, 59 pathology reports, 699 review articles, 10 letters, and 72 meta-analyses and systematic reviews. 2061 articles not related to this study were excluded by review of the title and abstract. Among the remaining 414 articles, 376 articles failed to obtain the full text or failed to meet the inclusion criteria, and finally 38 articles were included, with a total of 6285 people. (Figure 1). The articles we included all fulfilled the requirement that the study population was patients with LNM of PTC and that their diagnoses were shown preoperatively or postoperatively by two or more diagnostic methods, such as US, CT, and MRI, and the data for direct comparison in the results were evaluated (Table 2).

## 3.2 Quality assessment and publication bias

The 38 articles we included (11–13, 15–45, 47, 49) included 36 retrospective studies and 2 prospective studies. We performed a network meta-analysis using STATA 15.0 (14), and 38 articles were assessed for quality, risk of bias and applicability issues using QUADAS-2 (50). The overall quality of the articles was satisfactory, but there may be potential bias in personnel selection. Among the 38 articles, 8 articles had unclear risk of bias, and 2 articles had high risk. In terms of index testing, because some doctors did not strictly implement blinding in the processing of results, there were 12 literatures with unclear risk of bias. In terms of reference standard bias, there were 6 articles with unclear risk of bias, because it was not indicated whether there was an appropriate time interval between the trial to be

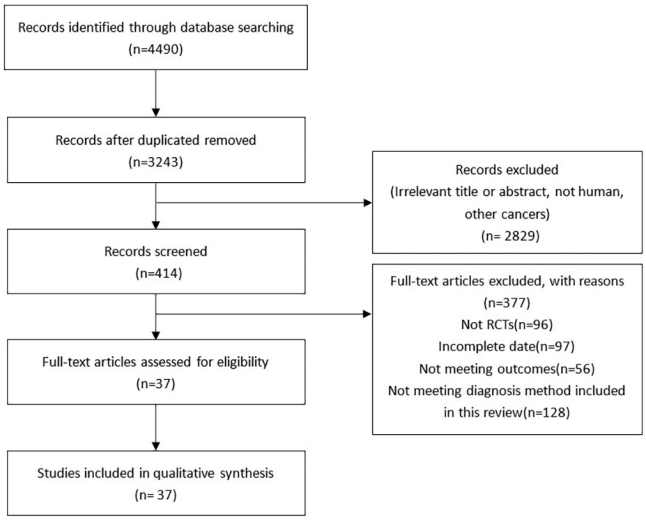



FIGURE 1
Flow diagram of literature selection.

TABLE 2 Characteristics of the studies included in the meta-analysis.

| First Author | Year | Country | Design | NO.of patients | Gender (M/F) | Histology | LN level | Diagnostic method | NO.of patients | Gender (M/F) | Histology |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Artür Salmaslıoğlu** (16) | 2010 | Turkey | RCT | 225 | 60/165 | 225PTC | ALL | ③ | ④ | | |
| **Ari Chong** (17) | 2017 | Korea | RCT | 85 | 71/14 | 85PTC | C, L | ⑦ | ⑥ | | |
| **Byung Hyun Byun** (18) | 2012 | Korea | RCT | 200 | 41/255 | 200PTC | ALL, C, L | ② | ⑥ | | |
| **C. De Crea** (19) | 2014 | Italy | RCT | 38 | 16/22 | 38PTC | C | ④ | ③ | | |
| **Cesur Samanci** (20) | 2019 | Turkey | RCT | 32 | 13/19 | 26PTC, 6FTC | ALL | ⑥ | ⑨ | | |
| **D.W.Lee** (21) | 2013 | Korea | RCT | 252 | 45/207 | 252PTC | C, L | ② | ⑦ | ②+⑦ | |
| **David Lesnik** (22) | 2013 | USA | RCT | 162 | NA | 162PTC | C, L | ② | ⑦ | ②+⑦ | ⑥ |
| **Enke Baldini** (23) | 2013 | Italy | RCT | 28 | 7/21 | 21PTC, 2MTC,2ATC | ALL | ③ | ④ | | |
| **Eunhee Kim** (13) | 2008 | Korea | RCT | 165 | 25/140 | 165PCT | ALL, C, L | ② | ⑦ | ②+⑦ | |
| **Eun NL** (24) | 2017 | Korea | RCT | 302 | 76/226 | 302PTC | ALL | ③ | ③+④ | | |
| **Farzana Alam** (25) | 2008 | Japan | RCT | 37 | 12/25 | 37PTC | ALL | ② | ① | | |
| **Han Sin Jeong** (26) | 2006 | Korea | RCT | 26 | 7/19 | 26PTC | ALL | ⑥ | ② | | |
| **Harry S. Hwang** (27) | 2009 | USA | RCT | 42 | NA | 42PTC | ALL | ⑥ | ② | | |
| **Jasna Mihailovic** (28) | 2010 | Serbia | RCT | 40 | 15/25 | 39PTC, 1FTC | ALL | ⑨ | ⑤ | | |
| **Jeon SJ** (29) | 2009 | Korea | RCT | 47 | NA | 47PTC | ALL | ③ | ④ | ③+④ | |
| **Ji Soo Choi** (30) | 2009 | Korea | RCT | 299 | 44/255 | 186PTMC, 113PTC | C, L | ② | ⑦ | ②+⑦ | |
| **Ji-Hoon Jung** (31) | 2015 | Korea | RCT | 193 | 25/168 | 151PTMC, 42PTC | ALL | ⑥ | ② | | |
| **Jia Zhan** (32) | 2019 | China | RCT | 56 | 16/40 | 56PTC | C | ② | ⑧ | | |
| **Johann-Martin Hempel** (33) | 2016 | Germany | RCT | 46 | 27/19 | 16FCT, 29PTC | ALL | ⑥ | ⑨ | | |
| **Jun Ho Lee** (21) | 2015 | Korea | RCT | 78 | 26/52 | 78PTC | ALL | ③ | ④ | | |
| **Kim MJ** (34) | 2012 | Korea | PRO | 68 | 16/52 | 68PTC | ALL | ④ | ③ | | |
| **Khadra** (35) | 2018 | USA | RCT | 138 | 31/106 | 138PTC | ALL | ③ | ④ | ③+④ | |
| **Lei Chen** (13) | 2019 | China | RCT | 46 | 6/40 | 46PTC | ALL | ② | ⑧ | ②+ ⑧ | |
| **Li J** (36) | 2016 | China | RCT | 124 | 23/101 | 124PTC | ALL | ③ | ④ | ③+④ | |
| **Shi JH** (37) | 2015 | China | RCT | 148 | NA | 148PTC | ALL | ④ | ③ | ③+④ | |
| **So Yeon Yang** (38) | 2019 | Korea | RCT | 453 | 108/345 | 453PTC | ALL | ② | ⑦ | ②+⑦ | |
| **Sohn YM** (39) | 2012 | Korea | RCT | 92 | NA | 92PTC | ALL | ④ | ③ | | |
| **Suh YJ** (40) | 2013 | Korea | RCT | 43 | NA | 43PTC | ALL | ④ | ③ | ③+④ | |
| **Tong Tong Liu** (41) | 2018 | China | RCT | 75 | 20/55 | 75PTC | C | ② | ① | ②+① | |
| **Yanfang Wang** (42) | 2021 | China | RCT | 120 | 37/83 | 120PTC,165PTMC | C | ② | ⑧ | | |
| **Ying Liu** (43) | 2021 | China | RCT | 600 | 211/389 | 600PTC | C | ② | ⑦ | ②+⑦ | |
| **Ying Wei** (44) | 2020 | China | RCT | 24 | 14/10 | 24PTC | ALL | ② | ⑧ | | |
| **Yoon Jung Choi** (45) | 2010 | Korea | RCT | 589 | 121/468 | 589PTC | C, L | ② | ⑦ | | |
| **Young Lan Seo** (12) | 2012 | Korea | RCT | 20 | 4/16 | 19PTC 1PTMC | ALL | ⑥ | ② | ⑦ | |

*(Continued)*

TABLE 2 Continued

| First Author | Year | Country | Design | NO.of patients | Gender (M/F) | Histology | LN level | Diagnostic method | NO.of patients | Gender (M/F) | Histology |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Yurong Hong** (46) | 2017 | China | RCT | 573 | 148/425 | 162PTC | All | ② | ⑧ | ②+⑧ | |
| **Zeming Liu** (47) | 2014 | China | RCT | 70 | NA | 70DTC | C, L | ⑨ | ② | | |
| **Zhao H** (48) | 2017 | China | PRO | 189 | NA | 189PTC | ALL | ④ | ③ | ③+④ | |

Lymph node level (LN level); All: All cervical LN; C: Central LN; L: Lateral LN.
PTC, papillary thyroid carcinoma; PTMC, papillary thyroid microcarcinoma; FTC, Follicular thyroid carcinoma; MTC, Medullary thyroid cancer;
① SE-US, strain elastography; ② US, Ultrasound.; ③ FNAC, fine-needle aspiration cytology; ④ FNA-Tg, fine-needle aspiration thyroglobulin; ⑤ IWBS, 131I whole-body; ⑥ 18F-FDG PET/CT, F-18 fluorodeoxyglucose positron emission tomography/computed tomography; ⑦ CT, Computed tomography; ⑧ CEUS, Contrast-Enhanced Ultrasonography; ⑨ MRI, Magnatic Resonance Imaging.

evaluated and the gold standard. Two articles had unclear risk of bias with respect to follow-up time. In terms of applicability, all literatures showed no high risk of bias in patient selection, index test, and reference standard. (Figure 2) The authors' assessment of each domain for included study. This study used funnel plots to detect possible publication bias, and the results showed that the distribution of funnel plots was roughly symmetric, suggesting that there was no publication bias or other bias in the study (Figure 3).

## 3.3 Pairwise meta-analysis

Through the results of SEN, SPE, PPV, NPV, and Accuracy of various diagnostic methods, a network meta-analysis graph can be made (Figure 4), in which the gridlines in the analysis diagram represented a direct comparison between the two groups of diagnostic methods, the thickness of the gridlines represented the number of articles included, and the size of the solid dots at both ends of the gridlines represented the sample size. Of the 38 studies compared, 12 were direct comparisons and 8 were indirect comparisons. We performed a direct pairwise comparison of SEN, SPE, PPV, NPV, and Accuracy for each method of diagnosing LNM of PTC. Meta-analyses of the results can be used to make direct comparisons. There were 3 studies comparing MRI with PET/CT, 2 studies comparing CT with FNA-Tg, 16 studies comparing US with CT, 6 studies comparing FNAC with FNA-Tg, 3 studies comparing US+CEUS with US, 5 studies comparing CT with PET/CT, 7 studies comparing US with CEUS, 2 studies comparing CT with FNAC, 2 studies comparing US with FNAC, 9 studies comparing US+CT with US, 6 studies comparing US with PET/CT, 3 studies comparing US+CEUS with CEUS, 9 studies comparing US+CT with CT, 2 studies comparing MRI with IWBC, 3 studies comparing US with SE-US, 3 studies comparing FNA-Tg with FNA-Tg+FNAC, and 3 studies comparing FNAC with FNA-Tg+FNAC. As shown in the figure.

## 3.4 Network meta-analysis

The OR values and 95% confidence intervals (CI) of SEN, SPE, PPV, NPV, and Accuracy measured by different diagnostic methods for LNM of PTC were statistically analyzed, and the statistical differences were judged by calculation.

### 3.4.1 SEN

According to the area under the cumulative ranking curve (SUCRA), the SEN of different diagnostic methods for lymph node metastasis can be obtained. The descending order was as follows: US+CT(86.8)>FNAC+FNA-Tg(86.6) >FNA-Tg(81.8) >CT(70.2)>US(51.8)>CEUS(44.5)>FNAC(42.5)>SE-US(42.2) >MRI(33.4)>IWBC(31.7)>PET/CT(27.5)>US+CEU (1.0). Among them, US+CT (OR=0.59, 95%CrI:(0.30,0.89)) and FNAC+FNA-Tg[OR=0.62,95%CI:(0.26,0.98)] ranked first and second, respectively, and had significant advantages compared with other diagnostic methods. There was significant heterogeneity between MRI and US, PET/CT in terms of sensitivity (P<0.05). The probability ranking of diagnosis methods to SEN was ranked first by US+CT in the SUCRA (SUCRA:86.8% as shown in Figure 5). A comparison between the two different diagnosis methods was shown in (Figure 6).

### 3.4.2 SPE

According to the area under the cumulative ranking curve (SUCRA), the SPE of different diagnostic methods for lymph node metastasis can be obtained. The descending order was as follows: FNAC(92.4)>FNAC+FNA-Tg(85.8)>FNA-Tg(80.5)>CEUS(75.4) >SE-US(64.4)>US(48.5)>CT(42.9)>PET/CT(34.5)>IWBC(32.6) >US+CT(27.1)>MRI(15.5)>US+CEUS(0.4). The analysis of the results showed that compared with US+CEUS, FNAC [OR=1.16,95%CI:(0.65,1.66)] and FNAC+FNA-Tg [OR=1.12,95%CI:(0.59,1.64)] had significant advantages. In terms of specificity (P<0.05), there was significant heterogeneity in the comparison between FNAC and US, FNA-Tg and US, and FNAC and CT. The probability ranking of diagnosis methods to


FIGURE 2
Bias risk of the included studies (QUADAS 2 criteria).

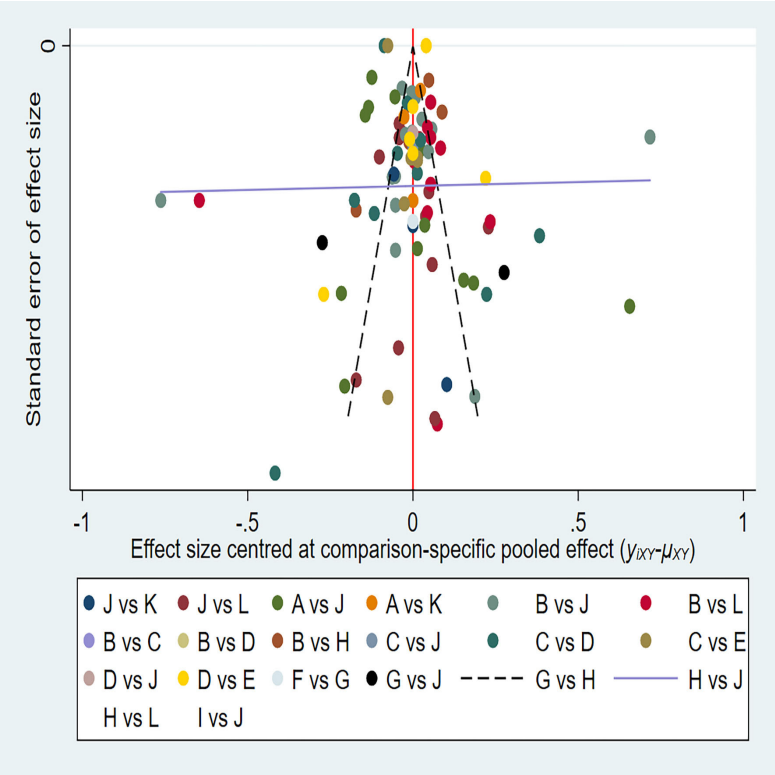



FIGURE 3
**A** = CEUS, **B** = CT, **C** = FNA-Tg, > **D** = FNAC, **E** = FNAC+FNA-Tg, **F** = IWBC, **G** = MRI, **H** = PET/CT, **I** = SE-US, **J** = US, **K** = US+CEUS, **L** = US+CT.

SPE was ranked first by FNAC in the SUCRA (SUCRA:92.4% as shown in Figure 7). A comparison between the two different diagnosis methods was shown in Figure 8.

### 3.4.3 PPV

According to the area under the cumulative ranking curve (SUCRA), the PPV of different diagnostic methods for lymph node metastasis can be obtained. The descending order was as follows: FNAC(89.4)>FNAC+FNA-Tg(81.6)>CEUS(77.6)>FNA-Tg(73.9)>SE-US(58.7)>US(51.3)>PET/CT(48.7)>CT(44.5)>US+CT(43.9)>IWBC(18.6)>MRI(11.3)>US+CEUS(0.3). The analysis of the results showed that compared with US+CEUS, FNAC[OR=1.01,95%CI:(0.63,1.38)] and FNAC+FNA-Tg[OR=0.98,95%CI:(0.59,1.37)] had significant advantages. There was no significant difference in specificity (P<0.05). The probability ranking of diagnosis methods to PPV was ranked first by FNAC in the SUCRA (SUCRA: 89.4% as shown in Figures 9). A comparison between the two different diagnosis methods was shown in Figure 10.

### 3.4.4 NPV

According to the area under the cumulative ranking curve (SUCRA), the NPV of different diagnostic methods for lymph

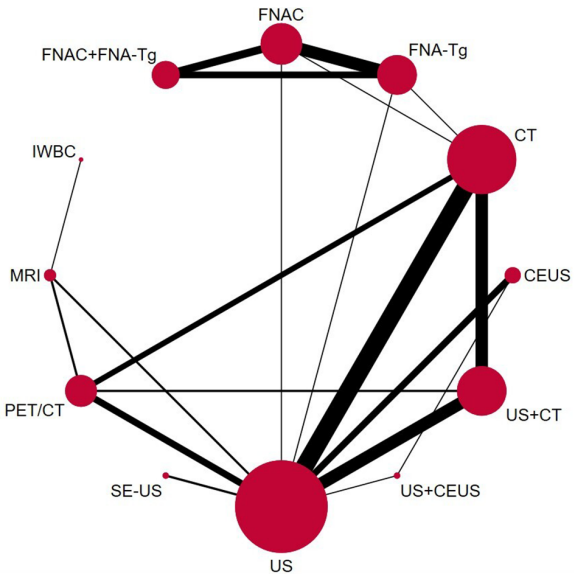



FIGURE 4
Network Mate-Analysis Figure.

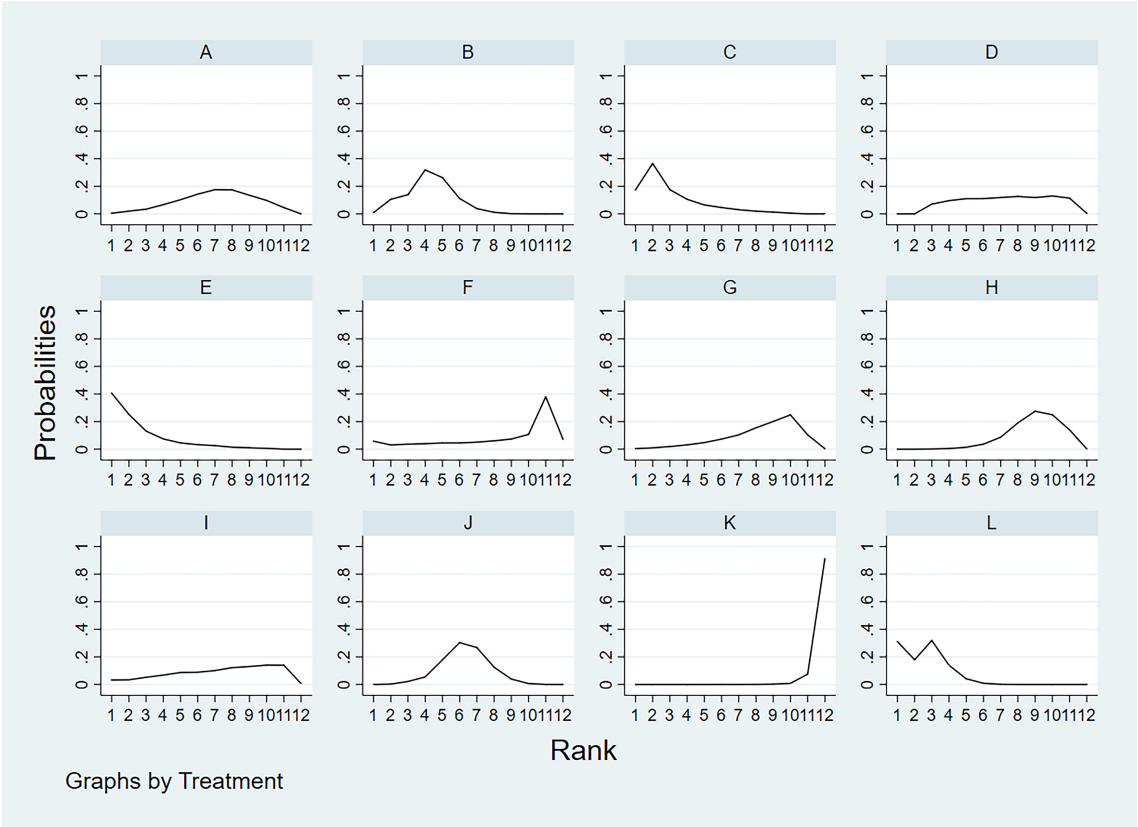



FIGURE 5
**A** = CEUS, **B** = CT, **C** = FNA-Tg, > **D** = FNAC, **E** = FNAC+FNA-Tg, **F** = IWBC, **G** = MRI, **H** = PET/CT, **I** = SE-US, **J** = US, **K** = US+CEUS, **L** = US+CT.

node metastasis can be obtained. The descending order was as follows: FNAC+FNA-Tg(99.4)>FNA-Tg(89.0)>US+CT(72.2)>CEUS(65.6)>CT(58.5)>US(52.0)>FNAC(49.9)>IWBC(31.4)>SE-US(28.8)>PET/CT(27.1)>MRI(25.6)>US+CEUS(0.6). The analysis of the results showed that compared with US+CEUS, FNAC+FNA-Tg (OR=0.64, 95%CrI: (0.38, 0.90) and FNA-Tg [OR=0.57, 95%CI: (0.31, 0.82)] had significant advantages. In terms of specificity (P<0.05), there was significant heterogeneity between MRI and US, PET/CT. The probability ranking of diagnosis methods to NPV was ranked first by FNAC+FNA-Tg in the SUCRA (SUCRA: 99.4% as shown in Figures 11). A comparison between the two different interventions was shown in Figure 12.

### 3.4.5 Accuracy

According to the area under the cumulative ranking curve (SUCRA), the Accuracy of different diagnostic methods for lymph node metastasis can be obtained. The descending order was as follows: FNAC+FNA-Tg(99.4)>CT(81.6)>FNA-Tg(76.0)>SE-US(61.6)>FNAC(59.8)>CEUS(57.8)>US+CT(52.4)>IWBC

| US+CT | FNAC+FNA-Tg | FNA-Tg | CT | US | CEUS | FNAC | SE-US | MRI | IWBC | PET/CT | US+CEUS |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **US+CT** | 0.02 (-0.21,0.26) | 0.00 (-0.22,0.22) | -0.07 (-0.15,0.02) | **-0.12 (-0.21,-0.04)** | -0.15 (-0.30,-0.00) | -0.16 (-0.38,0.07) | -0.17 (-0.40,0.06) | **-0.20 (-0.38,-0.02)** | -0.25 (-0.64,0.14) | **-0.21 (-0.33,-0.09)** | **-0.59 (-0.89,-0.30)** |
| -0.02 (-0.26,0.21) | **FNAC+FNA-Tg** | -0.02 (-0.13,0.08) | -0.09 (-0.32,0.14) | -0.15 (-0.38,0.08) | -0.18 (-0.43,0.08) | **-0.18 (-0.28,-0.08)** | -0.19 (-0.51,0.12) | -0.22 (-0.51,0.06) | -0.27 (-0.72,0.17) | -0.24 (-0.48,0.01) | **-0.62 (-0.98,-0.26)** |
| -0.00 (-0.22,0.22) | 0.02 (-0.08,0.13) | **FNA-Tg** | -0.07 (-0.28,0.14) | -0.13 (-0.34,0.09) | -0.15 (-0.40,0.09) | **-0.16 (-0.25,-0.07)** | -0.17 (-0.47,0.13) | -0.20 (-0.47,0.07) | -0.25 (-0.68,0.18) | -0.21 (-0.44,0.02) | **-0.59 (-0.95,-0.24)** |
| 0.07 (-0.02,0.15) | 0.09 (-0.14,0.32) | 0.07 (-0.14,0.28) | **CT** | -0.06 (-0.13,0.01) | -0.08 (-0.23,0.06) | -0.09 (-0.30,0.12) | -0.10 (-0.33,0.12) | -0.13 (-0.31,0.05) | -0.18 (-0.57,0.20) | **-0.14 (-0.25,-0.04)** | **-0.53 (-0.82,-0.24)** |
| **0.12 (0.04,0.21)** | 0.15 (-0.08,0.38) | 0.13 (-0.09,0.34) | 0.06 (-0.01,0.13) | **US** | -0.03 (-0.15,0.10) | -0.03 (-0.24,0.18) | -0.04 (-0.26,0.17) | -0.07 (-0.24,0.09) | -0.12 (-0.50,0.25) | -0.09 (-0.19,0.01) | **-0.47 (-0.75,-0.19)** |
| 0.15 (0.00,0.30) | 0.18 (-0.08,0.43) | 0.15 (-0.09,0.40) | 0.08 (-0.06,0.23) | 0.03 (-0.10,0.15) | **CEUS** | -0.01 (-0.25,0.24) | -0.02 (-0.27,0.23) | -0.05 (-0.26,0.16) | -0.10 (-0.50,0.30) | -0.06 (-0.22,0.10) | **-0.44 (-0.72,-0.16)** |
| 0.16 (-0.07,0.38) | **0.18 (0.08,0.28)** | **0.16 (0.07,0.25)** | 0.09 (-0.12,0.30) | 0.03 (-0.18,0.24) | 0.01 (-0.24,0.25) | **FNAC** | -0.01 (-0.31,0.29) | -0.04 (-0.31,0.23) | -0.09 (-0.53,0.34) | -0.06 (-0.29,0.17) | **-0.44 (-0.79,-0.09)** |
| 0.17 (-0.06,0.40) | 0.19 (-0.12,0.51) | 0.17 (-0.13,0.47) | 0.10 (-0.12,0.33) | 0.04 (-0.17,0.26) | 0.02 (-0.23,0.27) | 0.01 (-0.29,0.31) | **SE-US** | -0.03 (-0.30,0.24) | -0.08 (-0.52,0.36) | -0.04 (-0.28,0.19) | **-0.42 (-0.78,-0.07)** |
| **0.20 (0.02,0.38)** | 0.22 (-0.06,0.51) | 0.20 (-0.07,0.47) | 0.13 (-0.05,0.31) | 0.07 (-0.09,0.24) | 0.05 (-0.16,0.26) | 0.04 (-0.23,0.31) | 0.03 (-0.24,0.30) | **MRI** | -0.05 (-0.39,0.29) | -0.01 (-0.18,0.16) | **-0.39 (-0.72,-0.07)** |
| 0.25 (-0.14,0.64) | 0.27 (-0.17,0.72) | 0.25 (-0.18,0.68) | 0.18 (-0.20,0.57) | 0.12 (-0.25,0.50) | 0.10 (-0.30,0.50) | 0.09 (-0.34,0.53) | 0.08 (-0.36,0.52) | 0.05 (-0.29,0.39) | **IWBC** | 0.04 (-0.34,0.42) | -0.34 (-0.82,0.13) |
| **0.21 (0.09,0.33)** | 0.24 (-0.01,0.48) | 0.21 (-0.02,0.44) | **0.14 (0.04,0.25)** | 0.09 (-0.01,0.19) | 0.06 (-0.10,0.22) | 0.06 (-0.17,0.29) | 0.04 (-0.19,0.28) | 0.01 (-0.16,0.18) | -0.04 (-0.42,0.34) | **PET/CT** | **-0.38 (-0.68,-0.08)** |
| **0.59 (0.30,0.89)** | **0.62 (0.26,0.98)** | **0.59 (0.24,0.95)** | **0.53 (0.24,0.82)** | **0.47 (0.19,0.75)** | **0.44 (0.16,0.72)** | **0.44 (0.09,0.79)** | **0.42 (0.07,0.78)** | **0.39 (0.07,0.72)** | 0.34 (-0.13,0.82) | **0.38 (0.08,0.68)** | **US+CEUS** |

FIGURE 6
League table on SEN.

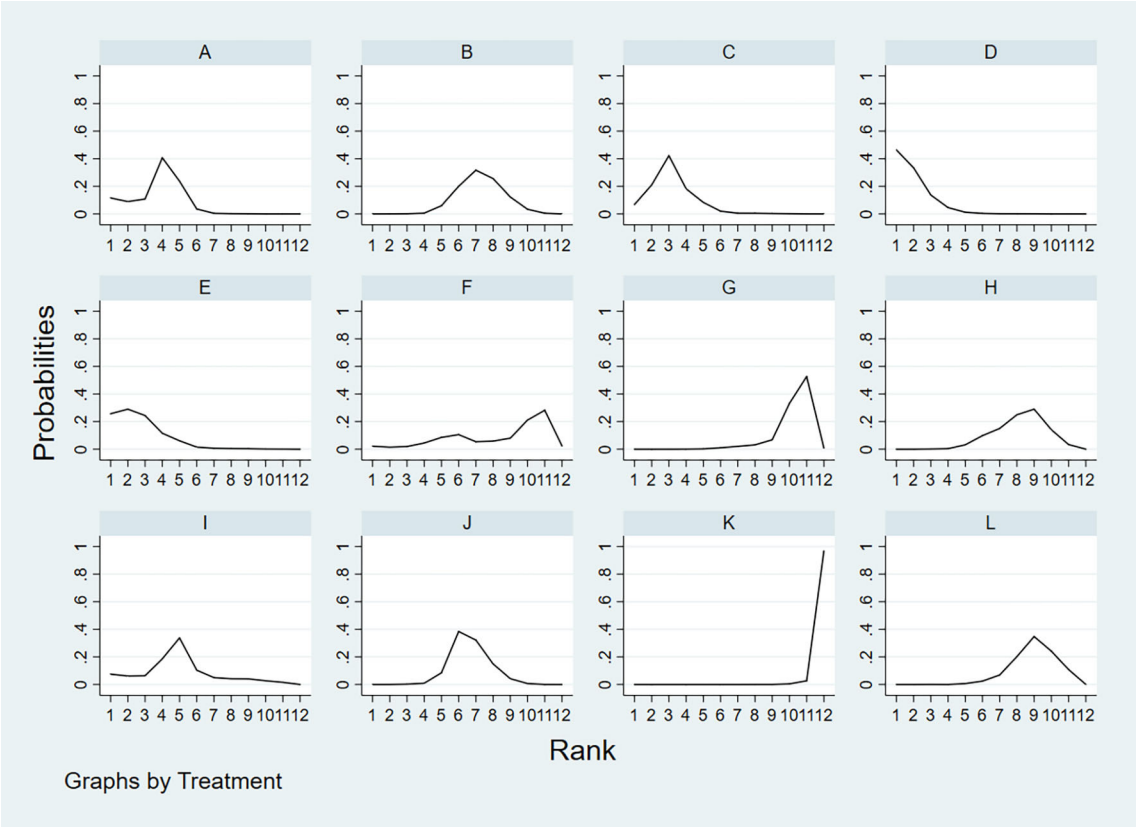



FIGURE 7
**A** = CEUS, **B** = CT, **C** = FNA-Tg, > **D** = FNAC, **E** = FNAC+FNA-Tg, **F** = IWBC, **G** = MRI, **H** = PET/CT, **I** = SE-US, **J** = US, **K** = US+CEUS, **L** = US+CT.

(39.0)>US(37.2)>MRI(28.6)>PET/CT(18.7)>US+CEUS(0.5). The analysis of the results showed that compared with US +CEUS, FNAC+FNA-Tg[OR=0.71,95%CI:(0.31,1.12)] and FNA-Tg[OR=0.67,95%CI:(0.35,0.99)] had significant advantages. In terms of specificity (P<0.05), there was significant heterogeneity between CT and US+CT, PET/CT and US, and PET/CT and US+CT. The probability ranking of diagnosis methods to ACC was ranked first by FNAC+FNA-Tg in the SUCRA (SUCRA: 99.4% as shown in Figure 13). A comparison between the two different diagnosis methods was shown in Figure 14.

## 4 Discussion

This meta-analysis aimed to evaluate the diagnostic performance of multiple diagnostic methods for predicting LNM of PTC. A total of 38 studies with 6285 patients were

| FNAC | FNAC+FNA-Tg | FNA-Tg | CEUS | SE-US | US | CT | PET/CT | IWBC | US+CT | MRI | US+CEUS |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FNAC | -0.04 (-0.22,0.15) | -0.08 (-0.23,0.07) | -0.17 (-0.54,0.19) | -0.26 (-0.71,0.18) | **-0.40 (-0.72,-0.08)** | **-0.43 (-0.75,-0.11)** | **-0.47 (-0.81,-0.12)** | -0.53 (-1.12,0.06) | **-0.50 (-0.84,-0.17)** | -0.63 (-1.03,-0.24) | **-1.16 (-1.66,-0.65)** |
| 0.04 (-0.15,0.22) | FNAC+FNA-Tg | -0.04 (-0.24,0.15) | -0.13 (-0.53,0.27) | -0.22 (-0.70,0.25) | **-0.36 (-0.72,-0.01)** | -0.39 (-0.74,-0.03) | **-0.43 (-0.81,-0.05)** | -0.49 (-1.11,0.12) | **-0.46 (-0.83,-0.09)** | **-0.59 (-1.02,-0.17)** | **-1.12 (-1.64,-0.59)** |
| 0.08 (-0.07,0.23) | 0.04 (-0.15,0.24) | FNA-Tg | -0.09 (-0.46,0.27) | -0.18 (-0.63,0.26) | -0.32 (-0.64,-0.00) | **-0.35 (-0.67,-0.03)** | **-0.39 (-0.73,-0.04)** | -0.45 (-1.04,0.14) | **-0.42 (-0.76,-0.09)** | **-0.55 (-0.95,-0.16)** | **-1.08 (-1.58,-0.57)** |
| 0.17 (-0.19,0.54) | 0.13 (-0.27,0.53) | 0.09 (-0.27,0.46) | CEUS | -0.09 (-0.45,0.27) | **-0.23 (-0.41,-0.05)** | **-0.25 (-0.46,-0.04)** | **-0.30 (-0.53,-0.06)** | -0.36 (-0.89,0.17) | **-0.33 (-0.55,-0.11)** | **-0.46 (-0.76,-0.16)** | **-0.98 (-1.37,-0.60)** |
| 0.26 (-0.18,0.71) | 0.22 (-0.25,0.70) | 0.18 (-0.26,0.63) | 0.09 (-0.27,0.45) | SE-US | -0.14 (-0.45,0.18) | -0.16 (-0.49,0.17) | -0.20 (-0.55,0.14) | -0.27 (-0.86,0.32) | -0.24 (-0.58,0.10) | -0.37 (-0.76,0.03) | **-0.89 (-1.39,-0.39)** |
| **0.40 (0.08,0.72)** | **0.36 (0.01,0.72)** | 0.32 (0.00,0.64) | **0.23 (0.05,0.41)** | 0.14 (-0.18,0.45) | US | -0.02 (-0.13,0.08) | -0.07 (-0.21,0.08) | -0.13 (-0.63,0.37) | -0.10 (-0.23,0.02) | -0.23 (-0.47,0.01) | **-0.75 (-1.14,-0.37)** |
| **0.43 (0.11,0.75)** | **0.39 (0.03,0.74)** | **0.35 (0.03,0.67)** | **0.25 (0.04,0.46)** | 0.16 (-0.17,0.49) | 0.02 (-0.08,0.13) | CT | -0.04 (-0.20,0.11) | -0.11 (-0.61,0.40) | -0.08 (-0.20,0.05) | -0.21 (-0.46,0.05) | **-0.73 (-1.13,-0.33)** |
| **0.47 (0.12,0.81)** | **0.43 (0.05,0.81)** | **0.39 (0.04,0.73)** | **0.30 (0.06,0.53)** | 0.20 (-0.14,0.55) | 0.07 (-0.08,0.21) | 0.04 (-0.11,0.20) | PET/CT | -0.06 (-0.56,0.44) | -0.03 (-0.21,0.14) | -0.16 (-0.40,0.08) | **-0.69 (-1.10,-0.27)** |
| 0.53 (-0.06,1.12) | 0.49 (-0.12,1.11) | 0.45 (-0.14,1.04) | 0.36 (-0.17,0.89) | 0.27 (-0.32,0.86) | 0.13 (-0.37,0.63) | 0.11 (-0.40,0.61) | 0.06 (-0.44,0.56) | IWBC | 0.03 (-0.48,0.54) | -0.10 (-0.54,0.34) | **-0.62 (-1.26,0.01)** |
| **0.50 (0.17,0.84)** | **0.46 (0.09,0.83)** | **0.42 (0.09,0.76)** | **0.33 (0.11,0.55)** | 0.24 (-0.10,0.58) | 0.10 (-0.02,0.23) | 0.08 (-0.05,0.20) | 0.03 (-0.14,0.21) | -0.03 (-0.54,0.48) | US+CT | -0.13 (-0.39,0.13) | **-0.65 (-1.06,-0.25)** |
| **0.63 (0.24,1.03)** | **0.59 (0.17,1.02)** | **0.55 (0.16,0.95)** | **0.46 (0.16,0.76)** | 0.37 (-0.03,0.76) | 0.23 (-0.01,0.47) | 0.21 (-0.05,0.46) | 0.16 (-0.08,0.40) | 0.10 (-0.34,0.54) | 0.13 (-0.13,0.39) | MRI | **-0.52 (-0.98,-0.07)** |
| **1.16 (0.65,1.66)** | **1.12 (0.59,1.64)** | **1.08 (0.57,1.58)** | **0.98 (0.60,1.37)** | **0.89 (0.39,1.39)** | **0.75 (0.37,1.14)** | **0.73 (0.33,1.13)** | **0.69 (0.27,1.10)** | 0.62 (-0.01,1.26) | **0.65 (0.25,1.06)** | **0.52 (0.07,0.98)** | US+CEUS |

FIGURE 8
League table on SPE.

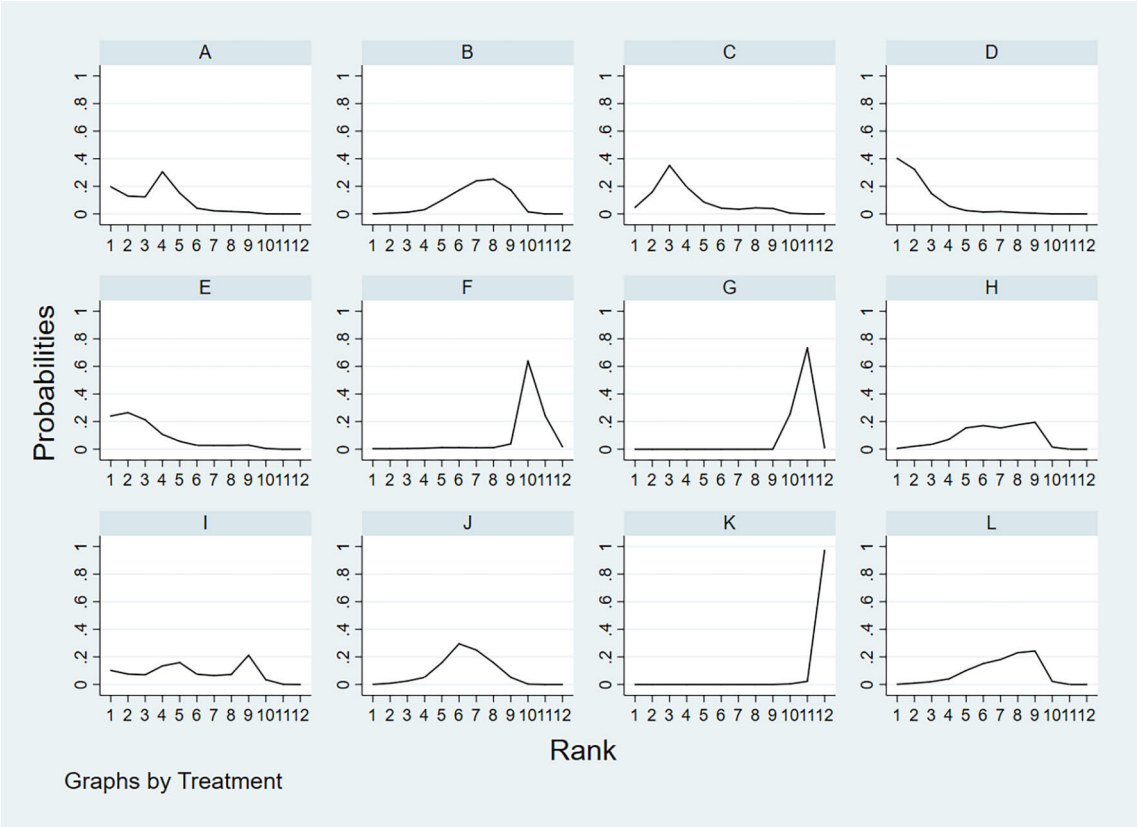



FIGURE 9
**A** = CEUS, **B** = CT, **C** = FNA-Tg, > **D** = FNAC, **E** = FNAC+FNA-Tg, **F** = IWBC, **G** = MRI, **H** = PET/CT, **I** = SE-US, **J** = US, **K** = US+CEUS, **L** = US+CT.

summarized and analyzed. Among the 12 diagnostic methods, we found that the combined diagnosis of CT+US and FNAC +FNA-Tg was the most effective diagnostic method for the recurrence and metastasis of LNM of PTC. A total of 6 articles were included in the comparison of US and CT, and the diagnostic performance of CT and US for LNM of PTC, such as sensitivity and specificity, were statistically analyzed. By analyzing the data we found that for the sensitivity of US and CT, the values of SUCRA were 51.8 and 70.2, respectively, while the value for the combined diagnosis of US+CT was 86.8. David (49) et al. found that the accuracy of US may be affected due to the difficulty in distinguishing between the metastatic lymph nodes immediately adjacent to the thyroid and the thyroid itself. In addition, some normal anatomical structures cause distinct acoustic shadows that effectively obscure the area behind them, including air-filled structures such as the larynx and trachea, as well as the clavicle and jawbone, areas deep in the sternum. This greatly limits the assessment of the mediastinal and retropharyngeal regions by ultrasound, regardless of the operator's interest and expertise. However, all of these areas

| FNAC | FNAC+FNA-Tg | CEUS | FNA-Tg | SE-US | US | PET/CT | CT | US+CT | IWBC | MRI | US+CEUS |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **FNAC** | -0.03 (-0.16,0.11) | -0.08 (-0.35,0.19) | -0.06 (-0.17,0.04) | -0.16 (-0.49,0.18) | -0.19 (-0.42,0.05) | -0.19 (-0.45,0.06) | -0.20 (-0.44,0.03) | -0.21 (-0.45,0.04) | **-0.51 (-0.95,-0.07)** | **-0.62 (-0.91,-0.32)** | **-1.01 (-1.38,-0.63)** |
| 0.03 (-0.11,0.16) | **FNAC+FNA-Tg** | -0.05 (-0.34,0.25) | -0.04 (-0.18,0.10) | -0.13 (-0.48,0.22) | -0.16 (-0.42,0.10) | -0.17 (-0.45,0.11) | -0.18 (-0.44,0.08) | -0.18 (-0.45,0.09) | **-0.48 (-0.94,-0.03)** | **-0.59 (-0.91,-0.28)** | **-0.98 (-1.37,-0.59)** |
| 0.08 (-0.19,0.35) | 0.05 (-0.25,0.34) | **CEUS** | 0.01 (-0.26,0.28) | -0.08 (-0.36,0.19) | -0.11 (-0.25,0.03) | -0.12 (-0.29,0.06) | -0.13 (-0.29,0.03) | -0.13 (-0.30,0.03) | **-0.43 (-0.83,-0.03)** | **-0.54 (-0.77,-0.32)** | **-0.93 (-1.22,-0.64)** |
| 0.06 (-0.04,0.17) | 0.04 (-0.10,0.18) | -0.01 (-0.28,0.26) | **FNA-Tg** | -0.09 (-0.43,0.24) | -0.12 (-0.36,0.11) | -0.13 (-0.38,0.12) | -0.14 (-0.37,0.09) | -0.14 (-0.39,0.10) | -0.45 (-0.89,-0.00) | **-0.56 (-0.85,-0.26)** | **-0.94 (-1.31,-0.57)** |
| 0.16 (-0.18,0.49) | 0.13 (-0.22,0.48) | 0.08 (-0.19,0.36) | 0.09 (-0.24,0.43) | **SE-US** | -0.03 (-0.27,0.21) | -0.04 (-0.30,0.23) | -0.05 (-0.30,0.20) | -0.05 (-0.31,0.21) | -0.35 (-0.80,0.09) | **-0.46 (-0.76,-0.16)** | **-0.85 (-1.22,-0.47)** |
| 0.19 (-0.05,0.42) | 0.16 (-0.10,0.42) | 0.11 (-0.03,0.25) | 0.12 (-0.11,0.36) | 0.03 (-0.21,0.27) | **US** | -0.01 (-0.12,0.10) | -0.02 (-0.10,0.06) | -0.02 (-0.11,0.07) | -0.32 (-0.70,0.05) | **-0.43 (-0.62,-0.25)** | **-0.82 (-1.11,-0.53)** |
| 0.19 (-0.06,0.45) | 0.17 (-0.11,0.45) | 0.12 (-0.06,0.29) | 0.13 (-0.12,0.38) | 0.04 (-0.23,0.30) | 0.01 (-0.10,0.12) | **PET/CT** | -0.01 (-0.13,0.11) | -0.01 (-0.14,0.12) | -0.31 (-0.69,0.06) | **-0.42 (-0.61,-0.24)** | **-0.81 (-1.12,-0.50)** |
| 0.20 (-0.03,0.44) | 0.18 (-0.08,0.44) | 0.13 (-0.03,0.29) | 0.14 (-0.09,0.37) | 0.05 (-0.20,0.30) | 0.02 (-0.06,0.10) | 0.01 (-0.11,0.13) | **CT** | -0.00 (-0.10,0.09) | -0.31 (-0.69,0.08) | **-0.42 (-0.61,-0.22)** | **-0.80 (-1.10,-0.50)** |
| 0.21 (-0.04,0.45) | 0.18 (-0.09,0.45) | 0.13 (-0.03,0.30) | 0.14 (-0.10,0.39) | 0.05 (-0.21,0.31) | 0.02 (-0.07,0.11) | 0.01 (-0.12,0.14) | 0.00 (-0.09,0.10) | **US+CT** | -0.30 (-0.69,0.08) | **-0.41 (-0.61,-0.21)** | **-0.80 (-1.10,-0.50)** |
| **0.51 (0.07,0.95)** | **0.48 (0.03,0.94)** | **0.43 (0.03,0.83)** | 0.45 (0.00,0.89) | 0.35 (-0.09,0.80) | 0.32 (-0.05,0.70) | 0.31 (-0.06,0.69) | 0.31 (-0.08,0.69) | 0.30 (-0.08,0.69) | **IWBC** | -0.11 (-0.44,0.22) | **-0.50 (-0.97,-0.02)** |
| **0.62 (0.32,0.91)** | **0.59 (0.28,0.91)** | **0.54 (0.32,0.77)** | **0.56 (0.26,0.85)** | **0.46 (0.16,0.76)** | **0.43 (0.25,0.62)** | **0.42 (0.24,0.61)** | **0.42 (0.22,0.61)** | **0.41 (0.21,0.61)** | 0.11 (-0.22,0.44) | **MRI** | **-0.39 (-0.73,-0.05)** |
| **1.01 (0.63,1.38)** | **0.98 (0.59,1.37)** | **0.93 (0.64,1.22)** | **0.94 (0.57,1.31)** | **0.85 (0.47,1.22)** | **0.82 (0.53,1.11)** | **0.81 (0.50,1.12)** | **0.80 (0.50,1.10)** | **0.80 (0.50,1.10)** | **0.50 (0.02,0.97)** | **0.39 (0.05,0.73)** | **US+CEUS** |

FIGURE 10
League table on PPV.

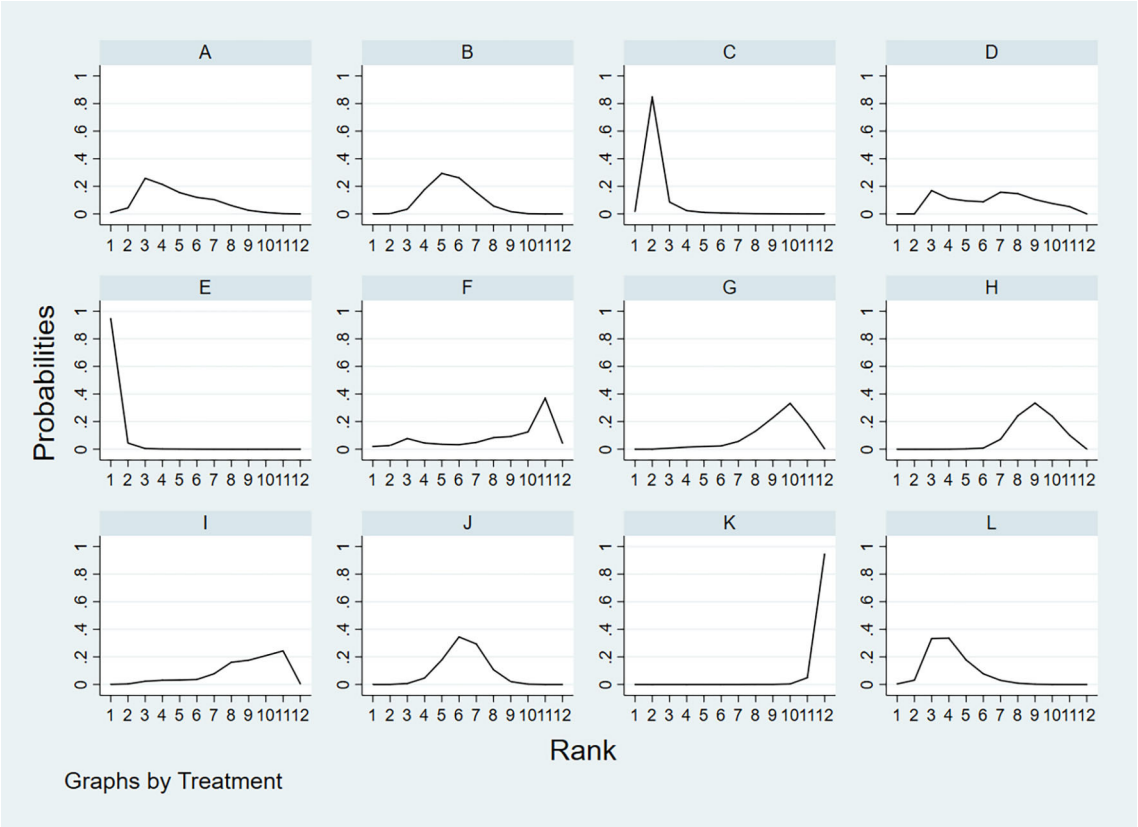



FIGURE 11
**A** = CEUS, **B** = CT, **C** = FNA-Tg, > **D** = FNAC, **E** = FNAC+FNA-Tg, **F** = IWBC, **G** = MRI, **H** = PET/CT, **I** = SE-US, **J** = US, **K** = US+CEUS, **L** = US+CT.

are commonly seen on CT, and CT scans in the central neck are twice as sensitive as US. Zhao H et al. (51) found that CT had higher sensitivity in the central region and throughout the lateral regions. However, ultrasound had a higher specificity in the central and whole lateral regions. CT can better reduce the missed diagnosis rate, while US can better avoid misdiagnosis. Therefore, this suggested that clinicians should consider the combination of CT and US when diagnosing LNM of PTC. This was consistent with our results that the combined diagnosis method of CT+US could more accurately diagnose LNM for better clinical decision-making.

This time, a variety of diagnostic methods for lymph node metastasis of thyroid cancer were compared. As far as we know, although some authors have published meta-analyses of two or more diagnostic methods for lymph node metastasis of thyroid cancer, our network meta-analysis was the first to comprehensively analyze the currently known diagnostic methods. As far as the retrieved literature is concerned, there are obviously more studies

| FNAC+FNA-Tg | FNA-Tg | US+CT | CEUS | CT | US | FNAC | IWBC | SE-US | PET/CT | MRI | US+CEUS |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FNAC+FNA-Tg | -0.08 (-0.15,-0.00) | **-0.19 (-0.36,-0.03)** | **-0.20 (-0.39,-0.02)** | **-0.22 (-0.38,-0.07)** | **-0.24 (-0.39,-0.08)** | **-0.25 (-0.32,-0.17)** | **-0.34 (-0.67,-0.01)** | **-0.32 (-0.54,-0.11)** | -0.31 (-0.48,-0.14) | **-0.33 (-0.53,-0.13)** | **-0.64 (-0.90,-0.38)** |
| 0.08 (0.00,0.15) | FNA-Tg | -0.12 (-0.27,0.03) | -0.13 (-0.30,0.04) | -0.15 (-0.29,-0.00) | **-0.16 (-0.31,-0.01)** | **-0.17 (-0.23,-0.11)** | -0.26 (-0.59,0.06) | **-0.25 (-0.45,-0.04)** | **-0.24 (-0.39,-0.08)** | **-0.25 (-0.45,-0.06)** | **-0.57 (-0.82,-0.31)** |
| **0.19 (0.03,0.36)** | 0.12 (-0.03,0.27) | US+CT | -0.01 (-0.12,0.10) | -0.03 (-0.09,0.02) | -0.04 (-0.10,0.01) | -0.05 (-0.21,0.10) | -0.15 (-0.44,0.15) | -0.13 (-0.28,0.03) | **-0.12 (-0.20,-0.04)** | -0.14 (-0.27,0.00) | **-0.45 (-0.66,-0.23)** |
| **0.20 (0.02,0.39)** | 0.13 (-0.04,0.30) | 0.01 (-0.10,0.12) | CEUS | -0.02 (-0.12,0.08) | -0.03 (-0.12,0.06) | -0.04 (-0.21,0.13) | -0.14 (-0.44,0.17) | -0.12 (-0.29,0.05) | -0.11 (-0.22,0.01) | -0.13 (-0.28,0.03) | **-0.44 (-0.64,-0.23)** |
| **0.22 (0.07,0.38)** | 0.15 (0.00,0.29) | **0.03 (-0.02,0.09)** | 0.02 (-0.08,0.12) | CT | -0.01 (-0.06,0.04) | -0.02 (-0.17,0.12) | -0.12 (-0.41,0.18) | -0.10 (-0.25,0.06) | **-0.09 (-0.16,-0.02)** | -0.11 (-0.24,0.03) | **-0.42 (-0.63,-0.21)** |
| **0.24 (0.08,0.39)** | **0.16 (0.01,0.31)** | 0.04 (-0.01,0.10) | 0.03 (-0.06,0.12) | 0.01 (-0.04,0.06) | US | -0.01 (-0.16,0.14) | -0.10 (-0.39,0.19) | -0.09 (-0.23,0.06) | **-0.08 (-0.14,-0.01)** | -0.09 (-0.22,0.04) | **-0.40 (-0.61,-0.20)** |
| **0.25 (0.17,0.32)** | **0.17 (0.11,0.23)** | 0.05 (-0.10,0.21) | 0.04 (-0.13,0.21) | 0.02 (-0.12,0.17) | 0.01 (-0.14,0.16) | FNAC | -0.09 (-0.42,0.23) | -0.08 (-0.28,0.13) | -0.07 (-0.22,0.09) | -0.08 (-0.28,0.11) | **-0.39 (-0.65,-0.14)** |
| **0.34 (0.01,0.67)** | 0.26 (-0.06,0.59) | 0.15 (-0.15,0.44) | 0.14 (-0.17,0.44) | 0.12 (-0.18,0.41) | 0.10 (-0.19,0.39) | 0.09 (-0.23,0.42) | IWBC | 0.02 (-0.30,0.34) | 0.03 (-0.26,0.32) | 0.01 (-0.25,0.27) | -0.30 (-0.66,0.05) |
| **0.32 (0.11,0.54)** | **0.25 (0.04,0.45)** | 0.13 (-0.03,0.28) | 0.12 (-0.05,0.29) | 0.10 (-0.06,0.25) | 0.09 (-0.06,0.23) | 0.08 (-0.13,0.28) | -0.02 (-0.34,0.30) | SE-US | 0.01 (-0.15,0.17) | -0.01 (-0.20,0.19) | **-0.32 (-0.57,-0.07)** |
| **0.31 (0.14,0.48)** | **0.24 (0.08,0.39)** | **0.12 (0.04,0.20)** | 0.11 (-0.01,0.22) | **0.09 (0.02,0.16)** | **0.08 (0.01,0.14)** | 0.07 (-0.09,0.22) | -0.03 (-0.32,0.26) | -0.01 (-0.17,0.15) | PET/CT | -0.02 (-0.15,0.11) | **-0.33 (-0.55,-0.11)** |
| **0.33 (0.13,0.53)** | **0.25 (0.06,0.45)** | 0.14 (-0.00,0.27) | 0.13 (-0.03,0.28) | 0.11 (-0.03,0.24) | 0.09 (-0.04,0.22) | 0.08 (-0.11,0.28) | -0.01 (-0.27,0.25) | 0.01 (-0.19,0.20) | 0.02 (-0.11,0.15) | MRI | **-0.31 (-0.55,-0.07)** |
| **0.64 (0.38,0.90)** | **0.57 (0.31,0.82)** | **0.45 (0.23,0.66)** | **0.44 (0.23,0.64)** | **0.42 (0.21,0.63)** | **0.40 (0.20,0.61)** | **0.39 (0.14,0.65)** | 0.30 (-0.05,0.66) | **0.32 (0.07,0.57)** | **0.33 (0.11,0.55)** | **0.31 (0.07,0.55)** | US+CEUS |

FIGURE 12
League table on NPV.

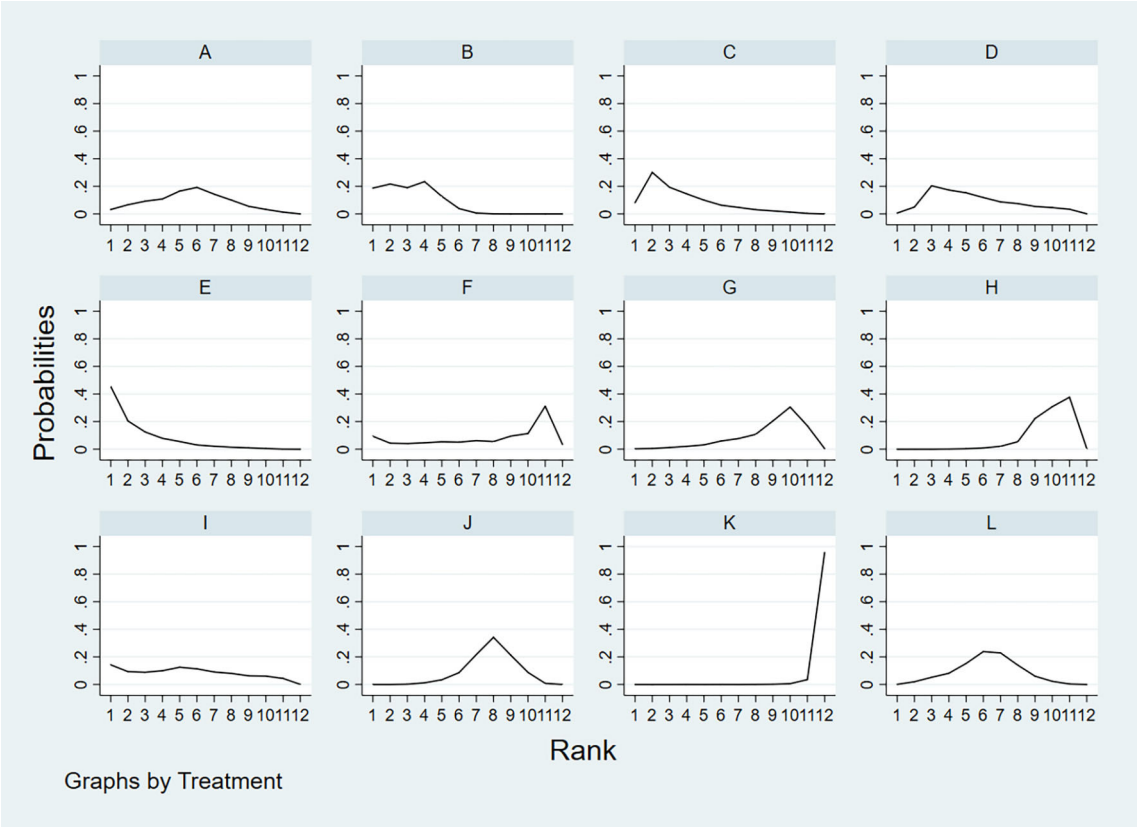



FIGURE 13
**A** = CEUS, **B** = CT, **C** = FNA-Tg, > **D** = FNAC, **E** = FNAC+FNA-Tg, **F** = IWBC, **G** = MRI, **H** = PET/CT, **I** = SE-US, **J** = US, **K** = US+CEUS, **L** = US+CT.

on US and CT than those on other diagnostic methods, indicating that these two diagnostic techniques are easier for statistical analysis through direct comparison, and at the same time indicating that US and CT are more widely accepted by everyone, while there are few studies that directly compare CEUS with FNA-Tg. It is recommended to carry out more original research on the comparative analysis of the diagnostic performance of these two diagnostic methods. At the same time, there are many studies comparing CT, US, and PET/CT in their diagnostic performance for lymph node metastasis of thyroid cancer, which also shows that these three studies have obvious comparative value (52). Through analysis, we found that such invasive diagnostic methods such as the combined application of FNAC,FNA-Tg and FNAC+FNA-Tg had better diagnostic performance. Although non-invasive diagnostic methods were more popular, except for the combined application of US+CT (19, 33), the diagnostic performance of other diagnostic methods was not ideal.

PTC is generally considered to be a tumor with a low probability of metastasis and recurrence. In some patients,

| FNAC+FNA-Tg | CT | FNA-Tg | SE-US | FNAC | CEUS | US+CT | IWBC | US | MRI | PET/CT | US+CEUS |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FNAC+FNA-Tg | -0.05 (-0.31,0.22) | -0.05 (-0.17,0.07) | -0.12 (-0.48,0.24) | -0.11 (-0.23,0.01) | -0.14 (-0.44,0.16) | -0.16 (-0.44,0.11) | -0.25 (-0.73,0.23) | -0.21 (-0.47,0.05) | -0.27 (-0.59,0.06) | **-0.30 (-0.58,-0.02)** | **-0.71 (-1.12,-0.31)** |
| 0.05 (-0.22,0.31) | CT | -0.00 (-0.25,0.24) | -0.08 (-0.33,0.18) | -0.06 (-0.31,0.18) | -0.09 (-0.26,0.07) | **-0.12 (-0.21,-0.02)** | -0.20 (-0.61,0.21) | **-0.17 (-0.25,-0.09)** | **-0.22 (-0.43,-0.02)** | **-0.26 (-0.38,-0.14)** | **-0.67 (-0.99,-0.35)** |
| 0.05 (-0.07,0.17) | 0.00 (-0.24,0.25) | FNA-Tg | -0.07 (-0.41,0.27) | -0.06 (-0.16,0.04) | -0.09 (-0.37,0.19) | -0.11 (-0.37,0.14) | -0.20 (-0.67,0.28) | -0.16 (-0.41,0.08) | -0.22 (-0.53,0.09) | -0.25 (-0.52,0.01) | **-0.66 (-1.06,-0.27)** |
| 0.12 (-0.24,0.48) | 0.08 (-0.18,0.33) | 0.07 (-0.27,0.41) | SE-US | 0.01 (-0.33,0.36) | -0.02 (-0.30,0.26) | -0.04 (-0.30,0.22) | -0.13 (-0.60,0.35) | -0.09 (-0.34,0.15) | -0.15 (-0.46,0.16) | -0.18 (-0.45,0.09) | **-0.59 (-0.99,-0.20)** |
| 0.11 (-0.01,0.23) | 0.06 (-0.18,0.31) | 0.06 (-0.04,0.16) | -0.01 (-0.36,0.33) | FNAC | -0.03 (-0.31,0.25) | -0.05 (-0.31,0.20) | -0.14 (-0.61,0.33) | -0.10 (-0.35,0.14) | -0.16 (-0.47,0.15) | -0.20 (-0.46,0.07) | **-0.60 (-1.00,-0.21)** |
| 0.14 (-0.16,0.44) | 0.09 (-0.07,0.26) | 0.09 (-0.19,0.37) | 0.02 (-0.26,0.30) | 0.03 (-0.25,0.31) | CEUS | -0.02 (-0.19,0.15) | -0.11 (-0.54,0.32) | -0.07 (-0.21,0.07) | -0.13 (-0.37,0.11) | -0.16 (-0.34,0.02) | **-0.57 (-0.88,-0.26)** |
| 0.16 (-0.11,0.44) | **0.12 (0.02,0.21)** | 0.11 (-0.14,0.37) | 0.04 (-0.22,0.30) | 0.05 (-0.20,0.31) | 0.02 (-0.15,0.19) | US+CT | -0.09 (-0.50,0.33) | -0.05 (-0.15,0.04) | -0.11 (-0.32,0.10) | **-0.14 (-0.27,-0.01)** | **-0.55 (-0.87,-0.22)** |
| 0.25 (-0.23,0.73) | 0.20 (-0.21,0.61) | 0.20 (-0.28,0.67) | 0.13 (-0.35,0.60) | 0.14 (-0.33,0.61) | 0.11 (-0.32,0.54) | 0.09 (-0.33,0.50) | IWBC | 0.04 (-0.37,0.44) | -0.02 (-0.38,0.34) | -0.06 (-0.46,0.35) | -0.46 (-0.98,0.05) |
| 0.21 (-0.05,0.47) | **0.17 (0.09,0.25)** | 0.16 (-0.08,0.41) | 0.09 (-0.15,0.34) | 0.10 (-0.14,0.35) | 0.07 (-0.07,0.21) | 0.05 (-0.04,0.15) | -0.04 (-0.44,0.37) | US | -0.06 (-0.25,0.14) | -0.09 (-0.20,0.02) | **-0.50 (-0.81,-0.19)** |
| 0.27 (-0.06,0.59) | **0.22 (0.02,0.43)** | 0.22 (-0.09,0.53) | 0.15 (-0.16,0.46) | 0.16 (-0.15,0.47) | 0.13 (-0.11,0.37) | 0.11 (-0.10,0.32) | 0.02 (-0.34,0.38) | 0.06 (-0.14,0.25) | MRI | -0.04 (-0.23,0.16) | **-0.44 (-0.81,-0.08)** |
| **0.30 (0.02,0.58)** | **0.26 (0.14,0.38)** | 0.25 (-0.01,0.52) | 0.18 (-0.09,0.45) | 0.20 (-0.07,0.46) | 0.16 (-0.02,0.34) | **0.14 (0.01,0.27)** | 0.06 (-0.35,0.46) | 0.09 (-0.02,0.20) | 0.04 (-0.16,0.23) | PET/CT | **-0.41 (-0.74,-0.08)** |
| **0.71 (0.31,1.12)** | **0.67 (0.35,0.99)** | **0.66 (0.27,1.06)** | **0.59 (0.20,0.99)** | **0.60 (0.21,1.00)** | **0.57 (0.26,0.88)** | **0.55 (0.22,0.87)** | 0.46 (-0.05,0.98) | **0.50 (0.19,0.81)** | **0.44 (0.08,0.81)** | **0.41 (0.08,0.74)** | US+CEUS |

FIGURE 14
League table on ACC.

PTC also occurs with lymph node metastasis and local recurrence (53). Studies have shown that tumor size, extrathyroidal extension and vascular invasion are independent risk factors for lymph node metastasis (54). However, extrathyroid extension and vascular infiltration are not helpful to surgeons undergoing surgery, and imaging features such as tumor size and shape can effectively improve the ability of preoperative diagnosis of PTC lymph node metastasis (55).

At present, most hospitals mostly use this non-invasive method of US and CT for preoperative or postoperative evaluation and examination of patients with LNM of PTC. However, US also has its own limitations. US is far from satisfactory in sensitivity and accuracy for lymph node metastasis. It cannot observe deeper tissues or lymph node metastasis, and the accuracy of the result often depends on the experience of the diagnostician, etc. CT is generally considered to be a diagnostic method that plays a key role in the diagnosis of lymph node metastasis of thyroid cancer, especially in the evaluation of deeper tissues (56). CT is slightly inferior to US in evaluating the diagnostic value of central and lateral cervical lymph nodes. Therefore, the combination of CT and US has been evaluated in 11 studies. In our study, regarding the sensitivity to thyroid nodules and lymph node infiltration, we found that the sensitivity of US+CT in the diagnosis of thyroid nodules and lymph node infiltration was superior to the sensitivity of FNAC and FNA-Tg alone.

FNAC was highly specific in the analysis, and Suh YJ et al. (36, 37, 39) found that the threshold values of FNA-Tg established in different experiments were different. If the threshold value of FNA-Tg was specified as 1ng/mL, and levels greater than 1ng/mL were accepted as the surgical indication, then NPV was 0%. Non-metastatic lymph nodes ranged in value from 2–32 ng/mL. The SPE, PPV, NPV and Accuracy of FNA-Tg were significantly lower than those of FNAC. When the critical value of FNA-Tg was 28.5ng/ml, the SPE, PPV, NPV and Accuracy were significantly higher than those of FNAC. Therefore, further experiments are required to determine the optimal critical value of FNA-Tg. Based on our analysis, the diagnostic performance of the combination of FNAC+FNA-Tg was significantly superior to that of FNAC and FNA-Tg alone. In the combination application of FNAC and FNA-Tg, the values of PPV, NPV and Accuracy in SUCRA were 81.6,99.4 and 86.8 respectively, which had obvious advantages in the diagnosis of LNM of PTC and also proved that the combination application of FNAC and FNA-Tg was one of the necessary diagnostic methods for accurate diagnosis of benign and malignant lymph node metastasis.

Although the combined diagnosis has higher diagnostic performance than the single diagnosis method, the price of the same part of the examination will be correspondingly increased, and the economic burden of different examination methods for patients is different. In China, some routine examinations such as CT,MRI,US,PET/CT are within the range of 100–800 RMB (57), which will not cause great economic pressure for most families. Moreover, the combination diagnosis of CT and US can also obtain relatively accurate results. However, the prices of FNAC and FNA-Tg are about 2000RMB, which is much more expensive compared with conventional examinations. Moreover, most of the patients undergoing such examinations are patients with LNM of PTC, who need to go to the hospital for regular reexamination in addition to the expenses required for daily treatment, so that they need to bear a heavier economic burden. Therefore, during the examination, we should not blindly pursue more accurate diagnosis performance of the disease, but we should be more concerned about their demands and the economic pressure they can bear from the perspective of patients. In the United States, the average cost of these routine imaging examinations in upper-level academic hospitals ranges from $686.13 to $1,390.12. The reason for the higher price in the United States may be that most residents have purchased health insurance, and they can only go to the designated hospitals with medical insurance for examinations, resulting in that they cannot go to the hospitals with lower costs for relevant examinations. At the same time, the price of related examinations is less transparent. Hospitals with lower prices may be stimulated by higher-priced hospitals, leading to price increases. All these have correspondingly increased the economic burden of consumers.

## 4.1 Limitation

At present, there are many diagnostic methods for lymph node metastasis of papillary thyroid carcinoma, but there are still differences in these diagnostic values, and further analysis and standardization are needed. (1) Compared with previous meta-analyses, we have included more diagnostic methods, but few articles have included some diagnostic methods, and their reliability required more experiments to verify. (2) It was found in the study that the results of direct comparison between interventions were less, and most of them were the results of indirect comparison. This would also affect the final quality evaluation results, which meant that further more direct comparison between different diagnostic methods was needed in the study. (3) Some results showed obvious heterogeneity in our study, but no source of heterogeneity was found after heterogeneity analysis. In view of the above shortcomings, it is recommended that the readers reasonably refer to and select the diagnostic methods in this study based on the clinical practice and actual results.

# 5 Results

In conclusion, we believe that the combined application of CT+US and FNAC+FNA-Tg is the optimal solution for the

diagnosis of lymph node metastasis in papillary thyroid carcinoma, but further prospective studies on its cost-effectiveness and clinical diagnostic performance are still needed.

## Author contributions

Study concept and design: S-RW, Q-LL. Acquisition of data: Q-LL, MC, TS, S-RW, FT. Analysis and interpretation of data: C-LC, S-RW, FT. Drafting of the manuscript: S-RW, FT. Critical revision of themanuscript for important intellectual content: JL. Approval of the final manuscript: JL, L-NS, TS, C-LC, S-RW, FT. Study supervision: JL.

## Funding

1.Supported by the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2020-PT330-003) 2.Supported by Open Research Fund of NHC Key Laboratory of Prevention and Treatment of Central Asia High Incidence Diseases 3.Key science and technology Project of Xinjiang Corps (2019DB012).

## 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.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Vaccarella S, Lortet-Tieulent J, Colombet M, Davies L, CA S, Schüz J, et al. Global patterns and trends in incidence and mortality of thyroid cancer in children. *Lancet Diabetes Endocrinol* (2021) 9(3):144–52144–52.

2. Olaleye O, Ekrikpo U, Moorthy R, Lyne O, Wiseberg J, Black M, et al. Increasing incidence of differentiated thyroid cancer in south East England: 1987-2006. *Eur Arch Otorhinolaryngol* (2011) 268(6):899–906. doi: 10.1007/s00405-010-1416-7

3. Lim H, Devesa SS, Sosa JA, Check D, Kitahara CM. Trends in thyroid cancer incidence and mortality in the united states, 1974-2013. *Jama* (2017) 317(13):1338–48. doi: 10.1001/jama.2017.2719

4. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LD, et al. Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma: A paradigm shift to reduce overtreatment of indolent tumors. *JAMA Oncol* (2016) 2(8):1023–9. doi: 10.1001/jamaoncol.2016.0386

5. Lamartina L, Grani G, Durante C, Filetti S. Recent advances in managing differentiated thyroid cancer. *F1000Res* (2018) 7:86. doi: 10.12688/f1000research.12811.1

6. Wunderbaldinger P, Harisinghani MG, Hahn PF, Daniels GH, Turetschek K, Simeone J, et al. Cystic lymph node metastases in papillary thyroid carcinoma. *AJR Am J Roentgenol* (2002) 178(3):693–7. doi: 10.2214/ajr.178.3.1780693

7. Li F, Pan D, He Y, Wu Y, Peng J, Li J, et al. Using ultrasound features and radiomics analysis to predict lymph node metastasis in patients with thyroid cancer. *BMC Surg* (2020) 20(1):315. doi: 10.1186/s12893-020-00974-7

8. Chasen NN, Wang JR, Gan Q, Ahmed S. Imaging of cervical lymph nodes in thyroid cancer: Ultrasound and computed tomography. *Neuroimaging Clin N Am* (2021) 31(3):313–26. doi: 10.1016/j.nic.2021.04.002

9. Schlumberger MJ, Torlantano M. Papillary and follicular thyroid carcinoma. *Baillieres Best Pract Res Clin Endocrinol Metab* (2000) 14(4):601–13. doi: 10.1053/beem.2000.0105

10. Grani G, Fumarola A. Thyroglobulin in lymph node fine-needle aspiration washout: a systematic review and meta-analysis of diagnostic accuracy. *J Clin Endocrinol Metab* (2014) 99(6):1970–82. doi: 10.1210/jc.2014-1098

11. Seo YL, Yoon DY, Baek S, Ku YJ, Rho YS, Eun-Jae Chung E-J, et al. Detection of neck recurrence in patients with differentiated thyroid cancer: comparison of ultrasound, contrast-enhanced CT and f-18-FDG PET/CT using surgical pathology as a reference standard: (ultrasound vs. CT vs. f-18-FDG PET/CT in recurrent thyroid cancer). *Eur Radiol* (2012) 22(10):2246–54.

12. Eunhee K, Park J, Son K, Kim J-H, Jeon SJ, Na DG. Preoperative diagnosis of cervical metastatic lymph nodes in papillary thyroid carcinoma: Comparison of ultrasound, computed tomography, and combined ultrasound with computed tomography. *Thyroid* (2008) 18(4):411–8.

13. Chen L, Chen L, Liu J, Wang B, Zhang H. Value of qualitative and quantitative contrast-enhanced ultrasound analysis in preoperative diagnosis of cervical lymph node metastasis from papillary thyroid carcinoma. *J Ultrasound Med* (2020) 39(1):73–81. doi: 10.1002/jum.15074

14. Kim MJ, Kim EK, Kim BM, Kwak JY, Lee EJ, Park CS, et al. Thyroglobulin measurement in fine-needle aspirate washouts: the criteria for neck node dissection for patients with thyroid cancer. *Clin Endocrinol (Oxf)* (2009) 70(1):145–51. doi: 10.1111/j.1365-2265.2008.03297.x

15. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* (2011) 155(8):529–36. doi: 10.7326/0003-4819-155-8-201110180-00009

16. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The cochrane collaboration's tool for assessing risk of bias in randomised trials. *Bmj* (2011) 343:d5928. doi: 10.1136/bmj.d5928

17. Salmaslioglu A, Erbil Y, Citlak G, Ersoz F, Sari S, Olmez A, et al. Diagnostic value of thyroglobulin measurement in fine-needle aspiration biopsy for detecting metastatic lymph nodes in patients with papillary thyroid carcinoma. *Langenbecks Arch Surg* (2011) 396(1):77–81. doi: 10.1007/s00423-010-0723-1

18. de Crea C, Raffaelli M, Maccora D, Carrozza C, Canu G, Fadda G, et al. Calcitonin measurement in fine-needle aspirate washouts vs. cytologic examination for diagnosis of primary or metastatic medullary thyroid carcinoma. *Acta Otorhinolaryngol Ital* (2014) 34(6):399–405.

19. Byun BH, Jeong U-G, Hong SP, Min JJ, Chong A, Song H-C, et al. Prediction of central lymph node metastasis from papillary thyroid microcarcinoma by f-18-fluorodeoxyglucose PET/CT and ultrasonography. *Ann Nucl Med* (2012) 26(6):471–7. doi: 10.1007/s12149-012-0594-3

20. de Crea C, Raffaelli M, Maccora D, Carrozza C, Canu G, Fadda G, et al. Calcitonin measurement in fine-needle aspirate washouts vs. cytologic examination for diagnosis of primary or metastatic medullary thyroid carcinoma. *Acta Otorhinolaryngol Ital* (2014) 34(6):399–405.

21. Samanci C, Onal Y, Sager S, Asa S, Ustabasioglu FE, Alis D, et al. Diagnostic capabilities of MRI versus f-18 FDG PET-CT in postoperative patients with thyroglobulin positive, I-131-negative local recurrent or metastatic thyroid cancer. *Curr Med Imaging* (2019) 15(10):956–64. doi: 10.2174/1573405614666180718124739

22. Lee JH, Lee HC, Yi HW, Kim BK, Bae SY, Lee SK, et al. Influence of thyroid gland status on the thyroglobulin cutoff level in washout fluid from cervical lymph nodes of patients with recurrent/metastatic papillary thyroid cancer. *Head Neck-Journal Sci Specialties Head Neck* (2016) 38:E1705–12. doi: 10.1002/hed.24305

23. David L, Cunnane ME, Zurakowski D, Acar G, Ecevit C, Mace A, et al. Papillary thyroid carcinoma nodal surgery directed by a preoperative radiographic map utilizing CT scan and ultrasound in all primary and reoperative patients. *Head Neck-Journal Sci Specialties Head Neck* (2014) 36(2):191–202.

24. Baldini E, Sorrent S, Di Gioia C, De Vito C, Antonelli A, Gnessi L, et al. Cervical lymph node metastases from thyroid cancer: Does thyroglobulin and calcitonin measurement in fine needle aspirates improve the diagnostic value of cytology? *BMC Clin Pathol* (2013) 13:7–7. doi: 10.1186/1472-6890-13-7

25. Eun NL, Son EJ, Kim JA, Gweon HM, Kang JH, Youk JH. Comparison of the diagnostic performances of ultrasonography, CT and fine needle aspiration cytology for the prediction of lymph node metastasis in patients with lymph node dissection of papillary thyroid carcinoma: A retrospective cohort study. *Int J Surg* (2018) 51:145–50. doi: 10.1016/j.ijsu.2017.12.036

26. Farzana A, Naito K, Horiguchi J, Fukuda H, Tachikake T, Ito K. Accuracy of sonographic elastography in the differential diagnosis of enlarged cervical lymph nodes: Comparison with conventional b-mode sonography. *Am J Roentgenology* (2008) 191(2):604–10.

27. Jeong HS, Baek CH, Son YI, Choi JY, Kim HJ, Ko YH, et al. Integrated 18F-FDG PET/CT for the initial evaluation of cervical node level of patients with papillary thyroid carcinoma: Comparison with ultrasound and contrast-enhanced CT. *Clin Endocrinol (Oxf)* (2006) 65(3):402–7. doi: 10.1111/j.1365-2265.2006.02612.x

28. Hwang H, Perez D, Orloff L. Comparison of positron emission Tomography/Computed tomography imaging and ultrasound in staging and surveillance of head and neck and thyroid cancer. *Laryngoscope* (2009) 119 (10):1958–65. doi: 10.1002/lary.20594

29. Mihailovic J, Prvulovic M, Ivkovic M, Markoski B, Martinov D. MRI Versus I-131 whole-body scintigraphy for the detection of lymph node recurrences in differentiated thyroid carcinoma. *Am J Roentgenology* (2010) 195(5):1197–203. doi: 10.2214/AJR.09.4172

30. Jeon SJ, Kim E, Park JS, Son KY, Baek J, Kim YS, et al. Diagnostic benefit of thyroglobulin measurement in fine-needle aspiration for diagnosing metastatic cervical lymph nodes from papillary thyroid cancer: Correlations with US features. *Korean J Radiol* (2009) 10(2):106–11. doi: 10.3348/kjr.2009.10.2.106

31. Choi YJ, Yun JS, Kook SH, Jung EC, Park YL. Clinical and imaging assessment of cervical lymph node metastasis in papillary thyroid carcinomas. *World J Surg* (2010) 34(7):1494–9. doi: 10.1007/s00268-010-0541-1

32. Jung JH, Kim CY, Son SH, Kim DH, Jeong SY, Lee SW, et al. Preoperative prediction of cervical lymph node metastasis using primary tumor SUVmax on 18F-FDG PET/CT in patients with papillary thyroid carcinoma. *PloS One* (2015) 10 (12):e0144152. doi: 10.1371/journal.pone.0144152

33. Zhan J, Diao X-H, Chen YC, Wang W-P, Ding H. Homogeneity parameter in contrast-enhanced ultrasound imaging improves the classification of abnormal cervical lymph node after thyroidectomy in patients with papillary thyroid carcinoma. *BioMed Res Int* (2019) 26(2019):9296010. doi: 10.1155/2019/9296010

34. Zhan J, Diao X-H, Chen YC, Wang W-P, Ding H. Homogeneity parameter in contrast-enhanced ultrasound imaging improves the classification of abnormal cervical lymph node after thyroidectomy in patients with papillary thyroid carcinoma. *BioMed Res Int* (2019) 2019. doi: 10.1155/2019/9296010

35. Lee DW, Ji YB, Sung ES, Park JS, Lee YJ, Park DW, et al. Roles of ultrasonography and computed tomography in the surgical management of cervical lymph node metastases in papillary thyroid carcinoma. *Ejso* (2013) 39 (2):191–6. doi: 10.1016/j.ejso.2012.07.119

36. Khadra H, Mohamed H, Al-Qurayshi Z, Sholl A M, Emad Kandil E. Superior detection of metastatic cystic lymphadenopathy in patients with papillary thyroid cancer by utilization of thyroglobulin washout. *Head Neck-Journal Sci Specialties Head Neck* (2019) 41(1):225–9. doi: 10.1002/hed.25488

37. Li J, Zhang K, Liu X, Hao F, Liu Z, Wang Z. Cervical lymph node thyroglobulin measurement in washout of fine-needle aspirates for diagnosis of papillary thyroid cancer metastases. *Br J BioMed Sci* (2016) 73(2):79–83. doi: 10.1080/09674845.2016.1173334

38. Shi JH, Xu YY, Pan QZ, Sui GQ, Zhou JP, Wang H. The value of combined application of ultrasound-guided fine needle aspiration cytology and thyroglobulin measurement for the diagnosis of cervical lymph node metastases from thyroid cancer. *Pak J Med Sci* (2015) 31(5):1152–5. doi: 10.12669/pjms.315.6726

39. Yang SY, Shin JH, Hahn SY, Lim Y, Hwang SY, Kim TH, et al. Comparison of ultrasonography and CT for preoperative nodal assessment of patients with papillary thyroid cancer: diagnostic performance according to primary tumor size. *Acta Radiol* (2020) 61(1):21–7. doi: 10.1177/0284185119847677

40. Sohn YM, Kim MJ, Kim EK, Kwak JY. Diagnostic performance of thyroglobulin value in indeterminate range in fine needle aspiration washout fluid from lymph nodes of thyroid cancer. *Yonsei Med J* (2012) 53(1):126–31. doi: 10.3349/ymj.2012.53.1.126

41. Suh YJ, Son EJ, Moon HJ, Kim EK, Han K, Kwak JY. Utility of thyroglobulin measurements in fine-needle aspirates of space occupying lesions in the thyroid bed after thyroid cancer operations. *Thyroid* (2013) 23(3):280–8. doi: 10.1089/thy.2011.0303

42. Liu T, Ge X, Yu J, Guo Y, Wang Y, Wang W, et al. Comparison of the application of b-mode and strain elastography ultrasound in the estimation of lymph node metastasis of papillary thyroid carcinoma based on a radiomics approach. *Int J Comput Assisted Radiol Surg* (2018) 13(10):1617–27. doi: 10.1007/s11548-018-1796-5

43. Wang Y, Nie F, Wang G, Liu T, Dong T, Sun Y. Value of combining clinical factors, conventional ultrasound, and contrast-enhanced ultrasound features in preoperative prediction of central lymph node metastases of different sized papillary thyroid carcinomas. *Cancer Manag Res* (2021) 13:3403–15. doi: 10.2147/CMAR.S299157

44. Liu Y, Li S, Yan C, He C, Yun M, Liu M, et al. Value of dual-phase, contrast-enhanced CT combined with ultrasound for the diagnosis of metastasis to central lymph nodes in patients with papillary thyroid cancer. *Clin Imaging* (2021) 75:5–11. doi: 10.1016/j.clinimag.2021.01.008

45. Wei Y, Yu M, Niu Y, Hao Y, Di J, Zhao Z, et al. Combination of Lymphatic and Intravenous Contrast-Enhanced Ultrasound for Evaluation of Cervical Lymph Node Metastasis from Papillary Thyroid Carcinoma: A Preliminary Study. *Ultrasound Med Biol* (2021) 47(2):252–60. doi: 10.1016/j.ultrasmedbio.2020.10.003

46. Hong YR, Luo ZY, Mo GQ, Wang P, Ye Q, Huang PT. Role of Contrast-Enhanced Ultrasound in the Pre-operative Diagnosis of Cervical Lymph Node Metastasis in Patients with Papillary Thyroid Carcinoma. *Ultrasound Med Biol* (2017) 43(11):2567–75. doi: 10.1016/j.ultrasmedbio.2017.07.010

47. Liu Z, Xun X, Wang Y, Mei L, He L, Zeng W, et al. MRI And ultrasonography detection of cervical lymph node metastases in differentiated thyroid carcinoma before reoperation. *Am J Trans Res* (2014) 6(2):147–54.

48. Zhao H, Wang Y, Wang MJ, Zhang ZH, Wang HR, Zhang B, et al. Influence of presence/absence of thyroid gland on the cutoff value for thyroglobulin in lymph-node aspiration to detect metastatic papillary thyroid carcinoma. *BMC Cancer* (2017) 17(1):296. doi: 10.1186/s12885-017-3296-3

49. Zhao H, Li H. Meta-analysis of ultrasound for cervical lymph nodes in papillary thyroid cancer: Diagnosis of central and lateral compartment nodal metastases. *Eur J Radiol* (2019) 112:14–21. doi: 10.1016/j.ejrad.2019.01.006

50. Morita S, Mizoguchi K, Suzuki M, Iizuka K. The accuracy of (18)[F]-fluoro-2-deoxy-D-glucose-positron emission tomography/computed tomography, ultrasonography, and enhanced computed tomography alone in the preoperative diagnosis of cervical lymph node metastasis in patients with papillary thyroid carcinoma. *World J Surg* (2010) 34(11):2564–9. doi: 10.1007/s00268-010-0733-8

51. Sheng L, Shi J, Han B, Lv B, Li L, Chen B, et al. Predicting factors for central or lateral lymph node metastasis in conventional papillary thyroid microcarcinoma. *Am J Surg* (2020) 220(2):334–40. doi: 10.1016/j.amjsurg.2019.11.032

52. Gu JH, Zhao YN, Xie RL, Xu WJ, You DL, Zhao ZF, et al. Analysis of risk factors for cervical lymph node metastasis of papillary thyroid microcarcinoma: A study of 268 patients. *BMC Endocr Disord* (2019) 19(1):124. doi: 10.1186/s12902-019-0450-8

53. Medas F, Canu GL, Cappellacci F, Boi F, Lai ML, Erdas E, et al. Predictive factors of lymph node metastasis in patients with papillary microcarcinoma of the thyroid: Retrospective analysis on 293 cases. *Front Endocrinol (Lausanne)* (2020) 11:551. doi: 10.3389/fendo.2020.00551

54. Loevner L, Kaplan S, Cunnane M-E, Moonis G. Cross-sectional imaging of the thyroid gland. *Neuroimaging Clinics North America* (2008) 18(3):445–61. doi: 10.1016/j.nic.2008.05.001

55. Yu XM, Yang XY, Qiu JC, Zhu YH, Zhou T, Lin J. Analysis of current situation of CT and MRI examination service price items. *China Med Equip* (2021) 18(04):168–171.

56. Paul AB, Oklu R, Saini S, Prabhakar AM. How much is that head CT? price transparency and variability in radiology. *J Coll Radiol* (2015) 12(5):453–7. doi: 10.1016/j.jacr.2014.12.016

57. Sinaiko AD, Rosenthal MB. Increased price transparency in health care–challenges and potential effects. *N Engl J Med* (2011) 364(10):891–4. doi: 10.1056/NEJMp1100041

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 01 December 2022
DOI 10.3389/fonc.2022.989595



OPEN ACCESS

EDITED BY
Sridhar Nimmagadda,
Johns Hopkins University,
United States

REVIEWED BY
Vetri Sudar Jayaprakasam,
Memorial Sloan Kettering Cancer Center, United States
Min Yang,
Jiangsu Institute of Nuclear Medicine, China

*CORRESPONDENCE
Jun Zhao
petcenter@126.com
Fengchun Hua
huafc@hotmail.com

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 08 July 2022
ACCEPTED 18 November 2022
PUBLISHED 01 December 2022

CITATION
Qi N, Wang H, Wang H, Ren S, You Z, Chen X, Guan Y, Xie F, Hua F and Zhao J (2022) Non-tumoral uptake of $^{68}$Ga-FAPI-04 PET: A retrospective study. *Front. Oncol.* 12:989595. doi: 10.3389/fonc.2022.989595



# Non-tumoral uptake of $^{68}$Ga-FAPI-04 PET: A retrospective study

Na Qi[1†], Hao Wang[2†], Haiyan Wang[1†], Shuhua Ren[3], Zhiwen You[1], Xing Chen[1], Yihui Guan[3], Fang Xie[3], Fengchun Hua[4*] and Jun Zhao[1*]

[1]Department of Nuclear Medicine, Shanghai East Hospital, Tongji University School of Medicine, Shanghai, China, [2]Department of General Surgery, Huashan Hospital, Fudan University, Shanghai, China, [3]PET Center, Huashan Hospital, Fudan University, Shanghai, China, [4]Department of Nuclear Medicine, Longhua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China

**Objective:** Fibroblast activation protein (FAP)-targeting radiopharmaceutical based on the FAP-specific inhibitor (FAPI) is considered as a potential alternative agent to FDG for tumor-specific imaging. However, FAP is also expressed in normal adult tissues. The aim of this study was to explore the image features of non-tumoral regions with high uptake of $^{68}$Ga-FAPI-04 in positron emission tomography (PET) imaging and to reveal the physiological mechanisms of these regions.

**Material:** A total of 137 patients who underwent whole-body $^{68}$Ga-FAPI-04 PET/MR (n=46) or PET/CT (n=91) were included in this retrospective study. Three experienced nuclear medicine physicians determined the non-tumoral regions according to other imaging modalities (CT, MRI, $^{18}$F-FDG PET, or ultrasound), clinical information, or pathological results. The regions of interest (ROIs) were drawn manually, and the maximum standardized uptake value ($SUV_{max}$) was measured.

**Results:** A total of 392 non-tumoral uptake regions were included in this study. The included physiological regions were uterus (n=38), submandibular gland (n=118), nipple (n=37), gingiva (n=65), and esophagus (n=31). The incidence of $^{68}$Ga-FAPI-04 uptake in physiological regions was independent of age, the tracer uptakes in the gingiva and esophagus were more common in male patients ($p$=0.006, 0.009), while that in the nipple was more common in female patients ($p < 0.001$). The included benign regions were inflammatory lymph node (n =10), pneumonia (n=13), atherosclerosis (n=10), pancreatitis (n=18), osteosclerosis (n=45), and surgical scar (n=7). No significant difference was observed in $SUV_{max}$ between physiological and benign regions.

**Conclusions:** A number of organs exhibit physiological uptakes of $^{68}$Ga-FAPI-04. Our study showed that regions with high $^{68}$Ga-FAPI-04 uptake did not necessarily represent malignancy. Being familiar with physiological and typical

benign $^{68}$Ga-FAPI-04 uptake regions can be helpful for physicians to interpret images and to make an accurate diagnosis.

KEYWORDS

68 Ga-FAPI-04, SUV, physiological uptake, benign uptake, multicenter retrospective study

# Introduction

Cancer-associated fibroblasts (CAFs) and extracellular fibrosis can account for 90% of the total tumor mass (1). Fibroblast activation protein (FAP), a type II membrane-bound glycoprotein of the dipeptidyl peptidase 4 family, is over-expressed in CAFs of many epithelial carcinomas and is involved in a variety of tumor-promoting activities, such as stromal remodeling, angiogenesis, chemotherapy resistance, and immunosuppression (1, 2). Since FAP is expressed at low levels in most normal organs, it is a promising target for imaging and radiation therapy (3). Radiopharmaceuticals targeting FAP have recently been developed based on FAP-specific inhibitors (FAPIs) (4). Among several recently developed tracers targeting FAP, $^{68}$Ga-FAPI-04 is regarded as a promising one for having high affinity towards FAP and suitable kinetics (5–7). Without the necessity of fasting in preparation before the scan and an equal or better tumor-to-background ratio compared with $^{18}$F-FDG PET scans, $^{68}$Ga-FAPI-04 is considered as a potential alternative agent to FDG for tumor-specific imaging (8).

Currently, most FAPI studies are focused on tumor imaging. Besides its high expression in epithelial carcinoma (8), FAP also plays a key role in normal development during embryo-genesis and tissue modeling (9). FAP can also be expressed in normal adult tissues such as active tissue damage, remodeling, inflammation, arthritis, atherosclerotic plaques, and fibrosis (3, 9, 10). Several non-oncology studies on FAPI revealed its unique values in IgG4-related diseases (11). Luo et al. (12) found that compared with $^{18}$F-FDG, $^{68}$Ga-FAPI-04 was more effective in detecting organs affected by IgG4-related disease. An animal study showed that joint FAPI concentration was correlated with arthritis scores in rats (13). A recent work reported that $^{68}$Ga-FAPI-04 focal non-tumoral uptake can occur in fibrous lesions, fibrous hyperplasia, and fibrous activity (14). $^{68}$Ga-FAPI-04 could also accumulate in some benign diseases of the bones and joints (15). Recent studies characterized the benign lesions with increased $^{68}$Ga-FAPI-04 uptake in PET/CT (16, 17). However, to the best of our knowledge, there are no systematic studies to reveal the pathophysiological mechanisms of non-tumoral $^{68}$Ga-FAPI-04 uptake regions. This study aimed to investigate the uptake characteristics in non-tumoral regions using $^{68}$Ga-FAPI-04 PET/CT or PET/MR with a relatively large sample size and provide a reference for imaging diagnosis.

# Materials and methods

## Patients

This retrospective analysis was performed on patients who underwent $^{68}$Ga-FAPI-04 PET/CT (Biograph mCT, Siemens Healthineers, Germany; Ingenuity TF, Philips Healthcare, USA; uMI510, United Imaging, China) or $^{68}$Ga-FAPI-04 PET/MR (uPMR790 TOF, United Imaging, China) from April 2020 to August 2021. The inclusion criteria were as follows: (i) patients who were able to sign informed consents for examination according to the guidelines of the Clinical Research Ethics Committee; (ii) patients with a predicted survival of more than 6 months. Exclusion criteria were (i) pregnancy, (ii) postmenopausal women with taking hormone replacement or related drugs, and (iii) patients with a predicted survival of <6 months.

## Radiopharmaceutical and imaging protocols

Good-manufacturing-practice (GMP)-grade precursors $^{68}$Ga-FAPI-04 was synthesized in the Radiochemistry Facility of the PET Center, Huashan Hospital, Fudan University, according to the protocol described previously (18). The radiochemical purity of $^{68}$Ga-FAPI-04 was over 95%, and the final product was sterile and pyrogen-free.

Whole-body PET/CT or PET/MR scans were performed 60 min after the injection of $^{68}$Ga-FAPI-04 with a dose of 150 ± 35 MBq (4.05 ± 0.95 mCi) from the vertex to the mid-thigh. For PET/CT, a PET scan was acquired after a low-dose CT scan, which was performed at 120 kV and 100–120 valid mAs. Brain PET scanning was performed 5 min/bed, and body PET scanning was performed 3 min/bed. PET/MR was performed with default clinical MRI sequences including T1w and T2w (TE = 2.24 ms, TR = 4.91 ms, flip angle = 10, echo train length = 30, FOV = 549 × 384, matrix = 256 × 329, slice thickness = 2 mm, slice spacing = 2 mm, transverse plane) (18). PET images were reconstructed by ordered subset expectation maximization 3D (OSEM 3D) method with 2 iterations and 20 subsets.

Since different scanners were used in this study, SUV measurements were normalized after data collection. A NEMA

IEC body phantom (Data Spectrum Corporation, Durham, NC, USA) with six simulated lesion spheres (diameters: 10, 13, 17, 22, 28, and 37 mm) was applied for SUV normalization with 2, 4, 8, and 16 times the background activity (background activity concentration =2 kBq/ml). A CT scan of the NEMA IEC body phantom was prepared for the attenuation correction of PET/MR. Correlation coefficients were obtained through this phantom study and used to standardize the SUV measurements as previously reported (18, 19).

## PET/CT and PET/MR imaging review

Three nuclear medicine physicians with 15, 10, and 8 years of experience in interpreting PET/CT and MR imaging determined the physiological and benign tracer uptake regions based on the patients' clinical data, imaging data (CT, MRI, $^{18}$F-FDG PET, or ultrasound), histopathology, and their own experiences in image interpretation. Physiological uptake refers to the slightly elevated uptake of $^{68}$Ga-FAPI-04 in generally normal tissues, which usually show no abnormal changes on other imaging modalities (20, 21). Benign uptake refers to inflammation, fibrosis, benign tumors, and other non-malignant tumor regions that may be abnormal on other imaging modalities (16, 22). For any differences in opinion, a consensus was reached by discussion together. The ROI was drawn manually for $SUV_{max}$ measurement.

## Statistical analyses

Shapiro–Wilk normality test was used to analyze the data distribution. Data were expressed as mean ± standard deviation (SD). Data of physiological and benign uptake regions were tested by independent sample T-test. Chi-square test and logistic regression analysis were used to investigate the influence of age and sex on the incidence of physiological uptake regions. Pearson correlation analysis was performed for $SUV_{max}$ of physiological regions and age. Statistical analysis was performed using SPSS 23.0 statistical software. Two-tailed $p < 0.05$ was considered statistically significant.

# Results

## Patient characteristics

Patient characteristics are summarized in Table 1. Briefly, 137 patients (84 male and 53 female; age, 58 ± 14 years; range from 18 to 86 years, mostly diagnosed with cancer) were included in this study. Non-tumoral regions were observed in the majority of patients (86.86%). A total of 392 non-tumoral regions were classified as physiological regions (n = 289, $SUV_{max}$ = 3.62 ± 2.86) or benign regions (n = 103, $SUV_{max}$ = 3.50 ± 2.25) according to other imaging features, clinical representations, or pathological results. T-test indicated no statistically significant difference between the physiological and benign groups ($p = 0.40$).

## Physiological uptake regions

The physiological uptakes are summarized in Table 2 and Figure 1. Elevated $^{68}$Ga-FAPI-04 uptakes in the head and neck were primarily observed in the submandibular gland (n = 118, $SUV_{max}$ range from 1.46 to 7.83) and gingiva (n = 65, $SUV_{max}$ range from 1.43 to 7.61), while in the chest, elevated uptake was mainly located in the nipple (n = 37, $SUV_{max}$ range from 1.12 to 4.88) and esophagus (n = 31, $SUV_{max}$ range from 1.33 to 3.87). Although the incidence of $^{68}$Ga-FAPI-04 uptake in the submandibular gland, gingiva, nipple, and esophagus was independent of age ($p$>0.05), the tracer uptakes in the gingiva and esophagus were more common in male patients ($p = 0.006$, 0.009), while uptakes in the nipple were more common in female patients ($p$<0.001) (Table 3). The uptake values of $^{68}$Ga-FAPI-04 ($SUV_{max}$) in the submandibular gland were positively correlated with age ($p = 0.01$) and higher in male patients ($p = 0.001$), while those in other physiological regions were independent of age and sex (all $p \geq 0.05$) (Table 3).

TABLE 1 Patient characteristics.

| N = 137 | Overall |
|---|---|
| Age | |
| Mean (SD) | 58 (14) |
| Median [min, max] | 60 [18,86] |
| Sex | |
| Male | 84 |
| Female | 53 |
| Diagnosis | |
| Liver cancer | 38 |
| Gastric cancer | 36 |
| Gynecological cancer | 11 |
| Lung cancer | 7 |
| Colorectal cancer | 6 |
| Pancreatic cancer | 5 |
| Neuroendocrine tumor | 2 |
| Duodenal tumor | 2 |
| Renal cancer | 2 |
| Breast cancer | 1 |
| Osteosarcoma | 1 |
| Esophagus cancer | 1 |
| Lymphoma | 1 |
| Prostate cancer | 1 |
| Other | 23 |

TABLE 2 Physiological uptake of $^{68}$Ga-FAPI-04.

| | Incidence (%) | | | SUVmax | | |
|---|---|---|---|---|---|---|
| | All | Male | Female | All | Male | Female |
| Submandibular gland | 45.26 | 51.19 | 35.85 | 2.81 ± 0.97 | 3.01 ± 1.05 | 2.33 ± 0.52 |
| Gingiva | 47.45 | 57.14 | 32.08 | 3.82 ± 1.33 | 3.94 ± 1.31 | 3.47 ± 1.36 |
| Nipple | 14.6 | 3.57 | 32.08 | 2.6 ± 1.17 | 1.94 ± 0.53 | 2.73 ± 1.22 |
| Esophagus | 22.63 | 30.95 | 9.43 | 2.31 ± 0.64 | 2.34 ± 0.59 | 2.13 ± 0.9 |
| Uterus | NA | NA | 71.7 | NA | NA | 7.82 ± 5.78 |

NA, Not Applicable.

High uptake of $^{68}$Ga-FAPI-04 in the uterus was also very common (n=38, mean $SUV_{max}$ = 7.82 ± 5.78, $SUV_{max}$ range from 1.26 to 24.33). Increased $^{68}$Ga-FAPI-04 uptake in the uterus was observed in 71.70% of female patients and occurred preferentially in premenopausal women (82.14%, $p$ = 0.07). The $SUV_{max}$ in the uterus did not correlate with the patients' age (r = −0.11, $p$ = 0.50). When comparing $SUV_{max}$ in the uterus between premenopausal and postmenopausal groups, no statistically significant difference was observed ($SUV_{max}$ = 8.40 ± 5.64 vs. 6.93 ± 6.06, $p$ = 0.50).

## Benign uptake regions

The benign regions included inflammatory lymph node (n = 10, mean $SUV_{max}$ = 2.75 ± 1.13, $SUV_{max}$ range from 1.37 to 4.91), pneumonia (n = 13, mean $SUV_{max}$ = 3.37 ± 1.22, $SUV_{max}$ range from 1.68 to 6.24), atherosclerosis (n = 10, mean $SUV_{max}$ = 2.85 ± 0.84, $SUV_{max}$ range from 1.49 to 4.65), pancreatitis (n = 18, mean $SUV_{max}$ = 3.41 ± 3.74, $SUV_{max}$ range from 1.27 to 18.11), osteosclerosis (n = 45, mean $SUV_{max}$ = 3.93± 2.22, $SUV_{max}$ range from 1.28 to 11.09), and surgical scar (n = 7, mean $SUV_{max}$ = 3.14 ± 0.98, $SUV_{max}$ range from 1.83 to 4.56). There was no significant difference in $SUV_{max}$ between these regions ($p$>0.05).

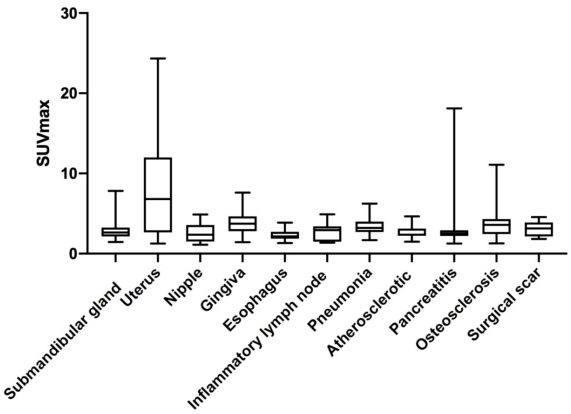



FIGURE 1
$SUV_{max}$ of non-tumoral tracer uptake regions.

We found some interesting cases with high $^{68}$Ga-FAPI-04 uptakes. A patient with a 30-year history of hepatitis B showed high $^{68}$Ga-FAPI-04 uptake in the liver (Figure 2A). High $^{68}$Ga-FAPI-04 uptake has also been found in the rectum of a patient with Crohn's disease (Figure 2B). A man diagnosed with disseminated non-tuberculous mycobacteriosis (tuber colectomy of the left chest wall and CT-guided percutaneous lung puncture biopsy found inflammatory granulomatous lesions; prostate puncture pathology revealed non-specific granulomatous prostatitis; second-generation DNA sequencing results suggested occasional mycobacterium infection) showed lesions throughout the body with high or mild uptake of $^{68}$Ga-FAPI-04 (Figure 2C). After anti-infective therapy, the intracranial lesions became smaller.

# Discussion

Due to the specific expression of FAP in tumor stromal fibrous tissues, FAP has received increasing attention as a specific marker of CAFs. Meanwhile, activated fibroblasts that undergo extracellular matrix (ECM) remodeling in the tissue due to chronic inflammation, fibrosis, and wound healing can also be observed by FAPI imaging (23–25). In this study, we described the $SUV_{max}$ of 392 non-tumoral uptake regions in 137 patients who underwent $^{68}$Ga-FAPI-04 PET/CT or PET/MR.

Consistent with previous studies (7, 26), physiological uptakes of $^{68}$Ga-FAPI-04 were observed in the submandibular gland, nipple, gingiva, and esophagus (Figures 3A–C). In our study, the incidence of $^{68}$Ga-FAPI-04 uptakes in the submandibular gland, gingiva, nipple, and esophagus were independent of age. Tracer uptakes in the gingiva and esophagus were more common in male patients, whereas uptake in the nipple was more common in female patients. It indicates that sex may have a more significant effect on physiological expression of FAPI than age. The uptake values of $^{68}$Ga-FAPI-04 ($SUV_{max}$) in the submandibular gland were positively correlated with age, suggesting that FAP activity in the submandibular gland may be affected by age.

TABLE 3 Main effects of age and sex on the $^{68}$Ga-FAPI-04 uptake.

| | Age | | | Sex | | |
|---|---|---|---|---|---|---|
| | Coef | [95% CI] | *p* | Coef | [95% CI] | *p* |
| Incidence | | | | | | |
| Submandibular gland | 0.01 | [−0.01, 0.04] | 0.31 | −0.53 | [−1.26, −0.20] | 0.16 |
| Gingiva | −0.003 | [−0.03, 0.02] | 0.8 | −1.07 | [−1.82, −0.31] | 0.006 |
| Nipple | −0.01 | [−0.05, 0.03] | 0.6 | 2.47 | [1.16, 3.79] | <0.001 |
| Esophagus | 0.006 | [−0.03, 0.04] | 0.75 | −1.42 | [−2.48, −0.36] | 0.009 |
| SUVmax | | | | | | |
| Submandibular gland | 0.006 | [0.001, 0.01] | 0.01 | −0.21 | [−0.32, −0.09] | 0.001 |
| Gingiva | −0.02 | [−0.05, −0.005] | 0.11 | −0.64 | [−1.40, 0.13] | 0.1 |
| Nipple | 0.006 | [−0.02, 0.03] | 0.68 | 0.84 | [−0.23, 1.92] | 0.12 |
| Esophagus | 0.00 | [−0.01, 0.01] | 0.94 | −0.13 | [−0.40, 0.14] | 0.34 |

The uptake of $^{68}$Ga-FAPI-04 in the uterus was significantly higher compared to other non-tumoral regions in our study (Figure 3C, red arrow). The high uptake in the uterus is considered to stem from the endometrial glandular cells, and its level is significantly lower than that of the malignant component in the uterus (27). Although a recent work suggested that tracer uptake decreases with age (28), in this study, we did not find a significant correlation between the $SUV_{max}$ of uterus and patient age, in line with a previous study (17) reporting that intense $^{68}$Ga-FAPI-04 uptake in the uterus was independent from menopause. High uterine FAP activity might be caused by tissue remodeling and angiogenesis during hormonal periodic changes in regeneration (29).

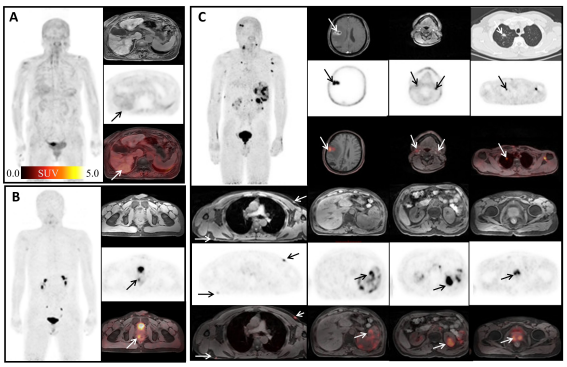



FIGURE 2
Interesting cases $^{68}$Ga-FAPI-04 imaging. **(A)** A 65-year-old woman with a history of hepatitis B over 30 years, arrows, cirrhosis of the liver, $SUV_{max}$ 3.24; **(B)** a 19-year-old man with a 2-year history of rectal Crohn's disease, arrows, rectal Crohn's disease, $SUV_{max}$ 5.22; **(C)** a 56-year-old man diagnosed with disease of disseminated non-tuberculous mycobacteriosis. The maximum intensity projection (MIP) image shows various FAPI-avid nodules: brain ($SUV_{max}$ =2.61), cervical lymph nodes ($SUV_{max}$ =1.81), upper lobe of right lung ($SUV_{max}$ = 1.74), subcutaneous nodule on the left chest ($SUV_{max}$ = 3.52), spleen ($SUV_{max}$= 2.03), left kidney ($SUV_{max}$ = 2.76), and prostate ($SUV_{max}$ = 4.02).

FAP can be induced by fibrosis foci during pulmonary fibrosis in ongoing tissue remodeling (30). In this study, elevated uptake of $^{68}$Ga-FAPI-04 was found in 13 pneumonia lesions (mean $SUV_{max}$ = 3.37) (Figures 4A, B), yet still lower than that in lung cancer lesions ($SUV_{max}$>12) according to the literature (31). Although $^{68}$Ga-FAPI-04 PET/CT is inferior to $^{18}$F-FDG PET/CT in detecting lymph nodes involved in IgG4-related diseases (12), Schmidkonz et al. reported high uptake of FAPI in lymph nodes infiltrated by a fibrotic process and decreased FAPI uptake in those after anti-fibrosis therapy (11). Inflammatory lymph nodes in our study also showed high uptake of $^{68}$Ga-FAPI-04 (Figure 4C), and $SUV_{max}$ was lower than that of fibrotic lymph nodes reported before (11). Mixed type of proliferative and fibrotic lymph nodes in our study may have led to such results.

FAP has recently been proposed as an inflammation-induced protease involved in the formation of vulnerable plaques (32). It has been reported that FAP expression was enhanced in the human atherosclerotic vessel and increased upon plaque progression (33). In our study, atherosclerotic plaques showed slightly high uptake of $^{68}$Ga-FAPI-04 with mean $SUV_{max}$ = 2.85 (Figure 5A). Forty-five joints in our study showed high $^{68}$Ga-FAPI-04 uptake (Figures 5B, C). In a study of the biological distribution of FAPI in cancer patients, mild low-grade uptake in the knee and shoulder was observed in three patients with no clinical symptoms of arthritis (34). FAP expression has been observed in synovial tissue samples of rheumatoid arthritis (35). In osteoarthritis, higher levels of FAP expression on the surface of the cartilage and on chondrocyte membranes were detected by Milner et al. (36). Terry Sy and his colleagues found that In-28H1 (anti-FAP antibody) radionuclide imaging could be used to evaluate the treatment response to etanercept in arthritic mice (13). Therefore, $^{68}$Ga-FAPI-04 might present a potential therapeutic

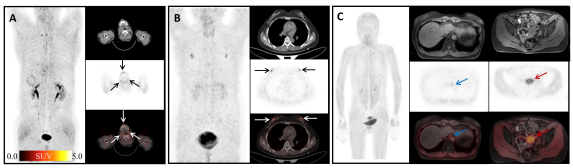



FIGURE 3
Physiological uptake of $^{68}$Ga-FAPI-04. **(A)** A 60-year-old man with hepatic hilar malignancy after the treatment of transcatheter arterial chemoembolization (TACE); the arrows indicate $^{68}$Ga-FAPI-04 uptake in the gingiva and submandibular gland ($SUV_{max}$ 4.45, 3.79, and 3.64). **(B)** A 65-year-old woman with microinvasive lung adenocarcinoma 6 months after surgery; arrows show physiological uptake in the nipples with $SUV_{max}$ 4.88. **(C)** A 55-year-old woman with signet ring cell carcinoma of stomach 2 months after endoscopic submucosal dissection (ESD); blue arrows show physiological uptake of the esophagus ($SUV_{max}$ =3.67), and red arrows show uterus ($SUV_{max}$=13.94).

target of arthritis, and $^{68}$Ga-FAPI-04 imaging has potential value in diagnosis and therapeutic efficacy evaluation in the future.

It has been reported that $^{68}$Ga-FAPI-04 could show focal high uptake in pancreatic fibrous lesions, fibroplasia, or fibrotic activity (14). Our study also found non-tumoral high uptake in pancreas caused by inflammation (mean $SUV_{max}$ = 2.55) (Figures 6A, B). Seven surgical scars in our study showed high uptake of $^{68}$Ga-FAPI-04 (mean $SUV_{max}$ = 3.34) (Figures 6C, D). Keloid is a fibroproliferative reticular dermal disorder characterized by inflammation, increased deposition of ECM protein, and invasion of the surrounding healthy skin (37). FAP expression is observed in keloid (37) and in the physiological process of wound healing (38). Consistent with its role in fibrosis, FAP has been found to be expressed in fibroblasts and hepatic stellate cells (HSCs) activated in cirrhosis but not in normal human livers (39, 40). Crohn's disease is a chronic inflammatory bowel disease in which myofibroblasts play a key role in the process of fibrosis. It

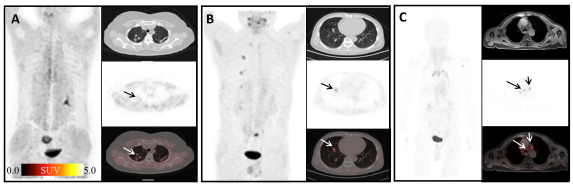



FIGURE 4
Physiological uptake of $^{68}$Ga-FAPI-04. **(A)** A 64-year-old female patient with gastric cancer 1 year after surgery; the arrows show organized pneumonia in the upper lobes of right lung with $SUV_{max}$ of 1.76. **(B)** A 60-year-old male patient after liver cancer surgery; the arrows indicate the inflammatory lesion in the middle lobe of the right lung with $SUV_{max}$ 3.24. **(C)** A 50-year-old male patient with weight loss of 10 kg in recent 6 months; arrows show mediastinal inflammatory lymph nodes with $SUV_{max}$ of 4.91.

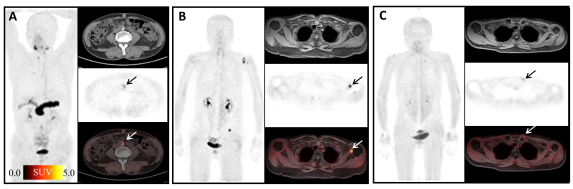



FIGURE 5
Physiological uptake of $^{68}$Ga-FAPI-04. **(A)** A 72-year-old man with duodenal papillary tumor; arrows show atherosclerosis of the abdominal aorta with $SUV_{max}$ =4.65. **(B)** A 70-year-old woman presented with adenocarcinoma at the descending colon–sigmoid junction; arrows show left shoulder arthritis with $SUV_{max}$ =11.09. **(C)** A 55-year-old woman with signet ring cell carcinoma of stomach 2 months after ESD; arrows show left sternoclavicular arthritis with $SUV_{max}$ = 8.21.

is worth mentioning that the myofibroblasts isolated from a colon specimen of a patient with stenosis were FAP positive. Tumor necrosis factor (TNF) and transforming growth factor (TGF) can further induce the expression of FAP (41). The systemic non-

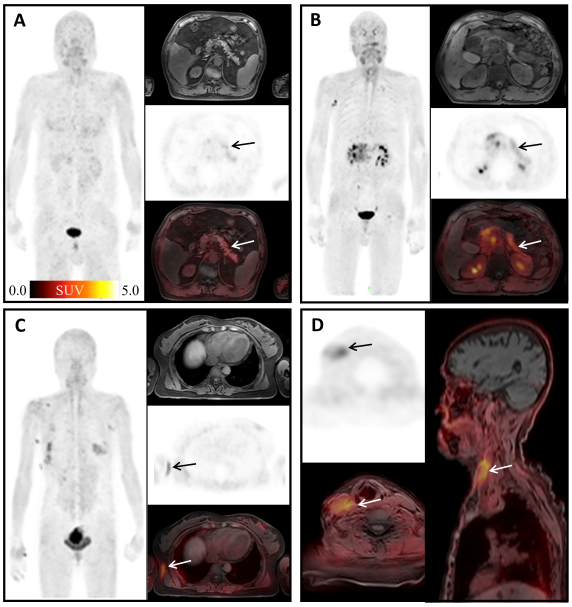



FIGURE 6
Physiological uptake of $^{68}$Ga-FAPI-04. **(A)** A 67-year-old man with mucinous adenocarcinoma of the right lung 1 year after surgery; arrows show pancreatic diffuse inflammatory uptake, $SUV_{max}$ =4.66. **(B)** A 63-year-old man diagnosed with duodenal malignancy; arrows show obstructive pancreatitis, $SUV_{max}$ = 18.11. **(C)** A 49-year-old woman, 3 months after surgery for early microinfiltrating adenocarcinoma of the right middle lobe and 2 months after surgery for left breast fibroma; arrows show surgical scar on the right chest with $SUV_{max}$ =3.89. **(D)** An 83-year-old woman diagnosed with gastric cancer 6 months ago; arrows show deep vein catheterization area of the right neck with $SUV_{max}$ =4.56.

tuberculous mycobacterium granuloma case suggests that $^{68}$Ga-FAPI-04 PET can be used as an effective imaging tool to detect the degree of infection and evaluate the therapeutic effect.

Similar to findings of FDG, our study showed that regions with high $^{68}$Ga-FAPI-04 uptake did not necessarily represent malignancy. A previous FAPI study analyzed $SUV_{max}$ of 28 different types of tumors (31) and reported that although $^{68}$Ga-FAPI-04 uptake was higher in malignant lesions than in benign lesions and physiological uptake regions, there was still some overlap. There was no statistically significant difference in $SUV_{max}$ between the benign uptake regions and the physiological regions. This suggests that $SUV_{max}$ cannot be used as a differential diagnostic index of physiological and benign uptake regions.

There were some limitations in our study. Since this study was retrospective, pathological verification of the lesions was challenging. Most of the diagnosis were based on the clinical history and the experience of the reviewers and with reference to other imaging modalities (CT, MRI, ultrasound, etc.), similar to previous studies. Although our sample size was relatively large, it was not possible to cover all non-tumor uptake regions. There is still a need to accumulate more cases in order to summarize the features of non-malignant lesion uptake in $^{68}$Ga-FAPI-04 as a way to improve the accuracy of diagnosis.

## Conclusions

This study evaluated the $SUV_{max}$ of $^{68}$Ga-FAPI-04 in non-tumoral uptake regions with a relatively large sample population and elaborated the possible pathophysiological mechanisms of these non-tumoral uptake regions. The results indicated that quite a few tissues exhibit physiological uptake of $^{68}$Ga-FAPI-04. Gender has a more significant effect on physiological expression of $^{68}$Ga-FAPI-04 than age. No statistical differences in $^{68}$Ga-FAPI-04 uptake were found between benign and physiological high uptake regions. Our study showed that regions with high $^{68}$Ga-FAPI-04 uptake did not necessarily represent malignancy, and therefore, being familiar with physiological $^{68}$Ga-FAPI-04 uptake and the uptake of typical benign lesions can be helpful for physicians to interpret images and diagnose disease.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

## Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

NQ participated in its design and coordination and drafted the manuscript. HW conducted statistical processing on the data. HYW and SR reviewed the images. ZY and XC contributed to data collection. YG, FX, FH, and JZ provided critical review and substantially revised the manuscript. All authors read and approved the final manuscript.

## Funding

The study was partially supported by the National Natural Science Foundation of China (81871388), Project of Science and Technology Commission of Shanghai Municipality (19DZ1930703).

## Acknowledgments

We would like to thank Qiaoyi Xue (Central Research Institute, UIH Group, Shanghai, China) for her linguistic assistance and her substantial revision during the preparation of this manuscript.

## 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.

## Publisher's note

# References

1. Jansen K, Heirbaut L, Cheng JD, Joossens J, Ryabtsova O, Cos P, et al. Selective inhibitors of fibroblast activation protein (FAP) with a (4-Quinolinoyl)-glycyl-2-cyanopyrrolidine scaffold. *ACS Med Chem Lett* (2013) 4(5):491–6. doi: 10.1021/ml300410d

2. Poplawski SE, Lai JH, Li Y, Jin Z, Liu Y, Wu W, et al. Identification of selective and potent inhibitors of fibroblast activation protein and prolyl oligopeptidase. *J Med Chem* (2013) 56(9):3467–77. doi: 10.1021/jm400351a

3. Pure E, Blomberg R. Pro-tumorigenic roles of fibroblast activation protein in cancer: Back to the basics. *Oncogene* (2018) 37(32):4343–57. doi: 10.1038/s41388-018-0275-3

4. Loktev A, Lindner T, Mier W, Debus J, Altmann A, Jäger D, et al. A tumor-imaging method targeting cancer-associated fibroblasts. *J Nucl Med* (2018) 59 (9):1423–9. doi: 10.2967/jnumed.118.210435

5. Giesel FL, Adeberg S, Syed M, Lindner T, Jimenez-Franco LD, Mavriopoulou E, et al. FAPI-74 PET/CT using either (18)F-AlF or cold-kit (68)Ga labeling: Biodistribution, radiation dosimetry, and tumor delineation in lung cancer patients. *J Nucl Med* (2021) 62(2):201–7. doi: 10.2967/jnumed.120.245084

6. Lindner T, Loktev A, Altmann A, Giesel F, Kratochwil C, Debus J, et al. Development of quinoline-based theranostic ligands for the targeting of fibroblast activation protein. *J Nucl Med* (2018) 59(9):1415–22. doi: 10.2967/jnumed.118.210443

7. Giesel FL, Kratochwil C, Lindner T, Marschalek MM, Loktev A, Lehnert W, et al. (68)Ga-FAPI PET/CT: Biodistribution and preliminary dosimetry estimate of 2 DOTA-containing FAP-targeting agents in patients with various cancers. *J Nucl Med* (2019) 60(3):386–92. doi: 10.2967/jnumed.118.215913

8. Niedermeyer J, Garin-Chesa P, Kriz M, Hilberg F, Mueller E, Bamberger U, et al. Expression of the fibroblast activation protein during mouse embryo development. *Int J Dev Biol* (2001) 45(2):445–7.

9. Rettig WJ, Garin-Chesa P, Beresford HR, Oettgen HF, Melamed MR, Old LJ. Cell-surface glycoproteins of human sarcomas: Differential expression in normal and malignant tissues and cultured cells. *Proc Natl Acad Sci U.S.A.* (1988) 85 (9):3110–4. doi: 10.1073/pnas.85.9.3110

10. Heckmann MB, Reinhardt F, Finke D, Katus HA, Haberkorn U, Leuschner F, et al. Relationship between cardiac fibroblast activation protein activity by positron emission tomography and cardiovascular disease. *Circ Cardiovasc Imaging* (2020) 13(9):e010628. doi: 10.1161/CIRCIMAGING.120.010628

11. Schmidkonz C, Rauber S, Atzinger A, Agarwal R, Gotz TI, Soare A, et al. Disentangling inflammatory from fibrotic disease activity by fibroblast activation protein imaging. *Ann Rheum Dis* (2020) 79(11):1485–91. doi: 10.1136/annrheumdis-2020-217408

12. Luo Y, Pan Q, Yang H, Peng L, Zhang W, Li F. Fibroblast activation protein-targeted PET/CT with (68)Ga-FAPI for imaging IgG4-related disease: Comparison to (18)F-FDG PET/CT. *J Nucl Med* (2021) 62(2):266–71. doi: 10.2967/jnumed.120.244723

13. Terry SY, Koenders MI, Franssen GM, Nayak TK, Freimoser-Grundschober A, Klein C, et al. Monitoring therapy response of experimental arthritis with radiolabeled tracers targeting fibroblasts, macrophages, or integrin Alphavbeta3. *J Nucl Med* (2016) 57(3):467–72. doi: 10.2967/jnumed.115.162628

14. Zhang X, Song W, Qin C, Liu F, Lan X. Non-malignant findings of focal (68) Ga-FAPI-04 uptake in pancreas. *Eur J Nucl Med Mol Imaging* (2021) 28(8):2635–41. doi: 10.1007/s00259-021-05194-6

15. Qin C, Song Y, Liu X, Gai Y, Liu Q, Ruan W, et al. Increased uptake of (68) Ga-DOTA-FAPI-04 in bones and joints: Metastases and beyond. *Eur J Nucl Med Mol Imaging* (2021) 49(2):709–20. doi: 10.1007/s00259-021-05472-3

16. Zheng S, Lin R, Chen S, Zheng J, Lin Z, Zhang Y, et al. Characterization of the benign lesions with increased (68)Ga-FAPI-04 uptake in PET/CT. *Ann Nucl Med* (2021) 35(12):1312–20. doi: 10.1007/s12149-021-01673-w

17. Gundogan C, Guzel Y, Can C, Kaplan I, Komek H. FAPI-04 uptake in healthy tissues of cancer patients in (68)Ga-FAPI-04 PET/CT imaging. *Contrast Media Mol Imaging* (2021) 2021:9750080. doi: 10.1155/2021/9750080

18. Jiang D, Chen X, You Z, Wang H, Zhang X, Li X, et al. Comparison of [(68) Ga]Ga-FAPI-04 and [(18)F]-FDG for the detection of primary and metastatic lesions in patients with gastric cancer: a bicentric retrospective study. *Eur J Nucl Med Mol Imaging* (2021) 49(2):732–42. doi: 10.1007/s00259-021-05441-w

19. Wang Q, Li YM, Li Y, Hua FC, Wang QS, Zhang XL, et al. (18)F-FDGPET/CT in fever of unknown origin and inflammation of unknown origin: a Chinese multi-center study. *Eur J Nucl Med Mol Imaging* (2019) 46(1):159–65. doi: 10.1007/s00259-018-4121-1

20. Kou Y, Jiang X, Yao Y, Shen J, Jiang X, Chen S, et al. Physiological tracer distribution and benign lesion incidental uptake of Al18F-NOTA-FAPI-04 on PET/CT imaging. *Nucl Med Commun* (2022) 43(7):847–54. doi: 10.1097/MNM.0000000000001563

21. Mu X, Huang X, Li M, Sun W, Fu W. Comparison of physiological uptake of normal tissues in patients with cancer using 18F-FAPI-04 and 18F-FAPI-42 PET/CT. *Front Nucl Med* (2022) 2:927843. doi: 10.3389/fnume.2022.927843

22. Dabir M, Novruzov E, Mattes-György K, Beu M, Dendl K, Antke C, et al. Distinguishing benign and malignant findings on [68 ga]-FAPI PET/CT based on quantitative SUV measurements. *Mol Imaging Biol* (2022) 1–10. doi: 10.1007/s11307-022-01759-5

23. Grinnell F. Fibroblasts, myofibroblasts, and wound contraction. *J Cell Biol* (1994) 124(4):401–4. doi: 10.1083/jcb.124.4.401

24. Yen TW, Aardal NP, Bronner MP, Thorning DR, Savard CE, Lee SP, et al. Myofibroblasts are responsible for the desmoplastic reaction surrounding human pancreatic carcinomas. *Surgery* (2002) 131(2):129–34. doi: 10.1067/msy.2002.119192

25. Sappino AP, Skalli O, Jackson B, Schurch W, Gabbiani G. Smooth-muscle differentiation in stromal cells of malignant and non-malignant breast tissues. *Int J Cancer* (1988) 41(5):707–12. doi: 10.1002/ijc.2910410512

26. Keane FM, Yao TW, Seelk S, Gall MG, Chowdhury S, Poplawski SE, et al. Quantitation of fibroblast activation protein (FAP)-specific protease activity in mouse, baboon and human fluids and organs. *FEBS Open Bio* (2013) 4:43–54. doi: 10.1016/j.fob.2013.12.001

27. Dolznig H, Schweifer N, Puri C, Kraut N, Rettig WJ, Kerjaschki D, et al. Characterization of cancer stroma markers: in silico analysis of an mRNA expression database for fibroblast activation protein and endosialin. *Cancer Immun* (2005) 5:10.

28. Kessler L, Ferdinandus J, Hirmas N, Zarrad F, Nader M, Kersting D, et al. Pitfalls and common findings in (68)Ga-FAPI-PET - a pictorial analysis. *J Nucl Med* (2021) 63(6):890–6. doi: 10.2967/jnumed.121.262808

29. Zhang X, Song W, Qin C, Song Y, Liu F, Hu F, et al. Uterine uptake of 68Ga-FAPI-04 in uterine pathology and physiology. *Clin Nucl Med* (2022) 47(1):7–13. doi: 10.1097/rlu.0000000000003968

30. Acharya PS, Zukas A, Chandan V, Katzenstein AL, Pure E. Fibroblast activation protein: a serine protease expressed at the remodeling interface in idiopathic pulmonary fibrosis. *Hum Pathol* (2006) 37(3):352–60. doi: 10.1016/j.humpath.2005.11.020

31. Kratochwil C, Flechsig P, Lindner T, Abderrahim L, Altmann A, Mier W, et al. (68)Ga-FAPI PET/CT: Tracer uptake in 28 different kinds of cancer. *J Nucl Med* (2019) 60(6):801–5. doi: 10.2967/jnumed.119.227967

32. Meletta R, Muller Herde A, Chiotellis A, Isa M, Rancic Z, Borel N, et al. Evaluation of the radiolabeled boronic acid-based FAP inhibitor MIP-1232 for atherosclerotic plaque imaging. *Molecules* (2015) 20(2):2081–99. doi: 10.3390/molecules20022081

33. Brokopp CE, Schoenauer R, Richards P, Bauer S, Lohmann C, Emmert MY, et al. Fibroblast aactivation protein is induced by inflammation and degrades type I collagen in thin-cap fibroatheromata. *Eur Heart J* (2011) 32(21):2713–22. doi: 10.1093/eurheartj/ehq519

34. Scott AM, Wiseman G, Welt S, Adjei A, Lee FT, Hopkins W, et al. A phase I dose-escalation study of sibrotuzumab in patients with advanced or metastatic fibroblast activation protein-positive cancer. *Clin Cancer Res* (2003) 9(5):1639–47.

35. Bauer S, Jendro MC, Wadle A, Kleber S, Stenner F, Dinser R, et al. Fibroblast activation protein is expressed by rheumatoid myofibroblast-like synoviocytes. *Arthritis Res Ther* (2006) 8(6):R171. doi: 10.1186/ar2080

36. Milner JM, Kevorkian L, Young DA, Jones D, Wait R, Donell ST, et al. Fibroblast activation protein alpha is expressed by chondrocytes following a pro-inflammatory stimulus and is elevated in osteoarthritis. *Arthritis Res Ther* (2006) 8 (1):R23. doi: 10.1186/ar1877

37. Dienus K, Bayat A, Gilmore BF, Seifert O. Increased expression of fibroblast activation protein-alpha in keloid fibroblasts: Implications for development of a novel treatment option. *Arch Dermatol Res* (2010) 302(10):725–31. doi: 10.1007/s00403-010-1084-x

38. Gao MQ, Kim BG, Kang S, Choi YP, Park H, Kang KS, et al. Stromal fibroblasts from the interface zone of human breast carcinomas induce an epithelial-mesenchymal transition-like state in breast cancer cells in vitro. *J Cell Sci* (2010) 123(Pt 20):3507–14. doi: 10.1242/jcs.072900

39. Rettig WJ, Su SL, Fortunato SR, Scanlan MJ, Raj BK, Garin-Chesa P, et al. Fibroblast activation protein: Purification, epitope mapping and induction by growth factors. *Int J Cancer* (1994) 58(3):385–92. doi: 10.1002/ijc.2910580314

40. Levy MT, McCaughan GW, Abbott CA, Park JE, Cunningham AM, Muller E, et al. Fibroblast activation protein: a cell surface dipeptidyl peptidase and gelatinase expressed by stellate cells at the tissue remodelling interface in human cirrhosis. *Hepatology* (1999) 29(6):1768–78. doi: 10.1002/hep.510290631

41. Rovedatti L, Di Sabatino A, Knowles CH, Sengupta N, Biancheri P, Corazza GR, et al. Fibroblast activation protein expression in crohn's disease strictures. *Inflammation Bowel Dis* (2011) 17(5):1251–3. doi: 10.1002/ibd.21446

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 16 December 2022
DOI 10.3389/fonc.2022.941744



OPEN ACCESS

EDITED BY
Vikram D Kodibagkar,
Arizona State University, United States

REVIEWED BY
Gianni Lazzarin,
Abano Terme Hospital, Italy
Xu Yan,
Siemens Healthineers, China

*CORRESPONDENCE
Zijie Zhang
sjtuzzj@163.com
Xiaohua Qian
xiaohua.qian@sjtu.edu.cn
Wei Chen
chensurg@aliyun.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 11 May 2022
ACCEPTED 21 November 2022
PUBLISHED 16 December 2022

CITATION
Dong Z, Chen X, Cheng Z, Luo Y, He M, Chen T, Zhang Z, Qian X and Chen W (2022) Differential diagnosis of pancreatic cystic neoplasms through a radiomics-assisted system. *Front. Oncol.* 12:941744. doi: 10.3389/fonc.2022.941744



# Differential diagnosis of pancreatic cystic neoplasms through a radiomics-assisted system

Zhenglin Dong[1,2†], Xiahan Chen[3†], Zhaorui Cheng[4], Yuanbo Luo[5], Min He[1], Tao Chen[1], Zijie Zhang[1,6*], Xiaohua Qian[3*] and Wei Chen[1*]

[1]Department of Biliary-Pancreatic Surgery, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China, [2]Department of orthopedics, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [3]School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China, [4]Department of Rheumatology, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China, [5]Department of Otorhinolaryngology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [6]Department of Liver Surgery, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

Pancreatic cystic neoplasms (PCNs) are a group of heterogeneous diseases with distinct prognosis. Existing differential diagnosis methods require invasive biopsy or prolonged monitoring. We sought to develop an inexpensive, non-invasive differential diagnosis system for PCNs based on radiomics features and clinical characteristics for a higher total PCN screening rate. We retrospectively analyzed computed tomography images and clinical data from 129 patients with PCN, including 47 patients with intraductal papillary mucinous neoplasms (IPMNs), 49 patients with serous cystadenomas (SCNs), and 33 patients with mucinous cystic neoplasms (MCNs). Six clinical characteristics and 944 radiomics features were tested, and nine features were finally selected for model construction using DXScore algorithm. A five-fold cross-validation algorithm and a test group were applied to verify the results. In the five-fold cross-validation section, the AUC value of our model was 0.8687, and the total accuracy rate was 74.23%, wherein the accuracy rates of IPMNs, SCNs, and MCNs were 74.26%, 78.37%, and 68.00%, respectively. In the test group, the AUC value was 0.8462 and the total accuracy rate was 73.61%. In conclusion, our research constructed an end-to-end powerful PCN differential diagnosis system based on radiomics method, which could assist decision-making in clinical practice.

KEYWORDS

**radiomics, computed tomography, pancreatic cystic neoplasm, differential diagnosis, ternary classification model**

# Introduction

Pancreatic cystic lesion (PCL) is one of the most common pancreatic diseases with a prevalence rate as high as 42% (1), and common pancreatic cystic neoplasms (PCN) accounts for 90.5% PCL (2). PCN is identified as a group of heterogeneous diseases with diverse characteristics and different prognosis (3). Serous cystadenoma (SCN), mucinous cystic neoplasm (MCN), and intraductal papillary mucinous neoplasm (IPMN) are the three main types of PCN, accounting for more than 85% of PCN (4). Because of the different rates of malignant transformation, the treatment principles of these three PCN subtypes recommended by the guidelines are variable (3).

Although modern modalities have been ubiquitously applied, the current diagnostic methods that can identify and evaluate PCN are still limited. In recent years, endoscopic ultrasonography–guided fine-needle aspiration (EUS-FNA) has been performed for cyst fluid analysis (5, 6). It is considered as a fairly sensitive tool for distinguishing PCN. However, EUS-FNA is an invasive method. Patients would suffer more pain than the non-invasive methods, such as computed tomography (CT) and MRI (7, 8). More importantly, CT and MRI have been widely applied as parts of health checkup. Compared with MRI, CT is even more widely used, is typically less expensive, and has less time for appointment. Unfortunately, SCNs, MCNs, and branch-duct IPMNs (BD-IPMNs) all demonstrate isolated cystic masses with low density cyst fluid and mild enhancement of the cyst wall on CT imaging, leading to a poor differential diagnosis and assessment even for an experienced radiologist (9). Moreover, the diagnostic accuracy rate of CT for PCN is between 40% and 81% (10–13). Thus, a higher diagnostic rate of contrast CT for PCN would extensively improve the total PCN screening rate.

Therefore, there is an urgent demand to develop a new non-invasive biomarker with a high accuracy in PCN diagnosis. Radiomics, sometimes referred to as "quantitative imaging", is an emerging field focusing on disease auxiliary diagnosis or prognosis prediction and always integrates information such as genomic, transcriptome, and clinical data (14). After automatic feature extraction from images based on different algorithms, radiomics model is validated by support vector machine (SVM) or other classifiers. Through high-dimensional quantifiable features from images, radiomics can effectively and quantitatively evaluate the heterogeneity of images. Radiomics-assisted CT scan–based systems have ameliorated the accuracy of differential diagnosis in several organs, such as lung, central nervous system, rectum, liver, and pancreas (15–19). However, until now, relatively few studies have focused on PCN by analyzing SCNs, MCNs, and IPMNs simultaneously.

Thus, the aim of our study is to develop a non-invasive differential diagnosis system for PCNs based on imaging features in conjunction with patient's clinical information.

# Materials and methods

## Workflow

The brief workflow of image processing and model development was illustrated in Figure 1.

## Patient selection

This retrospective study was approved by Human Research Ethics Committee of our hospital. From December 2013 to August 2022, a total of 253 patients, including 65 IPMN cases, 124 SCN cases, and 64 MCN cases, were pathologically confirmed as patients with PCN. All the patients were screened by the following exclusion criteria (shown in Figure 2.): (i) patients without an available contrast-enhanced CT (CECT) that could be downloaded from the hospital system; (ii) patients without a preoperative CECT 1 month before the surgery; (iii) patients without a complete CECT image that contains the whole pancreas and tumors; and (iv) patients lack one of the following tumor markers [CA50, CA125, CA199, and carcinoembryonic antigen (CEA)]. Finally, 129 patients, including 47 IPMN cases, 49 SCN cases, and 33 MCN cases, were included in this research.

## CT protocols

Patients were told to take quiet respiration to minimize the artifact disturbance. Then, CT examination was performed on a 64-detector spiral CT-system (Somantom Definition AS, Siemens, Erlangen, Germany), providing acquisition capability of 128 slices with the following scanning parameters: 120 kV, 280 mA, 0.33-mm slice thickness with an interval of 0.33 mm, 1.5 spiral pitch, and tube rotation cycle of 0.33 s. The four-phase CT images included (1) plain scan (2), arterial phase (20 to 25 s of delay) (3), venous phase (50 to 55 s of delay), and (4) delay phase.

## Segmentation

All the image segmentations were independently performed by two 5-year experienced pancreatic surgeons. The readers were both blinded to the pathologic result. Cases were divided equally to two readers. After all the work was done, images were exchanged between readers and checked again. Any disagreement would be sent to a director who would make the final decision. The three-dimensional (3D) Region of Interest (ROI) were completed by ITK-SNAP (version 3.8.0). According to the National Comprehensive Cancer Network (NCCN) guidelines (version 2.2019, 9 April 2019), venous phase images were recommended as the best phase for diagnosis (20). Examples of different patients in the venous phase are shown in Figure 3.

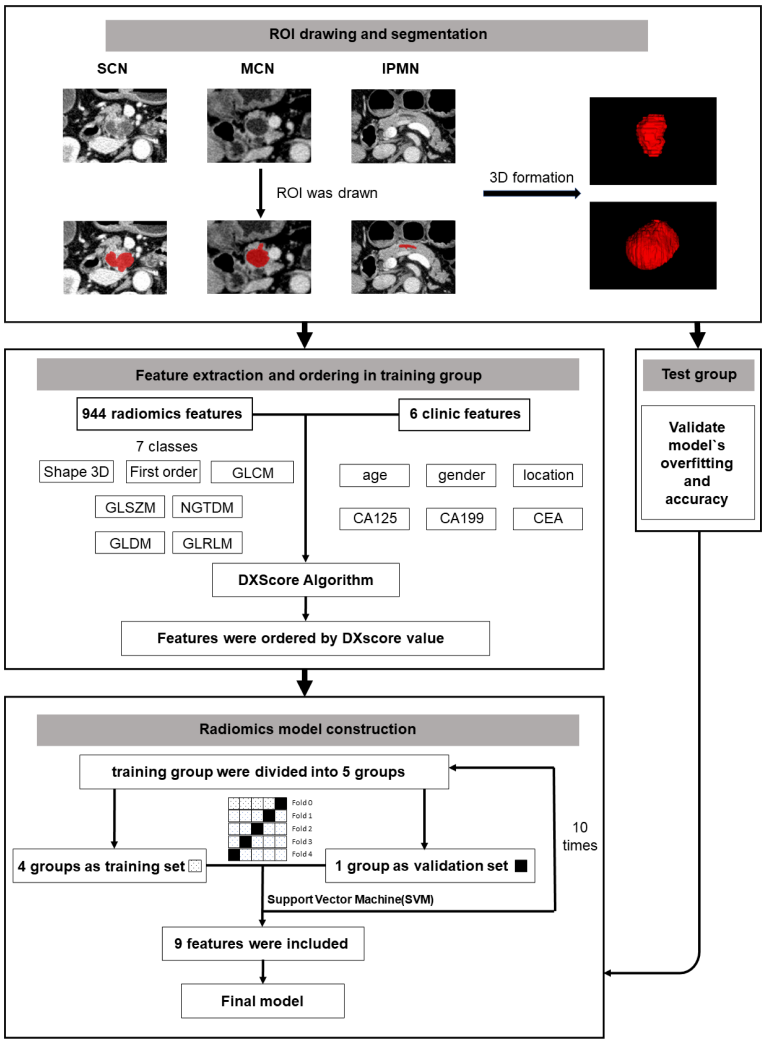



FIGURE 1
A general workflow of image processing and model development.

## Data preprocessing

First, truncate the gray value of the original data to [−100, 240] empirically, which still can fully cover the pancreas in the scans. Then, normalize it into the range [−1, 1]. Second, resample all samples to 1 × 1 × 1 mm according to the physical resolution of the original image. Third, perform the resample of CT image by bilinear interpolation from XY plane and XZ plane with python and OpenCV package.

## Radiomics feature extraction methodology

Five steps were applied to extract the radiomics feature.

1. Randomize the included cases into the training group and the test group. In each disease, 25% of the cases were randomly selected into the test group (Table 1). 2. Calculate the 3D images features by the Pyradiomics package (python 3.6) (21) and statistics clinical features of the dataset. 3. Use DXScore algorithm as follows to arrange the feature: ① Calculate each feature's DXScore. Its mathematical expression is $D(X) = (m_{positive} - m_{negative})^2/(d^2_{positive} + d^2_{negative})$ where $m$ and $d$ are the mean value and standard deviation of the feature X to positive (or negative) samples (22, 23); ② Arrange features in a descending order of the score. 4. Select features and construct model. The features were assessed with SVM in the training group with five-fold cross-validation, which is an effective method for validating overfitting (24, 25). Briefly, select the first N features, in turn, to assess the classification performance by five-fold cross-validation using SVM to obtain the relation of the number of features and the accuracy of classification (as shown in Figure 4).

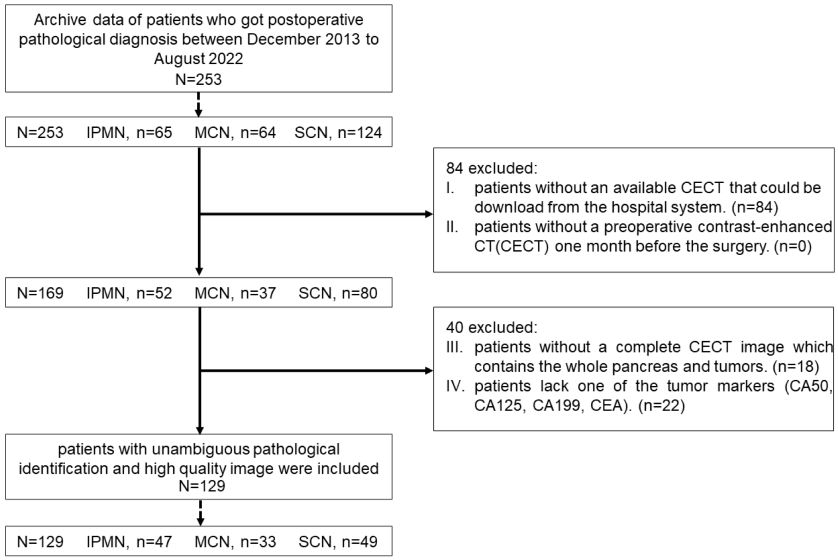



FIGURE 2
Patients' enrolment process for this research.

First, N features were selected under the maximum accuracy. Repeat the fourth step above 10 times to verify the stability of our method. 5. The test group was applied to verify the accuracy and robustness of the model again.

## Statistical analysis

Statistical analysis and graph drawing were performed using SPSS version 26.0 (IBM SPSS Inc., Chicago, IL, USA) and GraphPad prism version 9.4.1 (GraphPad Software., San Diego, CA, USA). *P*-value below 0.05 is considered statistically significant.

# Result

## Patient information

A total of 129 patients, including 47 IPMNs, 49 SCNs, and 33 MCNs, were analyzed in this study. The difference between age ($P = 9.3 \times 10^{-5}$), gender ($P = 4.0 \times 10^{-6}$), tumor location ($P = 1.3 \times 10^{-5}$), CA125 ($P_{SCNs/MCNs} = 5.1 \times 10^{-4}$), CA19-9 ($P = 0.044$), and CEA ($P = 0.001$) was statistically significant among the three categories. More detailed information about the characteristics of the patients is shown in Table 1.

## Selection of radiomics and clinical features

A total of 950 features, including six significant features (age, gender, tumor location, CA125, CA19-9, and CEA) in patients' clinical information, and 944 features from Pyradiomics documentation that could be categorized as seven classes (Figure S1) were extracted in the training group in Section 2.6. The features were ordered by a DXScore value in a descending order. More detailed information about ordered features is shown in Table S1.

## Construction and validation of radiomics model

We performed five-fold cross-validation 10 times to construct models. The standard deviation of area under the receiver operating characteristic curve (AUC) value (0.0152) and accuracy (ACC) value (0.0146) proved the models' robustness and reliability. Balanced with the number of the used features, diagnostic accuracy, and AUC value in each model, we finally chose Model 6 for the following analysis (Figure 5). The relationship between the number of the used features and the diagnostic accuracy in Model 6 is shown in Figure 4. All the selected features' information is shown in Figure 6.

And the respective AUC values of IPMN, SCN, and MCN have also reached 0.9083, 0.8622 and 0.8356 (Figure 7). We are were delighted to find that the overall predictive rate was 74.23% and predictive rate in IPMNs, SCNs and MCNs were 74.26%, 78.37% and 68.00%, respectively. Confusion matrix was also drawn in Figure 7. The sensitivity of IPMN, SCN, and MCN were 74.29%, 78.38% and 68.00%. The specificity of IPMN, SCN, and MCN were 91.94%,81.67% and 87.50%.

The classification model was conducted on the test group to test its robustness and generalization. The ROC curves and detailed information regarding diagnostic discrepancies between model's

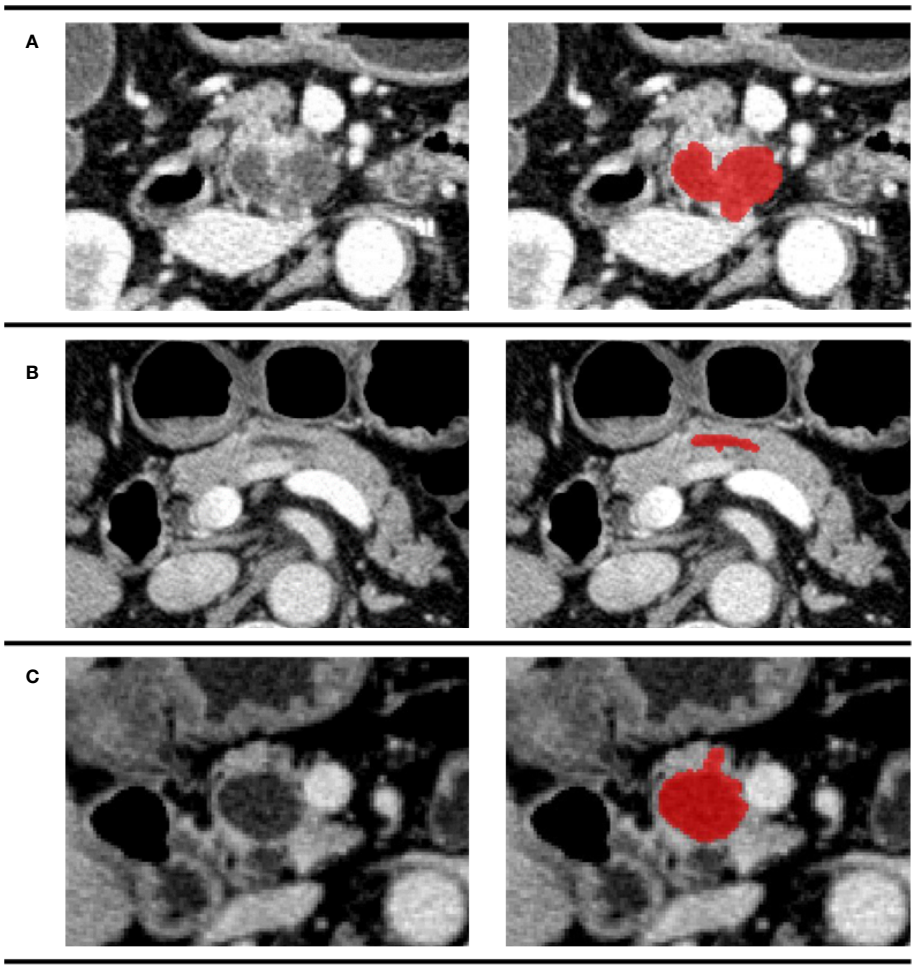



FIGURE 3
Examples of different patients in the venous phase. Tumor regions were drawn with red. **(A)** CECT image of a 74-year-old man pathologically diagnosed with IPMN; **(B)** CECT image of a 61-year-old woman pathologically diagnosed with SCN; **(C)** CECT image of a 70-year-old man pathologically diagnosed with MCN.

predictive result and definitive histological diagnosis in the test group are shown in Figure 8. In brief, the overall AUC value and predictive rate were 0.8462 and 73.61%, respectively, which illustrated the reliability of our model. The SEN rates of IPMN, SCN, and MCN in the test group were 83.33%, 75.00%, and 62.50%. The SPEC rates of IPMN, SCN, and MCN in the test group were 95.00%, 80.00%, and 87.50%.

Least absolute shrinkage selection operator (LASSO) regression is widely used because of its well performance (26–29). We compared DXScore algorithm with LASSO regression in the test group. ACC, AUC, SEN, and SPEC rates were all higher in DXScore algorithm than that in LASSO regression, especially 13.93% higher in ACC (Table 2).

To conduct a deeper and more detailed research, we also analyzed the relationship between selected features and each patient quantitatively. The distribution of the nine selected radiomics features in 129 cases, categorized as IPMNs, SCNs, and MCNs, is shown in Figure 6. For a better visualization, we normalized each value of radiomics features by mean and standard deviation. As is shown in the heat map, the higher the DXScore value of the features was, the deeper the color was used, and distinct patterns among three categories can be observed.

## Discussion

Patients with PCN have distinct treatment principles. Patients with SCNs demonstrate benign preponderance and a good prognosis (30, 31). Patients are recommended with regular monitoring and follow-up (32). MCNs, main pancreatic duct

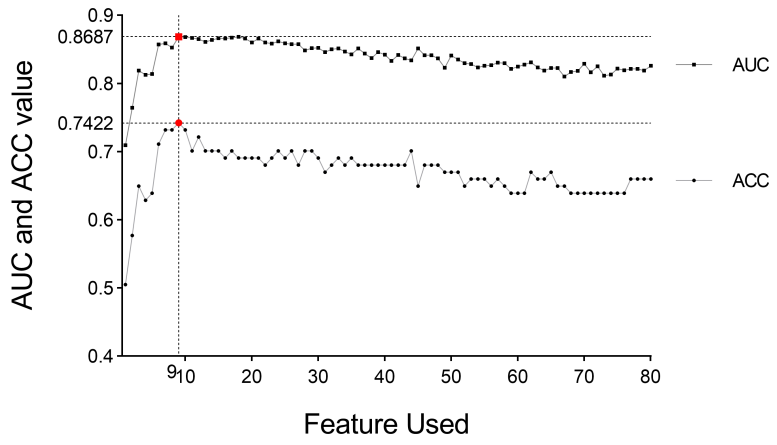



FIGURE 4
AUC value and diagnosis accuracy constructed with the number of used features. The optimal feature number was 9 with the highest ACC value of 74.22% and AUC value of 86.87% in Model 6. Moreover, the respective AUC values of IPMN, SCN, and MCN have also reached 0.9083, 0.8622, and 0.8356 (Figure 7). We find that the overall predictive rate was 74.23% and the predictive rates in IPMNs, SCNs, and MCNs were 74.26%, 78.37%, and 68.00%, respectively. Confusion matrix was also drawn in Figure 7. The sensitivity (SEN) rates of IPMN, SCN, and MCN were 74.29%, 78.38%, and 68.00%. The specificity (SPEC) rates of IPMN, SCN, and MCN were 91.94%, 81.67%, and 87.50%.

IPMNs, and mixed IPMNs show a higher risk of malignant transformation (33, 34). Surgery is recommended to those patients before the neoplasms progress to cancer (35). Although it is the most common type of IPMNs, BD-IPMN has a relatively low malignant tendency and it does not often invade the main pancreatic duct. Because of subtle difference in imaging characteristics, it is currently extremely difficult to perform differential diagnosis through traditional CT and MRI scans. According to the previous studies, the overall preoperative diagnostic accuracy rates of PCNs by clinicians were 37.3% and 61.0%, with SCN diagnostic accuracy rates of 30.4% and 24.2%, which is far from satisfactory (26, 36). Thus, we retrospectively analyzed 129 patients and constructed a DXScore algorithm–based model with an overall accuracy rate of 74.23%.

In our results, we have screened out a few important features. Among those radiomics features, "Original shape Sphericity" was

TABLE 1 Characteristics of patients with IPMNs, SCNs, and MCNs.

| | IPMNs | SCNs | MCNs | *P*-value |
|---|---|---|---|---|
| **Number of patients (training group, test group)** | 47 (35, 12) | 49 (37, 12) | 33 (25, 8) | |
| **Age (mean, range)** | 64.15 (36–83)$^{\#*}$ | 57.90 (27–77)$^{\&}$ | 52.54 (25–79) | $9.3 \times 10^{-5}$ |
| **Gender** | $^{\#*}$ | | | $4.0 \times 10^{-6}$ |
| Male | 30 (63.83%) | 14 (28.57%) | 4 (12.12%) | |
| Female | 17 (36.17%) | 35 (71.43%) | 29 (87.88%) | |
| **Tumor location** | $^{\#*}$ | | | $1.3 \times 10^{-5}$ |
| Head | 28 (59.57%) | 11 (22.45%) | 4 (12.12%) | |
| Neck | 11 (23.40%) | 11 (22.45%) | 5 (15.15%) | |
| Body | 5 (10.64%) | 9 (18.37%) | 9 (27.27%) | |
| Tail | 3 (6.38%) | 18 (36.73%) | 15 (45.45%) | |
| **Tumor marker (mean, range)** | | | | |
| CA50 | 25.05 (0.50, 267.10) | 15.75 (0.83, 74.67) | 14.34 (2.38, 87.79) | 0.192 |
| CA125 | 27.62 (5.02, 555.30) | 5.75$^{\&}$ (0.35, 25.60) | 19.29 (1.30, 95.01) | 0.087 |
| CA19-9 | 46.20 (2.00, 799.20) | 8.88$^{\&}$ (0.35, 80.90) | 23.99 (2.93, 117.00) | 0.044 |
| CEA | 3.16$^{\#}$ (0.91, 25.21) | 13.79$^{\&}$ (0.31, 180.0) | 1.69 (0.41, 4.30) | 0.001 |

# indicates significant difference between IPMNs and SCNs; * indicates significant difference between IPMNs and MCNs; & indicates significant difference between SCNs and MCNs. Age$^{\#}$, P = 0.012; Age*, P = $2.0 \times 10^{-5}$; Age$^{\&}$, P = 0.033; Gender$^{\#}$, P = $5.3 \times 10^{-4}$; Gender*, P = $4.0 \times 10^{-6}$; Tumor location$^{\#}$, P = $2.5 \times 10^{-4}$; Tumor location*, P = $4.0 \times 10^{-6}$; CA125$^{\&}$, P = $5.1 \times 10^{-4}$; CA19-9$^{\&}$, P = 0.016; CEA$^{\#}$, P = 0.022; CEA$^{\&}$, P = 0.007.

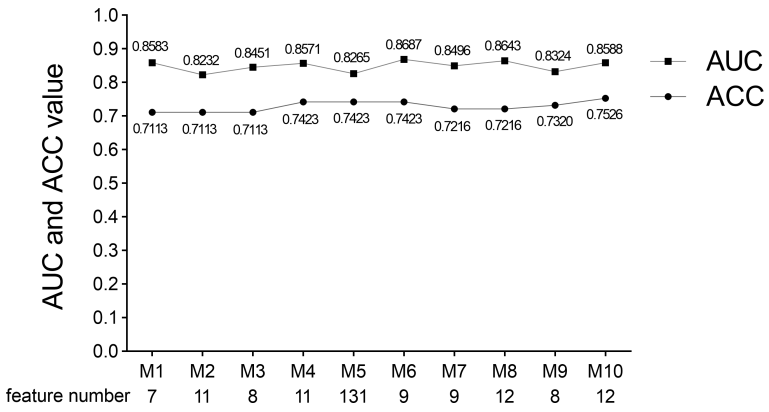



FIGURE 5
AUC value, ACC value, and the used feature number of 10 times five-fold cross-validation models. Model 6 was selected because of its suitable feature number and slightly higher AUC and ACC values. M indicates Model.

the most important one. Sphericity is a measure of the roundness of the shape of the tumor region relative to a sphere (37). In general, non-SCNs demonstrate a regular oval shape and usually hold a smooth contour, whereas SCNs tended to have a multicystic or lobulated shape. Sahani et al. and Kim et al. also discovered the importance of surface contours in diagnosing PCN (38, 39). Because the following seven radiomics features, included in model (Figure 6), were obtained through wavelet and principal component analysis that cannot be directly reflected in the original images, there was no intuitive clinical implication. As for the clinical features, tumor location was selected by DXScore algorithm, ranking first among the nine features. Indeed, the majority of IPMN is located in the head of the pancreas in traditional analysis, whereas SCN and MCN are often localized in the body or tail of the pancreas (40–43). Tumor location may have a more intuitive clinical implication than other clinical features in PCNs diagnosis. Whereas, five significant ones (age, gender, CA125, CA19-9, and CEA) were not included in the machine learning algorithm. Wei et al. found that gender was an important SCN diagnostic factor (26). Moreover, there was a controversial in the relationship between age and PCNs. Kim et al., Goh et al., and Wei et al. found that age was an insignificant differential diagnostic factor. Whereas, several other studies have considered age as an important factor (38, 44, 45). However, those were finally not selected by DXScore in our study. Therefore, this is a typical example that artificial intelligence is able to surpass people's inherent logic and achieve a better classification result.

For a more precise model and automatic process, there would be two steps: first is the automatic identification of the pancreas and second is the construction of models with higher accuracy. Manual segmentation takes time and effort. Although the accuracy of automatic pancreas segmentation has been up to 85%, there is a long way before automatic pancreas segmentation can be applied in clinical practices (46–48). For our sake, we will further explore in the field of automatic pancreas segmentation to accelerate the process of clinical application of automatic

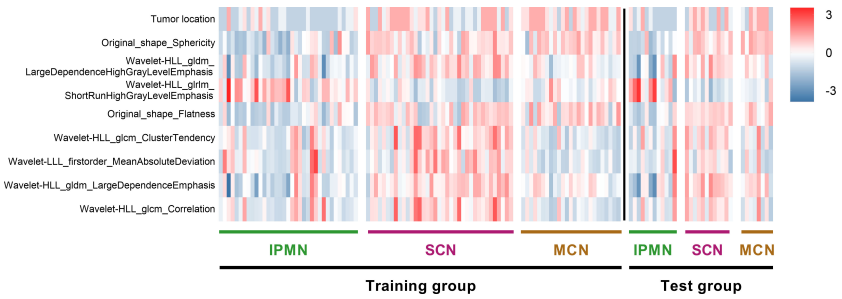



FIGURE 6
Heat map of radiomics features and cases in the three categories. The features on y-axis from top to bottom are arranged in a descending order according to the DXScore value. Right color bar represents color coding of normalized value of each radiomics feature.

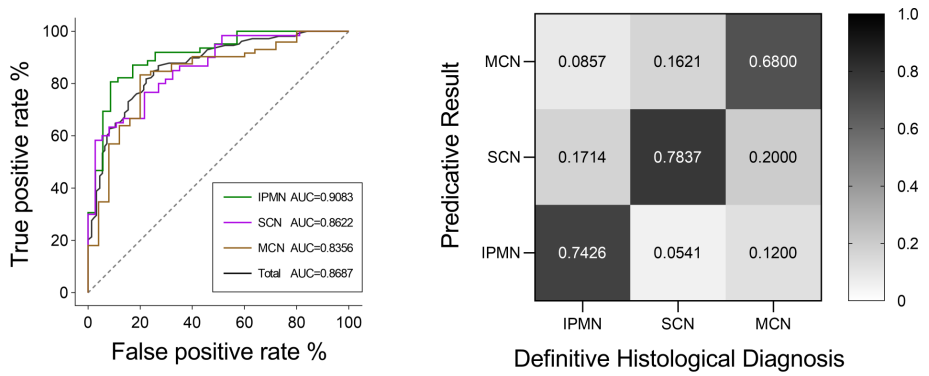



FIGURE 7
ROC curves of IPMN, SCN, and MCN and the total cases for the Model 6. The final AUC values of IPMN, SCN, and MCN were 0.9083, 0.8622, and 0.8356, respectively. In addition, the mean AUC value of five-fold cross-validation was 0.8687. ROC, receiver operating characteristic; AUC, area under the ROC curve. Relative confusion matrix of the three categories. Right color bar represents color coding of predicative rates from 0 to 1.

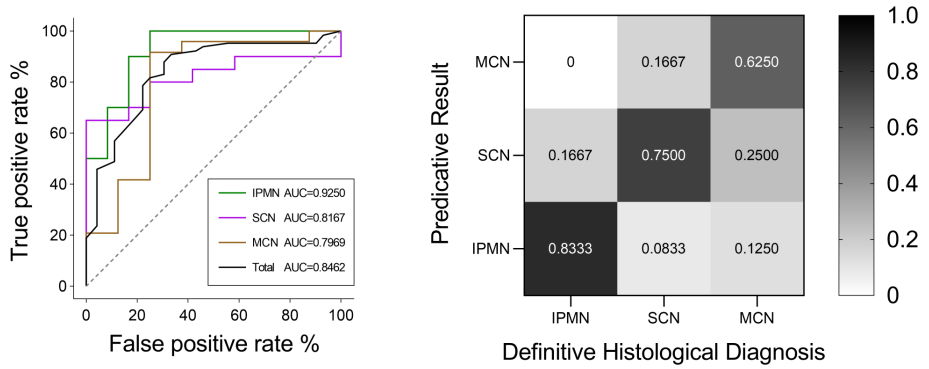



FIGURE 8
ROC curves of IPMN, SCN, and MCN and the total cases in the test group. The final AUC values of IPMN, SCN, and MCN were 0.9250, 0.8167, and 0.7969, respectively. In addition, the mean AUC value of test group was 0.8462. ROC, receiver operating characteristic; AUC, area under the ROC curve. Relative confusion matrix of the three categories. Right color bar represents color coding of predicative rates from 0 to 1.

pancreatic disease identification. As for classification algorithm, several methods have been widely applied in various research studies: Wilcoxon rank-sum test (WRST), relief, logistic regression, $X^2$-test, and LASSO. In our study, we applied a novel DXScore algorithm. Compared with an AUC value of 0.7590 for LASSO in our research, DXScore algorithm achieved an AUC value of 0.8462.

In the field of PCN differential diagnosis, most literatures constructed a binary classification model (26, 28, 49). Only Dmitriev et al. constructed a four-class classification model for the diagnosis of IPMNs, MCNs, SCNs, and solid pseudopapillary neoplasms with convolutional neural network and random forest classifier. This model reached a diagnostic rate of 95.9%, 64.3%, 51.7%, and 100%, respectively (50). In our ternary

TABLE 2 SVM classification performance of selected feature subsets with different methods in the test group. .

| Method of Feature Selection | Number of Selected Features | ACC | AUC | SEN | SPEC |
|---|---|---|---|---|---|
| DXScore | 9 | 0.7361 | 0.8462 | 0.6944 | 0.8522 |
| LASSO | 16 | 0.5938 | 0.7590 | 0.5694 | 0.7944 |

ACC, accuracy; AUC, area under the ROC curve; SEN, sensitivity; SPEC, specificity; LASSO, least absolute shrinkage selection operator.

classification model, the diagnostic rates in the five-fold cross-validation algorithm of IPMNs, MCNs, and SCNs have reached 74.26%, 78.37%, and 68.00%, respectively, and the diagnostic rates of IPMNs, MCNs, and SCNs in the test group have reached 83.33%, 75.00%, and 62.50%, respectively. Our model has shown a feasible performance in the differential diagnosis between SCNs and MCNs, which is the most difficult one according to the clinical experience.

Our article also has some limitations. Because of the ternary classification model, a relatively small number of patients were included in each category. Because most patients with pancreatic cysts have no clinical symptoms and not all the patients require surgical intervention, it is difficult for a single center to obtain a large number of pathologically identified cases (51, 52). In the next step, we will continue to recruit patients, improve the PCNs database capacity, and further test the extensibility of our model with multicenter data.

In conclusion, this research preliminarily verified the application value of radiomics in the differential diagnosis of pancreatic cystic tumors. With more intensive future research and the construction of more reliable prediction models, artificial intelligence technology will greatly help clinicians in the diagnosis and treatment of diseases. In the future, a large number of studies are still needed to conduct prospective studies to further confirm the diagnostic accuracy and application value of imaging omics in clinical practices.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The protocol and all amendments were approved by the Institutional Review Board of Renji Hospital, School of Medicine, Shanghai Jiao Tong University (Shanghai, China).

## Author contributions

ZD: Conceptualization;Methodology;Formal analysis; Writing - Original Draft;Visualization. XC: Methodology; Software;Visualization. ZC: Formal analysis;Data Curation. YL: Formal analysis;Data Curation. MH: Data Curation TC: Data Curation. ZZ: Conceptualization;Validation;Writing - Review and Editing;Supervision. XQ: Validation;Writing - Review and Editing;Supervision. WC: Validation;Writing - Review and Editing;Supervision;Funding acquisition. All authors contributed to the article and approved the submitted version.

## Conflict of interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.941744/full#supplementary-material

## References

1. Moris M, Bridges MD, Pooley RA, Raimondo M, Woodward TA, Stauffer JA, et al. Association between advances in high-resolution cross-section imaging technologies And increase in prevalence of pancreatic cysts from 2005 to 2014. *Clin Gastroenterol Hepatol* (2016) 14(4):585–593.e3. doi: 10.1016/j.cgh.2015.08.038

2. Falqueto A, Pelandré GL, Gouvêa da Costa MZ, Nacif MS, Marchiori E. Prevalence of pancreatic cystic neoplasms on imaging exams: association with signs of malignancy risk. *Radiol Bras* (2018) 51(4):218–24. doi: 10.1590/0100-3984.2017.0105

3. van Huijgevoort NCM, Del Chiaro M, Wolfgang CL, van Hooft JE, Besselink MG. Diagnosis and management of pancreatic cystic neoplasms: current evidence and guidelines. *Nat Rev Gastroenterol Hepatol* (2019) 16(11):676–89. doi: 10.1038/s41575-019-0195-x

4. Valsangkar NP, Morales-Oyarvide V, Thayer SP, Thayer SP, Wargo JA, Warshaw AL, et al. 851 resected cystic tumors of the pancreas: a 33-year experience at the Massachusetts general hospital. - *Surgery.* (2012) 152(3 Suppl 1):S4–12. doi: 10.1016/j.surg.2012.05.033

5. European Study Group on Cystic Tumours of the PancreasEuropean Evidence-based guidelines on pancreatic cystic neoplasms. *Gut* (2018) 67(5):789–804. doi: 10.1136/gutjnl-2018-316027

6. Tanaka M, Fernández-Del Castillo C, Kamisawa T, Jang JY, Levy P, Ohtsuka T, et al. Revisions of international consensus Fukuoka guidelines for the management of IPMN of the pancreas. *Pancreatology.* (2017) 17(5):738–53. doi: 10.1016/j.pan.2017.07.007

7. Lee YN, Moon JH, Kim HK, Choi HJ, Lee SH, Choi MH, et al. A triple approach for diagnostic assessment of endoscopic ultrasound-guided fine needle aspiration in pancreatic solid masses and lymph nodes. *Dig Dis Sci* (2014) 59 (9):2286–93. doi: 10.1007/s10620-014-3119-1

8. Matsubayashi H, Matsui T, Yabuuchi Y, Imai K, Tanaka M, Kakushima N, et al. Endoscopic ultrasonography guided-fine needle aspiration for the diagnosis of solid pancreaticobiliary lesions: Clinical aspects to improve the diagnosis. *World J Gastroenterol* (2016) 22(2):628–40. doi: 10.3748/wjg.v22.i2.628

9. Farrell JJ, Fernández-del Castillo C. Pancreatic cystic neoplasms: management and unanswered questions. *Gastroenterology* (2013) 144(6):1303–15. doi: 10.1053/j.gastro.2013.01.073

10. Sainani NI, Saokar A, Deshpande V, Fernández-del Castillo C, Hahn P, Sahani DV. Comparative performance of MDCT and MRI with MR cholangiopancreatography in characterizing small pancreatic cysts. *AJR Am J Roentgenol.* (2009) 193(3):722–31. doi: 10.2214/AJR.08.1253

11. Song SJ, Lee JM, Kim YJ, Kim SH, Lee JY, Han JK, et al. Differentiation of intraductal papillary mucinous neoplasms from other pancreatic cystic masses: comparison of multirow-detector CT and MR imaging using ROC analysis. *J Magn Reson Imaging.* (2007) 26(1):86–93. doi: 10.1002/jmri.21001

12. Lee HJ, Kim MJ, Choi JY, Hong HS, Kim KA. Relative accuracy of CT and MRI in the differentiation of benign from malignant pancreatic cystic lesions. *Clin Radiol* (2011) 66(4):315–21. doi: 10.1016/j.crad.2010.06.019

13. Visser BC, Muthusamy VR, Yeh BM, Coakley FV, Way LW. Diagnostic evaluation of cystic pancreatic lesions. *HPB (Oxford)* (2008) 10(1):63–9. doi: 10.1080/13651820701883155

14. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. - introduction to radiomics. *J Nucl Med* (2020) 61(4):488–95. doi: 10.2967/jnumed.118.222893

15. Thawani R, McLane M, Beig N, Prasanna P, Velcheti V, Madabhushi A, et al. - radiomics and radiogenomics in lung cancer: A review for the clinician. *Lung Cancer.* (2018) 115:34–41. doi: 10.1016/j.lungcan.2017.10.015

16. Fan Y, Feng M, Wang R. Application of radiomics in central nervous system diseases: a systematic literature review. *Clin Neurol Neurosurg* (2019) 187:105565. doi: 10.1016/j.clineuro.2019.105565

17. Horvat N, Bates DDB, Petkovska I. Novel imaging techniques of rectal cancer: what do radiomics and radiogenomics have to offer? a literature review. *Abdom Radiol (NY)* (2019) 44(11):3764–74. doi: 10.1007/s00261-019-02042-y

18. Wakabayashi T, Ouhmich F, Gonzalez-Cabrera C, Felli E, Saviano A, Agnus V, et al. - radiomics in hepatocellular carcinoma: a quantitative review. - *Hepatol Int* (2019) 13(5):546–59. doi: 10.1007/s12072-019-09973-0

19. Abunahel BM, Pontre B, Kumar H, Petrov MS. Pancreas image mining: a systematic review of radiomics. *Eur Radiol* (2021) 31(5):3447–67. doi: 10.1007/s00330-020-07376-6

20. Permuth JB, Choi J, Balarunathan Y, Kim J, Chen D-T, Chen L, et al. Combining radiomic features with a miRNA classifier may improve prediction of malignant pathology for pancreatic intraductal papillary mucinous neoplasms. *Oncotarget.* (2016) 7(52):85785–97. doi: 10.18632/oncotarget.11768

21. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* (2017) 77(21):e104–e7. doi: 10.1158/0008-5472.CAN-17-0339

22. Tan H, Bao J, Zhou X. A novel missense-mutation-related feature extraction scheme for 'driver'mutation identification. *Bioinformatics.* (2012) 28(22):2948–55. doi: 10.1093/bioinformatics/bts558

23. Wang JTL, Ma Q, Shasha D, Wu CH. New techniques for extracting features from protein sequences. *IBM Syst J* (2001) 40(2):426–41. doi: 10.1147/sj.402.0426

24. Moore AW. *Cross-validation for detecting and preventing overfitting.* (Pittsburgh, Commonwealth of Pennsylvania,USA: School of Computer Science Carneigie Mellon University) (2001).

25. Ghojogh B, Crowley M. The theory behind overfitting, cross validation, regularization, bagging, and boosting: tutorial. *arXiv preprint arXiv* (2019) 23. doi: 10.48550/arXiv.1905.12787

26. Wei R, Lin K, Yan W, Guo Y, Wang Y, Li J, et al. Computer-aided diagnosis of pancreas serous cystic neoplasms: A radiomics method on preoperative MDCT images. *Technol Cancer Res Treat* (2019) 18:1533033818824339. doi: 10.1177/1533033818824339

27. Bian Y, Guo S, Jiang H, Gao S, Shao C, Cao K, et al. Relationship between radiomics and risk of lymph node metastasis in pancreatic ductal adenocarcinoma. *Pancreas.* (2019) 48(9):1195–203. doi: 10.1097/MPA.0000000000001404

28. Tobaly D, Santinha J, Sartoris R, Dioguardi Burgio M, Matos C, Cros J, et al. CT-based radiomics analysis to predict malignancy in patients with intraductal papillary mucinous neoplasm (IPMN) of the pancreas. *Cancers (Basel)* (2020) 12 (11):3089. doi: 10.3390/cancers12113089

29. Dalal V, Carmicheal J, Dhaliwal A, Jain M, Kaur S, Batra SK. Radiomics in stratification of pancreatic cystic lesions: Machine learning in action. *Cancer Lett* (2020) 469:228–37. doi: 10.1016/j.canlet.2019.10.023

30. Galanis C, Zamani A, Cameron JL, Campbell KA, Lillemoe KD, Caparrelli D, et al. Resected serous cystic neoplasms of the pancreas: a review of 158 patients with recommendations for treatment. *J Gastrointest Surg* (2007) 11(7):820–6. doi: 10.1007/s11605-007-0157-4

31. Jais B, Rebours V, Malleo G, Salvia R, Fontana M, Maggino L, et al. Serous cystic neoplasm of the pancreas: a multinational study of 2622 patients under the auspices of the international association of pancreatology and European pancreatic club (European study group on cystic tumors of the pancreas). *Gut.* (2016) 65 (2):305–12. doi: 10.1136/gutjnl-2015-309638

32. Malleo G, Bassi C, Rossini R, Manfredi R, Butturini G, Massignani M, et al. Growth pattern of serous cystic neoplasms of the pancreas: observational study with long-term magnetic resonance surveillance and recommendations for treatment. *Gut.* (2012) 61(5):746–51. doi: 10.1136/gutjnl-2011-300297

33. Farrell JJ. Prevalence, diagnosis and management of pancreatic cystic neoplasms: Current status and future directions. *Gut Liver.* (2015) 9(5):571–89. doi: 10.5009/gnl15063

34. Lazzarin G, Romano L, Coletti G, Di Sibio A, Vicentini V, Fatayer MWA, et al. Branch duct - IPMN and PanIN, in IgG4-autoimmune pancreatitis: A case report. *Clin Case Rep* (2020) 8(11):2111–5. doi: 10.1002/ccr3.2641

35. Brugge WR, Lauwers GY, Sahani D, Fernandez-del Castillo C, Warshaw AL. Cystic neoplasms of the pancreas. *N Engl J Med* (2004) 351(12):1218–26. doi: 10.1056/NEJMra031623

36. Chiaro MD, Segersvärd R, Mucelli RP, Rangelova E, Kartalis N, Ansorge C, et al. Comparison of preoperative conference-based diagnosis with histology of cystic tumors of the pancreas. *Annals Surg Oncol* (2014) 21(5):1539–44. doi: 10.4291/wjgp.v6.i2.29

37. Parekh V, Jacobs MA. Radiomics: a new application from established techniques. *Expert Rev Precis Med Drug Dev* (2016) 1(2):207–26. doi: 10.1080/23808993.2016.1164013

38. Sahani DV, Kambadakone A, Macari M, Takahashi N, Chari S, Fernandez-del Castillo C. Diagnosis and management of cystic pancreatic lesions. - *AJR Am J Roentgenol.* (2013) 200(2):343–54. doi: 10.2214/AJR.12.8862

39. Kim SY, Lee JM, Kim SH, Shin K-S, Kim YJ, An SK, et al. - macrocystic neoplasms of the pancreas: CT differentiation of serous oligocystic adenoma from mucinous cystadenoma and intraductal papillary mucinous tumor. - *AJR Am J Roentgenol.* (2006) 187(5):1192–8. doi: 10.2214/AJR.05.0337

40. Tanaka M, Fernández-del Castillo C, Adsay V, Chari S, Falconi M, Jang JY, et al. International consensus guidelines 2012 for the management of IPMN and MCN of the pancreas. *Pancreatology.* (2012) 12(3):183–97. doi: 10.1016/j.pan.2012.04.004

41. Furuta K, Watanabe H, Ikeda S. Differences between solid and duct-ectatic types of pancreatic ductal carcinomas. *Cancer.* (1992) 69(6):1327–33. doi: 10.1002/1097-0142(19920315)69:6<1327::AID-CNCR2820690605>3.0.CO;2-N

42. Scholten L, van Huijgevoort NCM, van Hooft JE, Besselink MG, Del Chiaro M. Pancreatic cystic neoplasms: Different types, different management, new guidelines. *Visc Med* (2018) 34(3):173–7. doi: 10.1159/000489641

43. Capella C, Solcia E, Sobin L, Arnold R. WHO classification of tumours of the digestive system. In: *Endocrine tumours of the colon and rectum.* Lyon: IARC Press (2000).

44. Lennon AM, Wolfgang C. Cystic neoplasms of the pancreas. *J Gastrointest Surg* (2013) 17(4):645–53. doi: 10.1007/s11605-012-2072-6

45. Buerke B, Domagk D, Heindel W, Wessling J. Diagnostic and radiological management of cystic pancreatic lesions: Important features for radiologists. *Clin Radiol* (2012) 67(8):727–37. doi: 10.1016/j.crad.2012.02.008

46. Zhou Y, Xie L, Shen W, Wang Y, Fishman EK, Yuille AL. editors. A fixed-point model for pancreas segmentation in abdominal CT scans. In: *International conference on medical image computing and computer-assisted intervention.* (Springer Cham, Switzerland: Springer) (2017). p. 693–701.

47. Roth H, Oda M, Shimizu N, Oda H, Hayashi Y, Kitasaka T, et al. Towards dense volumetric pancreas segmentation in CT using 3D fully convolutional networks. In: *Medical imaging 2018: Image processing.* (Houston, TX: International Society for Optics and Photonics) (2018). p. 105740B.

48. Zhu Z, Xia Y, Shen W, Fishman EK, Yuille ALIEEE, et al. editors. A 3D coarse-to-fine framework for volumetric medical image segmentation. In: *2018 international conference on 3D vision (3DV).* (Verona, ITALY: IEEE) (2018). p. 682–90.

49. Hanania AN, Bantis LE, Feng Z, Wang H, Tamm EP, Katz MH, et al. Quantitative imaging to evaluate malignant potential of IPMNs. *Oncotarget.* (2016) 7(52):85776–84. doi: 10.18632/oncotarget.11769

50. Dmitriev K, Kaufman AE, Javed AA, Hruban RH, Fishman EK, Lennon AM, et al. - classification of pancreatic cysts in computed tomography images using a random forest and convolutional neural network ensemble. *Med Image Comput Comput Assist Interv.* (2017) 10435:150–8. doi: 10.1007/978-3-319-66179-7_18

51. Zárate X, Williams N, Herrera MF. Pancreatic incidentalomas. *Best Pract Res Clin Endocrinol Metab* (2012) 26(1):97–103. doi: 10.1016/j.beem.2011.06.005

52. Edirimanne S, Connor SJ. Incidental pancreatic cystic lesions. *World J Surg* (2008) 32(9):2028–37. doi: 10.1007/s00268-008-9633-6

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 11 January 2023
DOI 10.3389/fonc.2022.1022305



OPEN ACCESS

EDITED BY
Carl D. Malchoff,
University of Connecticut Health Center, United States

REVIEWED BY
Fajin Dong,
Jinan University, China
Ilker Sengul,
Giresun University, Turkey

*CORRESPONDENCE
Xin-Wu Cui
cuixinwu@live.cn
Ai-Jiao Yi
938186822@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 18 August 2022
ACCEPTED 19 December 2022
PUBLISHED 11 January 2023

CITATION
Wang B, Ou X, Yang J, Zhang H, Cui X-W, Dietrich CF and Yi A-J (2023) Contrast-enhanced ultrasound and shear wave elastography in the diagnosis of ACR TI-RADS 4 and 5 category thyroid nodules coexisting with Hashimoto's thyroiditis. *Front. Oncol.* 12:1022305. doi: 10.3389/fonc.2022.1022305



# Contrast-enhanced ultrasound and shear wave elastography in the diagnosis of ACR TI-RADS 4 and 5 category thyroid nodules coexisting with Hashimoto's thyroiditis

Bin Wang[1†], Xiaoyan Ou[1†], Juan Yang[1], Haibo Zhang[1], Xin-Wu Cui[2*], Christoph F. Dietrich[3] and Ai-Jiao Yi[1*]

[1]Department of Medical Ultrasound, Yueyang Central Hospital, Yueyang, China, [2]Department of Medical Ultrasound, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [3]Department Allgemeine Innere Medizin, Kliniken Hirslanden Beau Site, Salem und Permanence, Bern, Switzerland

**Objective:** This study aims to evaluate the value of contrast-enhanced ultrasound (CEUS), shear wave elastography (SWE), and their combined use in the differentiation of American College of Radiology (ACR) thyroid imaging reporting and data system (TI-RADS) 4 and 5 category thyroid nodules coexisting with Hashimoto's thyroiditis (HT).

**Materials and methods:** A total of 133 pathologically confirmed ACR TI-RADS 4 and 5 category nodules coexisting with HT in 113 patients were included; CEUS and SWE were performed for all nodules. The sensitivity, specificity, negative predictive value (NPV), positive predictive value (PPV), accuracy, and the area under the receiver operating characteristic curve (AUC) of the SWE, CEUS, and the combined use of both for the differentiation of benign and malignant nodules were compared, respectively.

**Results:** Using CEUS alone, the sensitivity, specificity, PPV, NPV, and accuracy were 89.2%, 66.0%, 81.3%, 78.6%, and 80.5%, respectively. Using SWE alone, Emax was superior to Emin, Emean, and Eratio for the differentiation of benign and malignant nodules with the best cutoff Emax >46.8 kPa, which had sensitivity of 65.1%, specificity of 90.0%, PPV of 91.5%, NPV of 60.8%, and accuracy of 74.4%, respectively. Compared with the diagnostic performance of qualitative CEUS or/and quantitative SWE, the combination of CEUS and SWE had the best sensitivity, accuracy, and AUC; the sensitivity, specificity, PPV, NPV, accuracy, and AUC were 94.0%, 66.0%, 82.1%, 86.8%, 83.5%, and 0.80 (95% confidence interval: 0.713, 0.886), respectively.

**Conclusion:** In conclusion, CEUS and SWE were useful for the differentiation of benign and malignant ACR TI-RADS 4 and 5 category thyroid nodules

coexisting with HT. The combination of CEUS and SWE could improve the sensitivity and accuracy compared with using CEUS or SWE alone. It could be a non-invasive, reliable, and useful method to differentiate benign from malignant ACR TI-RADS 4 and 5 category thyroid nodules coexisting with HT.

KEYWORDS

shear wave elastography, Hashimoto's thyroiditis, thyroid nodule, American College of Radiology, thyroid imaging reporting and data system, contrast-enhanced ultrasound

## 1 Introduction

Thyroid nodules are a common disease and have shown a prevalence of 5% to 7% in the adult population with physical examination alone, while ultrasound examination shows a prevalence of 20% to 76% in this same population (1–3). Thyroid nodules are the most commonly found tumors in the cervical region, with nearly 10% being malignant nodules (4). It has been reported that one-third of thyroid cancer cases are prone to lymph node metastasis; thus, precise diagnosis and early treatment are significantly important for the recovery and better outcome of patients (5).

Hashimoto's thyroiditis (HT) is the most common autoimmune thyroid disease, and the link between HT and thyroid cancer has been debated and remains controversial (6). The diagnosis of thyroid nodules with HT may be confusing in clinical work. Conventional ultrasound can differentiate benign from malignant nodules based on some characteristics, including marked hypo-echogenicity, a shape which is taller than wide, irregular margin, and micro-calcification (7). However, these characteristics have overlaps between benign and malignant nodules, especially those coexisting with HT. Some studies found that the irregular or micro-lobulated margins of benign nodules were more frequent in HT patients with a heterogeneous echogenicity background, and some conventional ultrasound characteristics were difficult to identify in the heterogeneous thyroid gland coexisting with HT, such as margin and calcification (8, 9). Thus, it is important to find a method to differentiate benign from malignant nodules coexisting with HT.

Currently, fine-needle aspiration (FNA) is the most effective and practical method used to reach a definitive diagnosis (10). However, most nodules are benign, and some malignant nodules frequently present an indolent behavior; therefore, not all thyroid nodules need FNA. To identify the most clinically significant malignant nodules and reduce the number of biopsies, the American College of Radiology (ACR) presents the thyroid imaging reporting and data system (TI-RADS) (7). However, the cancer risk levels were 5%–20% for ACR TI-RADS 4 category and at least 20% for ACR TI-RADS 5 category (11). There was a great overlap between benign and malignant nodules in ACR TI-RADS 4 and 5 categories, especially for nodules coexisting with TH. Thus, it is important to find a reliable and noninvasive method to differentiate benign from malignant ACR TI-RADS 4 and 5 category nodules, especially for nodules without recommendation of FNA.

In recent years, shear wave elastography (SWE) and contrast-enhanced ultrasound (CEUS) have been widely used in the diagnosis of thyroid nodules. SWE can quantitatively measure real-time tissue elasticity, along with a color-coded elasticity map (12). Many studies have found that the elasticity index of thyroid cancer is higher than that of benign thyroid nodules (13–17). CEUS can qualitatively or quantitatively evaluate the macro- and micro-vascularization patterns of thyroid nodules compared with the surrounding tissue, which is a promising noninvasive method for differentiating benign from malignant nodules. However, there were overlapping parameters and patterns with CEUS qualitative and quantitative evaluation in the differentiation of benign and malignant nodules, which indicate a limitation in the interpretation of tumor vascularization (18).

ACR TI-RADS 4 and 5 category thyroid nodules coexisting with HT were difficult to differentiate, and using CEUS or SWE alone had its limitation. The study aimed to explore the diagnostic performance of SWE, CEUS, and the combination of CEUS and SWE in ACR TI-RADS 4 and 5 category thyroid nodules coexisting with HT and find the reliable and noninvasive method to differentiate benign from malignant nodules, which would be beneficial to manage patients and improve their prognosis.

**Abbreviations:** HT, Hashimoto's thyroiditis; FNA, fine-needle aspiration; ACR, American College of Radiology; TI-RADS, thyroid imaging reporting and data system; SWE, shear wave elastography; CEUS, contrast-enhanced ultrasound; EI, elasticity index; RFA, radiofrequency ablation; MI, mechanical index; SE, stain elastography; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the curve; ROC, receiver operating characteristic curve.

# 2 Materials and methods

This prospective study was approved by the ethics committee of Yueyang Central Hospital.

## 2.1 Patients

From April 2020 to July 2021, a total of 113 patients with 133 ACR TI-RADS 4 and 5 category thyroid nodules were recruited. The inclusion criteria were listed as follows: (a) patients aged 18 years or older with at least a thyroid nodule detected on the conventional ultrasound, (b) all patients gave signed informed consent before the SWE and CEUS examinations, (c) the pathology of all the thyroid nodules was confirmed *via* surgery according to standard clinical protocols, and (d) all patients were pathologically diagnosed as HT. Patients were excluded if they previously had a FNA, previously had radiofrequency ablation, had a contraindication of CEUS, or had unsatisfactory images.

## 2.2 Ultrasound examination

All ultrasound examinations including conventional ultrasound, SWE, and CEUS were performed with a high frequency transducer (L15-4 or L10-5 Aixplorer, Supersonic Imaging, France).

Conventional ultrasound was performed. When a target thyroid nodule was detected, the general characteristics were observed, including composition, echogenicity, shape, margin, and echogenic foci (7). Each lesion was classified into ACR TI-RADS 4 category (moderately suspicious) or ACR TI-RADS 5 category (highly suspicious).

The SWE imaging examinations were induced by the L10-5 transducer. In order to overcome the effect of artery pulsation on SWE measurement, a longitudinal section is often selected to conduct the SWE imaging. The probe was lightly applied while the patients were asked to hold their breath. The region of interest included the whole thyroid nodules, and in order to obtain satisfactory SWE imaging, several tips were suggested: (1) the upper edge of the region of interest is more than 1 cm away from the skin, (2) the depth of the lower edge of the region of interest should not exceed 4 cm, and (3) the length of the region of interest is two to three times larger than the nodules. The stiffness range of color map was from blue to red (0–180 kPa); the standard SWE imaging was obtained with several seconds of immobilization. The SWE measurement used quantification box (Q-box), and the Q-box should contain the nodules, excluding the surrounding organizations, which can automatically obtain Young's modulus of Emin, Emean, and Emax. Eratio was defined as the ratio of Young's modulus, which was obtained from the ratio of two Q-box in the same depth. The first Q-box was placed in the hardest region of the nodules, while the second Q-box was placed in the surrounding normal thyroid tissues. The diameter of the Q-box was 2 mm. The median was taken for five measurements to obtain more accurate results.

After the SWE imaging examination, CEUS was performed by using the L10-5 transducer. The patients were asked to lie down in the supine position, and the double-contrast mode was used to display nodules clearly. To avoid micro-bubble disruption, the focus was placed slightly deeper than the nodules with a low mechanical index that ranged from 0.06 to 0.08. A total of 25 mg sulfur hexafluoride (SonoVue®, Bracco International, Milan, Italy) was dissolved in 5 ml 0.9% sodium chloride and was injected with an intravenous bolus of 2.4 ml per patient, followed by a 5-ml saline flush. Each contrast imaging acquisition that lasted at least 120 s was stored in the machine hard disk.

All sonographic examinations were performed by the same investigator who had more than 10 years of experience in thyroid ultrasound and 5 years of experience in SWE and CEUS.

## 2.3 Statistical analysis

SPSS 23.0 and MedCalc 19.0 were used for all statistical analyses. The parameters of SWE and the enhancement patterns of CEUS were compared with the *t*-test, Kappa analysis, or Fisher's exact test. A receiver operating characteristic (ROC) curve differentiating benign from malignant thyroid nodules was drawn according to Young's modulus for each nodule. The optimal cutoff value and area under the curve (AUC) were calculated. The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and AUC of SWE, CEUS, or the combination of SWE and CEUS were calculated and compared, respectively. $P$ <0.05 was considered to indicate statistical significance.

# 3 Results

## 3.1 General information between the benign and malignant nodules

Of the 133 nodules, the surgical pathological results showed that 50 nodules were benign and 83 nodules were malignant; all nodules were coexisting with HT (Figure 1). There was no significant difference in sex, age, the maximum diameter of nodules, and the number of nodules between benign and malignant nodules ($P$ >0.05) (Table 1).

## 3.2 CEUS in the differentiation of the benign and malignant nodules

The enhancement patterns of all thyroid nodules are summarized in Table 2. There were significant differences in peak

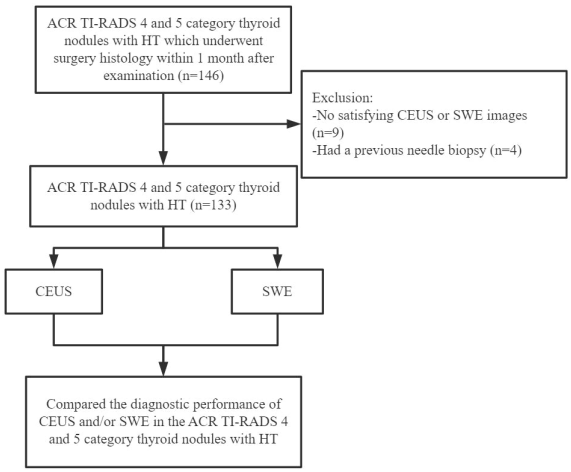



FIGURE 1
Flow chart for the selection of thyroid nodules.

enhancement, enhancement evenness, and ring enhancement with qualitative CEUS between benign and malignant thyroid nodules ($P < 0.05$). The benign thyroid nodules mostly manifested no enhancement, hyper-enhancement, iso-enhancement, homogeneous enhancement, or ring enhancement, while the malignant thyroid nodules mostly manifested hypo-enhancement, heterogeneous enhancement, or without ring enhancement. Using CEUS alone, the sensitivity, specificity, PPV, NPV, and accuracy were 89.2%, 66.0%, 81.3%, 78.6%, and 80.5%, respectively.

## 3.3 SWE in the differentiation of the benign and malignant nodules

A ROC curve was drawn based on Emin, Emean, Emax, and Eratio to determine the optimal cutoff point for discriminating benign from malignant nodules. The optimal cutoff point was 27.8 kPa for Emin, 34.1 kPa for Emean, 46.8 kPa for Emax, and 1.25 for Eratio, respectively. Compared with the diagnostic performance of Emin, Emean, Emax, and Eratio (Table 3), Emax was superior to Emin, Emean, and Eratio, which had sensitivity of 65.1%, specificity of 90.0%, PPV of 91.5%, NPV of 60.8%, and accuracy of 74.4%.

## 3.4 Comparing the diagnostic performance of CEUS, SWE, and the combination of CEUS and SWE in differentiating benign and malignant nodules

The standard of the combination of qualitative CEUS and quantitative SWE was that the nodules were recognized as

TABLE 1 General information between benign and malignant nodules.

| Group | Case | Sex | | Age | Number of nodules | | Size |
|---|---|---|---|---|---|---|---|
| | | Male | Female | | Solitary nodule | Multiple nodules | |
| Benign group | 50 | 6 | 44 | 46.16 ± 11.71 | 19 | 31 | 9.00 ± 5.56 |
| Malignant group | 83 | 13 | 70 | 46.21 ± 10.57 | 34 | 49 | 7.82 ± 3.27 |
| $\chi^2$/Fisher/$t$ | | 0.342 | | 0.485 | 0.114 | | 1.540 |
| $P$-value | | 0.559 | | 0.629 | 0.735 | | 0.125 |

TABLE 2 Enhancement patterns of all thyroid nodules.

| Group | Case | Peak intensity | | | Enhancement evenness | | Ring enhancement | |
|---|---|---|---|---|---|---|---|---|
| | | No enhancement or hyper-enhancement | Iso-enhancement | Hypo-enhancement | Homogeneous | Heterogeneous | With | Without |
| Benign | 50 | 17 | 16 | 17 | 32 | 18 | 10 | 40 |
| Malignant | 83 | 1 | 10 | 72 | 7 | 76 | 1 | 82 |
| $\chi^2$/Fisher | | 44.124 | | | 46.487 | | – | |
| $P$-value | | 0.000 | | | 0.000 | | 0.000 | |

TABLE 3 Diagnostic performance of Emin, Emean, Emax, and Eratio in differentiating benign and malignant nodules.

| Elastography | Pathology | Benign | Malignant | Sensitivity | Specificity | Positive predictive value | Negative predictive value | Accuracy |
|---|---|---|---|---|---|---|---|---|
| Emin | Benign | 47 | 3 | 45.80% | 94.0% | 92.7% | 51.1% | 63.9% |
| | Malignant | 45 | 38 | | | | | |
| Emean | Benign | 41 | 9 | 62.7% | 82.0% | 85.2% | 56.9% | 69.6% |
| | Malignant | 31 | 52 | | | | | |
| Emax | Benign | 45 | 5 | 65.1% | 90.0% | 91.5% | 60.8% | 74.4% |
| | Malignant | 29 | 54 | | | | | |
| Eratio | Benign | 30 | 20 | 79.5% | 60.0% | 76.7% | 63.8% | 72.2% |
| | Malignant | 17 | 66 | | | | | |

benign when qualitative CEUS manifested benign enhancement patterns and Emax <46.8 kPa simultaneously (Figure 2). Compared with using CEUS or SWE alone, combination of CEUS and SWE had the best sensitivity, accuracy, and AUC (Figure 3). The sensitivity, specificity, PPV, NPV, accuracy, and AUC were 94.0%, 66.0%, 82.1%, 86.8%, 83.5%, and 0.80 (95% confidence interval: 0.713, 0.886), respectively (Table 4).

## 3.5 Management and recommendation of ACR TI-RADS 4 and 5 category thyroid nodules coexisting with HT

According to the ACR TI-RADS, TI-RADS 4 category nodules with the maximal diameter ≧1.5 cm or TI-RADS 5 category nodules with the maximal diameter ≧1 cm were

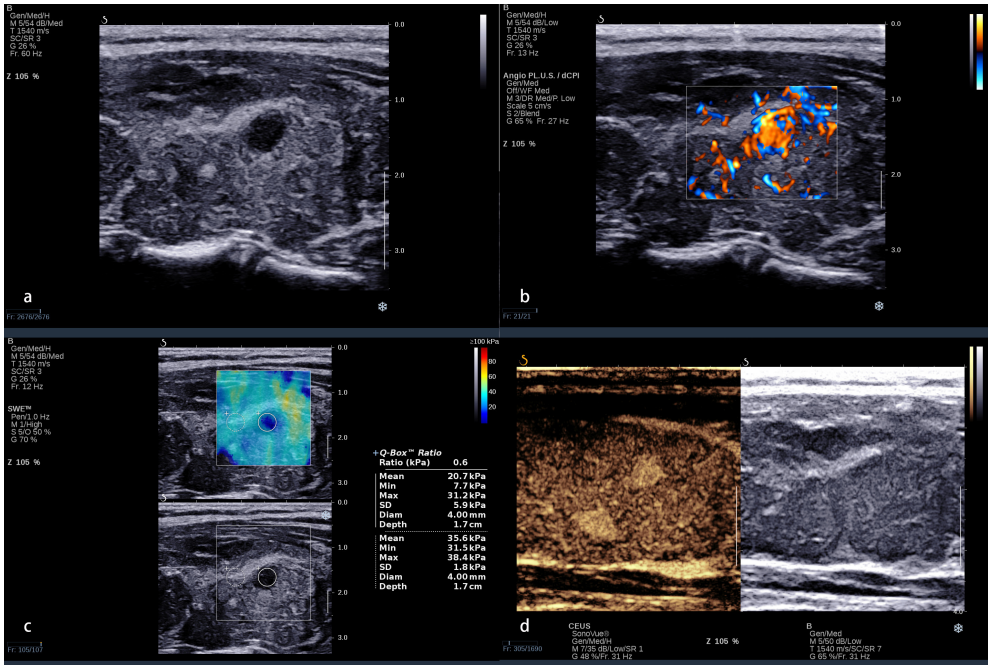



FIGURE 2
A 32-year-old woman with a nodule in the left lobe of the thyroid. **(A)** Conventional ultrasound revealed a 5 × 6-mm solid hypo-echoic nodule with a smooth margin, shape that was taller-than-wide, and without obvious calcification; this was categorized ACR TI-RADS 5. **(B)** The nodule presented a rich blood flow signal. **(C)** Quantitative shear wave elastography (SWE) revealed 7.7 kPa for Emin, 20.7 KPa for Emean, 31.2 kPa for Emax, and 0.6 for Eratio. **(D)** Qualitative contrast-enhanced ultrasound (CEUS) of the nodule revealed homogenous, hyper-enhancement. The combination of CEUS and SWE revealed that this nodule may be benign. The surgical pathology was focal Hashimoto's thyroiditis.

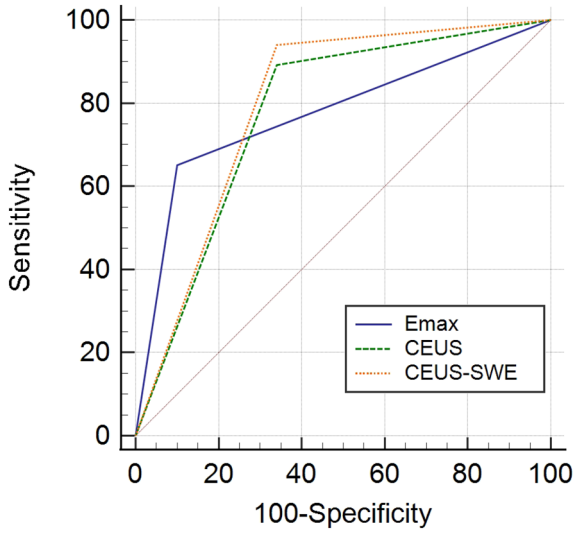



FIGURE 3
ROC of contrast-enhanced ultrasound (CEUS), shear wave elastography (SWE), and the combination of CEUS and SWE in differentiating benign and malignant nodules. ROC, receiver operating characteristic; AUC, area under the ROC curve.

recommended for FNA. In this study, the malignant rate of TI-RADS 4 and 5 category nodules coexisting with HT and recommended for FNA was 50% (13/26), while the malignant rate of TI-RADS 4 and 5 category nodules coexisting with HT recommended for follow-up was 64.5% (70/107).

## 4 Discussion

Our study found that CEUS could evaluate tumor vascularization with enhancement patterns, and SWE could provide additional stiffness information, which was useful for the differentiation of ACR TI-RADS 4 and 5 category thyroid nodules coexisting with HT. The combination of CEUS and SWE could improve the sensitivity and accuracy compared with using CEUS or SWE alone.

ACR TI-RADS (7) presents a system for risk stratification of thyroid nodules, which was widely used to identify the most clinically significant malignant nodules and recommend them for FNA. According to the ACR TI-RADS, TI-RADS 4 category nodules with maximal diameter ≧1.5 cm or TI-RADS 5 category nodules with maximal diameter ≧1 cm were recommended for FNA; however, small thyroid cancers with maximal diameter ≤10 mm might mainly cause the "epidemic" of thyroid carcinoma (19). In this study, many TI-RADS 4 or 5 category nodules coexisting with HT and recommended for follow-up might be malignant. The malignant rate accounted for 65.4%. These small thyroid carcinomas made the patients endure a great psychological burden. Moreover, the ACR TI-RADS 4 and 5 category nodules coexisting with HT in patients were difficult to differentiate because of the heterogeneous and coarse thyroid parenchyma caused by the repetitive damage of chronic inflammation (20). Thus, it is a major challenge in the management of thyroid nodules with indeterminate cytology or suspicious conventional ultrasound features.

There were many studies on the qualitative CEUS features in differentiating benign thyroid nodules from malignant ones. According to the EFSUMB guidelines and recommendations for the clinical practice of contrast-enhanced ultrasound in non-hepatic applications (18), hypo-enhancement and heterogeneous enhancement are the predictors of malignancy on CEUS (21–26). A hypo-enhancement pattern was the most precise predictor of malignancy, which had sensitivity, specificity, and accuracy of 82%, 85%, and 84%, respectively (18, 23), while a heterogeneous enhancement pattern had sensitivity, specificity, and accuracy of 88.2%, 92.5%, and 90.4%, respectively (24–26). Some studies (27) found that a ring enhancement pattern was a very strong indicator of benign nodules, which had sensitivity, specificity, and accuracy of 83.0%, 94.1%, and 88.5%, respectively (24, 25). In this study, we observed CEUS enhancement patterns in ACR TI-RADS 4 and 5 category thyroid nodules coexisting with HT; it was consistent with a previous study. HT did not seem to affect the CEUS enhancement patterns in thyroid nodules.

When CEUS was used alone, 80.5% ACR TI-RADS 4 and 5 category nodules coexisting with HT could be accurately diagnosed. However, there were 34.0% (17/50) benign nodules that presented heterogeneous and hypo-enhancement, including 11 nodules that were nodular goiter with fibrosis or calcification,

TABLE 4 Diagnostic performance of contrast-enhanced ultrasound (CEUS), shear wave elastography (SWE), and the combination of CEUS and SWE in differentiating benign and malignant nodules.

| | Sensitivity (%) | Specificity (%) | Positive predictive value (%) | Negative predictive value (%) | Accuracy (%) | Area under the curve | P* |
|---|---|---|---|---|---|---|---|
| CEUS | 89.2 | 66.0 | 81.3 | 78.6 | 80.5 | 0.776 (0.687–0.864) | – |
| Emax | 65.1 | 90.0 | 91.5 | 60.8 | 74.4 | 0.775 (0.695–0.856) | 0.99 |
| CEUS + Emax | 94 | 66.0 | 82.1 | 86.8 | 83.5 | 0.800 (0.713–0.886) | 0.04 |

Data are expressed as percentage (numbers). *Comparison of diagnostic performance of using CEUS alone with using SWE alone and the combination of CEUS and SWE.

five nodules that were focal HT, and one nodule that was subacute thyroiditis. The malignant CEUS enhancement patterns of focal HT may be related to focal hypothyroidism with severe follicular degeneration (28), while that of subacute thyroiditis may be caused by the heterogeneous distribution of inflammatory cells with focal fiber hyperplasia. There were 10.8% (9/83) malignant nodules that present hyper-enhancement, iso-enhancement, or homogeneous enhancement, while the maximal diameter of 88.9% (eight out of nine) of these nodules was <10 mm, which may be because it was difficult to observe the enhancement patterns in small nodules or the neo-vascularization of some small thyroid nodules was not obvious. Therefore, using CEUS alone was insufficient to differentiate benign from malignant ACR TI-RADS 4 and 5 category nodules coexisting with HT.

In recent years, elastography has become available for thyroid nodule evaluation as reported in many studies, which is emerging as a potential method for the differentiation of benign and malignant thyroid nodules and may provide additional information to support clinical decision-making (29). Some studies (30–33) reported that strain elastography (SE) was useful for the prediction of malignancy and differentiation of thyroid nodules with indeterminate FNA cytology. Sengul et al. (31) found that SE score could affect the clinical decision-making for patients with indeterminate FNA cytology. Zhu et al. (32) reported that a high SE score was a significant predictor for malignancy. However, SE was challenged and criticized due to its operator dependency (34).

Compared with SE, SWE is less influenced by the experience and operation of the operator. Zhang et al. (35) found that SWE had high diagnostic efficiency for ACR TI-RADS 4 and 5 category thyroid nodules; the accuracy was 76.1%, with the best cutoff of Emax being 40.9 kPa. Chen et al. (36) conducted a meta-analysis with 4,296 thyroid nodules and found that Supersonic shear imaging showed high accuracy in the differentiation between benign and malignant thyroid nodules, which could serve as a noninvasive and important tool for thyroid nodule evaluation. Liao et al. (37) found that SWE could be an independent predictor for malignant thyroid nodules; the sensitivity and specificity were 81% and 65%, respectively, with the best cutoff of Emean being 32 kPa. In this study, we found that SWE was useful in the differentiation of benign and malignant ACR TI-RADS 4 and 5 category thyroid nodules, which was consistent with a previous study (35–37), and HT did not seem to affect the stiffness of thyroid nodules. Compared with Emin, Emean, and Eratio, Emax >46.8 kPa had the best diagnostic performance for ACR TI-RADS 4 and 5 category thyroid nodules coexisting with HT in this study.

When using Emax alone, there were 10% (five out of 50) benign nodules with Emax >46.8 kPa, which may be because some nodular goiter nodules with fibrosis and calcification could significantly improve the Emax. There were 34.9% (29/83) malignant nodules with Emax <46.8 kPa, and the maximal diameter of 100% (29/29) of these nodules was <10 mm (Figure 4). In general, the characteristics of thyroid nodules may be determined already during the initial formation, such as the growth of malignant nodules, the normal follicle damage, the interstitial components being reduced, fibrosis in nodules which increased, and partial nodules presenting calcium and salt deposition. All the above-mentioned features could improve the hardness of thyroid nodules. Therefore, the tissue components gradually change, and the hardness increases, accompanied by the nodules' increase in size. Thus, the elastography in the differentiation of benign and malignant thyroid nodules might be affected by nodule size. Sengul et al. (38) found that nodule size over 15 mm might strengthen the prediction among high SE scores. Li et al. (39) found that thyroid nodule size affects the optimal Emax cutoff value of SWE. Shang et al. (40) found that Emax was significantly correlated with the size of the nodules. Wang et al. (41) found that conventional ultrasound combined with SWE had higher specificity for nodules smaller than 10 mm and higher sensitivity for nodules larger than 10 mm. Therefore, it was insufficient to identify the characteristics of ACR TI-RADS 4 and 5 category thyroid nodules coexisting with HT, especially for those with maximal diameter <10 mm, which may be related to the small malignant nodules with inconspicuous fibrosis or calcification.

When CEUS was combined with SWE, the sensitivity and accuracy increased compared with using CEUS or SWE alone. Only 6.0% (five out of 83) malignant thyroid nodules were misdiagnosed due to benign enhancement patterns on CEUS and SWE <46.8 kPa. The maximal diameter of all these nodules was <10 mm, including 80% (four out of five) of them which were 5 to 6 mm, which may be caused by inconspicuous neo-vascularization and atypical fibrosis or calcification. Thus, the combination of CEUS and SWE should be carefully used for these small thyroid nodules.

There were several limitations in this study. First, it was a preliminary study in one center with a small sample. Second, qualitative CEUS was used in this study, and no comparison was made with quantitative CEUS. Third, some final pathology results of patients were not available, which may have caused a selection bias of enrolment.

## 5 Conclusion

In conclusion, CEUS and SWE were useful for the differentiation of benign and malignant ACR TI-RADS 4 and 5 category thyroid nodules coexisting with HT. The combination of CEUS and SWE could improve the sensitivity and accuracy compared with using CEUS or SWE alone, which could be a

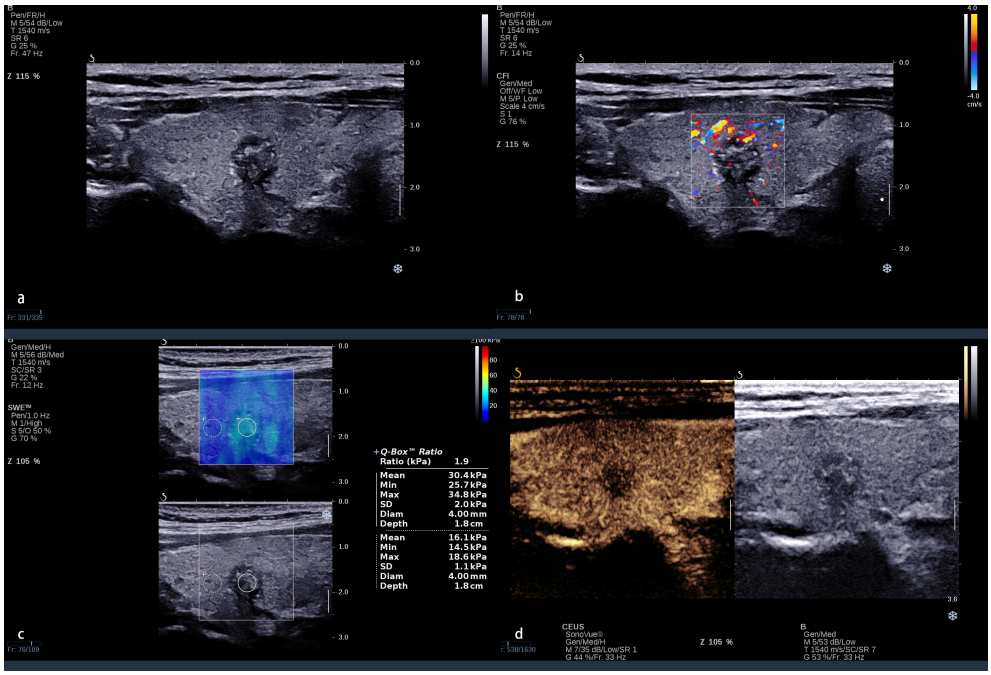



FIGURE 4
A 38 year-old-woman with a nodule in the right lobe of the thyroid. **(A)** Conventional ultrasound revealed an 8 × 9-mm solid hypo-echoic nodule with an irregular margin, shape that was taller-than-wide, and with punctate echogenic foci; this was categorized ACR TI-RADS 5. **(B)** The nodule presented a spot blood flow signal. **(C)** Quantitative shear wave elastography (SWE) revealed 25.7 kPa for Emin, 30.4 kPa for Emean, 34.8 kPa for Emax, and 1.9 for Eratio. **(D)** Qualitative contrast-enhanced ultrasound (CEUS) of the nodule revealed heterogenous, hypo-enhancement. Even the Emax was <46.8 kPa; the CEUS presented malignant enhancement patterns. The combination of CEUS and SWE recognized this nodule as malignant. The surgical pathology was papillary thyroid carcinoma with Hashimoto's thyroiditis.

non-invasive, reliable, and useful method to differentiate benign from malignant ACR TI-RADS 4 and 5 category thyroid nodules coexisting with HT, and it might be beneficial to manage patients and improve their prognosis.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors without undue reservation.

## Ethics statement

The studies involving human participants were reviewed and approved by the ethics committee of Yueyang Central Hospital. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

Conception and design: BW, A-JY, X-WC, and CD. Drafting of the article: BW, XO, and JY. Critical revision of the article for important intellectual content: BW, XO, and HZ. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the National Natural Science Foundation of China (nos. 82071953 and 21878237) and the Science and Technology Department of Hunan Province (no. 2020SK52705).

## 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.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Bomeli SR, LeBeau SO, Ferris RL. Evaluation of a thyroi d nodule. *Otolaryngol Clin North Am* (2010) 43(2):229–38. doi: 10.1016/j.otc.2010.01.002

2. Durante C, Costante G, Lucisano G, Bruno R, Meringolo D, Paciaroni A, et al. The natural history of benign thyroid nodules. *JAMA* (2015) 313(9):926–35. doi: 10.1001/jama.2015.0956

3. Popoveniuc G, Jonklaas J. Thyroid nodules. *Med Clin North Am* (2012) 96 (2):329–49. doi: 10.1016/j.mcna.2012.02.002

4. Durante C, Grani G, Lamartina L, Filetti S, Mandel SJ, Cooper DS. The diagnosis and management of thyroid nodules: A review. *JAMA* (2018) 319 (9):914–24. doi: 10.1001/jama.2018.0898

5. Sherman SI, Angelos P, Ball DW, Beenken SW, Byrd D. Clark OH et al; national comprehensive cancer network. *Thyroid carcinoma J Natl Compr Canc Netw* (2005) 3(3):404–57. doi: 10.6004/jnccn.2005.0021

6. Noureldine SI, Tufano RP. Association of hashimoto's thyroiditis and thyroid cancer. *Curr Opin Oncol* (2015) 27(1):21–5. doi: 10.1097/CCO.0000000000000150

7. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, et al. ACR thyroid imaging, reporting and data system (TI-RADS): White paper of the ACR TI-RADS committee. *J Am Coll Radiol* (2017) 14(5):587–95. doi: 10.1016/j.jacr.2017.01.046

8. Wang D, Du LY, Sun JW, Hou XJ, Wang H, Wu JQ, et al. Evaluation of thyroid nodules with coexistent hashimoto's thyroiditis according to various ultrasound-based risk stratification systems:A retrospective research. *Eur J Radiol* (2020) 131:109059. doi: 10.1016/j.ejrad.2020.109059

9. Park M, Park SH, Kim EK, Yoon JH, Moon HJ, Lee HS, et al. Heterogeneous echogenicity of the underlying thyroid parenchyma: How does this affect the analysis of a thyroid nodule? *BMC Cancer* (2013) 13:550. doi: 10.1186/1471-2407-13-550

10. Singh Ospina N, Brito JP, Maraka S, Espinosa de Ycaza AE, Rodriguez-Gutierrez R, Gionfriddo MR, et al. Diagnostic accuracy of ultrasound-guided fine needle aspiration biopsy for thyroid malignancy: Systematic review and meta-analysis. *Endocrine* (2016) 53(3):651–61. doi: 10.1007/s12020-016-0921-x

11. Nguyen XV, Choudhury KR, Eastwood JD, Lyman GH, Esclamado RM, Werner JD, et al. Incidental thyroid nodules on CT: Evaluation of 2 risk-categorization methods for work-up of nodules. *AJNR Am J Neuroradiol* (2013) 34(9):1812–7. doi: 10.3174/ajnr.A3487

12. Bamber J, Cosgrove D, Dietrich CF, Fromageau J, Bojunga J, Calliada F, et al. EFSUMB guidelines and recommendations on the clinical use of ultrasound elastography. part 1: Basic principles and technology. *Ultraschall Med* (2013) 34 (2):169–84. doi: 10.1055/s-0033-1335205

13. Xu HX, Yan K, Liu BJ, Liu WY, Tang LN, Zhou Q, et al. Guidelines and recommendations on the clinical use of shear wave elastography for evaluating thyroid nodule1. *Clin Hemorheol Microcirc* (2019) 72(1):39–60. doi: 10.3233/CH-180452

14. Sebag F, Vaillant-Lombard J, Berbis J, Griset V, Henry JF, Petit P, et al. Shear wave elastography: a new ultrasound imaging mode for the diferential diagnosis of benign and malignant thyroid nodules. *J Clin Endocrinol Metab* (2010) 95 (12):5281–8. doi: 10.1210/jc.2010-0766

15. Liu B, Liang J, Zheng Y, Xie X, Huang G, Zhou L, et al. Two-dimensional shear wave elastography as promising diagnostic tool for predicting malignant thyroid nodules: A prospective single-centre experience. *Eur Radiol* (2015) 25 (3):624–34. doi: 10.1210/jc.2010-0766

16. Park AY, Son EJ, Han K, Youk JH, Kim JA, Park CS. Shear wave elastography of thyroid nodules for the prediction of malignancy in a large scale study. *Eur J Radiol* (2015) 84(3):407–12. doi: 10.1016/j.ejrad.2014.11.019

17. Kim HJ, Kwak MK, Choi IH, Jin SY, Park HK, Byun DW, et al. Utility of shear wave elastography to detect papillary thyroid carcinoma in thyroid nodules: Efcacy of the standard deviation elasticity. *Korean J Intern Med* (2018) 34(4):850–7. doi: 10.3904/kjim.2016.326

18. Sidhu PS, Cantisani V, Dietrich CF, Gilja OH, Saftoiu A, Bartels E, et al. The EFSUMB guidelines and recommendations for the clinical practice of contrast-enhanced ultrasound (CEUS) in non-hepatic applications: Update 2017 (Long version). *Ultraschall Med* (2018) 39(2):e2–e44. doi: 10.1055/a-0586-1107

19. Sengul I, Sengul D. Hermeneutics for evaluation of the diagnostic value of ultrasound elastography in TIRADS 4 categories of thyroid nodules. *Am J Med Case Rep* (2021) 9(11):538–9. doi: 10.12691/ajmcr-9-11-5

20. Takashima S, Matsuzuka F, Nagareda T, Tomiyama N, Kozuka T. Thyroid nodules associated with hashimoto's thyroiditis: Assessment with US. *Radiology* (1992) 185:125–30. doi: 10.1148/radiology.185.1.1523294

21. Zhang Y, Zhou P, Tian SM, Zhao YF, Li JL, Li L. Usefulness of combined use of contrast-enhanced ultrasound and TI-RADS classification for the differentiation of benign from malignant lesions of thyroid nodules. *Eur Radiol* (2017) 27:1527–36. doi: 10.1007/s00330-016-4508-y

22. Wang Y, Nie F, Liu T, Yang D, Li Q, Li J, et al. Revised value of contrast-enhanced ultrasound for solid hypo-echoic thyroid nodules graded with the thyroid imaging reporting and data system. *Ultrasound Med Biol* (2018) 44:930–40. doi: 10.1016/j.ultrasmedbio.2017.12.018

23. Deng J, Zhou P, Tian SM, Zhang L, Li JL, Qian Y. Comparison of diagnostic efficacy of contrast-enhanced ultrasound, acoustic radiation force impulse imaging, and their combined use in differentiating focal solid thyroid nodules. *PloS One* (2014) 9(3):e90674. doi: 10.1371/journal.pone.0090674

24. Zhang B, Jiang YX, Liu JB, Yang M, Dai Q, Zhu QL, et al. Utility of contrast-enhanced ultrasound for evaluation of thyroid nodules. *Thyroid* (2010) 20(1):51–7. doi: 10.1089/thy.2009.0045

25. Cantisani V, Consorti F, Guerrisi A, Guerrisi I, Ricci P, Di Segni M, et al. Prospective comparative evaluation of quantitative-elastosonography (Q-elastography) and contrast-enhanced ultrasound for the evaluation of thyroid nodules: preliminary experience. *Eur J Radiol* (2013) 82(11):1892–8. doi: 10.1016/j.ejrad.2013.07.005

26. Ma X, Zhang B, Ling W, Liu R, Jia H, Zhu F, et al. Contrast-enhanced sonography for the identification of benign and malignant thyroid nodules: Systematic review and meta-analysis. *J Clin Ultrasound* (2016) 44(4):199–209. doi: 10.1002/jcu.22311

27. Petrasova H, Slaisova R, Rohan T, Stary K, Kyclova J, Pavlik T, et al. Contrast-enhanced ultrasonography for differential diagnosis of benign and malignant thyroid lesions: Single-institutional prospective study of qualitative and quantitative CEUS characteristics. *Contrast Media Mol Imaging* (2022) 2022:8229445. doi: 10.1155/2022/8229445

28. Fu X, Guo L, Zhang H, Ran W, Fu P, Li Z, et al. "Focal thyroid inferno" on color Doppler ultrasonography: A specific feature of focal hashimoto's thyroiditis. *Eur J Radiol* (2012) 81:3319–25. doi: 10.1016/j.ejrad.2012.04.033

29. Săftoiu A, Gilja OH, Sidhu PS, Dietrich CF, Cantisani V, Amy D, et al. The EFSUMB guidelines and recommendations for the clinical practice of elastography in non-hepatic applications: Update 2018. *Ultraschall Med* (2019) 40(4):425–53. doi: 10.1055/a-0838-9937

30. Bojunga J, Herrmann E, Meyer G, Weber S, Zeuzem S, Friedrich-Rust M. Real-time elastography for the differentiation of benign and malignant thyroid nodules: A meta-analysis. *Thyroid* (2010) 20(10):1145–50. doi: 10.1089/thy.2010.0079

31. Sengul D, Sengul I, Van Slycke S. Risk stratification of the thyroid nodule with Bethesda indeterminate cytology, category III, IV, V on the one surgeon-performed US-guided fine-needle aspiration with 27-gauge needle, verified by histopathology of thyroidectomy: The additional value of one surgeon-performed elastography. *Acta Chir Belg* (2019) 119(1):38–46. doi: 10.1080/00015458.2018.1551769

32. Zhu L, Chen Y, Ai H, Gong W, Zhou B, Xu Y, et al. Combining real-time elastography with fine-needle aspiration biopsy to identify malignant thyroid nodules. *J Int Med Res* (2020) 48(12):300060520976027. doi: 10.1177/0300060520976027

33. Sengul D, Sengul I. Reassessing combining real-time elastography with fine-needle aspiration biopsy to identify malignant thyroid nodules. *Am J Med Case Rep* (2021) 9(11):552–3. doi: 10.12691/ajmcr-9-11-9

34. Moon HJ, Sung JM, Kim EK, Yoon JH, Youk JH, Kwak JY. Diagnostic performance of gray-scale US and elastography in solid thyroid nodules. *Radiol* (2012) 262(3):1002–13. doi: 10.1148/radiol.11110839

35. Zhang WB, Deng WF, Mao L, He BL, Liu H, Chen J, et al. Comparison of diagnostic value of SWE, FNA and BRAF gene detection in ACR TI-RADS 4 and 5 thyroid nodules. *Clin Hemorheol Microcirc* (2022) 81(1):13–21. doi: 10.3233/CH-211280

36. Chen Y, Dong B, Jiang Z, Cai Q, Huang L, Huang H. SuperSonic shear imaging for the differentiation between benign and malignant thyroid nodules: A

meta-analysis. *J Endocrinol Invest* (2022) 45(7):1327–39. doi: 10.1007/s40618-022-01765-y

37. Liao LJ, Chen HW, Hsu WL, Chen YS. Comparison of strain elastography, shear wave elastography, and conventional ultrasound in diagnosing thyroid nodules. *J Med Ultrasound* (2019) 27(1):26–32. doi: 10.4103/JMU.JMU_46_18

38. Sengul D, Sengul I, Egrioglu E, Ozturk T, Aydin I, Kesicioglu T, et al. Can cut-off points of 10 and 15 mm of thyroid nodule predict malignancy on the basis of three diagnostic tools: i) strain elastography, ii) the Bethesda system for reporting thyroid cytology with 27-gauge fine-needle, and iii) histopathology? *J BUON* (2020) 25(2):1122–9.

39. Li H, Kang C, Xue J, Jing L, Miao J. Influence of lesion size on shear wave elastography in the diagnosis of benign and malignant thyroid nodules. *Sci Rep* (2021) 11(1):21616. doi: 10.1038/s41598-021-01114-8

40. Shang H, Wu B, Liu Z, Liu Y, Cheng W. The effectiveness of shear wave elastography in the diagnosis of PTMC. *Technol Health Care* (2020) 28(2):221–6. doi: 10.3233/THC-191895

41. Wang F, Chang C, Chen M, Gao Y, Chen YL, Zhou SC, et al. Does lesion size affect the value of shear wave elastography for differentiating between benign and malignant thyroid nodules? *J Ultrasound Med: Off J Am Institute Ultrasound Med.* (2018) 37(3):601–9. doi: 10.1002/jum.14367

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 27 January 2023
DOI 10.3389/fonc.2023.1007464



OPEN ACCESS

EDITED BY
Rosaria Maddalena Ruggeri,
University of Messina, Italy

REVIEWED BY
Fajin Dong,
Jinan University, China
Gustavo Rubio,
Jackson Health System, United States

*CORRESPONDENCE
Jun Li
1287424798@qq.com
Xin Wu Cui
cuixinwu@live.cn

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 30 July 2022
ACCEPTED 09 January 2023
PUBLISHED 27 January 2023

CITATION
Li J, Zhang Y-R, Ren J-Y, Li Q-L, Zhu P-S, Du T-T, Ge X-Y, Chen M and Cui XW (2023) Association between diagnostic efficacy of acoustic radiation force impulse for benign and malignant thyroid nodules and the presence or absence of non-papillary thyroid cancer: A meta-analysis. *Front. Oncol.* 13:1007464. doi: 10.3389/fonc.2023.1007464



# Association between diagnostic efficacy of acoustic radiation force impulse for benign and malignant thyroid nodules and the presence or absence of non-papillary thyroid cancer: A meta-analysis

Jun Li[1,2]*, Yu-Rui Zhang[1], Jia-Yu Ren[3], Qiao-Li Li[1], Pei-Shan Zhu[1], Ting-Ting Du[1], Xiao-Yan Ge[1], Ming Chen[1] and Xin Wu Cui[3]*

[1]Department of Ultrasound, The First Affiliated Hospital of Medical College, Shihezi University, Shihezi, China, [2]NHC Key Laboratory of Prevention and Treatment of Central Asia High Incidence Diseases, First Affiliated Hospital, School of Medicine, Shihezi University, Shihezi, China, [3]Department of Medical Ultrasound, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

**Purpose:** The aim of this study was to investigate the diagnostic efficacy of Acoustic Radiation Force Impulse (ARFI) for benign and malignant thyroid nodules in the presence and absence of non-papillary thyroid cancer (NPTC) and to determine the cut-off values of Shear Wave Velocity (SWV) for the highest diagnostic efficacy of Virtual Touch Quantification (VTQ) and Virtual Touch Tissue Imaging and Quantification (VTIQ).

**Methods:** The diagnostic accuracy of ARFI for benign and malignant thyroid nodules was assessed by pooling sensitivity, specificity and area under the curve (AUC) in each group in the presence and absence of both non-papillary thyroid glands, using histology and cytology as the gold standard. All included studies were divided into two groups according to VTQ and VTIQ, and each group was ranked according to the magnitude of the SWV cutoff value to determine the SWV cutoff interval with the highest diagnostic efficacy for VTQ and VTIQ.

**Results:** A total of 57 studies were collected on the evaluation of ARFI for the diagnosis of benign and malignant thyroid nodules. The results showed that the presence of non-papillary thyroid carcinoma led to differences in the specificity of VTIQ for the identification of benign and malignant thyroid nodules, and the differences were statistically significant. In addition, the diagnostic efficacy of VTQ was best when the cutoff value of SWV was in the interval of 2.48-2.55 m/s, and the diagnostic efficacy of VTIQ was best when the cutoff value of SWV was in the interval of 3.01-3.15 m/s.

**Conclusion:** VTQ and VTIQ have a high diagnostic value for benign and malignant thyroid nodules; however, when the malignant nodules in the study contain non-papillary thyroid carcinoma occupying the thyroid gland, the findings should be viewed in a comprehensive manner.

KEYWORDS

non-papillary thyroid cancer (NPTC), virtual touch quantification (VTQ), virtual touch tissue imaging and quantification (VTIQ), shear wave velocity (SWV), meta-analysis

# Introduction

Thyroid nodules are a very common thyroid disorder and the incidence of thyroid nodules has shown an increasing trend year by year over the last few decades (1). Thyroid cancer accounts for 5% of thyroid nodules (2). There are four main types of thyroid cancer pathology: papillary, follicular, medullary and interstitial. The most common of these pathological types is papillary thyroid cancer (PTC), which also has the best prognosis among thyroid cancers, while the others have a poor prognosis (3). Among them, interstitial thyroid cancer, although less common, is one of the most dangerous tumors and is an associated cause of death in nearly half of thyroid cancer patients (4). Therefore, the first prerequisite for clinical diagnosis is to identify the benign and malignant thyroid nodules and then to develop the most appropriate treatment plan based on this, in order to reduce unnecessary surgeries and surgical complications, and ultimately to improve the quality of life as well as the health status of patients.

Ultrasonography is the test of choice for thyroid disease. Preoperative ultrasound examination of thyroid nodules is the most commonly used clinical method. However, conventional ultrasonography, including color Doppler ultrasound, cannot accurately differentiate between benign and malignant thyroid nodules, even when combined with CT and MRI examinations (1).

Currently, fine-needle aspiration biopsy (FNAB) is one of the recommended adjuncts for the diagnosis of thyroid nodules, but studies have shown that the sensitivity and specificity of FNAB for the diagnosis of thyroid nodules are 65-98% and 72-100%, respectively, and 20%-30% of samples cannot be diagnosed pathologically, with a certain rate of underdiagnosis (5–10). Moreover, FNAB is an invasive procedure with potential complications that have a negative impact on the patient's health.

In recent years, acoustic radiation force pulse elastography (ARFI) has been widely used in the examination of thyroid diseases, which can reflect the different hardness characteristics of benign and malignant lesions and is very useful for the identification of benign and malignant lesions (11). ARFI includes virtual touch tissue imaging and quantification (VTIQ) and virtual touch tissue quantification (VTQ) techniques, which are based on the principle of measuring the shear wave velocity (SWV) of the regions of interest (ROI) of the tissue. SWV is used to quantify the stiffness of the tissue. In a tissue lesion, the faster the shear wave velocity, the harder the lesion; the slower the shear wave velocity, the softer the lesion (2). Tissue stiffness is a characteristic that can reflect the nature of the nodule. The degree of fibrosis and the number of tumor cells vary among different histologic types of thyroid nodules, resulting in different stiffness in different histologic types of thyroid nodules. Compared to papillary carcinomas, other types of thyroid carcinomas, such as follicular, medullary, and undifferentiated carcinomas exhibit relatively soft structures (12).

In the past, a meta-analysis was performed to evaluate the diagnostic efficacy of ARFI in identifying benign and malignant thyroid nodules, and the results of the study showed that ARFI performed well in the differential diagnosis of benign and malignant thyroid nodules, and that ARFI may help guide the clinical choice of surgery for patients with thyroid nodules (13). However, this study only made a simple benign-malignant distinction between thyroid nodules and did not further delineate the pathological types of thyroid cancer.

The main objective of the present study, taken together with previous studies, was to assess whether the presence of nonpapillary thyroid cancer affects the diagnostic efficacy of ARFI for benign and malignant thyroid nodules and to determine the cut-off interval of SWV with optimal diagnostic efficacy for VTQ and VTIQ.

# Material and methods

## Search strategy

The search databases web of science, PubMed, and Embase were searched for relevant studies published up to May 1, 2022, with the search terms "(Acoustic Radiation Force Impulse or ARFI or VTIQ or VTQ or Virtual Touch tissue imaging and quantification or Virtual Touch tissue quantification) and (thyroid or thyroid nodules)". The search language was English. In order to search as much relevant literature as possible, the search method of this paper was subject terms combined with free words, web search combined with manual search, and secondary search of the retrieved relevant literature was conducted.

## Study selection

Inclusion criteria: (i) the literature study must include the diagnostic analysis of thyroid nodules by ARFI; (ii) there is a gold standard for diagnosing the pathology of thyroid nodules in the literature, and the number of benign and malignant nodules must be given directly or indirectly; (iii) the number of patients must be ≥30;

(iv) the literature should provide raw data and calculate the sensitivity, specificity, false positives and false negatives can be calculated directly or indirectly.

Exclusion criteria: (i) diagnostic criteria were not described; (ii) data could not be extracted; (iii) duplicate literature; (iv) pathological histology was not used as the gold standard; (v) cutoff values for SWV were not indicated; (vi) editorials, letters, case reports, review articles, commentaries, case-control studies, and conference articles.

## Literature inclusion

Two authors independently searched and read the titles, abstracts, and keywords of the detected literature to initially identify eligible literature that could be selected, and then carefully read the full text of the literature to finalize the eligible literature that could be included.

If the 2 authors disagreed on whether the literature should be included, a third author helped to suggest a solution.

## Data extraction

Relevant database literature was screened by 2 independent authors using a blinded method and in strict accordance with the inclusion and exclusion criteria of the literature; those that met the requirements were included and those that did not were excluded. The extracted literature included the authors' names, the location of the study, the time of publication, the number of included lesions, the number of benign and malignant lesions, the pathology of malignant nodules, and the sensitivity, specificity, and accuracy of the test to be evaluated.

## Statistical analysis

Both stata 16.0 software and RevMan 5.3 software were used for statistical analysis in this study.

The statistical software was used to produce summary receiver operating characteristic (SROC) curves, publication bias funnel plots, and to calculate the sensitivity, specificity, and area under curve (AUC) of the diagnosis, respectively.

## Literature quality evaluation

All included literature was evaluated for quality using RevMan 5.3, a revised tool for quality assessment of diagnostic accuracy studies, including patient selection, index tests, reference standards, processes, and timelines.

# Results

## Literature search results

In this meta-analysis, 1405 original articles were retrieved based on the search terms "(Acoustic Radiation Force Impulse or ARFI or VTIQ or VTQ or Virtual Touch tissue imaging and quantification or Virtual Touch tissue quantification) and (thyroid or thyroid nodules)". By carefully reading the titles and abstracts of the articles, 134 papers were initially included, and then the papers were strictly screened and excluded according to the inclusion and exclusion criteria, and finally 57 papers met the conditions of meta-analysis. The specific inclusion process of the literature is shown in Figure 1.

## Characteristics of the eligible studies

A total of 8802 thyroid nodules were included in the pooled 57 studies. The nature of all thyroid nodules in all included studies was histologically confirmed. Benign thyroid nodules included nodular goiter, eosinophilia, Hashimoto's thyroiditis, subacute thyroiditis, and thyroid adenoma, while malignant thyroid nodules included papillary, follicular, undifferentiated, metastatic, and medullary carcinomas.

## Data analysis

The pooled 57 papers were divided into two groups according to the different pathological characteristics of malignant nodes, with group A being studies in which all included malignant nodes were papillary carcinomas, 21 in total, and group B being studies in which included malignant nodes included medullary carcinomas, follicular carcinomas, undifferentiated carcinomas, and other metastatic carcinomas in addition to papillary carcinomas, 36 in total (see Tables 1 and 2 for the specific data of the two groups, respectively), and then classified according to VTQ and VTIQ two techniques were classified again, and groups A and B were divided into $A_{VTQ}$ group, $A_{VTIQ}$ group,$B_{VTQ}$ group and $B_{VTIQ}$ group, respectively. Regression analysis was done for each of the four data groups, and the data showed that the sensitivity, specificity, and AUC of the $A_{VTQ}$ group were 0.82 (95CI%, 0.76-0.87), 0.84 (95CI%, 0.78-0.89), and 0.90 (95CI %, 0.87-0.92) (Figure 2), respectively; the sensitivity, specificity, and AUC of the $A_{VTIQ}$ group were 0.75 (95CI%, 0.69-0.80), 0.83 (95CI%, 0.75-0.89) and 0.79 (95CI%, 0.75-0.82), respectively (Figure 3); the sensitivity, specificity and AUC of $B_{VTQ}$ group were 0.82 (95CI%, 0.77-0.85), (95CI%, 0.82-0.90) and 0.90 (95CI%, 0.87-0.93) (Figure 4); the sensitivity, specificity, and AUC of the $B_{VTIQ}$ group were 0.81 (95CI%, 0.76-0.86), 0.85 (95CI%, 0.74-0.91), and 0.89 (95CI%, 0.85-0.91), respectively (Figure 5). The data from the $A_{VTQ}$ group were compared with the $B_{VTQ}$ group and the $A_{VTIQ}$ group with the $B_{VTIQ}$ group. The results of data analysis showed that the difference in sensitivity and specificity between the $A_{VTQ}$ and $B_{VTQ}$ groups was small and not statistically significant ($p < 0.01$ for both sensitivity and specificity), and there was no difference in sensitivity but a difference in specificity between the $A_{VTIQ}$ and $B_{VTIQ}$ groups and the difference was statistically significant ($p = 0.08 > 0.05$).

Then, the 58 papers were divided into two groups according to the two techniques of VTIQ and VTQ, and each group was sorted according to the size of the cut-off value from smallest to largest, and then the sensitivity, specificity and AUC of each group were calculated in every three groups. The sensitivity, specificity, and AUC of VTQ were 0.91 (95CI%, 0.80-0.97), 0.88

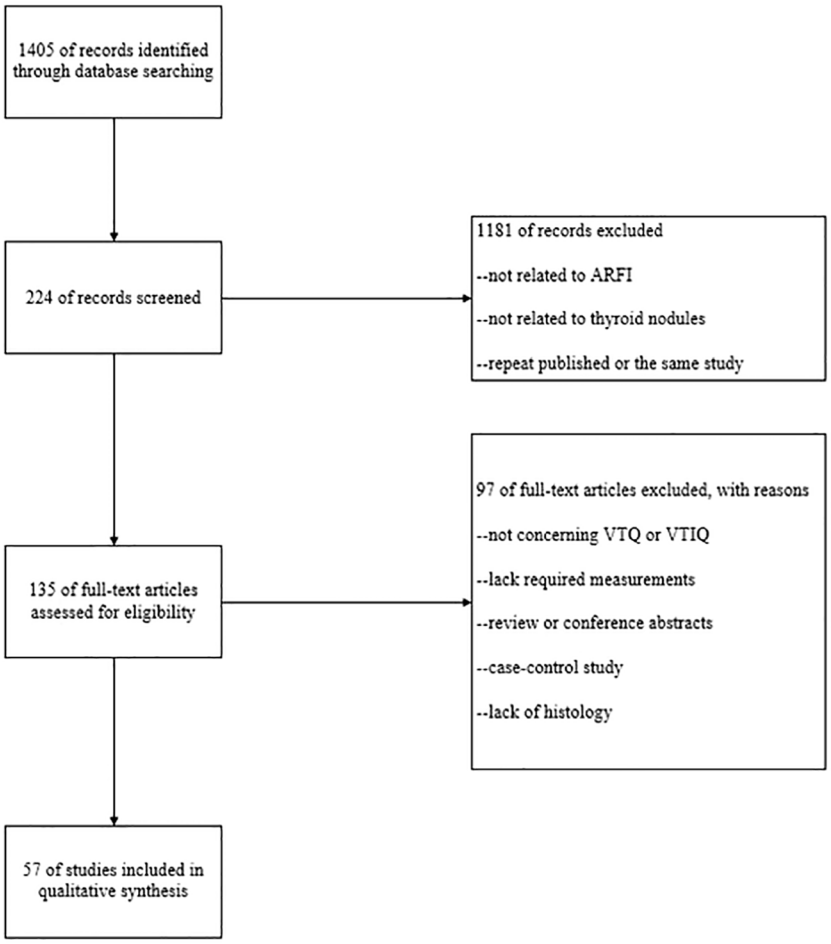



FIGURE 1
Inclusion and exclusion process of the literature.

(95CI%, 0.73-0.95), and 0.96 (95CI%, 0.93-0.97), respectively; when the cut-off value was in the interval of 3.01-3.15m/s, the diagnostic efficacy of VTIQ was the best. The best diagnostic performance of VTIQ was achieved when the cut-off value was in the interval of 3.01-3.15m/s, with sensitivity, specificity and AUC of 0.74 (95CI%, 0.59-0.58), 0.92 (95CI%, 0.75-0.98) and 0.88 (95CI%, 0.84-0.90), respectively.

## Publication bias

When using meta-analysis in diagnostic trials, Deeks funnel plots are usually chosen to assess publication bias, and the results of Deeks funnel plots are shown in Figure 6. $p > 0.05$, suggesting no publication bias in this study.

## Literature quality assessment

All included literature was evaluated for quality using RevMan 5.3, and the results of the literature quality evaluation are shown in Figures 7 and 8.

# Discussion

In this study, the included literature was divided into four groups according to whether all malignant nodules were papillary thyroid carcinomas and the difference between VTQ and VTIQ. From the results, it is clear that there was a statistically significant difference in specificity between group A and group B only when VTIQ was used to identify benign and malignant thyroid nodules, and the specificity of diagnosis was better when non-papillary thyroid carcinomas were included in malignant thyroid nodules.

From past studies, it is known that non-thyroidal papillary carcinomas such as follicular and medullary carcinomas are pathologically different from papillary thyroid carcinomas, with follicular and medullary carcinomas having less fibrous content and more cellular components compared to papillary carcinomas. Papillary carcinomas are often accompanied by sand-like calcification formation, so the pathological specimens of papillary carcinomas are harder, while follicular and medullary carcinomas are softer in texture (11). However, the results of this study showed that there was no difference in the diagnostic efficacy of VTQ for malignant nodules regardless of whether they contained non-papillary thyroid carcinoma, whereas the specificity of VTIQ was

TABLE 1 Characteristics of Group A literature.

| Author | Study region | year | Number of nodules | Benign/ Malignant | Malignant % | Sensitive | Specificity | TP | FP | FN | TN | Cut-off m/s | Type |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Xiao,LL (14) | China | 2012 | 67 | 28/39 | 58.2 | 76.9 | 78.6 | 30 | 6 | 9 | 22 | 2.78 | VTQ |
| Hou,XJ (15) | China | 2013 | 85 | 65/20 | 23.5 | 80.0 | 89.2 | 16 | 7 | 4 | 58 | 2.42 | VTQ |
| Dong,FJ (16) | China | 2015 | 55 | 28/27 | 49.1 | 88.9 | 96.4 | 24 | 1 | 3 | 27 | 2.42 | VTQ |
| Liu,BJ (17) | China | 2017 | 141 | 70/71 | 50.4 | 76.1 | 70.0 | 54 | 21 | 17 | 49 | 2.58 | VTQ |
| Wu,L (18) | China | 2018 | 88 | 56/32 | 36.4 | 78.1 | 91.1 | 25 | 5 | 7 | 51 | 2.80 | VTQ |
| Chen,L (19) | China | 2013 | 78 | 50/28 | 35.9 | 71.4 | 86.0 | 20 | 7 | 8 | 43 | 3.18 | VTQ |
| Zhang,YF (20) | China | 2014 | 173 | 77/96 | 55.5 | 56.2 | 79.2 | 54 | 16 | 42 | 61 | 3.10 | VTQ |
| Huang,R (21) | China | 2018 | 51 | 17/34 | 66.7 | 76.5 | 94.1 | 26 | 1 | 8 | 16 | 2.19 | VTQ |
| Zhang,FJ (22) | China | 2013 | 155 | 93/62 | 40.0 | 96.8 | 95.7 | 60 | 4 | 2 | 89 | 2.84 | VTQ |
| Jiang,LY (23) | China | 2016 | 195 | 103/92 | 47.2 | 72.8 | 77.7 | 67 | 23 | 25 | 80 | 2.98 | VTQ |
| Sha,YM (24) | China | 2017 | 95 | 24/71 | 74.7 | 88.7 | 95.8 | 63 | 1 | 8 | 23 | 2.67 | VTQ |
| Song,HY (25) | China | 2014 | 193 | 136/57 | 29.5 | 97.0 | 81.0 | 55 | 26 | 2 | 110 | 2.48 | VTQ |
| Xing,P (26) | China | 2016 | 90 | 54/36 | 40.0 | 80.6 | 74.1 | 29 | 14 | 7 | 40 | 2.57 | VTQ |
| Ha,seung Mi (27) | Korea | 2016 | 198 | 168/30 | 15.2 | 86.7 | 50.6 | 26 | 83 | 4 | 85 | 2.37 | VTQ |
| Ke,K (28) | China | 2017 | 69 | 37/32 | 46.4 | 87.5 | 86.5 | 28 | 5 | 4 | 32 | 2.55 | VTQ |
| Zhang,YF (29) | China | 2012 | 173 | 129/44 | 25.4 | 75.0 | 82.2 | 33 | 23 | 11 | 106 | 2.87 | VTQ |
| Zhang,Y (30) | China | 2019 | 62 | 22/40 | 64.5 | 77.5 | 63.6 | 31 | 8 | 9 | 14 | 3.00 | VTIQ |
| Wu,SN (31) | China | 2016 | 51 | 16/35 | 68.6 | 88.6 | 93.7 | 31 | 1 | 4 | 15 | 2.49 | VTIQ |
| Li,DX (32) | China | 2017 | 186 | 82/104 | 55.9 | 72.1 | 87.8 | 75 | 10 | 29 | 72 | 2.91 | VTIQ |
| He,YP (33) | China | 2017 | 75 | 49/26 | 34.7 | 65.4 | 83.7 | 17 | 8 | 9 | 41 | 3.51 | VTIQ |
| Peng,QH (34) | China | 2019 | 85 | 36/49 | 57.6 | 73.5 | 80.6 | 36 | 7 | 13 | 29 | 3.20 | VTIQ |

superior for the group of malignant nodules containing non-papillary thyroid carcinoma.

This result was unexpected, for which several speculations were made: one, it may be because some malignant nodules such as follicular carcinoma and medullary carcinoma have more distinct ultrasound features due to their poor differentiation. According to the latest European Thyroid Association guidelines, when a lesion has one of the above features of irregular shape, irregular border, microcalcifications and deep hypoechogenicity, the nodule may be malignant up to 26-87%. The more malignant features a tumor has, the highest its risk of malignancy. In a study by Zhao,J 2020, it was shown that some medullary carcinomas have more obvious malignant ultrasound features, specifically the irregular morphology of the tumor, poor demarcation with surrounding tissues, solid hypoechoic or very hypoechoic, and intra-nodular calcification (69); secondly, it is also possible that there are many microscopic thyroid papillary carcinomas among the papillary thyroid carcinomas, and The ROI range of ARFI is 6mm×5mm, which is not suitable for the diagnosis of smaller nodules, and this may also be the reason for this result (70). For example, in the included study by Chen, SH in 2014, they included a total of 275 nodules and 23 microscopic papillary thyroid carcinomas out of 60 papillary thyroid carcinomas. The sensitivity of VTQ for thyroid nodules in that article was 75% and the specificity was 70% (44); in addition, all nodules included in Zhu,J's article in 2015 were microscopic nodules, and the

10.3389/fonc.2023.1007464

TABLE 2 Characteristics of Group B literature.

| Author | Study region | year | Number of nodules | Benign/ Malignant | Malignant % | Sensitive | Specificity | TP | FP | FN | TN | Cut-off m/s | Type |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cao,DM (35) | China | 2019 | 148 | 120/28 | 18.9 | 81.2 | 64.7 | 23 | 42 | 5 | 78 | 3.30 | VTQ |
| Deng,J (36) | China | 2014 | 175 | 119/56 | 32.0 | 80.4 | 84.0 | 45 | 19 | 11 | 100 | 2.59 | VTQ |
| Wang,R (37) | China | 2015 | 129 | 70/59 | 45.7 | 67.8 | 91.4 | 40 | 6 | 19 | 64 | 2.43 | VTQ |
| Ning,CP (38) | China | 2014 | 179 | 64/115 | 64.2 | 74.8 | 73.4 | 86 | 17 | 29 | 47 | 2.47 | VTQ |
| Zhou,J (1) | China | 2014 | 191 | 122/69 | 36.1 | 96.3 | 96.2 | 66 | 5 | 3 | 117 | 2.55 | VTQ |
| Gu,JY (39) | China | 2012 | 98 | 76/22 | 22.4 | 86.4 | 93.4 | 19 | 5 | 3 | 71 | 2.56 | VTQ |
| Wang,XY (40) | China | 2016 | 88 | 59/29 | 33.0 | 75.9 | 94.9 | 22 | 3 | 7 | 56 | 2.57 | VTQ |
| Hamidi,C (41) | Italian | 2015 | 95 | 62/33 | 34.7 | 100.0 | 82.3 | 33 | 11 | 0 | 51 | 2.75 | VTQ |
| Bojunga,J (42) | Germany | 2012 | 158 | 137/21 | 13.3 | 57.0 | 85.0 | 12 | 20 | 9 | 117 | 2.57 | VTQ |
| Zhan,J (43) | China | 2015 | 170 | 102/68 | 40.0 | 79.4 | 84.3 | 54 | 16 | 14 | 86 | 2.75 | VTQ |
| Chen,SH (44) | China | 2014 | 125 | 62/63 | 50.4 | 75.0 | 70.0 | 47 | 19 | 16 | 43 | 2.50 | VTQ |
| Hou,JX (45) | China | 2014 | 44 | 19/25 | 56.8 | 88.0 | 94.7 | 22 | 1 | 3 | 18 | 2.76 | VTQ |
| Zou,X (46) | China | 2014 | 144 | 65/79 | 54.9 | 84.8 | 75.4 | 67 | 16 | 12 | 49 | 2.79 | VTQ |
| Xu,JM (47) | China | 2014 | 441 | 325/116 | 26.3 | 71.6 | 83.4 | 83 | 54 | 33 | 271 | 2.87 | VTQ |
| Chen,Q (48) | China | 2018 | 271 | 162/109 | 40.2 | 76.2 | 86.4 | 83 | 22 | 26 | 140 | 2.81 | VTQ |
| Yang,YP (49) | China | 2017 | 107 | 87/20 | 18.7 | 70.0 | 95.4 | 14 | 4 | 6 | 83 | 2.83 | VTQ |
| Zhang,FJ (50) | China | 2017 | 152 | 97/55 | 36.2 | 78.2 | 83.5 | 43 | 16 | 12 | 81 | 2.87 | VTQ |
| Li,J (51) | China | 2015 | 100 | 77/23 | 23.0 | 91.3 | 85.7 | 21 | 11 | 2 | 66 | 2.88 | VTQ |
| Xu,JM (52) | China | 2014 | 183 | 117/66 | 36.1 | 68.2 | 76.9 | 45 | 27 | 21 | 90 | 2.87 | VTQ |
| Zhang,FJ (53) | China | 2014 | 113 | 67/46 | 40.7 | 91.3 | 85.1 | 42 | 10 | 4 | 57 | 2.90 | VTQ |
| Du,YR (54) | China | 2018 | 142 | 70/72 | 50.7 | 91.7 | 60.0 | 66 | 28 | 6 | 42 | 2.31 | VTQ |
| Zhang,HP (55) | China | 2014 | 71 | 39/32 | 45.1 | 71.9 | 100.0 | 23 | 0 | 9 | 39 | 2.91 | VTQ |
| Ni,JN (56) | China | 2013 | 275 | 152/123 | 44.7 | 91.1 | 82.3 | 112 | 27 | 12 | 125 | 2.35 | VTQ |
| Jung,WS (57) | Korea | 2016 | 127 | 95/32 | 25.2 | 75.0 | 91.0 | 24 | 9 | 8 | 86 | 3.28 | VTQ |
| Pandey, NN (58) | India | 2017 | 40 | 26/14 | 35.0 | 85.7 | 96.2 | 12 | 1 | 2 | 25 | 2.53 | VTQ |
| Grazhdani H (59) | Italy | 2014 | 82 | 60/22 | 26.8 | 95.0 | 75.0 | 21 | 15 | 1 | 45 | 2.46 | VTQ |
| Tong,J (60) | China | 2020 | 98 | 45/53 | 54.1 | 84.9 | 86.7 | 45 | 6 | 8 | 39 | 2.96 | VTIQ |

*(Continued)*

TABLE 2 Continued

| Author | Study region | year | Number of nodules | Benign/ Malignant | Malignant % | Sensitive | Specificity | TP | FP | FN | TN | Cut-off m/s | Type |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Xu,L (61) | China | 2020 | 922 | 405/517 | 56.1 | 86.3 | 80.5 | 446 | 79 | 71 | 326 | 3.55 | VTIQ |
| Zhao,N (62) | China | 2022 | 212 | 69/143 | 67.5 | 83.9 | 96.7 | 120 | 2 | 23 | 67 | 2.66 | VTIQ |
| Mao,F (63) | China | 2016 | 109 | 65/44 | 40.4 | 79.5 | 83.1 | 35 | 11 | 9 | 54 | 2.92 | VTIQ |
| Xu,L (64) | China | 2018 | 117 | 43/74 | 63.2 | 86.7 | 82.3 | 64 | 8 | 10 | 35 | 3.03 | VTIQ |
| Li,DX (65) | China | 2019 | 204 | 83/121 | 59.3 | 71.1 | 65.1 | 86 | 29 | 35 | 54 | 2.74 | VTIQ |
| Li,X (66) | China | 2019 | 130 | 57/73 | 56.2 | 92.3 | 63.2 | 67 | 21 | 6 | 36 | 2.80 | VTIQ |
| Zhou,H (67) | China | 2017 | 302 | 237/65 | 21.5 | 84.6 | 70.0 | 55 | 71 | 10 | 166 | 2.60 | VTIQ |
| Sun,CY (68) | China | 2017 | 388 | 238/150 | 38.7 | 64.7 | 86.6 | 97 | 32 | 53 | 206 | 3.15 | VTIQ |
| Yang,YP (49) | China | 2017 | 107 | 87/20 | 18.7 | 70.0 | 98.8 | 14 | 1 | 6 | 86 | 3.01 | VTIQ |

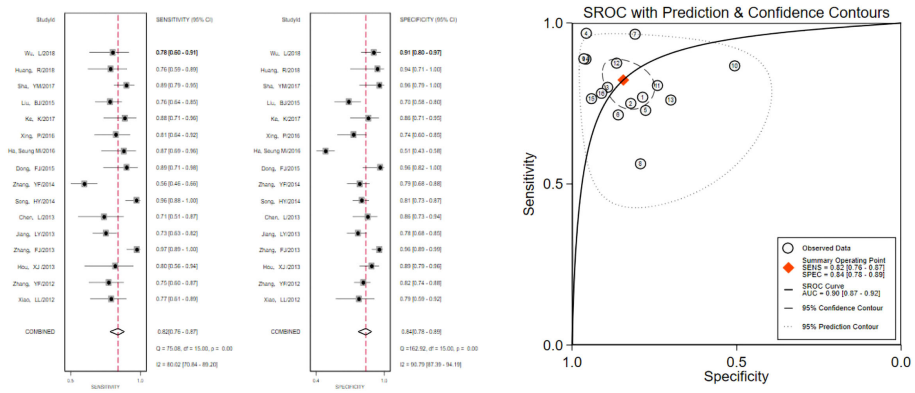



FIGURE 2
The sensitivity and specificity of the $A_{VTQ}$ group in the diagnosis of thyroid nodules and the summary ROC (Summary ROC) curve of the $A_{VTQ}$ group were analyzed. the AUC indicates the area under the curve.

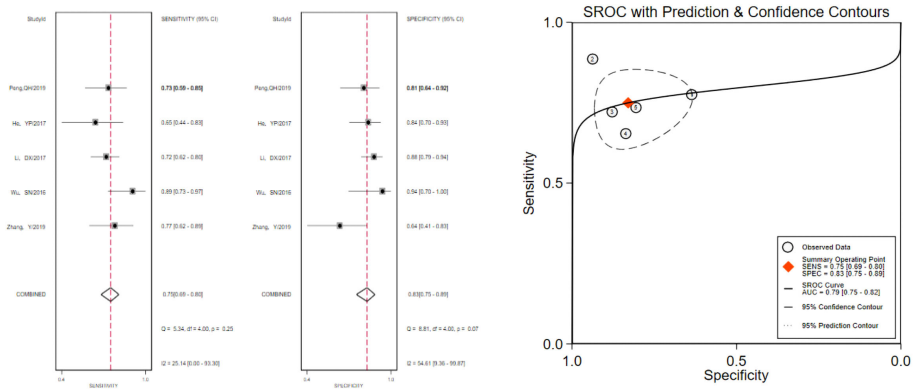



FIGURE 3
The sensitivity and specificity of the $A_{VTIQ}$ group in the diagnosis of thyroid nodules and the summary ROC (Summary ROC) curve of the $A_{VTIQ}$ group were analyzed. the AUC indicates the area under the curve.

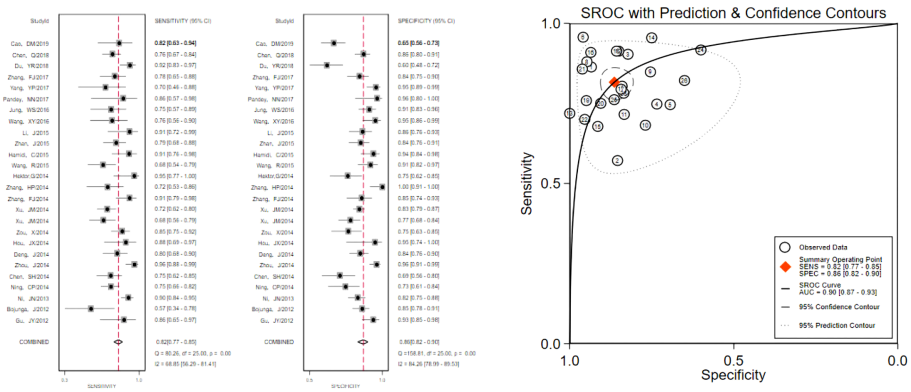



FIGURE 4
The sensitivity and specificity of the $B_{VTQ}$ group in the diagnosis of thyroid nodules and the summary ROC (Summary ROC) curve of the $B_{VTQ}$ group were analyzed. the AUC indicates the area under the curve.

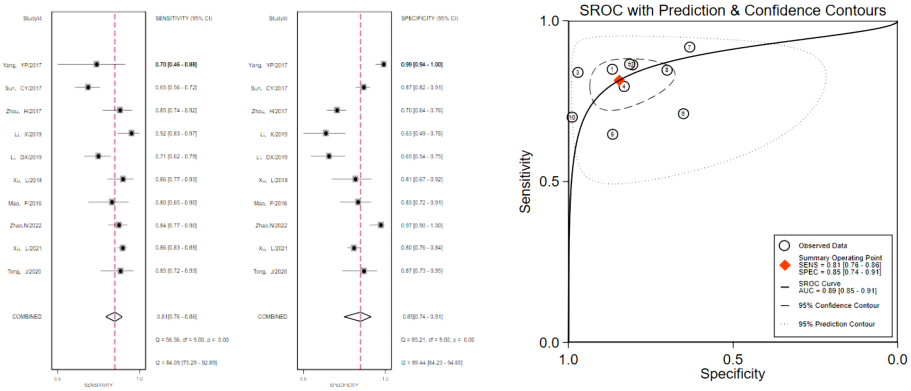



FIGURE 5
The sensitivity and specificity of the $B_{VTIQ}$ group in the diagnosis of thyroid nodules and the summary ROC (Summary ROC) curve of the $B_{VTIQ}$ group were analyzed. the AUC indicates the area under the curve.

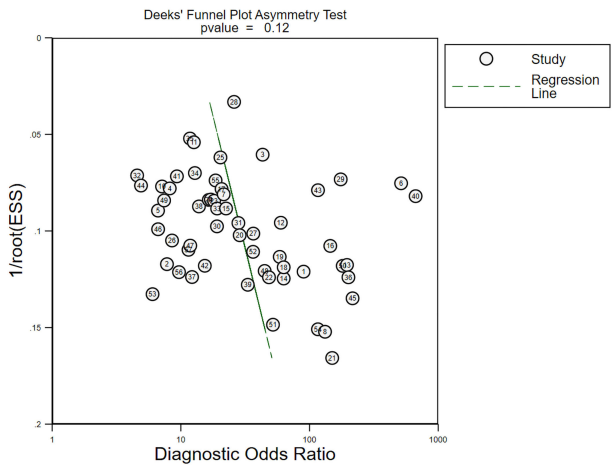



FIGURE 6
Publication bias assessment of all included literature.

sensitivity of VTQ for diagnosing benign and malignant thyroid nodules in that study was 76.4% and the specificity was 75.8% (70).

In addition, the measurement range of SWV is 0.5-8.4 m/s, a characteristic that makes ARFI unable to achieve satisfactory measurement results for extremely hard or soft tissues, and the quality of imaging is difficult to guarantee (2). Therefore, it will have an unavoidable impact on the quality of the article. In addition to this, there are some studies that did not exclude nodules with a background of diffuse thyroid lesions when they were included. Pathologically, diffuse thyroid lesions are caused by infiltration of thyroid follicular cells by diffuse lymphocytes, destruction of follicles by atrophy and fibrosis, and these factors can make the texture of the thyroid gland harder (71). It is also possible that although the malignant nodules in group B included nonpapillary carcinomas, the proportion of nonpapillary carcinomas in thyroid cancer was so low that the results in this section were biased, but because of the low incidence of nonpapillary thyroid carcinomas and the paucity of data, there is not a large body of data to support the exclusion of this speculation.

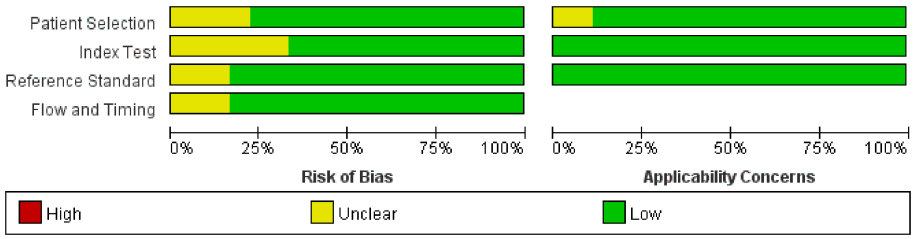



FIGURE 7
Literature quality evaluation of group A.

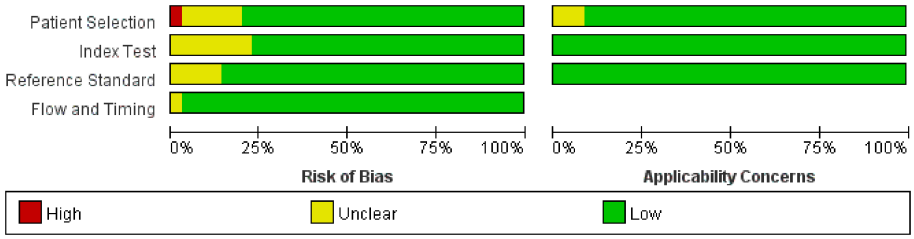



FIGURE 8
Literature quality evaluation of group B.

The sensitivity, specificity and AUC of each group were calculated by comparing the data of each group and found that the best diagnostic efficacy was achieved when the cut-off values were in the range of 2.48-2.55m/s. The sensitivity, specificity and AUC of VTQ were 0.91 (95CI%, 0.80-0.97), 0.88 (95CI%, 0.73-0.95) and 0.96 (95CI %, 0.93-0.97), respectively, and the diagnostic efficacy of VTIQ was best when the cut-off value was in the range of 3.01-3.15m/s. The sensitivity, specificity and AUC of VTQ were best when the cut-off value was in the range of 3.01-3.15m/s. The diagnostic efficacy of VTIQ was best when the cut-off value was in the range of 3.01-3.15m/s, with sensitivity, specificity and AUC of 0.74 (95CI%, 0.59-0.58), 0.92 (95CI%, 0.75-0.98) and 0.88 (95CI%, 0.84-0.90), respectively.

## Limitation

This meta-analysis has several limitations. We searched only three databases, PubMed, Web of science, and Embase, suggesting that there may be relevant studies that were missed. Also, this meta-analysis included only English-language literature, so there may be language bias.

## Conclusion

In summary, ARFI imaging is a highly effective imaging tool to identify benign and malignant thyroid nodules. There is no difference in the diagnostic effectiveness of VTQ for malignant nodules with or without non-papillary thyroid cancer, while VTIQ has a better specificity for the diagnosis of malignant nodules with non-papillary thyroid cancer. Therefore, ARFI imaging of benign and malignant thyroid nodules must take into account various clinical information of the patient and be analyzed critically to make a more accurate diagnosis.

## Author contributions

JL, Y-RZ and X-WC contributed to the conception and design of the study. Y-RZ and P-SZ searched and reviewed studies, extracted and analyzed the data, and wrote the first draft of the manuscript. Q-LL, MC and J-YR reviewed and edited the manuscript. X-YG and T-TD directed the project and contributed to discussion as well as reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

(No. 2020-PT330-003):Supported by Open Research Fund of NHC Key Laboratory of Prevention and Treatment of Central Asia High Incidence Diseases. Supported by the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences. (No. 2019DB012): The Corps Science and Technology Key Project.

## 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.

## Publisher's note

# References

1. Zhuo J, Ma Z, Fu WJ, Liu SP. Differentiation of benign from malignant thyroid nodules with acoustic radiation force impulse technique. *Br J Radiol* (2014) 87 (1035):20130263. doi: 10.1259/bjr.20130263

2. Liu BJ, Li DD, Xu HX, Guo LH, Zhang YF, Xu JM, et al. Quantitative shear wave velocity measurement on acoustic radiation force impulse elastography for differential diagnosis between benign and malignant thyroid nodules: A meta-analysis. *Ultrasound Med Biol* (2015) 41(12):3035–43. doi: 10.1016/j.ultrasmedbio.2015.08.003

3. Xu J, Xu X, Xu H, Zhang Y, Guo L, Liu L, et al. Prediction of cervical lymph node metastasis in patients with papillary thyroid cancer using combined conventional ultrasound, strain elastography, and acoustic radiation force impulse (Arfi) elastography. *Eur Radiol* (2016) 26(8):2611–22. doi: 10.1007/s00330-015-4088-2

4. Luo H, Xia X, Kim G, Xu H. Characterizing dedifferentiation of thyroid cancer by integrated analysis. *Sci Adv* (2021) 7(31):eabf3657. doi: 10.1126/sciadv.abf3657

5. La Rosa GL, Belfiore A, Giuffrida D, Sicurella C, Ippolito O, Russo G, et al. Evaluation of the fine needle aspiration biopsy in the preoperative selection of cold thyroid nodules. *Cancer* (1991) 67(8):2137–41. doi: 10.1002/1097-0142(19910415)67:8<2137::aid-cncr2820670822>3.0.co;2-y

6. Bojunga J, Herrmann E, Meyer G, Weber S, Zeuzem S, Friedrich-Rust M. Real-time elastography for the differentiation of benign and malignant thyroid nodules: A meta-analysis. *Thyroid* (2010) 20(10):1145–50. doi: 10.1089/thy.2010.0079

7. Ogilvie JB, Piatigorsky EJ, Clark OH. Current status of fine needle aspiration for thyroid nodules. *Adv Surg* (2006) 40:223–38. doi: 10.1016/j.yasu.2006.06.003

8. Al-azawi D, Mann GB, Judson RT, Miller JA. Endocrine surgeon-performed us guided thyroid fnac is accurate and efficient. *World J Surg* (2012) 36(8):1947–52. doi: 10.1007/s00268-012-1592-2

9. Bohacek L, Milas M, Mitchell J, Siperstein A, Berber E. Diagnostic accuracy of surgeon-performed ultrasound-guided fine-needle aspiration of thyroid nodules. *Ann Surg Oncol* (2012) 19(1):45–51. doi: 10.1245/s10434-011-1807-z

10. Tee YY, Lowe AJ, Brand CA, Judson RT. Fine-needle aspiration may miss a third of all malignancy in palpable thyroid nodules: A comprehensive literature review. *Ann Surg* (2007) 246(5):714–20. doi: 10.1097/SLA.0b013e3180f61adc

11. Shuzhen C. Comparison analysis between conventional ultrasonography and ultrasound elastography of thyroid nodules. *Eur J Radiol* (2012) 81(8):1806–11. doi: 10.1016/j.ejrad.2011.02.070

12. Chang N, Zhang X, Wan W, Zhang C, Zhang X. The preciseness in diagnosing thyroid malignant nodules using shear-wave elastography. *Med Sci Monit* (2018) 24:671–7. doi: 10.12659/msm.904703

13. Zhan J, Jin JM, Diao XH, Chen Y. Acoustic radiation force impulse imaging (Arfi) for differentiation of benign and malignant thyroid nodules–a meta-analysis. *Eur J Radiol* (2015) 84(11):2181–6. doi: 10.1016/j.ejrad.2015.07.015

14. Xiao L, Zhao Y, Gao L. Value of virtual touch tissue quantification technique in the diagnosis of small solid thyroid nodules. *Chin J Ultrasonogr* (2012) 21(9):771–4. doi: 10.3760/cma.j.issn.1004-4477.2012.09.012

15. Hou XJ, Sun AX, Zhou XL, Ji Q, Wang HB, Wei H, et al. The application of virtual touch tissue quantification (Vtq) in diagnosis of thyroid lesions: A preliminary study. *Eur J Radiol* (2013) 82(5):797–801. doi: 10.1016/j.ejrad.2012.12.023

16. Dong FJ, Xu JF, Liu HY. Virtual touch tissue quantification in differential diagnosis of benign and malignant thyroid nodules. *Chin J Med Imaging Technol* (2015) 31(3):347–50. doi: 10.13929/j.1003-3289.2015.03.007

17. Liu BJ, Zhao CK, Xu HX, Zhang YF, Xu JM, Li DD, et al. Quality measurement on shear wave speed imaging: Diagnostic value in differentiation of thyroid malignancy and the associated factors. *ONCOTARGET* (2017) 8(3):4948–59. doi: 10.18632/oncotarget.13996

18. Wu L, Li SY, Ni ZL. Virtual touch tissue quantification imaging in differential diagnosis of thyroid imaging report and data system 4 thyroid nodules. *Chin J Interv Imaging Ther* (2018) 15(12):732–5. doi: 10.13929/j.1672-8475.201807005

19. Chen L, Chen Y, Chen L, Zhan J. Comparison of ultrasound elastography imaging area ratio with acoustic radiation force impulse in differential diagnosis of thyroid nodules. *Chin J Ultrasound* (2013) 29(9):772–4. doi: 10.3969/j.issn.1002-0101.2013.09.002

20. Zhang YF, Liu C, Xu HX, Xu JM, Zhang J, Guo LH, et al. Acoustic radiation force impulse imaging: A new tool for the diagnosis of papillary thyroid microcarcinoma. *BioMed Res Int* (2014) 2014:416969. doi: 10.1155/2014/416969

21. Huang R, Wang XT, Wang R, Sha YM. Shear wave elastography and acoustic radiation force impulse imaging in differential diagnosis of benign and malignant thyroid nodules. *Chin J Interv Imaging Ther* (2018) 15(5):277–81. doi: 10.13929/j.1672-8475.201710050

22. Zhang FJ, Han RL. The value of acoustic radiation force impulse (Arfi) in the differential diagnosis of thyroid nodules. *Eur J Radiol* (2013) 82(11):e686–90. doi: 10.1016/j.ejrad.2013.06.027

23. Jiang LY, Cai YY, He YQ, Gu JY, Du LF, Yang YR. Grey-scale sonography and arfi-vtq in diagnosis of papillary thyroid microcarcinoma smaller than 1 Cm. *TUMOR* (2016) 36:202–7. doi: 10.3781/j.issn.1000-7431.2016.33.988

24. Sha YM, Wang XT, Wang R. Acoustic radiation force impulse imaging in differential diagnosis of small solid thyroid nodules coexisting with hashimoto's thyroiditis. *Chin J Interv Imaging Ther* (2017) 14(8):504–8. doi: 10.13929/j.1672-8475.201703044

25. Song HY, Huang DZ, Yang JZ. Virtual touch tissue quantification in differential diagnosis of thyroid benign and malignant nodules and its infouence factors. *Chin J Ultrasonogr* (2014) 23(3):227–30. doi: 10.3760/cma.j.issn.1004-4477.2014.03.017

26. Xing P, Chen Q, Yang ZW, Liu CB, Wu CJ. Combination of conventional ultrasound and tissue quantification using acoustic radiation force impulse technology for differential diagnosis of small thyroid nodules. *Int J Of Clin And Exp Med* (2016) 9 (5):8288–95. doi: 10.1055/s-0042-1757423

27. Ha SM, Cho SW. Virtual touch tissue quantification in the differential diagnosis of benign and malignant thyroid nodules. *J Korean Soc Radiol* (2016) 74(6):365. doi: 10.3348/jksr.2016.74.6.365

28. Ke K, Zhang QX, Wang ZG. Thyroid imaging reporting and data system, virtual tough tissues quantification technique and ceus in differential diagnosis of benign and malignant thyroid nodules. *Chin J Interv Imaging Ther* (2017) 14(5):287–91. doi: 10.13929/J.1672-8475.201610034

29. Zhang YF, Xu HX, He Y, Liu C, Guo LH, Liu LN, et al. Virtual touch tissue quantification of acoustic radiation force impulse: A new ultrasound elastic imaging in the diagnosis of thyroid nodules. *PloS One* (2012) 7(11):e49094. doi: 10.1371/journal.pone.0049094

30. Zhang YL, Yu K, Yang M, Zhang Y, Li JL, Tang J. Comparison of elasticity imaging and virtual touchtm tissue imaging and quantification in the diagnosis of thyroid nodules. *Zhongguo yi xue ke xue yuan xue bao Acta Academiae Medicinae Sinicae* (2019) 41 (1):383–7. doi: 10.3881/j.issn.1000-503X.10772

31. Wu SN, He JD, Jiang TA, Zhong LY, Zhang XF Virtual touch tissue imaging quantification in differential diagnosis of benign and malignant Ti-rads 4 thyroid nodules. *Chin J Ultrasonography* (2016) 25(7):573–8. doi: 10.3760/cma.j.issn.1004-4477.2016.07.006

32. Li DX, Liu QQ, Guo BB, Tian J, Liu Q, Wu CJ. Study on the diagnostic values of virtual touch tissue imaging quantification and virtual touch tissue quantification techniques in differentiating thyroid benign and malignant nodules. *Chin J Med Ultrasound(Electronic Edition)* (2017) 14(10):749–54. doi: 10.3877/cma.j.issn.1672-6448.2017.10.007

33. He Y-P, Xu H-X, Li X-L, Li D-D, Bo X-W, Zhao C-K, et al. Comparison of virtual touch tissue imaging & quantification (Vtiq) and Toshiba shear wave elastography (T-swe) in diagnosis of thyroid nodules: Initial experience. *Clin Hemorheology Microcirculation* (2017) 66(1):15–26. doi: 10.3233/ch-16217

34. Peng Q, Niu C, Zhang M, Peng Q, Chen S. Sonographic characteristics of papillary thyroid carcinoma with coexistent hashimoto's thyroiditis: Conventional ultrasound, acoustic radiation force impulse imaging and contrast-enhanced ultrasound. *Ultrasound Med Biol* (2019) 45(2):471–80. doi: 10.1016/j.ultrasmedbio.2018.10.020

35. Cao DM, Lin MF, Lin DM. Diagnosis of benign and malignant thyroid nodules by acoutic pulsed radiation elastography combined with conventional ultrasound. *Chin J Ultrasound* (2019) 35(7):577–80. doi: CNKI:SUN:ZGCY.0.2019-07-001

36. Deng J, Zhou P, Tian SM, Zhang L, Li JL, Qian Y. Comparison of diagnostic efficacy of contrast-enhanced ultrasound, acoustic radiation force impulse imaging, and their combined use in differentiating focal solid thyroid nodules. *PloS One* (2014) 9(3): e90674. doi: 10.1371/journal.pone.0090674

37. Wang R, Wang XT, Hou XJv. Virtual touch tissue quantification in differential diagnosis of solid thyroid nodules. *Chin J Med Imaging Technol* (2015) 31(10):1506–9. doi: 10.13929/j.1003-3289.2015.10.014

38. Ning CP, Xu LH, Fang SB, Sun YM, Wang JH. Clinical value and impact factors of virtual touch tissue quantification techniques in evaluating benign and malignant thyroid nodules. *Chin J Ultrasonogr* (2014) 23(7):594–7. doi: 10.3760/cma.j.issn.1004-4477.2014.07.013

39. Gu J, Du L, Bai M, Chen H, Jia X, Zhao J, et al. Preliminary study on the diagnostic value of acoustic radiation force impulse technology for differentiating between benign and malignant thyroid nodules. *J Of Ultrasound In Med* (2012) 31(5):763–71. doi: doi :10.7863/jum.2012.31.5.763

40. Wang XY, Chen XX, Ling B, Huang XL, He D, Zhang B. Value of contrast-enhanced ultrasound and acoustic radiation force impulse-imaging in differential diagnosis of benign and malignant thyroid nodules. *Chin J Ultrasound Med* (2016) 32 (8):673–6. doi: 10.3969/j.issn.1002-0101.2016.08.001

41. Hamidi C, Goya C, Hattapoglu S, Uslukaya O, Teke M, Durmaz MS, et al. Acoustic radiation force impulse (Arfi) imaging for the distinction between benign and malignant thyroid nodules. *Radiol Med* (2015) 120(6):579–83. doi: 10.1007/s11547-014-0495-8

42. Bojunga J, Dauth N, Berner C, Meyer G, Holzer K, Voelkl L, et al. Acoustic radiation force impulse imaging for differentiation of thyroid nodules. *PloS One* (2012) 7 (8):e42735. doi: 10.1371/journal.pone.0042735

43. Zhan J, Diao XH, Wan M, Wang HM, Chen Y. The re-evaluation of united stiffness score system for thyroid nodules. *Chin J Ultrasound Med* (2015) 31(11):964–7. doi: CNKI:SUN:ZGCY.0.2015-11-002

44. Chen S, Li Q, Li DD. Acoustic radiation force impulse technology in identification of benign and mlignant solid thyroid nodules. *Chin J Med Imaging Technol* (2014) 30 (5):711–4. doi: 10.13929/j.1003-3289.2014.05.020

45. Hou JX, Wang XT, Wang R, et al. Applicative value of acoustic radiation force impluse imaging in differential diagnosis of small hypoechoic thyroid nodules. *Chin J Ultrasonogr* (2014) 23(1):27–30. doi: 10.3760/cma.j.issn.1004-4477.2014.01.008

46. Zou X, Li QS, Li QS, Chen SH, Luo CR, Xiong HH. Value of acoustic radiation force impulse imaging vtq in the differential diagnosis between benign and malignant thyroid nodules. *Chin J Ultrasound Med* (2014) 30(7):588–91. doi: CNKI:SUN:ZGCY.0.2014-07-005

47. Xu J, Xu X, Xu H, Zhang Y, Zhang J, Guo LH, et al. Conventional us, us elasticity imaging, and acoustic radiation force impulse imaging for prediction of malignancy in thyroid nodules. *Radiology* (2014) 272(2):577–86. doi: 10.1148/radiol.14132438

48. Chen Q, Xing P, Liu Q, Wu C-J. Malignancy risk assessment of category 4 nodules by applying acoustic radiation force impulse in thyroid imaging reporting and data system. *Int J Of Clin And Exp Med* (2018) 11(4):3473–83. doi: 10.2174/1573405619666221115135842

49. Yang Y-P, Xu X-H, Bo X-W, Liu B-J, Guo L-H, Xu J-M, et al. Comparison of virtual touch tissue imaging & quantification (Vtiq) and virtual touch tissue quantification (Vtq) for diagnosis of thyroid nodules. *Clin Hemorheology Microcirculation* (2017) 65 (2):137–49. doi: 10.3233/ch-16142

50. Zhang FJ, Zhao XM, Han RL, Du M, et al. Comparison of acoustic radiation force impulse imaging and strain elastography in differentiating malignant from benign thyroid nodules. *J Ultrasound Med* (2017) 36(12):2533–43. doi: 10.1002/jum.14302

51. Li J, Xu SZ, Du TT, et al. Conventional ultrasound, ultrasonic elastography and acoustic radiation force impulse elastography in the diagnostic tests of benign and malignant thyroid nodules. *Chin Gen Pract* (2015) 18(6):720–3. doi: 10.3969/j.issn.1007-9572.2015.06.028

52. Xu JM, Xu HX, Xu XH, Liu C, Zhang YF, Guo LH, et al. Solid hypo-echoic thyroid nodules on ultrasound: The diagnostic value of acoustic radiation force impulse elastography. *Ultrasound Med Biol* (2014) 40(9):2020–30. doi: 10.1016/j.ultrasmedbio.2014.04.012

53. Zhang FJ, Han RL, Zhao XM. The value of virtual touch tissue image (Vti) and virtual touch tissue quantification (Vtq) in the differential diagnosis of thyroid nodules. *Eur J Radiol* (2014) 83(11):2033–40. doi: 10.1016/j.ejrad.2014.08.011

54. Du Y-R, Ji C-L, Wu Y, Gu X-G. Combination of ultrasound elastography with Ti-rads in the diagnosis of small thyroid nodules (≤10 mm): A new method to increase the diagnostic performance. *Eur J Radiol* (2018) 109:33–40. doi: 10.1016/j.ejrad.2018.10.024

55. Zhang HP, Shi QS, Gu JY, Jiang LY, Bai M, Liu L, et al. Combined value of virtual touch tissue quantification and conventional sonographic features for differentiating benign and malignant thyroid nodules smaller than 10 mm. *J Ultrasound Med* (2014) 33 (2):257–64. doi: 10.7863/ultra.33.2.257

56. Ni JN, Huang PT, Zhang HW, Su N, et al. Diagnostic value of virtual touch tissue quantification in diacriminating thyroid benign and malignant nodule. *Chin J Ultrasonogr* (2013) 22(2):137–40. doi: 10.3760/cma.j.issn.1004-4477.2013.02.015

57. Jung WS, An YY, Ihn YK, Park YH. The diagnostic performance of acoustic radiation force impulse elasticity imaging to differentiate malignant from benign thyroid nodules: Comparison with conventional b-mode sonographic findings. *J Korean Soc Radiol* (2016) 74(2):96. doi: 10.3348/jksr.2016.74.2.96

58. Pandey NN, Pradhan GS, Manchanda A, Garg A. Diagnostic value of acoustic radiation force impulse quantification in the differentiation of benign and malignant thyroid nodules. *Ultrason Imaging* (2017) 39(5):326–36. doi: 10.1177/0161734617706170

59. Grazhdani H, Cantisani V, Lodise P, Di Rocco G, Proietto MC, Fioravanti E, et al. Prospective evaluation of acoustic radiation force impulse technology in the differentiation of thyroid nodules: Accuracy and interobserver variability assessment. *J Ultrasound* (2014) 17(1):13–20. doi: 10.1007/s40477-013-0062-5

60. Tong J, Huang L, Li J, Cao CL, Du TT Virtual touch tissue imaging quantification combined with ceus in differential diagnosis of benign and malignant Ti-rads 4 thyroid nodules. *Chin J Med Imaging Technol* (2020) 36(6):828–33. doi: 10.13929/j.issn.1003-3289.2020.06.006

61. Xu L, Zhou Y, Li Y, Lu B, Liu T Reducing unnecessary biopsy of American college of radiology thyroid imaging reporting and data system category 4 nodules. *J Ultrasound Med* (2020) 40(2):227–36. doi: 10.1002/jum.15391

62. Zhao N, Chen J, Yao M, Han R. The diagnostic value of the different acoustic radiation force impulse technique for solid thyroid nodules with different diameters: A case–control study. *Asian J Surg* (2022) 45(11):2246–52. doi: 10.1016/j.asjsur.2021.12.056

63. Mao F, Xu HX, Zhang SM, Zhang YF, Bo XY, Li XL, et al Study on the application of virtual touch tissue imaging quantification in combination with the 2015 guidelines of American thyroid association for differentiating benign thyroid nodules from malignant thyroid nodules. *Chin Gen Pract* (2016) 19(21):2585–90. doi: 10.3969/j.issn.1007-9572.2016.21.019

64. Xu L, Zhou YB, Xu C, Tian G, Jiang TA The value of virtual touch tissue imaging quantification and real-time elastography techniques in the differentiation of thyroid imaging reporting and data syatem 4 nodules of thyroid. *Chin J Med Ultrasound (Electronic Edition)* (2018) 15(1):53–8. doi: 10.3877/cma.j.issn.1672-6448.2018.01.010

65. Li DX, Chen Q, Liu QQ, Tian J, Wu CJ Study on the diagnostic vaiue of acoustic radiation for impulse imaging combined with acr-tirads in differentiating thyroid benign and malignant noduies Co-exists with hashimoto's thyroiditis. *Chin J Ultrasonic Med* (2019) 35(11):961–4. doi: CNKI:SUN:ZGCY.0.2019-11-034

66. Li X, Hou X-J, Du L-Y, Wu J-Q, Wang L, Wang H, et al. Virtual touch tissue imaging and quantification (Vtiq) combined with the American college of radiology thyroid imaging reporting and data system (Acr Ti-rads) for malignancy risk stratification of thyroid nodules. *Clin Hemorheology Microcirculation* (2019) 72(3):279–91. doi: 10.3233/ch-180477

67. Zhou H, Zhou X-L, Xu H-X, Li D-D, Liu B-J, Zhang Y-F, et al. Virtual touch tissue imaging and quantification in the evaluation of thyroid nodules. *J Ultrasound Med* (2017) 36(2):251–60. doi: 10.7863/ultra.15.12070

68. Sun C-Y, Lei K-R, Liu B-J, Bo X-W, Li X-L, He Y-P, et al. Virtual touch tissue imaging and quantification (Vtiq) in the evaluation of thyroid nodules: The associated factors leading to misdiagnosis. *Sci Rep* (2017) 7(1):41958. doi: 10.1038/srep41958

69. Zhao J, Yang F, Wei X, Mao Y, Mu J, Zhao L, et al. Ultrasound features value in the diagnosis and prognosis of medullary thyroid carcinoma. *Endocrine* (2020) 72(3):727–34. doi: 10.1007/s12020-020-02510-2

70. Zhu J, Zhan J, Diao XH, Wan M. The value of combined stiffness score system for the diagnosis of thyroid micronodules. *Chin J Med Ultrasound(Electronic Edition)* (2015) 12(10):768–72. doi: 10.3877/cma.j.issn.1672-6448.2015.10.006

71. Liu B-J, Xu H-X, Zhang Y-F, Xu J-M, Li D-D, Bo X-W, et al. Acoustic radiation force impulse elastography for differentiation of benign and malignant thyroid nodules with concurrent hashimoto's thyroiditis. *Med Oncol* (2015) 32(3):50. doi: 10.1007/s12032-015-0502-5

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 09 February 2023
DOI 10.3389/fonc.2023.953524



OPEN ACCESS

EDITED BY
Min Wu,
Sichuan University, China

REVIEWED BY
Andrzej Szuba,
Wroclaw Medical University, Poland
Giusto Trevisan,
University of Trieste, Italy

*CORRESPONDENCE
Yanfang Jin
cjr.jinyanfang@vip.163.com
Jun Ma
13159543160@163.com
Xuemei Du
duxuemei@mail.ccmu.edu

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 26 May 2022
ACCEPTED 17 January 2023
PUBLISHED 09 February 2023

CITATION
Li B, Li J, Hao K, Jin Y, Ma J and Du X (2023) Magnetic resonance findings of Stewart–Treves Syndrome in primary limb lymphedema compared with pathology: A retrospective single-center study. *Front. Oncol.* 13:953524. doi: 10.3389/fonc.2023.953524



# Magnetic resonance findings of Stewart–Treves Syndrome in primary limb lymphedema compared with pathology: A retrospective single-center study

Bin Li[1], Jiyuan Li[1], Kun Hao[2], Yanfang Jin[1]*, Jun Ma[3]* and Xuemei Du[4]*

[1]Department of MRI, Beijing Shijitan Hospital, Capital Medical University, Beijing, China, [2]Department of Lymphatic Surgery, Beijing Shijitan Hospital, Capital Medical University, Beijing, China, [3]Department of Radiology, Chuiyangliu Hospital Affiliated to Tsinghua University, Beijing, China, [4]Department of Pathology, Beijing Shijitan Hospital, Capital Medical University, Beijing, China

**Background:** Stewart–Treves Syndrome in Primary Limb Lymphedema (STS-PLE) is an extremely rare malignant tumor. A retrospective analysis was conducted to elucidate the relationship between magnetic resonance imaging (MRI) findings and signs compared to pathology.

**Methods:** Seven patients with STS-PLE were enrolled at Beijing Shijitan Hospital, Capital Medical University, from June 2008 to March 2022. All cases were examined by MRI. The surgical specimens were subjected to histopathological and immunohistochemical staining for CD31, CD34, D2-40, and Ki-67.

**Results:** There were two different types of MRI findings. One was mass shape (STS-PLE I type) in three male patients, and the other was the "trash ice" d sign (STS-PLE II type) observed in four female patients. The average duration of lymphedema (DL) of STS-PLE I type (18 months) was shorter than that of STS-PLE II type (31 months). The prognosis for the STS-PLE I type was worse than that for the STS-PLE II type. Regarding overall survival (OS), the STS-PLE I type (17.3 months) was three times shorter than that of the STS-PLE II type (54.5 months). For STS-PLE I type, the older the STS-PLE onset, the shorter the OS. However, there was no significant correlation in STS-PLE II type. MRI was compared to histological results to provide an explanation for the differences in MR signal changes, especially on T2WI. Against a background of dense tumor cells, the richer the lumen of immature vessels and clefts, the higher the T2WI MRI signal (taking muscle signal as the internal reference standard) and the worse the prognosis, and vice versa. We also found that younger patients with a lower Ki-67 index (<16%) had better OS, especially for the STS-PLE I type. Those with stronger positive expression of CD31 or CD34 had shorter OS. However, the expression of D2-40 was positive in nearly all cases, and seemed not to be associated with prognosis.

**Conclusions:** In lymphedema, the richer the lumen of immature vessels and clefts based on dense tumor cells, the higher the T2WI signal on the MRI. In adolescent patients, the tumor often showed a "trash ice" sign (STS-PLE II-type) and prognosis was better than for the STS-PLE I type. While in middle-aged and older patients,

10.3389/fonc.2023.953524

tumors showed a mass shape (STS-PLE I type). The expression of immunohistochemical indicators (CD31, CD34, and KI-67) correlated with clinical prognosis, especially decreased Ki-67 expression. In this study, we determined it was possible to predict prognosis comparing MRI findings with pathological results.

KEYWORDS

Stewart-Treves Syndrome, MRI, primary, lymphedema, low limb

# 1 Introduction

Stewart–Treves Syndrome (STS) is as an aggressive malignant tumor with a poor prognosis (1). To date, only a few cases have been reported (2). The majority of reported cases were the result of postoperative lymphedema, mainly (90%) after mastectomy (3–5). In clinical practice, STS has often been misdiagnosed and underestimated due to the following factors (6–8): 1) abnormal skin thickening caused by fibrous hyperplasia and hyperkeratosis related to chronic lymphedema, 2) STS lesions mimicking benign tumors with a clear outline, and 3) difficulty distinguishing STS lesions latently distributed in the subcutaneous tissue where lymphedema occurs.

To our knowledge, magnetic resonance imaging (MRI) findings of STS-Primary Limb Lymphedema (STS-PLE) have not been reported, especially in cases involving the lower limb. Moreover, a systematic assessment of STS is lacking. In this retrospective study, a systematic analysis that included clinical and pathological data was performed to investigate MRI findings in the diagnosis and evaluation of STS-PLE.

# 2 Methods

## 2.1 Ethical approval

The study was approved by the Human Research Ethics Committee of the Shijitan Hospital of Capital Medical University of Beijing. Written informed consent was obtained from the individuals(s) for the publication of any identifiable images or data included in the article. All procedures were conducted in accordance with the ethical standards of human experimentation (institutional or regional) and with the Helsinki Declaration of 1975, as revised in 2000. Because this was a retrospective study and the data analysis was performed anonymously, this study was exempted from ethical approval and informed consent from each patient.

## 2.2 Patient population

Between June 2008 and March 2022, a total of 9831 cases of limb lymphedema (age range, 1–78 years; mean ± standard deviation, 48.0 ± 31.7 years) were diagnosed at Beijing Shijitan Hospital. In total, seven STS-PLE patients (age range, 15–66 years; median age 39 years) were clinically and pathologically verified. The following clinical data were collected: duration of lymphedema (LD), specific skin findings (appearance, color, temperature, ulcer and hemorrhagic discharge, and surgical methods), and survival time.

## 2.3 Magnetic resonance imaging

MRI was performed with a 1.5T MRI unit (Ingenia, Philips Medical Systems, The Netherlands) using a 16-channel body coil. The checked limbs were placed in the correct anatomical location and as centered in the coil as possible for the patient. MRI was performed with the following sequences: STIR/axial and coronal planes (TR 5200 ms, TE 80 ms IT 160 ms, Slice 5 mm, GAP 0.5 mm, Voxel 1.5 $mm^2$, FOV 220 mm×200 mm/320 mm×240mm, NEX=2) and mDixon (TR 5.5 ms, $TE_{1\ =}$ 1.8 ms, $TE_{2\ =}$ 4.0 ms, Slice 5 mm, GAP 0, Voxel 1.5 $mm^2$, and FOV 220 mm×200 mm/320 mm×240 mm, NEX=2).

To determine infiltrative growth (6), a supplementary explanation of the "trash ice" sign was established, which refers to the appearance of scattered patches, with intermediately low and heterogeneous signals on T2WI, against the background of lymphedema without recognizable signs of it (6, 7). All MRI images were interpreted by two lymphedema radiologists who had 5 and 10 years of experience, respectively. The MRI findings including location (derma, subcutaneous tissue, and superficial fascia of muscle), shape, outline, and signal intensity of STS-PLE lesions were identified and analyzed, based on the consensus of two radiologists.

## 2.4 Histopathological analysis

Resected surgical specimens were fixed in 10% phosphate-buffered, neutral formaldehyde solution at room temperature for 24 h and dehydrated in an ascending series of ethanol. The samples were routinely embedded in paraffin, made transparent with xylene, rehydrated in a descending series of ethanol, washed with distilled water, and stained with hematoxylin and eosin (HE) for 30 min at room temperature (VENTANA HE600 System, Roche). Sections (4 μm thick) were observed under a light microscope (ZEISS Axio Scope.A1) with magnifications of ×40, ×100, ×200, and ×40.

## 2.5 Immunohistochemistry

Tissue sections (4 μm thick) were deparaffinized, rehydrated, and subjected to antigen retrieval with FLEX High Ph Target Retrieval

solution High Ph (50×) according to the manufacturer's protocol (EnVision FLEX+, Mouse, high Ph (Link) HRP; cat. no. K8002; Dako; Agilent Technologies, Inc., Santa Clara, CA, USA) in PT Link (cat. number PT100; Dako; Agilent Technologies, Inc.) at 95°C for 20 min and washed in distilled water (9). Endogenous peroxidase was blocked by DAKO Envision flex peroxidase blocking reagent for 10 min and washed again three times in PBS wash buffer (Origene Technologies, Inc., Wuxi, China). All slides were incubated for 20–30 min at room temperature in a humidity chamber with appropriate dilutions of primary antibodies (Table 1).

The sections were then incubated with secondary antibody (MA-2000, Origene Technologies, Inc., Wuxi, China) for the coupling reaction for 20–30 min at room temperature. The substrate (EnVision FLEX DAB+ Chromogen) was used to produce a crisp brown color at the target antigen. Hematoxylin (1 or 2 dips) was used as a counter stain.

The sections were observed under a light microscope with magnifications of ×40, ×100, ×200, and ×400. Two senior pathologists independently scored immunohistochemical staining. The number of stained nuclei was expressed as a percentage (index) of positively immunoreactive cells. The proportion of Ki-67 positive cells was counted in at least 1000 tumor cells in the highest labeled area (defined as the hot spot) (10).

# 3 Results

## 3.1 Clinical data

STS-PLE was detected in 0.07% (7/9831) of the study population. The average LD, median survival time (MST), and 5-year survival rate were 25.4 years, 45 months, and 28.6% (2/7), respectively. Of the seven cases (Table 2), the primary locations of STS-PLE were the following: thigh in one case, hand in one case, ankle in two cases, and foot in three cases. The ratio of the left to the right limb was 4:3. Negative skin findings were observed in one young female patient, but positive in the other three female patients. Three male patients presented positive findings of all four items. The ratio of wide excision to amputation was 4:3, in which OS of the former (34.75 months) was nearly 9 months less than the latter (43.67 months).

The male to female ratio was 3:4. The average LD (18 months) of three male patients was shorter than that of four female patients (31 months). The average survival time in male patients (17.3 months) was three times shorter than in female patients (54.5 months). For male patients, the older the STS-PLE onset, the shorter the survival. However, there were no significant differences in female patients (Table 3).

## 3.2 Magnetic resonance findings

Two tumor signs were recognized by MRI: a mass shape (classified into STS-PLE I type) and "trash ice" sign (classified into STS-PLE II type). The average survival time of the STS-PLE I type (Figure 1D) and the STS-PLE II type (Figure 2D) was 17.3 months and 54.5 months, respectively. The outline of the STS-PLE I type and the STS-PLE II type was clear and indistinct, respectively.

There was a recognizable difference in T2WI signals between the STS-PLE I type and the STS-PLE II type. The STS-PLE I type lesions showed a heterogeneous increase in signal intensity in T2WI (Figure 1B), while those of the STS-PLE II type exhibited a slightly or moderately reduced signal in T2WI (Figure 2B). There was no significant difference in the intensity of the signal on T1WI, which presented as a slightly or moderately decreased signal (Figures 1C, 2C). There was no significant difference in the range of invasion, as the tumors invaded more than two locations.

## 3.3 Pathological results

For the STS-PLE I type, the histopathological characteristics were as follows. The tumor was predominantly composed of atypical epithelioid cells displaying hyperchromatism. Spindle endothelial cells lined the immature vessels. The tumor-adjacent tissue was rich in collagen fiber (Figure 1A). The histopathological characteristics of the STS PLE II type were as follows. The atypical epithelioid cells displayed as slit-like anastomosing spaces. The tumor was composed of loosely spindle-shaped epithelioid cells. The tumor showed a lack of vessels. Dysplastic collagen fibers were distributed in a scattered pattern (Figure 2A).

The Ki-67 index of the STS-PLE I type (mean 70%) was nearly three times higher than that of the STS PLE II type (mean 26.75%). Patients presenting a Ki-67 index >40% died within 5 years, in which the shortest survival time of two patients was 2 months and 3 months, respectively. There was no significant difference in the expression of D2-40. The lower the Ki-67 index (< 16%) or the earlier the onset, the longer the survival, especially for the STS-PLE I type. The stronger positive expression of CD31 and CD34 was correlated with a shorter survival time. The expression of D2-40 was positive in almost all cases and the distribution of P53 was scattered and not associated with a prognosis (Table 4).

# 4 Discussion

STS is an extremely rare malignant tumor with poor prognosis. STS based on secondary chronic limb lymphedema is an irreversible

TABLE 1 Primary antibodies used for immunohistochemistry.

| Target | Supplier | Catalog number | Dilution |
|---|---|---|---|
| CD31 | Origene Technologies, Inc. | ZA-0568 | Ready to use |
| CD34 | Origene Technologies, Inc. | ZA-0550 | Ready to use |
| D2-40 | Gene Tech Co., Ltd. | GM361929 | 1:60 |
| Ki67 | Origene Technologies, Inc. | UM870033 | 1:100 |

TABLE 2 Clinical characteristics of the patients with STS-PLE.

| Number | Sex | Age (years) | DL* (years) | Side/ Location | Skin findings† Protuberance | color‡ | temperature ulcer | HD | Surgical method§ | Survival (months) |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Male | 63 | 6 | Right/thigh | 1 | dark red | 1 | 1 | A | 2 |
| 2 | Male | 39 | 25 | Left/foot | 1 | dark red | 1 | 1 | A | 5 |
| 3 | Male | 23 | 23 | Left/hand | 1 | dark red | 1 | 1 | B | 45 |
| 4 | Female | 37 | 37 | Right/foot | 1 | dark red | 1 | 0 | B | 26 |
| 5 | Female | 15 | 6 | Left/ankle | 1 | dark red | 1 | 0 | A | 60 |
| 6 | Female | 21 | 21 | Left/ankle | 0 | 0 | 0 | 0 | A | 72‖ |
| 7 | Female | 66 | 60 | Right/foot | 1 | purple | 1 | 0 | B | 60 ¶ |

* Calculation method in duration of lymphedema: starting point is the year of diagnosis of limb lymphedema, end point is the year of diagnosis of STS, taking the whole year as the counting unit. †, 0, negative; 1, positive. ‡, the color of spot in the skin. §, A represents wide resection; B represents amputation. ||, alive with lung metastasis. ¶, alive with lung metastasis and bone metastasis. HD refers to hematological discharge.

consequence of postoperative care. The causes include infection, filariasis, and injury (described as secondary lymphedema), of which 90% occurs postmastectomy. Although usually associated with mastectomy, the term 'Stewart–Treves syndrome' can be broadly applied. However, STS also develops from primary chronic lymphedema. To date, almost all reports and clinical data published on STS have been associated with secondary lymphedema (STS-SLE), including the average LD. Time before the tumor occurs ranges from 3.25 to 13 years (5, 8), with MST reported as 19 months and a 5-year survival rate of 5–13.6% (5, 8). The incidence of STS varies between 0.07% and 0.45% in the arm (4, 5). However, there have been no previous relevant reports of STS in primary limb lymphedema (STS-PLE). Furthermore, MRI has not been reported in the lower limb.

## 4.1 Clinical data in STS-PLE

In our study, the average survival rate of LD and MST at 5 years was 25.4 years and 45 months in 28.6% (2/7) of patients, respectively. The average LD and MST of STS-PLE were longer than those of STS-SLE (5, 8), by at least two-fold. Interestingly, MST in male patients (17.3 months) was three times shorter than that in female patients (54.5 months). The prognosis for male patients has not previously been described to be worse than that for female patients. The 5-year survival rate of STS-PLE was consistent with previous STS-SLE reports (5, 8). We speculate that it is possible that female patients with a better prognosis than male patients could be positively related with female sex hormone levels, that is, female sex hormone may act as a protective factor.

In our study, the STS-PLE location was found to be significant in the distal part of the limb (6/7). This finding was significantly different from previous STS-SLE reports (3–8). We consider it possible that the distribution trend is related to the two following factors: 1) the reduced lymphatic circulation was poorly correlated with gravity and 2) the frequency of immunodeficiency is higher (11). As a result, mutations in tumor cells should more likely occur in the distal part of the lower extremity.

In terms of skin findings, positive results (skin protuberance, dark red spots, and abnormally increased temperature) were observed in six cases. A young female patient presented with negative skin findings. If skin findings are present in PLE, especially with ulcer or hemorrhagic discharge, the diagnosis of STS should be considered.

The ratio between wide excision and amputation was 4:3, in which the MST of the former (34.75 months) was nearly 9 months less than the latter (43.67 months). In our study, there was a trend showing that the greater the scope of surgical resection or the earlier

TABLE 3 MR findings of the patients with STS-PLE.

| Number | Shape | Outline | T1WI** | T2WI†† | Scope of invasion |
|---|---|---|---|---|---|
| 1 | Mass | Unclear | Moderate | Heterogeneous | 1.2 |
| 2 | Mass | Unclear | Slight | Heterogeneous | 1.2.3 |
| 3 | Mass | Unclear | Slight | Heterogeneous | 1.2.3 |
| 4 | Trash Ice | Unrecognized | Moderate | Intermediate | 1.2 |
| 5 | Trash Ice | Unrecognized | Moderate | Intermediate | 1.2 |
| 6 | Trash Ice | Unrecognized | Moderate | Intermediate | 1.2 |
| 7 | Trash Ice | Unrecognized | Moderate | Intermediate | 1.2 |

**The muscle signal in the same layer was taken as medium signal of the internal reference. †† 1 represents skin and dermis, 2 represents subcutaneous tissue, and 3 represents superficial fascia.

10.3389/fonc.2023.953524

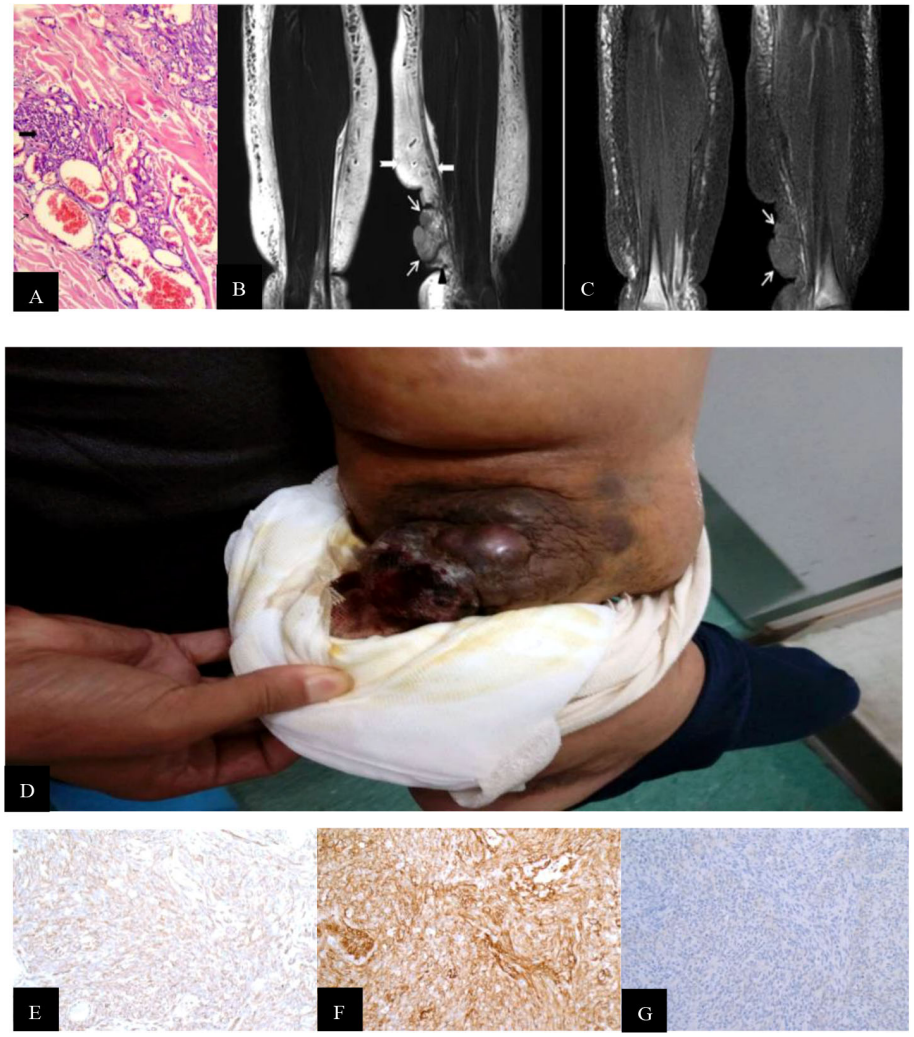



FIGURE 1
**(A)** Hematoxylin-and-eosin-stained histopathology of the M200 STS-PLE I type: the tumor nests were composed predominantly of densely atypical epithelioid cells displaying hyperchromatism (wide black arrowhead). The spindle endothelial cells lined immature vessels, with erythrocytes detected inside the vessels (fine black arrowhead) and distributed in collagen fiber stroma (+). **(B)** Coronal T2WI identified a mass-like shape (fine white arrowhead) located in the subcutaneous soft tissue of the left medial malleolus. Its signal was heterogeneously decreased against the background of lymphedema. If compared with normal muscle, the tumor signals are increased heterogeneously on the T2WI. The extensive lymphedema (swallow-tail white arrowhead) is indicated by an abnormally thickened superficial fascia (wide white arrowhead) and adjacent thickened skin. **(C)** Corresponding T1WI signal of STS was slightly decreased compared with the normal muscle (fine white arrowhead). **(D)** A large mass was located and protruded in the surface of the lower leg, and presented as a dark red superficial ulcer with hemorrhagic discharge, dark-red spots, and increased skin temperature. **(E, F)** CD31 and CD34 staining were positive (magnification, x200). **(G)** D2 40 staining was partially positive (magnification, x200).

the operation, the better the prognosis (3–6). This finding might be helpful for clinical decision-making.

## 4.2 STS-PLE magnetic resonance findings compared with histological results

MRI is a useful tool to establish a modality for exploring the accurate anatomical definition and classification of peripheral lymphedema (6–8). There are few previous imaging reports of STS-SLE evaluated by computed tomography (12) and MRI (8, 13–15). An MRI study of STS-PLE has not been previously reported.

In our study, there were two different types of MRI findings. One was a mass shape (STS-PLE I type) observed in three male patients, and the other was the 'trash ice' sign d (STS-PLE II type) detected in four female patients. There was a discernible shape difference based on sex. The average survival time of STS-PLE I type (17.3 months) was three times shorter than that of STS-PLE II type (54.5 months). For the STS-PLE I type, the older the patient or the later the lymphedema occurred, the shorter the survival. However, there was no significant difference in STS-PLE II type. It seemed that the shape of the tumor on MRI was related to the prognosis. The survival of the STS-PLE I type was worse than that of the STS-PLE II type. Although there was no significant difference in the scope of invasion between the two types, MRI revealed tumor cell infiltration beyond the subcutaneous tissues, including the abnormally thickened dermis and superficial fascia adjacent to the tumor, as well as the disappearance of regular lymphedema signs referring to tumor nests distributed in the subcutaneous tissues (6–8).This finding may be helpful in surgical decision-making.

10.3389/fonc.2023.953524

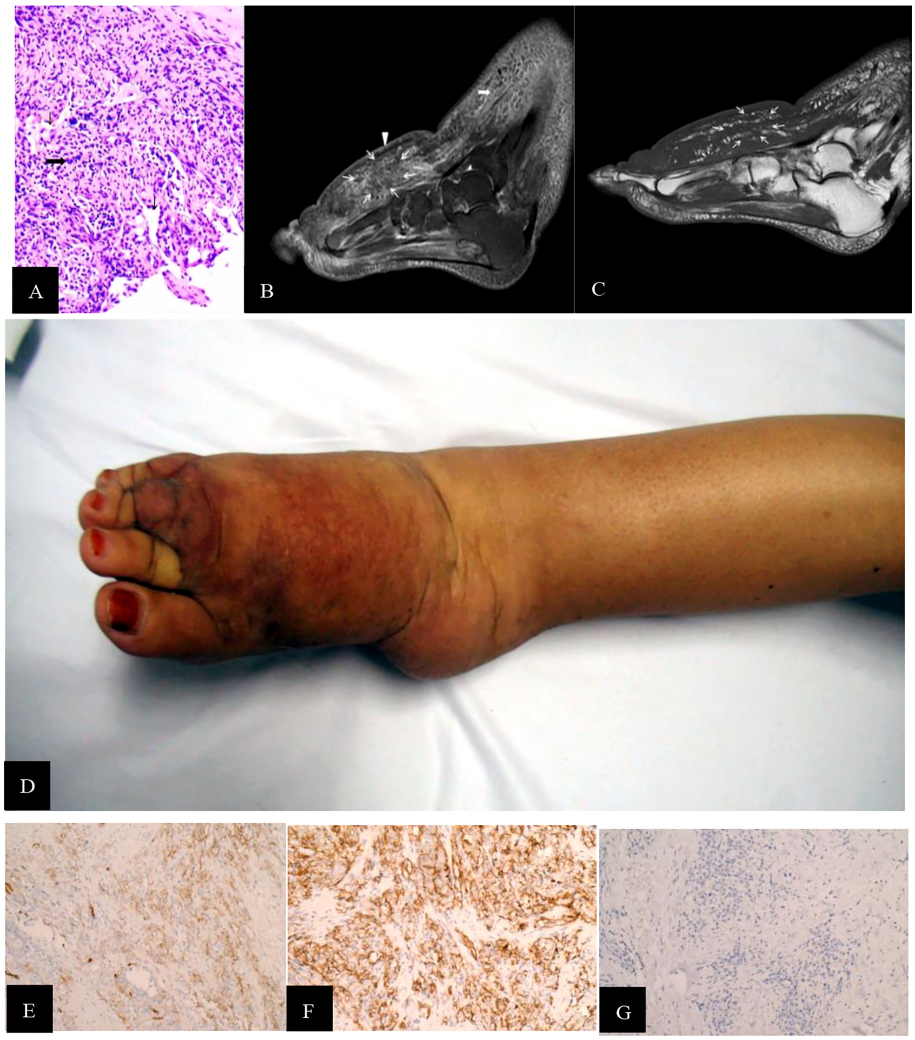



FIGURE 2
**(A)** STS PLE II type (histopathology, M200). Atypical epithelioid present as slit-like anastomosing spaces (fine black arrowhead) and are composed of loose spindle-type epithelioid cells (wide black arrowhead). There is a lack of vessels. The dysplastic collagen fibers show a scattered distribution. The underlying muscle had a normal signal intensity. **(B)** Sagittal T2WI was intermediately and heterogeneously decreased and scattered, revealing a patch-shape in the lymphedema (fine white arrowhead). If compared with the normal muscle tissues, the signal was slightly increased heterogeneously without any recognizable lymphedema signs. One of the regular signs of lymphedema, a Honeycomb sign, presented in the front subcutaneous tissue (wide white arrowhead). **(C)** Corresponding T1WI signal of STS was slightly decreased compared with the normal muscle (fine white arrowhead). **(D)** The dorsal skin and subcutaneous soft tissue of the right foot, which showed dark-red spots, increased skin temperature, and non-ulcer. **(E, F)** CD31 and CD34staining were positive (magnificatin, x200). **(G)** D2-40 staining was partially positive (magnification, x200).

TABLE 4 Immunohistochemical results of the patients with STS-PLE.

| Number | Ki-67 (%) | CD31 | CD34 D2-40 | |
|---|---|---|---|---|
| 1 | 90 | ++ | ++ | – |
| 2 | 90 | ++ | ++ | + |
| 3 | 30 | + | + | + |
| 3 | 50 | + | + | + |
| 4 | 40 | + | + | + |
| 6 | 2 | + | + | + |
| 7 | 15 | + | + | + |

+, focal positive; + +, diffuse positive; - negative.

Stewart and Treves (1) first used the term lymphangiosarcoma (LAS) as an independent pathological diagnostic term. In 2011, when Yu and Yang (16) clearly explained the pathological diagnostic requirements of pure LAS and the mechanisms of angiosarcoma (AS), LAS was no longer used as a diagnostic term. In our study, we found that tumor shape and distribution could be easily recognized by the change in the T2WI signal (Figures 1B, 2B), especially for the STS-PLE II type. The STS-PLE I type signal was heterogeneously decreased based on a background of lymphedema. With reference to signals of normal muscle tissue, the STS-PLE I type signal was intermediately increased heterogeneously on T2WI. The STS-PLE II type signal showed a slight and heterogeneous decrease and a scattered patch shape in lymphedema. With regard to the signal of normal muscle tissue, the STS-PLE II type signal showed a slightly increased heterogeneous in T2WI (6–8).

The histopathological findings revealed STS-PLE tumor cells were predominantly composed of dense atypical epithelioid cells with oval nuclear contours, prominent nucleoli, and eosinophilic cytoplasm in the two types (Figures 1A, 2A). In the STS-PLE II type, collagen fibrils were more lavishly distributed between tumor nests than in the STS-PLE I type (Figure 2A). Otherwise, in the STS-PLE I type, there were many dilated clefts and channels with free erythrocytes. Moreover, several immature vessels in the tumor nests and collagen fibrils were distributed in the STS-PLE II type (Figure 1A). According to the above findings, we considered two factors could explain the correlation of both histopathological results and T2WI signal levels: 1) the higher the tumor cell-density and the more fibrous the stroma were, the lower the T2WI signal intensity, and 2) the richer proportion of immature and small vessels in STS-PLE, the higher T2WI signal intensity. Similar results (8, 13, 14) were found in the previous STS-SLE reports.

### 4.3 Immunohistochemical results and survival time in STS-PLE

CD31 and CD34 were used as specific vascular endothelial markers and D2-40 was used as an available selective marker for the lymphatic endothelium in immunohistochemistry assay. Yu and Yang (16) considered that the typical pure LAS should meet the following criteria: CD31 and CD34 being weakly positive or negative and D2-40 being positive. In our study, CD31 and CD 34 were positive in all cases. This finding meant that no cases were classified as pure LAS. One of seven patients exhibited negative expression of D2-40, and CD31 and CD34 showed strongly positive. We classified it as pure AS. Further, six out of seven cases were classified to a subset of LAS according to the previous reports (6–8, 16–19).

In our study, the expression of CD31 and CD34 was strongly positive in two male patients. In another five cases, there was no significantly different expression of CD31, CD34, and D2-40 (Figures 1E–G, 2E–G). Of all the seven cases, those having a Ki-67 index of more than 40% died within 5 years, in which the shortest survivals reported were 2 months and 5 months. We found that irrespective of STS-PLE I type or STS-PLE II type, all cases shared the following basic features: 1) the lower the Ki-67 index (<16%) and the younger the patient, the longer the survival was, especially in the case of STS-PLE I type; 2) the stronger the positive expression of CD31, CD34, c-myc, and ERG were, the shorter the ST was; and 3) The expression of D2-40 was positive in nearly all cases and the distribution of P53 was scattered, so it could not show association between both types. A similar finding was reported previously (8).

Of course, when considering the prognostic factors of STS-PLE in this study, we only evaluated clinical data: MRI findings, pathological results, and immunohistochemical results. Notably, two other factors are easily ignored in clinical practice (17–20). One is the correlation between the evolution of lymphedema and the reduced surveillance of the immune system. The other is the correlation between the occurrence of local immune deficiency and evolution of STS-PLE. Although we did not include these two factors in our study, we should pay continuous attention to such issues in future research.

## 5 Conclusion

Pathological histopathological findings are closely related to MRI signs. Against the background of dense tumor cells, the richer the lumen of immature vessels and clefts, the higher the T2WI signal on the MRI (taking muscle signal as the internal reference standard), the worse the prognosis, and vice versa. In the adolescent age group, a tumor nest finding often shows the “trash ice” sign (STS-PLE II type) in a PLE, while in middle-aged and older patients, tumor nests show a mass shape (STS-PLE I type). The indicators (CD31, CD34, D2-40, and KI-67) of pathological immunohistochemistry are significantly correlated with clinical prognosis, with Ki-67 being especially prominent.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

## Ethics statement

The study was approved by Human Research Ethics Committee at Beijing Shijitan Hospital of Capital Medical University and the procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional or regional) and with the Helsinki Declaration of 1975, as revised in 2000. Written informed consent from the [patients/participants OR patients/participants legal guardian/next of kin] was not required to participate in this study in accordance with the national legislation and the institutional requirements.

## Author contributions

BL, JL, and KH performed the experiment and collected, analyzed, and drafted the manuscript. JL and YJ interpreted the data involved in the study and preprocessed image data. YJ and JM designed the study and ensured the questions related to all aspects of the work. YJ, JM, and XD contributed equally to this work and share

corresponding authorship. All authors contributed to the article and approved the submitted version.

## Funding

This study was supported by grants from the National Natural Science Foundation of China (No. 61876216), the Beijing Municipal Administration of Hospitals Incubating Program of China (No. PX2020030), and the Special open Project of Lymphatic Surgery (No. 2019-LB04).

## Acknowledgments

The authors are grateful to Pro. Yunlong Song for fruitful discussions.

## 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.

## Publisher's note



## References

1. Stewart FW, Treves N. Lymphangiosarcoma in post-mastectomy lymphedema: A report of six cases in elephantiasis chirurgica. *Cancer* (1948) 1:64–81. doi: 10.1002/1097-0142(194805)1:1<64::AID-CNCR2820010105>3.0.CO;2-W

2. Wang LL, Zhou Q, Cui LF, Shi F, Gao Y, Meng Y, et al. Stewart–Treves syndrome: A clinicopathological analysis of thirteen cases. *J Diag Pathol* (2018) 25:742–5. doi: 10.3969/j.issn.1007-8096.2018.11.002

3. Trattner A, Shamai-Lubovitz O, Segal R, Zelikovski A. Stewart–Treves angiosarcoma of arm and ipsilateral breast in post -traumatic lymphedema. *Lymphology* (1996), 29.

4. Fitzpatrick PJ. Lymphangiosarcoma in breast cancer. *Can J Surg* (1969) 12:172–7.

5. Woodward A, Ivins J, Soule E. Lymphangiosarcoma arising in chronic lymphedematous extremities. *Cancer* (2006) 30:562–72. doi: 10.1002/1097-0142(197208)30:2<562::AID-CNCR2820300237>3.0.CO;2-V

6. Li B, Wang ZC, Shen WB, Chang K, Zuo LL, Zhang J. MR findings of secondary lymphedema in the upper-limb following mastectomy. *J Pract Radio* (2016) l32:1736–9. doi: 10.3969/j.issn.1002-1671.2016.11.022

7. Li B, Wang ZC, Jin YF, Liu HL, Zhang N. The comparative study of MR findings in mild and moderate upper-limb lymphedema. *Chin J Med* (2017) 52:92–8. doi: 10.3969/j.issn.1008-1070.08.027

8. Li B, Wang ZC. Stewart–Treves syndrome: Magnetic resonance imaging data compared with pathological results from a single center. *Oncol Lett* (2018) 15:1113–8. doi: 10.3892/ol.2017.7363

9. Hashmi AA, Hussain ZF, Faridi N, Khurshid A. Distribution of Ki67 proliferative indices among WHO subtypes of non-hodgkin's lymphoma: Association with other clinical parameters. *Asian. Pac. J Cancer. Prev* (2014) 15:8759–63. doi: 10.7314/apjcp.2014.15.20.875 9

10. Klöppel G, Rosa SL. Ki67 labeling index: Assessment and prognostic role in gastroenteropancreatic neuroendocrine neoplasms. *Virchows. Arch* (2018) 472:341–9. doi: 10.1007/s00428-017-22582258-0

11. Ruocco V, Brunetti G, Puca RV and Ruocco E. The immunocompromised district: A unifying concept for lymphoedematous, herpes-infected and otherwise damaged sites. *J Eur Acad Dermatol Venereol* (2009) 23:1364–73. doi: 10.1111/j.1468-3083.2009.03345.x

12. Kazerooni E, Hessler C. CT appearance of angiosarcoma associated with chronic lymphedema. *AJR Am J Roentgenol* (1991) 156:543–4. doi: 10.2214/ajr.156.3.1899753

13. Gelmetti W, Michalik-Himmelmann R, Leppek R, Klose K-J. Lymphedemaassociated angiosarcoma (the Stewart treves syndrome) of the lower extremity: Its MR imaging. *Fortschr Rontgenstr* (1996) 164:168–70. doi: 10.1055/s-2007-1015632

14. Schindera ST, Streit M, Kaelin U, Stauffer E, Steinbach L, Anderson SE. Stewart–Treves syndrome: MR imaging of a postmastectomy upper-limb chronic lymphedema with angiosarcoma. *Skeletal Radiol* (2005) 34:156–60. doi: 10.1007/s00256-004-0807-5

15. Nakazono T, Kudo S, Matsuo Y, Matsubayashi R, Ehara S, Narisawa H, et al. Angiosarcoma associated with chronic lymphedema (Stewart–treves syndrome) of the leg: MR imaging. *Skeletal Radiol* (2000) 29:413–6. doi: 10.1007/s002560000225

16. Yu L, Yang SJ. Lymphangiosarcoma of the vocal cord: A rare entity defned by a D2-40 immunohistochemical and ultrastructural study. *J Clin Oncol* (2011) 29:e57–61. doi: 10.1200/JCO.2010.31.2967

17. Fukunaga M. Expression of D2-40 in lymphatic endothelium of normal tissues and in vascular tumours. *Histopathology* (2005) 46:396–402. doi: 10.1111/j.1365-2559.2005.02098.x

18. Ohsawa M, Naka N, Tomita Y, Kawamori D, Kanno H, Aozasa K. Use of immunohistochemical procedures in diagnosing angiosarcoma. evaluation of 98 cases. *Cancer* (1995) 75:2867–74. doi: 10.1002/1097-0142(19950615)75:12<2867::AID-CNCR2820751212>3.0.CO;2-8

19. Mankey CC, McHugh JB, Thomas DG, Lucas DR. Can lymphangiosarcoma be resurrected? a clinicopathological and immunohistochemical study of lymphatic differentiation in 49 angiosarcomas. *Histopathology* (2010) 56:364–71. doi: 10.1111/j.1365-2559.2010.03484.x

20. Di Meo N, Drabeni M, Gatti A. Trevisan g. a Stewart–treves syndrome of the lower limb. *Dermatol Online J* (2012) 18:14. doi: 10.5070/D38D1463NQ

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 24 February 2023
DOI 10.3389/fonc.2023.1013085



OPEN ACCESS

EDITED BY
Chintan Parmar,
Dana–Farber Cancer Institute,
United States

REVIEWED BY
Nguyen Minh Duc,
Pham Ngoc Thach University of Medicine, Vietnam
Sikandar Shaikh,
Shadan Hospital and Institute of Medical Sciences, India

*CORRESPONDENCE
Sheng Lin
linsheng@swmu.edu.cn
Haowen Pang
haowenpang@foxmail.com

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 06 August 2022
ACCEPTED 13 February 2023
PUBLISHED 24 February 2023

CITATION
Li Y, Gao X, Tang X, Lin S and Pang H (2023) Research on automatic classification technology of kidney tumor and normal kidney tissue based on computed tomography radiomics.
*Front. Oncol.* 13:1013085.
doi: 10.3389/fonc.2023.1013085



# Research on automatic classification technology of kidney tumor and normal kidney tissue based on computed tomography radiomics

Yunfei Li†, Xinrui Gao†, Xuemei Tang†, Sheng Lin* and Haowen Pang*

Department of Oncology, The Affiliated Hospital of Southwest Medical University, Luzhou, China

**Purpose:** By using a radiomics-based approach, multiple radiomics features can be extracted from regions of interest in computed tomography (CT) images, which may be applied to automatically classify kidney tumors and normal kidney tissues. The study proposes a method based on CT radiomics and aims to use extracted radiomics features to automatically classify of kidney tumors and normal kidney tissues and to establish an automatic classification model.

**Methods:** CT data were retrieved from the 2019 Kidney and Kidney Tumor Segmentation Challenge (KiTS19) in The Cancer Imaging Archive (TCIA) open access database. Arterial phase-enhanced CT images from 210 cases were used to establish an automatic classification model. These CT images of patients were randomly divided into training (168 cases) and test (42 cases) sets. Furthermore, the radiomics features of gross tumor volume (GTV) and normal kidney tissues in the training set were extracted and screened, and a binary logistic regression model was established. For the test set, the radiomic features and cutoff value of P were consistent with the training set.

**Results:** Three radiomics features were selected to establish the binary logistic regression model. The accuracy (ACC), sensitivity (SENS), specificity (SPEC), area under the curve (AUC), and Youden index of the training and test sets based on the CT radiomics classification model were all higher than 0.85.

**Conclusion:** The automatic classification model of kidney tumors and normal kidney tissues based on CT radiomics exhibited good classification ability. Kidney tumors could be distinguished from normal kidney tissues. This study may complement automated tumor delineation techniques and warrants further research.

KEYWORDS

**computed tomography (CT), radiomics, kidney, kidney tumor, automatic classification**

# Introduction

Radiomics refers to the high-throughput extraction of a large amount of information from medical images, such as computed tomography (CT), magnetic resonance imaging (MRI), and positron emission computed tomography (PET), to extract features and establish models. In general, visual image information is converted into digital feature variables for quantitative research. Extensive research and analysis of massive image data information can assist doctors in providing a more accurate diagnosis of a patient's condition (1–3). Compared to biopsy, radiomics has the technical advantage of obtaining non-invasive and repeatable radiological images, thus providing a safer and more way for conducting patient follow-ups and prognosis prediction. Radiomics method can also be used for pathological tumor classification and grading (4, 5). Traditional radiomics, which involves the extraction and screening of high-throughput features of regions of interest from medical images, is used primarily for the diagnosis of benign and malignant diseases, prognosis evaluation, and survival prediction (6–13). Radiomics combined with machine learning and deep learning performs well for differentiating, grading and staging kidney tumors (14–16).

Traditional classification methods using computed tomography (CT) images mainly rely on the pixel value of the image and less on other parameters; this makes it difficult to accurately distinguish the tumor area from the surrounding normal organs (17–22). However, radiomics-based methods can be used to extract more than 800 radiomics features from regions of interest in CT images. In the present study, we proposed a radiomics research method based on the automatic classification technology of radiomics which has the potential to supplement the deep learning automatic delineation technology. This study is the first to report such a method. Therefore, the purpose of this study is to establish a preliminary classification model based on CT radiomics to automatically classify kidney tumors and normal kidney tissues.

# Materials and methods

## Data collection

The study was reviewed and approved by Ethics Committee of the Affiliated Hospital of Southwest Medical University (18 January 2017, KY2021023). CT data were retrieved from the 2019 Kidney and Kidney Tumor Segmentation Challenge (KiTS19) (https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=61081171) of The Cancer Imaging Archive (TCIA) open access database. Arterial phase-enhanced images of 210 patients from the database were used to establish an automatic classification model. The kidneys and tumors were already manually segmented. These CT images were randomly divided into a training set (168 cases) and test set (42 cases).

## Radiomics feature extraction

The CT images were preprocessed using wavelet-based methods. Before feature extraction, all images were resampled according to a voxel size of $1 \times 1 \times 1$ mm$^3$. The gross tumor volume (GTV) and normal kidney tissue were regarded as the regions of interest. Feature extraction was based on a three-dimensional (3D) slicer platform and performed using the pyradiomics package; the package is available at http://PyRadiomics.readthedocs.io/en/latest/ (last accessed on June 30, 2019). The eigenvalue data of all radiomic features were processed using z-score standardization. Figure 1 shows a CT sectional view of a patient in the training set. The CT radiomic feature variables of the GTV and normal kidney tissue were extracted. A total of 837 radiomics features were extracted, including first-order statistics, gray level co-occurrence matrix (GLCM), gray level dependence matrix (GLDM), gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), and neighboring gray tone difference matrix (NGTDM). Shape features were removed in this study. First-order features describes single pixel or voxel within the ROI. GLCM defines different combination of gray levels of an image area. GLDM quantifies the gray level dependencies in an image. GLRLM provides information about runs of consecutive pixels with the same gray level. GLSZM quantifies gray level zones in an image. And wavelet-based features were transformed based on above features.

## Screening of radiomics features of the training set

### (A) Univariate feature screening

By considering the radiomics features of the GTV and normal kidney tissue as independent variables and the GTV and normal kidney tissue as binary variables, the area under curve (AUC; defined as the area surrounded by the receiver operation characteristic (ROC) curve and the abscissa and ordinate axis) corresponding to each radiomics feature was calculated. The AUC ranged from 0.5 to 1. The closer the value is to 1.0, the higher the authenticity, whereas an AUC value equal to 0.5 indicates that the radiomics have no application value in the study. The ROC curve takes the false positive rate (FPR) as the abscissa and true positive rate (TPR) as the ordinate. The curve is mostly used for the evaluation of binary classification problems. Radiomics features with an AUC less than 0.7 were excluded after the univariate screening.

### (B) LASSO logistic regression feature screening

The principle of least absolute shrinkage and selection operator (LASSO) regression is to compress the original eigenvalue coefficients; more specifically, it involves directly compressing the original small coefficients to 0 and treating the eigenvalue variables corresponding to these coefficients as nonsignificant variables. Such

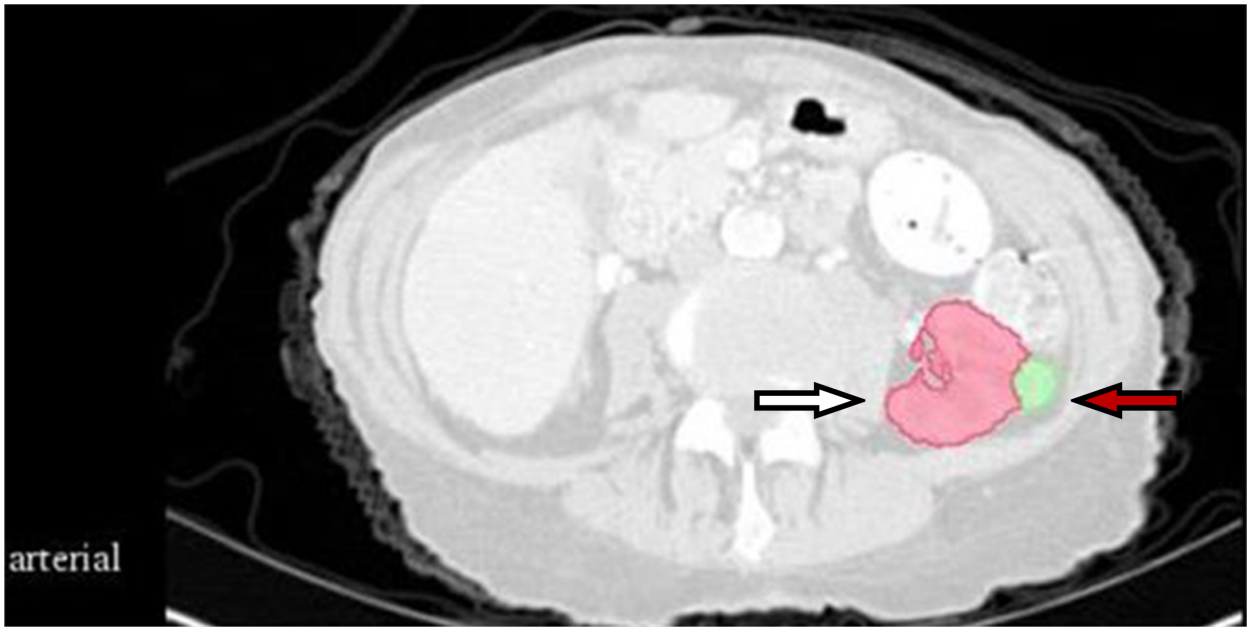



FIGURE 1
CT section of one patient in the training set. The white arrow indicates normal kidney tissue, and the red arrow indicates gross tumor volume.

nonsignificant variables have little or no impact on the final classification results; thus, such variables can be directly discarded, which results in variable screening. LASSO logistic regression was conducted using the method of five-fold cross-validation method to select the radiomics features. In the LASSO regression analysis, the L1 regularization term is added based on the least-squares fit to improve the accuracy of the linear regression model. Its penalty function is the absolute value of the regression coefficient, which guarantees that the parameter estimation results equal to zero. Thus, it is helpful for feature selection. This study is a binary classification problem, and logistic regression analysis is a generalized linear model commonly used in binary classification or one-to-many classifications. It normalizes the response of simple linear regression to zero and one. Therefore, the linear regression in the LASSO regression model can be replaced by logistic regression to select the characteristics. The objective function of LASSO logistic regression optimization is as follows:

$$\min_{\omega,b} \sum_{i=1}^{n} \log\left(\exp\left(-y_i\left(X_i^T\omega + b\right)\right) + 1\right) + \lambda \parallel \omega \parallel_1$$

where n is the total number of samples; Xi is an m × n-size raw data (each sample has m eigenvalues); yi is the corresponding response value of each sample; ω is the linear regression coefficient; b is the cutoff value of linear regression; and λ is a nonnegative regularization parameter used to control the sparsity of regression coefficients.

The extracted radiomics features were input into the LASSO logistic regression model; subsequently, the lambda (λ) value with the smallest model deviation was calculated and the radiomics features were screened.

### (C) Model collinearity detection

The variance expansion factor (VIF) of the independent variable of the logistic regression model was calculated after screening the variables using the LASSO logistic regression. The VIF measures the severity of multicollinearity in multiple regression models. This represents the ratio of the variance of the estimator of the regression coefficient to the variance when no linear correlation between the independent variables is assumed.

The VIF can be calculated as follows.

$$\mathrm{VIF} = \frac{1}{1 - R_i^2},$$

where Ri is the negative correlation coefficient of the regression analysis for other independent variables. The larger the VIF, the greater the possibility of collinearity between the independent variables. Generally, multicollinearity is assumed when the VIF value is greater than five; thus, removing the radiomics features with a VIF value greater than five is necessary.

## Model establishment

The final binary logistic regression model was established based on the final radiomics features (X):

$$\mathrm{logitP} = g(x_i),$$

$$\mathrm{P} = 1/\left(1 + \exp(-\mathrm{logitP})\right),$$

where P is the probability that GTV is positive. The ROC curve of the model was plotted and its AUC value was calculated.

The sensitivity (SENS) and specificity (SPEC) corresponding to each point on the ROC curve were used to calculate the point that maximized SENS + SPEC, which is the cutoff value of P.

## Model diagnosis

The accuracy (ACC), SENS, SPEC, AUC, and the Youden index of the model were used for evaluating the effectiveness of the model. The Youden index, also known as the correct index, was used to evaluate the authenticity of screening tests. It is calculated as Youden index=SENS+ SPEC -1.

## Model validation

For the test set, the radiomics features extracted concerning GTV and normal kidney tissue were consistent with those of the training set. The cutoff value of P in the test set is also consistent with the training set. If the GTV of the test set is positive, the ROC curve is drawn to calculate the AUC value. Thereafter, the ACC, SENS, SPEC, AUC, and Youden index of the test model were again calculated. The flowchart of the proposed method is shown in Figure 2.

All statistical analyses were performed using R software, version 4.1.2 (R Foundation for Statistical Computing, Vienna, Austria).

# Results

Altogether, 837 radiomics features were extracted. After the univariate screening, 217 radiomic features were identified. Using LASSO logistic regression for variable screening, according to the calculation, the deviation of the model was the smallest when the minimum value was 0.1715; Figure 3 shows the model deviation and lambda. Finally, three radiomics features were extracted: dependence entropy of the GLDM of the original (Feature 1), zone entropy of GLSZM of the wavelet-HLL (H = high-frequency band, L = low-frequency band; Feature 2), and gray level non-uniformity of GLSZM of the wavelet-LLL (Feature 3). Table 1 lists the VIF values corresponding to these features. The VIFs were all less than five, which indicates that no multicollinearity existed among the three radiomics features.

The final binary logistic regression model was established based on the final radiomics features as follows.

$$\text{logitP} = -0.0156(\text{Feature 1}) - 0.0523(\text{Feature 2}) - 0.0004(\text{Feature 3})$$

$$P = 1/(1 + \exp(-\text{logitP})),$$

where P is the probability that GTV is positive. The cutoff value in this study was 0.4851. It was used as a radiomics marker to determine the tumor area and was equivalent to the critical point. If the detected value is greater than the cutoff value, it is the GTV; if the detected value is less than the cutoff value, it is the normal kidney tissue.

Table 2 lists the diagnostic parameters of the training set model, and Figure 4 shows the ROC curve of the training set model.

Table 3 lists the model diagnostic parameters of the test set, and Figure 5 shows the ROC curve of the test set model.

# Discussion

Confirming the GTV is a challenge when treating a tumor. It tests the patience and proficiency of clinicians in relevant clinical

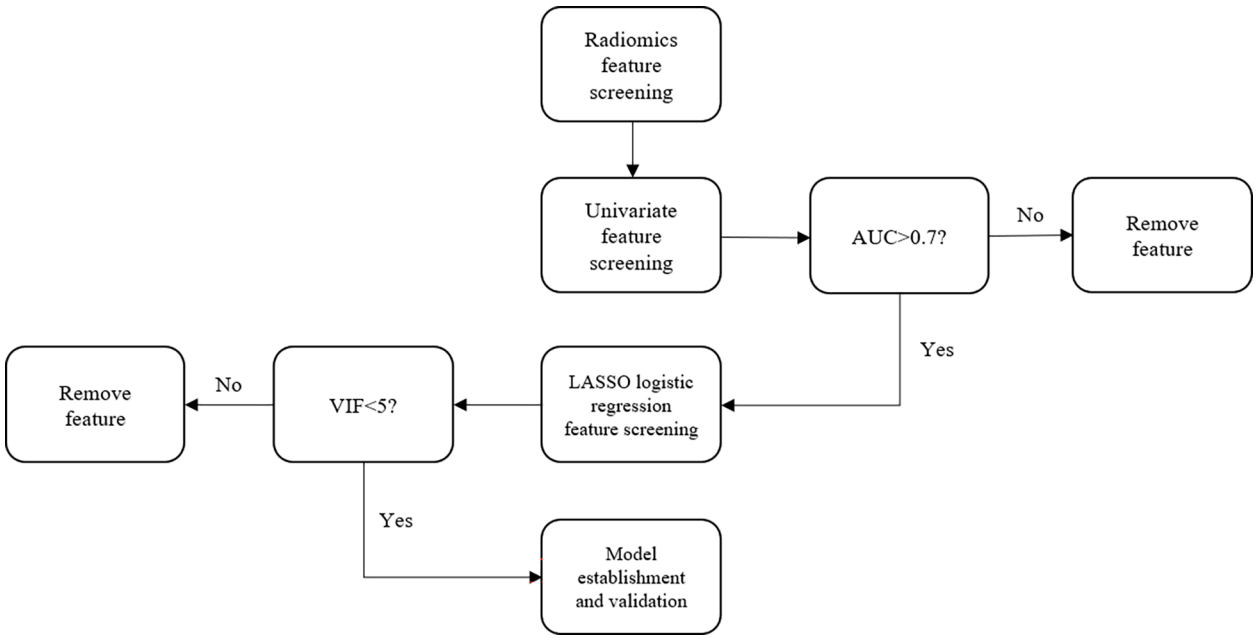



FIGURE 2
Flowchart of the proposed method.

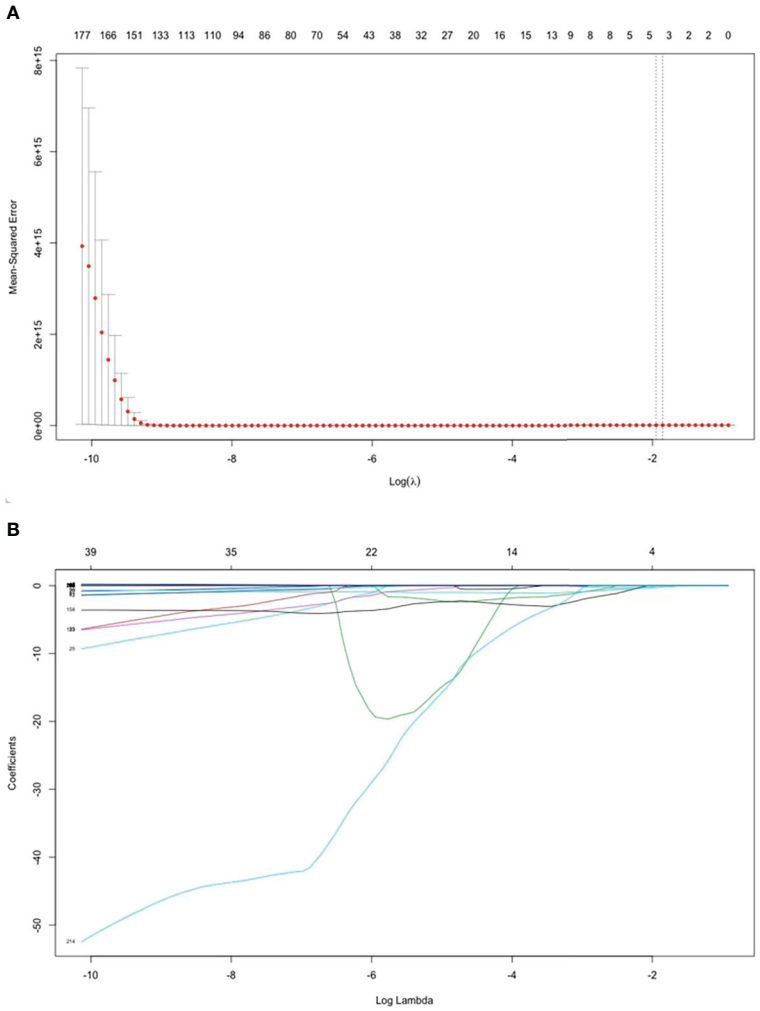



FIGURE 3
Features selected by the LASSO regression model. **(A)** Selection of the tuning parameter (*λ*) in the LASSO model *via* 5-fold cross-validation. The optimal *λ* are indicated by the dotted vertical lines, and a value of 0.1715 was selected. **(B)** LASSO coefficient profiles of 217 radiomics features. A coefficient profile plot was generated versus the selected log λ value using five-fold cross-validation. Three radiomics features with non-zero coefficients were selected.

knowledge. Furthermore, tumors of different shapes are difficult to delineate. A qualified clinician must have systematic learning and continuous practice to be competent in sketching. In this study, based on the CT radiomics method, the extracted CT radiomics features were used to automatically distinguish kidney tumor areas from normal kidney tissues.

The cutoff value, which is the critical point, was used as a radiomics marker for determining the GTV. If the detection value was greater than the cutoff value, it was considered positive; if the detection value was less than the cutoff value, it was considered negative. In this study, a positive detection value represented the kidney tumor part, and a negative value represented normal tissues and organs. The AUC values of the training and test sets obtained from the ROC curve were 0.9798 and 0.9841, respectively, which were significantly greater than 0.7. The ACC, SENS, SPEC, and Youden indices for the training and test sets based on the CT radiomics classification model were all greater than 0.85. This indicates that automatic classification technology based on radiomics had achieved good application results. Further studies need to extract only selected radiomic features from the GTV and kidney training sets instead of all radiomic features to improve efficiency.

TABLE 1 Model collinearity analysis.

| Features | VIF |
|---|---|
| Feature 1 | 1.125322 |
| Feature 2 | 1.052620 |
| Feature 3 | 1.084162 |

In recent years, artificial intelligence technologies such as machine learning and neural networks have been widely used in the field of automatic tumor and organ delineation in radiotherapy, such as automatic segmentation technology based on convolutional neural network and automatic classification technology based on the U-NET model. These have greatly reduced clinician workload and increased productivity (23–26). Artificial intelligence models have achieved great success in the automatic delineation of organs, but the

TABLE 2 Training set model diagnosis parameters.

| Cutoff (Predicted P value) | ACC | SENS | SPEC | AUC | Youden |
|---|---|---|---|---|---|
| 0.4851 | 0.9405 | 0.9583 | 0.9226 | 0.9798 | 0.881 |

accuracy of the automatic delineation of tumor regions is still a problem. For future research, we can limit the region of interest to the entire kidney region, grid it, and extract only the radiomics features filtered through the training set to improve efficiency. A binary logistic regression model can be established based on the final radiomics features extracted from the training set, and the cutoff value calculated from the training set can be used as a radiomics marker for classifying tumor regions. When the P value is greater than the Cutoff value, the grid is considered to be a tumor region; when the P value is less than the Cutoff value, the grid is considered to be a normal organ. After clustering the tumor region or substituting P-values for pixel values, it is possible to automatically delineate the tumor region, which may be a supplement to the automatic delineation technology of deep learning.

In conclusion, automatic classification technology based on radiomics can be feasibly applied to distinguish between GTV and normal kidney tissue in patients. Nevertheless, our study has a few limitations. First, this study used a limited number of samples. Second, the establishment and optimization of the model were affected by the quality of the CT images and the accuracy of manual tumor segmentation. Third, different clinicians have a different understanding of GTV boundaries, and some CT images still have problems such as a fuzzy boundary of the tumor target area. Forth, the pre-malignant state and benign kidney lesions should be

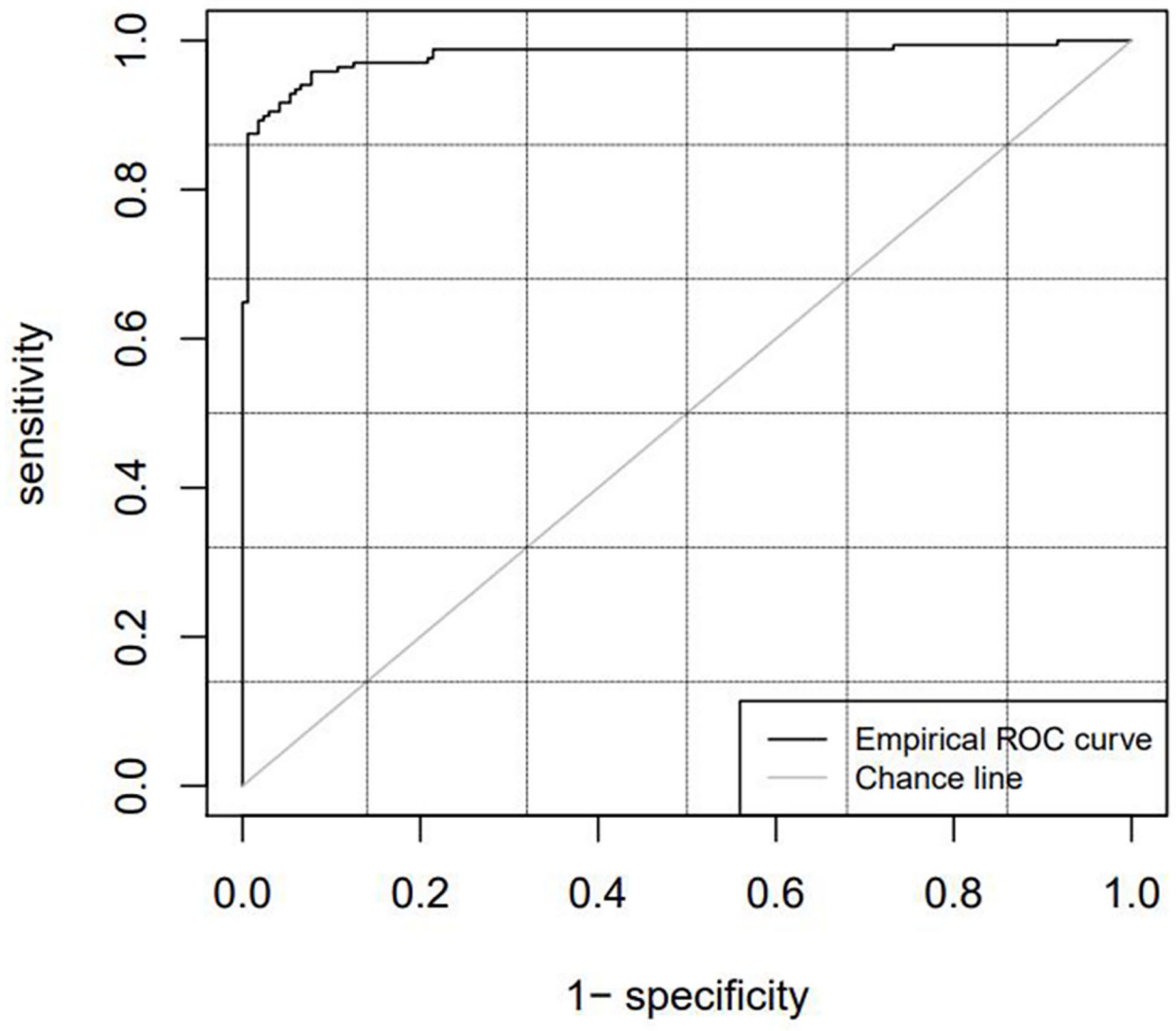



FIGURE 4
ROC curve of the training set.

TABLE 3 Test set model diagnosis parameters.

| Cutoff(Predicted P value) | ACC | SENS | SPEC | AUC | Youden |
|---|---|---|---|---|---|
| 0.4851 | 0.9286 | 0.9285 | 0.9285 | 0.9841 | 0.8571 |

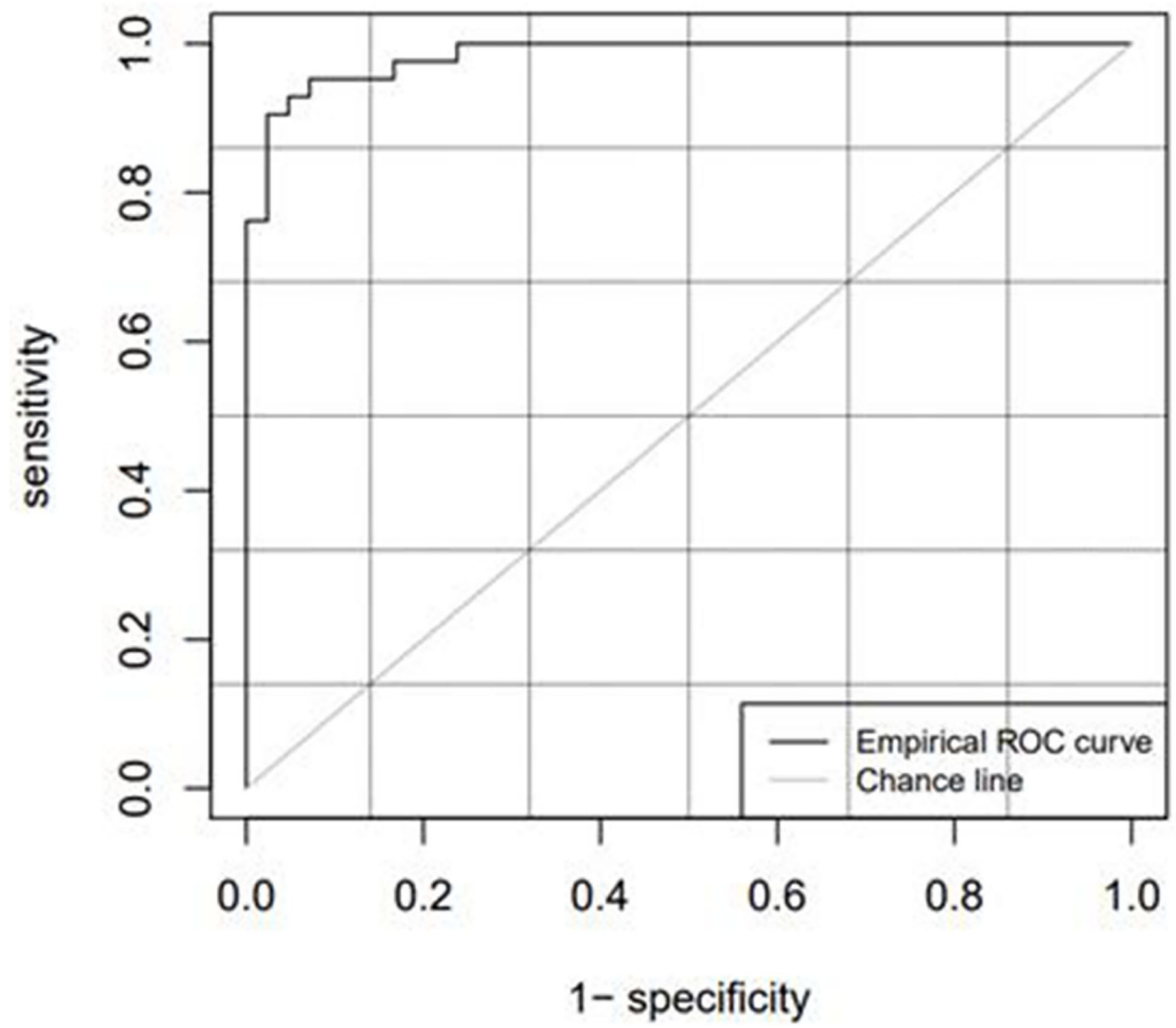



FIGURE 5
ROC curve of the test set.

analyzed and differentiated. Thus, the data may have biases. For future clinical application, we plan to train and test CT image data of more patients with different tumors, to obtain better automatic classification results. If the technology is further matured and developed, it may reduce the risk of inaccurate drawing due to a lack of experience, save valuable processing time, and benefit doctors, patients, and radiotherapy technology.

Radiomics methods are generally used in the diagnosis of benign and malignant diseases, prognosis evaluation, and survival prediction. The preliminary method of automatic classification technology based on radiomics proposed in this study aims to enrich research on radiomics and may be to supplement to deep-learning automatic rendering technology to realize more accurate GTV determination. Future research should focus on different diseases and increase the number of samples to further improve the accuracy of this automatic classification technology.

## Conclusion

The automatic classification model of kidney tumors and normal kidney tissue based on CT radiomics exhibited good classification ability. Tumorous kidney tissues could be distinguished from normal kidneys, with these observations worthy of further study.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

## Author contributions

HP and SL conceived the proposed concept and verified the underlying data. HP, YL, XG and XT took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis, and manuscript. All authors contributed to the article and approved the submitted version.

## Funding

We acknowledge funding from the Gulin County People's Hospital, Southwest Medical University Affiliated Hospital Science and Technology Strategic Cooperation Project (project number 2022GLXNNYDFY05). We acknowledge funding from the Sichuan Medical Association Scientific Research Project (project number S21005).

## 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.

## Publisher's note



## References

1. Avanzo M, Stancanello J, Pirrone G, Sartor G. Radiomics and deep learning in lung cancer. *Strahlenther Onkol* (2020) 196(10):879–87. doi: 10.1007/s00066-020-01625-9

2. Tagliafico AS, Piana M, Schenone D, Lai R, Massone AM, Houssami N. Overview of radiomics in breast cancer diagnosis and prognostication. *Breast* (2020) 49:74–80. doi: 10.1016/j.breast.2019.10.018

3. Zhou J, Zhang Y, Chang KT, Lee KE, Wang O, Li J, 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(3):798–809. doi: 10.1002/jmri.26981

4. E L, Lu L, Li L, Yang H, Schwartz LH, Zhao B. Radiomics for classification of lung cancer histological subtypes based on nonenhanced computed tomography. *Acad Radiol* (2019) 26(9):1245–52. doi: 10.1016/j.acra.2018.10.013

5. Wang S, Wei Y, Li Z, Xu J, Zhou Y. Development and validation of an mri radiomics-based signature to predict histological grade in patients with invasive breast cancer. *Breast Cancer (Dove Med Press)* (2022) 14:335–42. doi: 10.2147/bctt.S380651

6. Hao W, Gong J, Wang S, Zhu H, Zhao B, Peng W. Application of mri radiomics-based machine learning model to improve contralateral bi-rads 4 lesion assessment. *Front Oncol* (2020) 10:531476. doi: 10.3389/fonc.2020.531476

7. Zheng X, Yao Z, Huang Y, Yu Y, Wang Y, Liu Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun* (2020) 11(1):1236. doi: 10.1038/s41467-020-15027-z

8. Wu K, Shui Y, Sun W, Lin S, Pang H. Utility of radiomics for predicting patient survival in hepatocellular carcinoma with portal vein tumor thrombosis treated with stereotactic body radiotherapy. *Front Oncol* (2020) 10:569435. doi: 10.3389/fonc.2020.569435

9. Wu Q, Wang S, Chen X, Wang Y, Dong L, Liu Z, et al. Radiomics analysis of magnetic resonance imaging improves diagnostic performance of lymph node metastasis in patients with cervical cancer. *Radiother Oncol* (2019) 138:141–8. doi: 10.1016/j.radonc.2019.04.035

10. Wang T, Gao T, Yang J, Yan X, Wang Y, Zhou X, et al. Preoperative prediction of pelvic lymph nodes metastasis in early-stage cervical cancer using radiomics nomogram developed based on T2-weighted mri and diffusion-weighted imaging. *Eur J Radiol* (2019) 114:128–35. doi: 10.1016/j.ejrad.2019.01.003

11. Dercle L, Lu L, Schwartz LH, Qian M, Tejpar S, Eggleton P, et al. Radiomics response signature for identification of metastatic colorectal cancer sensitive to therapies targeting egfr pathway. *J Natl Cancer Inst* (2020) 112(9):902–12. doi: 10.1093/jnci/djaa017

12. Chang N, Cui L, Luo Y, Chang Z, Yu B, Liu Z. Development and multicenter validation of a ct-based radiomics signature for discriminating histological grades of pancreatic ductal adenocarcinoma. *Quant Imaging Med Surg* (2020) 10(3):692–702. doi: 10.21037/qims.2020.02.21

13. Liza H, Anais S, Julia S, Florent C, Isabelle B-A, Jean Marc C, et al. Radiomics method for the differential diagnosis of radionecrosis versus progression after fractionated stereotactic body radiotherapy for brain oligometastasis. *Radiat Res* (2020) 193(5):471–80. doi: 10.1667/RR15517.1

14. Weszka JS. A survey of threshold selection techniques. In: *Computer graphics and image processing*, vol. 7. (1978). p. 259–65. doi: 10.1016/0146-664X(78)90116-8

15. Al-Amri SS, Kalyankar N, Khamitkar S. Image segmentation by using edge detection. *Int J Comput Sci Eng* (2010) 2(3):804–7. doi: 10.1016/j.ijleo.2013.10.049

16. Lu X, Wu J, Ren X, Zhang B, Li Y. The study and application of the improved region growing algorithm for liver segmentation. *Optik* (2014) 125(9):2142–7. doi: 10.1016/j.ijleo.2013.10.049

17. Uhlig J, Leha A, Delonge LM, Haack AM, Shuch B, Kim HS, et al. Radiomic features and machine learning for the discrimination of renal tumor histological subtypes: A pragmatic study using clinical-routine computed tomography. *Cancers (Basel)* (2020) 12(10):3010. doi: 10.3390/cancers12103010

18. Hussain MA, Hamarneh G, Garbi R. Learnable image histograms-based deep radiomics for renal cell carcinoma grading and staging. *Comput Med Imaging Graph* (2021) 90:101924. doi: 10.1016/j.compmedimag.2021.101924

19. Zhao Y, Chang M, Wang R, Xi IL, Chang K, Huang RY, et al. Deep learning based on mri for differentiation of low- and high-grade in low-stage renal cell carcinoma. *J Magn Reson Imaging* (2020) 52(5):1542–9. doi: 10.1002/jmri.27153

20. Cabezas M, Oliver A, Llado X, Freixenet J, Cuadra MB. A review of atlas-based segmentation for magnetic resonance brain images. *Comput Methods Programs BioMed* (2011) 104(3):e158–77. doi: 10.1016/j.cmpb.2011.07.015

21. Aljabar P, Heckemann RA, Hammers A, Hajnal JV, Rueckert D. Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. *Neuroimage* (2009) 46(3):726–38. doi: 10.1016/j.neuroimage.2009.02.018

22. Sharp G, Fritscher KD, Pekar V, Peroni M, Shusharina N, Veeraraghavan H, et al. Vision 20/20: Perspectives on automated image segmentation for radiotherapy. *Med Phys* (2014) 41(5):050902. doi: 10.1118/1.4871620

23. Mohammadi R, Shokatian I, Salehi M, Arabi H, Shiri I, Zaidi H. Deep learning-based auto-segmentation of organs at risk in high-dose rate brachytherapy of cervical cancer. *Radiother Oncol* (2021) 159:231–40. doi: 10.1016/j.radonc.2021.03.030

24. Ahn SH, Yeo AU, Kim KH, Kim C, Goh Y, Cho S, et al. Comparative clinical evaluation of atlas and deep-Learning-Based auto-segmentation of organ structures in liver cancer. *Radiat Oncol* (2019) 14(1):213. doi: 10.1186/s13014-019-1392-z

25. Vrtovec T, Močnik D, Strojan P, Pernuš F, Ibragimov B. Auto-segmentation of organs at risk for head and neck radiotherapy planning: From atlas-based to deep learning methods. *Med Phys* (2020) 47(9):e929–e50. doi: 10.1002/mp.14320

26. Zhang X, Liang M, Yang Z, Zheng C, Wu J, Ou B, et al. Deep learning-based radiomics of b-mode ultrasonography and shear-wave elastography: Improved performance in breast mass classification. *Front Oncol* (2020) 10:1621. doi: 10.3389/fonc.2020.01621

frontiers | Frontiers in Oncology

TYPE Systematic Review
PUBLISHED 05 April 2023
DOI 10.3389/fonc.2023.1029936



OPEN ACCESS

EDITED BY
Min Tang,
Jiangsu University, China

REVIEWED BY
Weiwei Zhan,
Shanghai Jiaotong University School of Medicine, China
Xiao-Yan Zhang,
Chinese Academy of Medical Sciences and Peking Union Medical College Hospital, China

*CORRESPONDENCE
Lixue Yin
yinlixue_cardiac@163.com

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 28 August 2022
ACCEPTED 27 March 2023
PUBLISHED 05 April 2023

CITATION
Luo H and Yin L (2023) Diagnostic value of superb microvascular imaging and color doppler for thyroid nodules: A meta-analysis.
*Front. Oncol.* 13:1029936.
doi: 10.3389/fonc.2023.1029936



# Diagnostic value of superb microvascular imaging and color doppler for thyroid nodules: A meta-analysis

Haorou Luo[1] and Lixue Yin[2]*

[1]Chengdu Women's and Children's Central Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, Sichuan, China, [2]Sichuan Academy of Medical Sciences, Sichuan Provincial People's Hospital, School of Medicine University of Electronic Science and Technology of China (UESTC), Chengdu, Sichuan, China

**Objective:** Superb micro-vascular imaging (SMI) is a new noninvasive modality for the diagnosis of thyroid nodules. However, the performance of SMI in differentiating malignant and benign thyroid nodules has not been systematically evaluated. This meta-analysis was performed to assess the accuracy of SMI in diagnosing thyroid nodules.

**Methods:** PubMed, Cochrane Library, Embase, Web of Science, Sinomed, Scopus were searched. We recorded the characteristics of the included studies and assessed the quality of each study using the QUADAS-2 tool. The pooled sensitivity, specificity, positive likelihood ratio (LR), negative LR, diagnostic odds ratio (DOR), and area under the curve (AUC) were calculated. We also evaluated the publication bias.

**Results:** This meta-analysis included 10 studies with a total of 1083 thyroid nodules. The pooled the sensitivity, specificity, and positive and negative LR were 0.84, 0.86, 6.2, and 0.18, respectively. The DOR and AUC were 33 and 0.91, respectively. Heterogeneity existed between the included studies. No significant publication bias was observed.

**Conclusion:** Compared with CDFI, Superb micro-vascular imaging (SMI) has higher diagnostic sensitivity and specificity, better diagnostic efficiency, and could be used to diagnose benign and malignant nodules in the display of blood flow distribution capabilities of thyroid nodules; at the same time, Fagan plot showed that the SMI technique had a good clinical application value, and it could supplement the deficiencies of color Doppler imaging in the diagnosis of thyroid nodules.

KEYWORDS

**thyroid nodules, superb microvascular imaging, ultrasonography, meta-analysis, SMI = superb microvascular imaging**

**Abbreviations:** SMI, superb microvascular imaging; CDFI, color doppler flow imaging; DOR, diagnostic odds ratio; AUC, area under the curve; ROC, receiver operating curve; CI, confidence interval.

## Introduction

Thyroid nodules are discrete lesions in the thyroid gland that are distinct from the surrounding parenchyma (1). Recently, the incidence of thyroid nodule has been increasing year by year, making it one of the most common diseases of the thyroid system, with occult thyroid nodule accounting for 65% of the total population (2). According to the report, approximately 5% of thyroid nodules are at risk of developing thyroid cancer (3), with thyroid carcinoma becoming the most common tumour in the head and neck region (4). In clinical practice, ultrasound is the most sensitive imaging method for diagnosing thyroid nodules (5). According to research, it is necessary to evaluate the blood supply in thyroid nodules and distinguish benign from malignant nodules (6).

As a new angiography technology, superb microvascular imaging (SMI) can visualise low-speed blood flow and depict vascular flow in more detail than colour Doppler flow imaging, allowing for higher-quality microvascular flow images without the use of contrast media (7, 8). In recent years, with the widespread use of SMI technology in breast, liver and other fields, an increasing number of research teams have begun to pay attention to the diagnostic value of SMI in thyroid nodules. Despite numerous studies comparing ultrasound microangiography to conventional two-dimensional ultrasound diagnosis, there are still internal debates regarding the comparison of sensitivity, specificity and other indicators. Therefore, in order to provide a relatively objective and comprehensive evidence-based diagnostic basis for clinical thyroid management, this study collected all relevant literatures from both domestic and international sources, summarised the data of the two diagnostic methods using meta-analysis and comprehensively compared the diagnostic efficacy with pathological diagnosis as the reference standard.

## Review methods

This present study has been conducted and reported in accordance with the PRISMA statement. Institutional review board approval was not required as our study was a meta-analysis and consent was not required for this research.

## Literature search

We conducted a comprehensive search of all literatures in PubMed, Cochrane Library, Embase, Web of Science, Sinomed, Scopus and other databases until January 2021. Searches were conducted using the following MeSH heading and key words: “thyroid nodules[MeSH] OR thyroid disease OR thyroid” and”Superb Microvascular Imaging OR SMI”. In order to expand the scope of our search, we also screened the references of the retrieved articles. Language restrictions are not applied.

## Inspection methods and study selection

Routine two-dimensional ultrasonography was performed on the patients, and the thyroid nodule was observed in terms of boundary, size, internal echo, morphology, calcification, etc., and was preliminarily classified according to the TI-RADS classification system (1). Superb Microvascular Imaging (SMI) and color Doppler flow imaging (CDFI) were used to evaluate the nodules on the basis of conventional two-dimensional imaging. The blood supply and blood flow distribution pattern of the nodules were mainly evaluated. The classification standard of blood flow distribution pattern was based on Kim semi-quantitative method (9), and the blood flow signal grading was based on Adler classification method (10).

The selection of the literature by the two personnel was carried out according to the inclusion and exclusion criteria for diagnostic trials recommended by Cochrane. The process was ensured to be independent and separate. Finally, the two data were obtained, and then the data were cross-checked.

Inclusion criteria: (1) The pathological diagnosis results were considered as the “gold standard”, and the disease diagnosis was clear; (2) Subjects included patients with thyroid nodule with other diseases, whose nature was not defined before thyroid nodule examination; (3) Data sample size ≥30 cases; (4) Microangiography and conventional two-dimensional ultrasonography were in the same group; (5) The literatures of TP, FP, FN and TN can be directly extracted or indirectly calculated according to sensitivity, specificity and number of cases.

Exclusion criteria: (1) No literature on pathological diagnosis; (2) Literature that was directly included in the study with cases after diagnosis; (3) Reviews, conferences, case reports; (4) Incomplete data, TP, FP, FN and TN cannot be directly extracted or indirectly calculated according to sensitivity, specificity and number of cases; (5) Duplicate literature searched in different databases.

## Data extraction and quality assessment

The data for qualified full-text articles were independently extracted by two researchers (Haorou, Luo and Lixue, Yin), and different comments were solved by discussion and consensus. The recorded data are as follows: First author, language, year, mean age of patients, number of malignant lesions, “gold standard” diagnosis, type of study. The diagnostic trials were analyzed according to the Cochrane recommended diagnostic accuracy study quality assessment tool (QUADAS-2) (11), and then the risk of bias and clinical applicability were evaluated summatively.

## Statistical analyses

The heterogeneity of the included literatures were evaluated using the inconsistency index ($I^2$) and Cochrane Q-tests($x^2$). If $I^2$ >

is 50% or P < 0.10, high heterogeneity exists among the results. $I^2$ < 25% indicates that there is little heterogeneity among the results. 25%≤$I^2$ ≤ 50%, indicating moderate heterogeneity of the results. At the same time, the Bivariate boxplot can be used to make rough judgment, and then guide the sensitivity analysis. If the heterogeneity result is large, try to trace the source of heterogeneity by means of meta-regression. The positive likelihood ratio, negative likelihood ratio, sensitivity, specificity and diagnostic odds ratio of the two methods were analyzed and compared, and then the SROC curve was drawn. Sensitivity analysis and meta-regression subgroup analysis were performed. Deek funnel plot was drawn to evaluate whether there was publication bias, and the funnel plot was observed to be symmetrical, combined with P value to make a comprehensive evaluation. Fagan plot can show the correlation between prior probability, likelihood ratio and post-test probability. The greater the difference between prior and post-test probability, the more important the diagnostic test is, and it can provide a reference for clinical application evaluation. All analyses were completed using StataV 15.0.

# Results

## Literature search and selection

As shown in Figure 1, the initial search yielded 684 records. 229 records were excluded after removing duplicates. and 400 were excluded after screening the titles and abstracts. Depending on the inclusion and exclusion criteria, 45 records were deleted, ten studies ultimately are included in our meta-analysis (12–21).

## Study characteristics

The included 10 records included patients who were examined by both ultrasound microvascular imaging and color Doppler blood flow imaging, with a total of 1083 lesions, of which 518 were malignant lesions. Details of the eligible studies are shown in Table 1. All of them were diagnostic studies published from 2016 through 2020, 5 of which were recorded in Chinese and 5 in English;

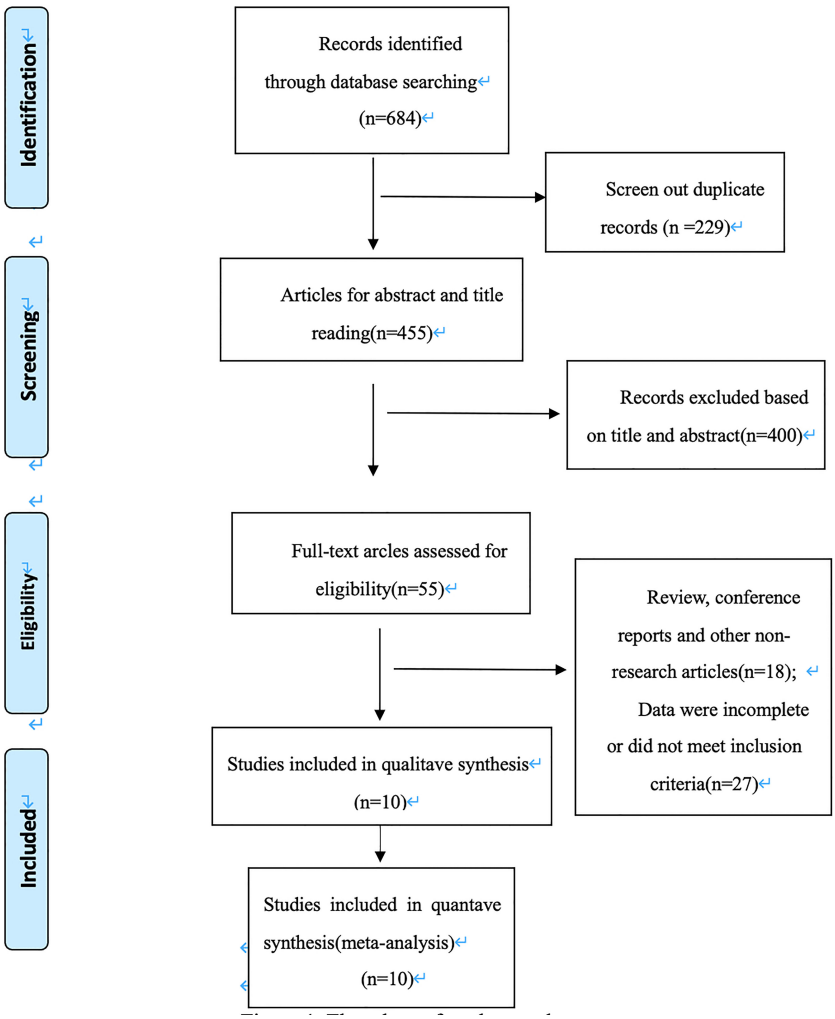



FIGURE 1
Flowchart of study search.

TABLE 1 characteristics of the included studies.

| Author | Year | Country | Language | average age | Total Lesions | Malignant Lesions | Male/ Female | Gold Standard | Study design |
|---|---|---|---|---|---|---|---|---|---|
| **Tian J** | 2017 | China | Chinese | 47.3 ± 10.8 | 45 | 25 | 8/26 | surgical Pathology | Prospective cohort |
| **Ouyang** | 2017 | China | Chinese | 39.2 ± 11.5 | 76 | 41 | 16/41 | surgical Pathology | Prospective cohort |
| **Diao XH** | 2016 | China | Chinese | 44.8 ± 17.6 | 68 | 27 | 11/36 | surgical Pathology | Prospective cohort |
| **Yang HX** | 2017 | China | Chinese | 47.1 ± 8.2 | 83 | 28 | 18/42 | surgical Pathology | Prospective cohort |
| **Wang H** | 2020 | China | Chinese | 38.9 ± 12.2 | 120 | 72 | 13/91 | surgical Pathology | Retrospective cohort |
| **Li YH** | 2017 | China | English | 39.0 ± 16.5 | 254 | 73 | 53/188 | surgical Pathology +FNA | Prospective cohort |
| **Kong J** | 2017 | China | English | 42 | 113 | 79 | 48/44 | surgical Pathology | Prospective cohort |
| **Zhu YC** | 2018 | China | English | 49.6 ± 13.2 | 76 | 29 | 35/41 | surgical Pathology +FNA | Prospective cohort |
| **Hye Shin** | 2017 | South Korea | English | 51.6 ± 11.2 | 52 | 26 | 9/48 | surgical Pathology +FNA | Prospective cohort |
| **Pei SF** | 2019 | China | English | – | 196 | 118 | 58/112 | surgical Pathology | Retrospective cohort |

Nine records were prospective studies, and only one record was retrospective studies; The gold standard for all the selected studies was pathological tissue study, of which 3 records were surgical pathology plus fine needle aspiration pathology. The average age of all included patients was under 55. The number of lesions varied from 45 to 254, and the number of patients varied from 34 to 241.

## Quality assessment

The quality assessment of the included studies were evaluated as follows (Table 2). All included studies were in compliance with the index test standards, with pathological evaluation as the reference standard.

TABLE 2 Quality assessment of the included studies using QUADAS-2 tool.

| References | Risk of bias | | | | Applicability concerns | | |
|---|---|---|---|---|---|---|---|
| | Patient selection | Index test | Reference standard | Flow and timing | Patient selection | Index test | Reference standard |
| **Tian J** | U | L | L | L | L | L | L |
| **Ouyang** | L | L | U | L | L | L | L |
| **Diao XH** | U | L | L | L | U | L | L |
| **Yang HX** | L | L | U | L | L | L | L |
| **Wang H** | L | H | H | L | U | L | L |
| **Li YH** | L | L | L | H | L | L | L |
| **Kong J** | L | L | L | L | L | L | L |
| **Zhu YC** | L | L | L | H | L | L | L |
| **Hye Shin** | L | L | L | L | L | L | L |
| **Pei SF** | L | L | L | L | L | L | L |

*. L, low risk of bias; H, high risk of bias; U, unclear risk of bias.

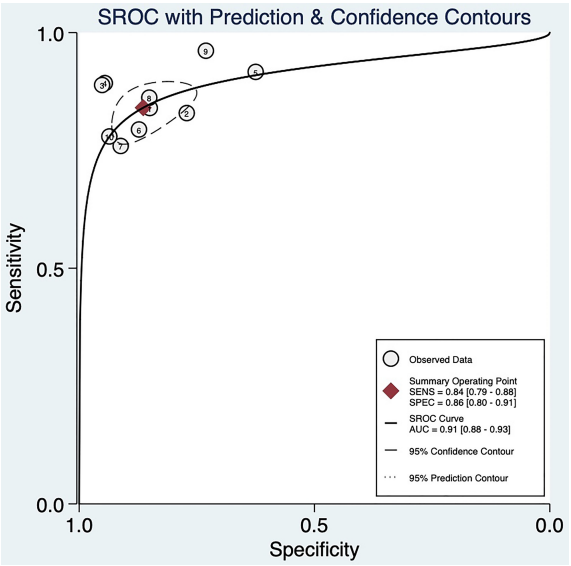



FIGURE 2
SROC curve and 95% confidence region.

## Diagnostic performance

Random-effects model was used to generate the pooled result. The pooled sensitivity, specificity, positive LR, and negative LR of SMI in differentiating malignant and benign thyroid nodules were 0.84 (95% confidence interval [CI], 0.79-0.88), 0.86 (95% CI, 0.80-0.91), 6.2 (95% CI, 4.1-9.2), and 0.18 (95% CI, 0.14-0.24), respectively. The area under the curve (AUC) was 0.91 (Figure 2). The forest plots of sensitivity, specificity, positive LR, negative LR, and DOR of the included studies were described in Figures 3–5. The detailed comparison of the combined effect size SMI and color Doppler was shown in Table 3.

## Heterogeneity analyses

When all 10 included studies were analyzed, heterogeneity was found both in sensitivity (heterogeneity, Q=14.26, p =0.11; $I^2$ = 36.88) and specificity (heterogeneity, Q =38.90, p < 0.0001; $I^2$ = 76.86%). To trace the sources of heterogeneity, we primarily analyzed the threshold effect. But there was no shoulder-arm

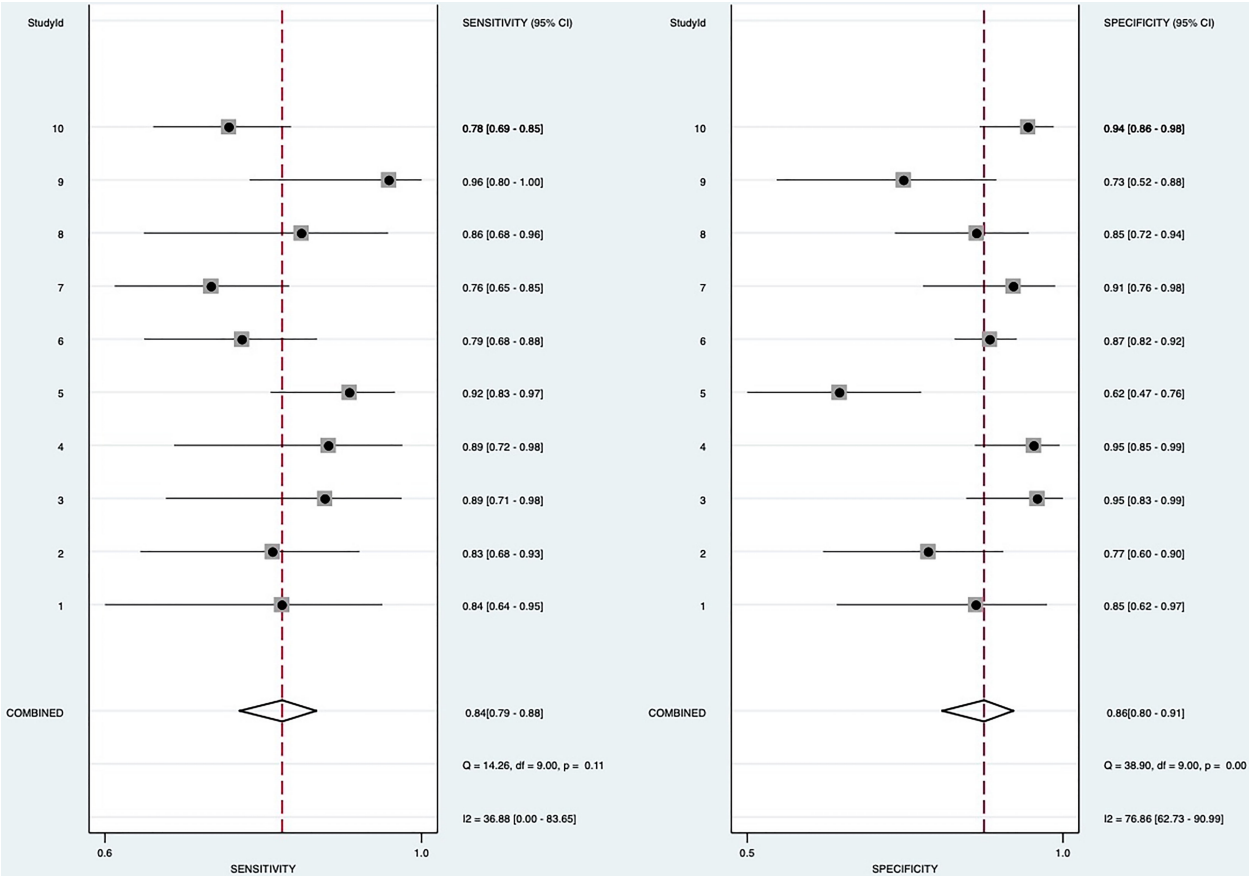



FIGURE 3
Forest plot for sensitivity and specificity.

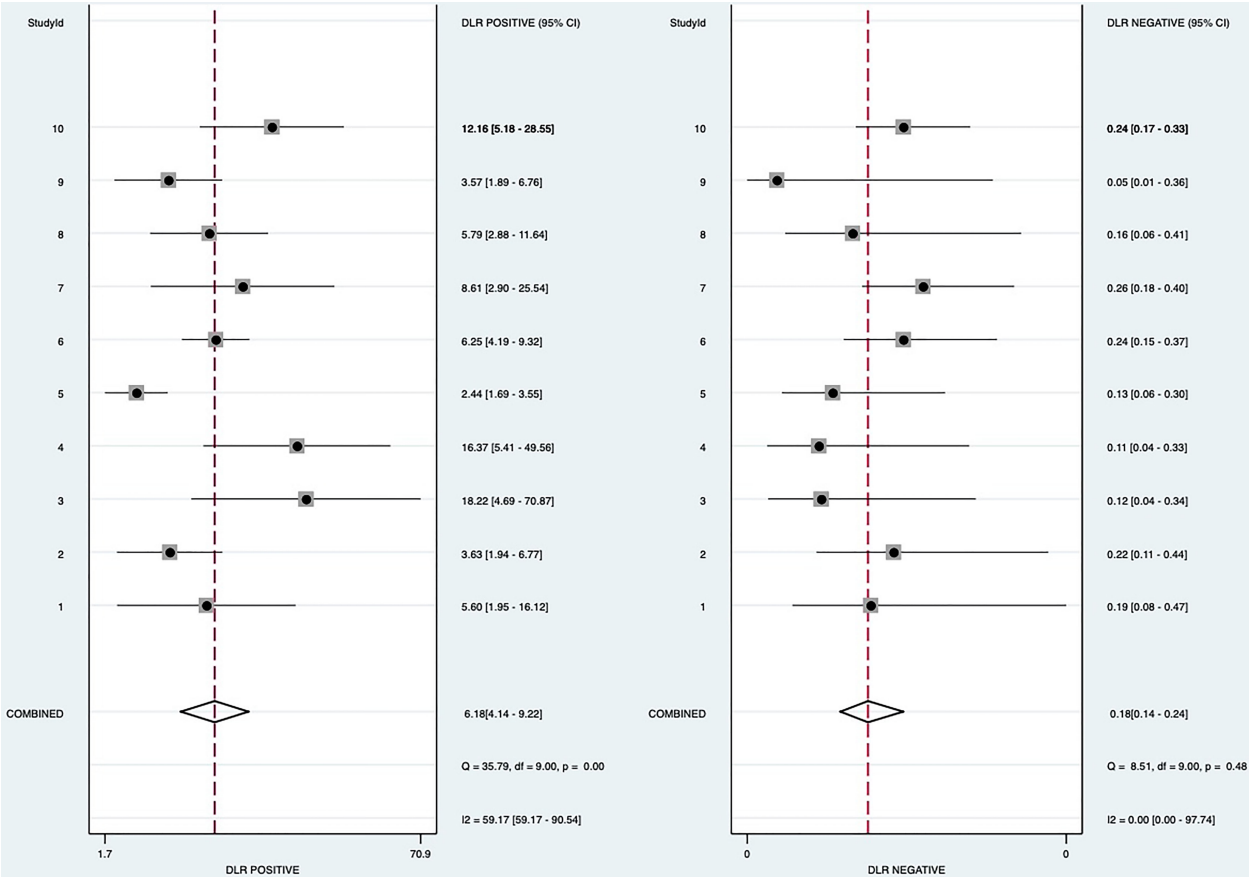



FIGURE 4
Forest plot for Positive likelihood ratio and negative likelihood ratio.

sign shown on the sROC space (Figure 2), and the Spearman correlation coefficient was 0.349 (P=0.324), indicating that all records had no significant threshold effect, and the effect size could be combined. It also indicated that there were other factors leading to heterogeneity in the inclusion analysis. The results of the subgroup analyses showed in Table 4. For the bivariate box chart, the documents outside the central area were deleted and then meta-analysed. It was found that the combined effect of various indicators had little change, indicating that the included documents were stable

## Publication bias

By Deeks' funnel plot asymmetry test, we found there was no significant publication bias among the eligible studies (P = 0.43). The Deeks' funnel plot was shown in Figure 6.

## Clinical application evaluation

According to Fagan's plot, the post-test probability of SMI was significantly higher than that of CDFI (Figure 7).

# Discussion

Thyroid nodules have been detected in up to 65% of the total population, indicating that thyroid nodules have become a very common disease (22). The majority of nodules detected by us are benign and can be safely managed through monitoring procedures. The purpose of initial screening and long-term follow-up of thyroid nodules is to better manage the most critical thyroid malignancies (2).

Fine needle aspiration biopsy has always been used to determine the nature of nodules related to clinical malignancies, and its status as the gold standard is unshakable (23). However, routine physical examination requires a non-invasive diagnostic method to assess the characteristics of blood flow distribution and understand the characteristics of blood supply in order to guide puncture (24).

Thyroid ultrasound is considered to be the best imaging method for assessing thyroid nodules or diffuse disease (25), Over the past 25 years, significant advances have been made in the evaluation of the thyroid by ultrasound and in the identification of suspicious and non-suspicious thyroid nodules (26). However, in general, the current conventional color Doppler blood flow imaging still has the disadvantages of insufficient accuracy, a large number of misdiagnosis, and missed diagnosis in the identification of the

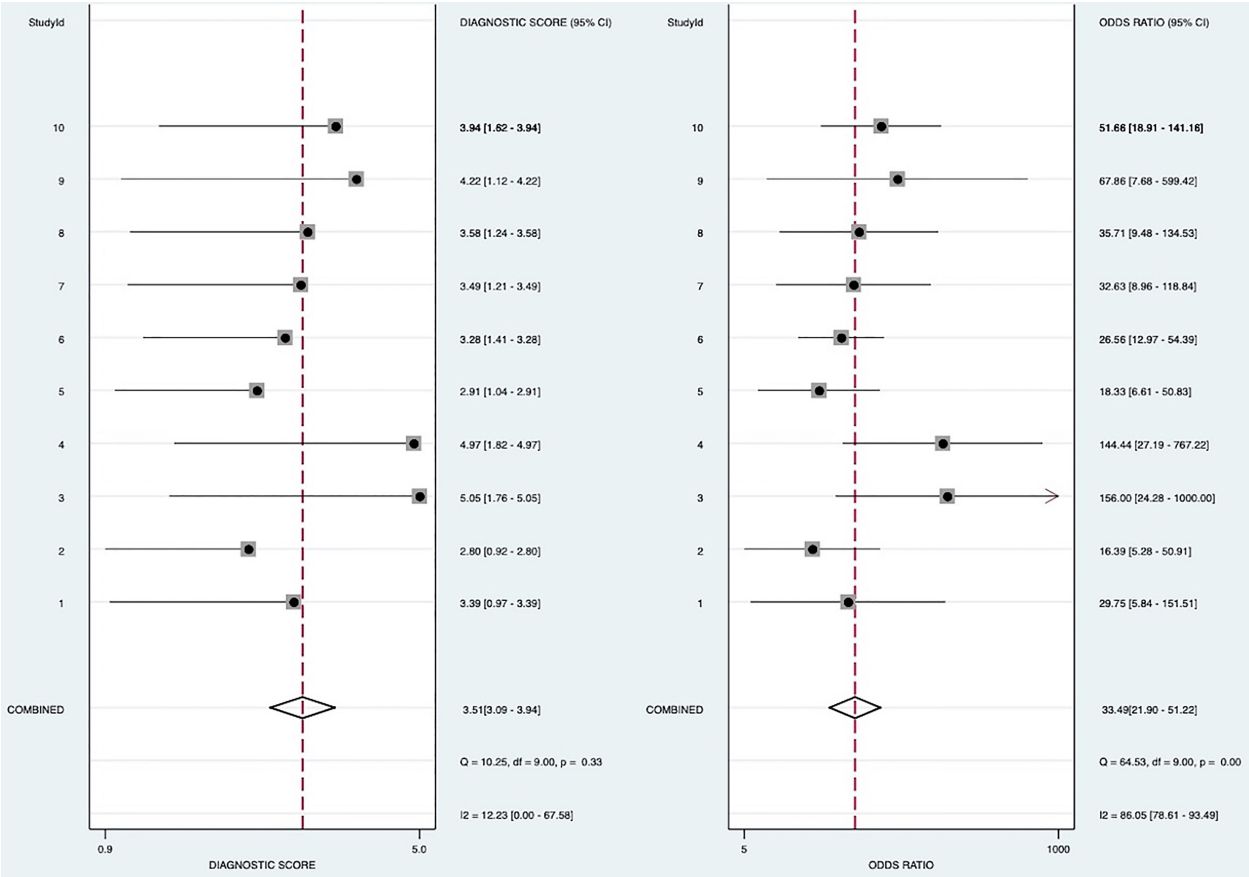



FIGURE 5
Forest plot for Odds ratio.

TABLE 3 Comparison of combined effects.

| | sensitivity | specificity | PLR | NLR | DOR | AUROC |
|---|---|---|---|---|---|---|
| **SMI** | 0.84(0.79-0.88) | 0.86(0.80-0.91) | 6.2(4.1-9.2) | 0.18(0.14-0.24) | 33(22-51) | 0.91(0.88-0.93) |
| **CDFI** | 0.64(0.56-0.71) | 0.78(0.72-0.83) | 2.8(2.3-3.5) | 0.47(0.39-0.56) | 6(4-9) | 0.77(0.74-0.81) |

TABLE 4 Meta regression analysis results.

| Subgroup | Number of studies | Pooled sensitivity | P1 | Pooled specificity | P2 | $I^2$ value,% |
|---|---|---|---|---|---|---|
| **Prospective** | 8 | 0.84(0.79-0.89) | 0.00 | 0.87(0.81-0.93) | 0.46 | 0 |
| **Retrospective** | 2 | 0.85(0.78-0.93) | – | 0.82(0.68-0.97) | – | – |
| **Lesions≥100** | 4 | 0.82(0.76-0.88) | 0.00 | 0.86(0.78-0.95) | 0.05 | 30 |
| **Lesions ≤ 100** | 6 | 0.88(0.82-0.93) | – | 0.87(0.79-0.94) | – | – |
| **English** | 5 | 0.81(0.76-0.87) | 0.00 | 0.88(0.80-0.95) | 0.12 | 6 |
| **Chinese** | 5 | 0.87(0.82-0.93) | – | 0.85(0.76-0.94) | – | 77 |
| **Surgical pathology** | 7 | 0.83(0.78-0.89) | 0.00 | 0.88(0.82-0.94) | 0.30 | 0 |
| **Surgical pathology+FNA** | 3 | 0.85(0.78-0.93) | – | 0.83(0.71-0.95) | – | – |
| **Chinese** | 9 | 0.83(0.79-0.87) | 0.14 | 0.87(0.82-0.93) | 0.53 | 48 |
| **South Koreans** | 1 | 0.96(0.89-1.00) | – | 0.74(0.45-1.00) | – | – |

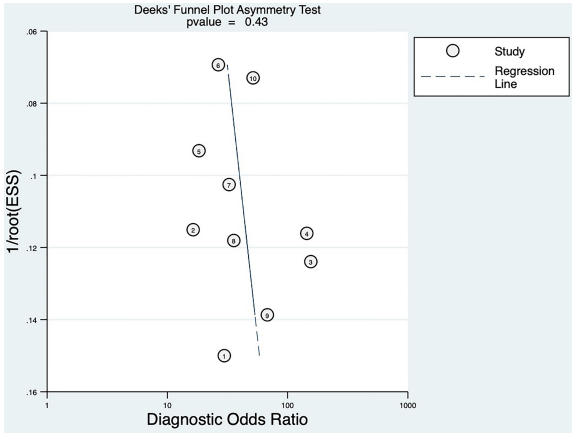



FIGURE 6
Deek's funnel plot of publication bias.

nature of thyroid nodules. And there are technical limitations to detecting small blood vessels and low blood flow.

Angiogenesis and the growth of irregular vascular structures have been established as the salient features of malignancy. Several previously published studies had shown that peri-nodular blood flow was more associated with benign thyroid nodules, while intra - nodular blood flow tended to be present in malignancies. Therefore, changes in vascular characteristics may provide more valuable diagnostic clues (9, 27).

In recent years, the rapid development of Superb micro-vascular imaging aimed at the deficiency of two-dimensional gray scale ultrasound and color Doppler blood flow imaging, and provided information about low speed blood flow, detection of small vessels (8). It reveals more microvascular branches, and shows the blood flow distribution in the nodules and adjacent thyroid parenchyma in more detail, which has become a more effective detection method (28). Although Superb micro-vascular imaging could be used to distinguish malignant and benign thyroid nodules, the value of its differential diagnosis is still inconclusive, and there is currently no Meta-analysis on the diagnostic performance of SMI technology. Therefore, the method of systematic review in this study was adopted to compare the value of SMI and CDFI in differentiating benign and malignant thyroid nodules. To provide an exact evidence-based basis for the differential diagnosis of benign and malignant thyroid nodules by SMI technology, so as to better guide the clinical diagnosis and treatment.

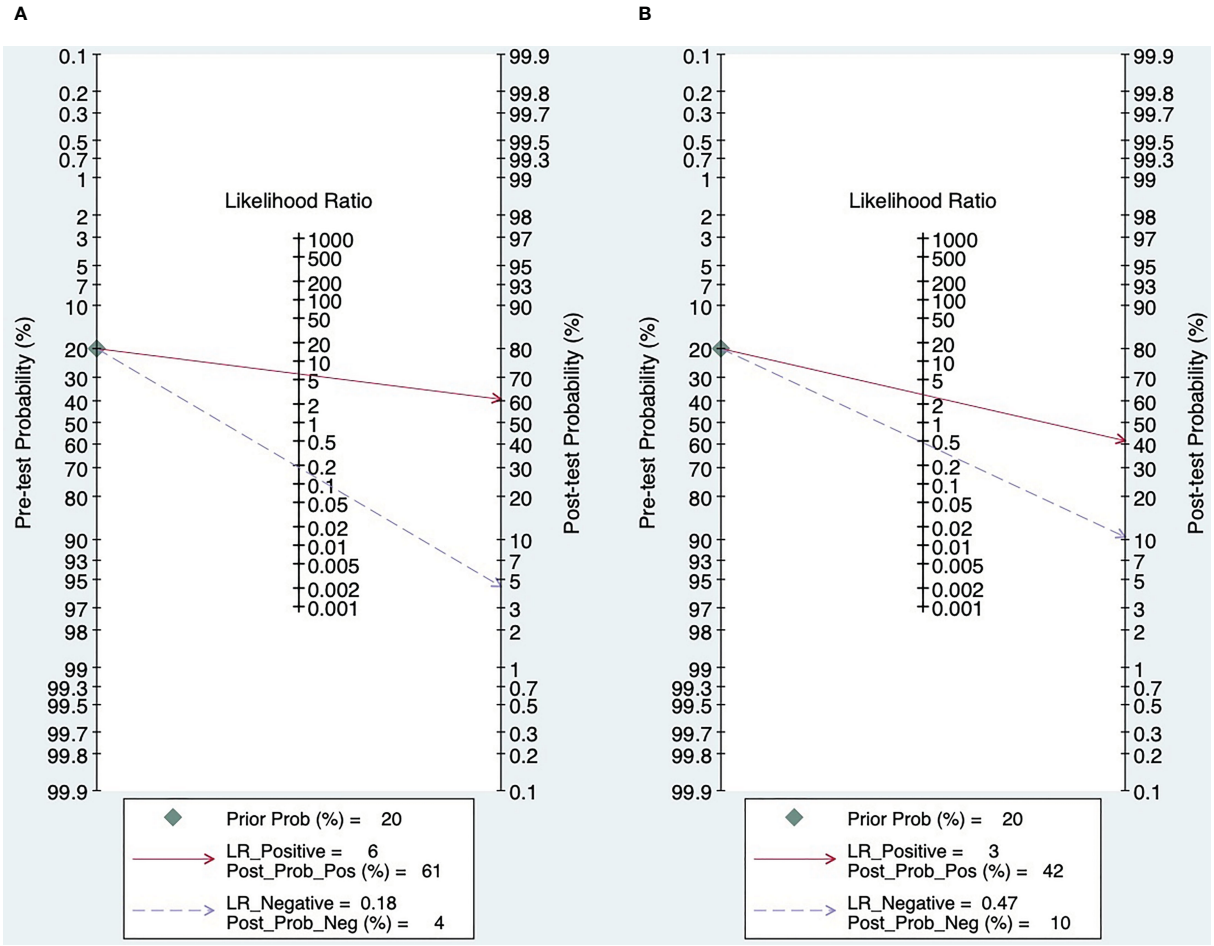



FIGURE 7
Fagan's plot (**A**. SMI; **B**. CDFI).

The pooled results showed that SMI could diagnose thyroid nodules, and the pooled sensitivity, specificity, positive LR and negative LR values were 0.84, 0.86, 6.2, and 0.18, respectively. These results indicated SMI had high sensitivity and specificity in diagnosing thyroid nodules. The AUC was 0.91, and the DOR was 33, which were the overall evaluation indicators of diagnostic tests, which proved the high accuracy of SMI in the differentiation of thyroid malignant nodules and benign nodules.

The heterogeneity test of our study showed moderate heterogeneity. We discovered that there was no significant threshold effect among the included records. Therefore, subgroup analyses were performed to find the source of heterogeneity. The results of the subgroup analyses showed that the heterogeneity of the combined estimates was reduced after excluding the Chinese study. The reassessed pooled results still suggested that SMI was more accurate in the diagnosis of thyroid nodules (AUC = 0.91, DOR = 32). In the 10 literatures included in this paper, although the "gold standard" was pathological examination, some were surgical biopsy histological pathology and some were cell puncture pathological results, which may also be one of the sources of heterogeneity.

Limitations of our meta-analysis: (1). Only 10 records were included in our meta-analysis. A small number of studies might reduce the effectiveness of the tests of heterogeneity and publication bias. (2). Different diagnostic criteria were used in the included studies, and neither of the two diagnostic methods was quantified and there was no unified diagnostic standard.

In conclusion, Superb micro-vascular imaging (SMI) is more sensitive, specific and effective than color doppler flow imaging in differentiating malignant and benign thyroid nodules. At the same time, SMI has good clinical value and can supplement the deficiency of CDFI in the diagnosis of thyroid nodule. SMI technology could truly reflect the blood flow distribution without the presence of blood flow signal spillover, and could be very sensitive to detect low blood flow and small blood vessels. Meanwhile, the examination cost is low, and compared with other imaging examinations, SMI technology has the advantages of non-trauma and non-radiation.

At present, as an emerging diagnostic technology, SMI still needs to develop a more detailed classification evaluation to quantify the form of blood flow distribution. At the same time, a unified diagnostic standard is also needed to guide clinical diagnosis, treatment and prognosis, To better tap the application potential of SMI diagnosis, future large-scale, controlled, multicenter studies are recommended.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Author contributions

HL: conception and design, acquisition of data, analysis and interpretation of data, drafting and revising the article; final approval of the version to be published. LY: the corresponding author of this paper, acquisition of data, analysis and interpretation of data. revise the article. All authors contributed to the article and approved the submitted version.

## Funding

Funding: The fundamental Research Funds for the central Universities, Grant/Award Number: ZYGX2020ZB038. Funding belongs to the corresponding author's laboratory.

## Acknowledgments

The authors express their gratitude to Sichuan provincial people's hospital and UESTC for their cooperation and assistance.

## 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.

## Publisher's note



## References

1. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American thyroid association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid* (2016) 26(1):1–133. doi: 10.1089/thy.2015.0020

2. Durante C, Grani G, Lamartina L, Filetti S, Mandel SJ, Cooper DS. The diagnosis and management of thyroid nodules: A review. *Jama* (2018) 319(9):914–24. doi: 10.1001/jama.2018.0898

3. Grani G, Sponziello M, Pecce V, Ramundo V, Durante C. Contemporary thyroid nodule evaluation and management. *J Clin Endocrinol Metab* (2020) 105(9):2869–83. doi: 10.1210/clinem/dgaa322

4. Lim H, Devesa SS, Sosa JA, Check D, Kitahara CM. Trends in thyroid cancer incidence and mortality in the united states, 1974-2013. *Jama* (2017) 317(13):1338–48. doi: 10.1001/jama.2017.2719

5. Iwen A. [Ultrasound examination of the thyroid gland - step by step]. *Dtsch Med Wochenschr* (2020) 145(4):260–7. doi: 10.1055/a-0674-7845

6. Cansu A, Ayan E, Kul S, Eyüboğlu İ, Oğuz Ş, Mungan S. Diagnostic value of 3D power Doppler ultrasound in the characterization of thyroid nodules. *Turk J Med Sci* (2019) 49(3):723–9. doi: 10.3906/sag-1803-92

7. Ma Y, Li G, Li J, Ren WD. The diagnostic value of superb microvascular imaging (SMI) in detecting blood flow signals of breast lesions: A preliminary study comparing SMI to color Doppler flow imaging. *Med (Baltimore)* (2015) 94(36):e1502. doi: 10.1097/MD.0000000000001502

8. Machado P, Segal S, Lyshchik A, Forsberg F. A novel microvascular flow technique: Initial results in thyroids. *Ultrasound Q* (2016) 32(1):67–74. doi: 10.1097/RUQ.0000000000000156

9. Kim DW, In HS, Choo HJ, Jung SJ, Kim YW. Solid and isoechoic thyroid nodules without malignant sonographic features: comparison of malignancy rate according to nodule size, shape and color Doppler pattern. *Ultrasound Med Biol* (2013) 39(2):269–74. doi: 10.1016/j.ultrasmedbio.2012.09.018

10. Adler DD, Carson PL, Rubin JM, Quinn-Reid D. Doppler Ultrasound color flow imaging in the study of breast cancer: Preliminary findings. *Ultrasound Med Biol* (1990) 16(6):553–9. doi: 10.1016/0301-5629(90)90020-D

11. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* (2003) 3:25. doi: 10.1186/1471-2288-3-25

12. Tian J, Yong Q, Li C. Preliminary study on the application value of ultramicroangiography in differentiating benign and malignant thyroid nodules. *Cancer progression* (2017) 15(04):450–3.doi: 10.11877/j.issn.1672-1535.2017.15.04.28

13. Ou Y. Study on the value of microangiography and contrast-enhanced ultrasound in differentiating benign and malignant thyroid nodule. *Chin J Ultrasound Med* (2017) 4(33):293–6. doi: 10.3969/j.issn.1002-0101.2017.04.003

14. Diao X, Zhan J, Chen L, Chen Y, Cao H. Application value of super microangiography in differential diagnosis of solid thyroid nodules. *Chin J Med Ultrasound* (2016) 16(08):622–6. doi: 10.3877/cma.j.issn.1672-6448.2016.08.013

15. Yang H, Zhang L, Li H. Analysis of the value of super microangiography in the differential diagnosis of solid thyroid nodules. *J Clin Exp Med* (2017) 16(10):1029–32. doi 10.3969/j.issn.1671-4695.2017.10.030

16. Wang H, Qiu J, Luo S. Evaluation of benign and malignant thyroid nodules by grayscale model microangiography. *J Appl Radiol* (2020) 5(36):714–7. doi: 10.3969/j.issn.1002-1671.2020.05.007

17. Li Y, Wen D, Li C, Li XJ, Xue G. The role of ATA (2015) guidelines, superb microvascular imaging, and spectral Doppler in differentiation between malignant and benign thyroid nodules. *Lin chuang er bi yan hou tou jing wai ke za zhi= J Clin otorhinolaryngol head Neck Surg* (2017) 31(15):1152–6. doi: 10.13201/j.issn.1001-1781.2017.15.003

18. Kong J, Li J-C, Wang H-Y, Wang YH, Zhao RN, Zhang Y, et al. Role of superb micro-vascular imaging in the preoperative evaluation of thyroid nodules: Comparison with power Doppler flow imaging. *J ultrasound Med* (2017) 36(7):1329–37. doi: 10.7863/ultra.16.07004

19. Zhu YC, Zhang Y, Deng SH, Jiang Q. A prospective study to compare superb microvascular imaging with grayscale ultrasound and color Doppler flow imaging of vascular distribution and morphology in thyroid nodules. *Med Sci Monit* (2018) 24:9223–31. doi: 10.12659/MSM.911695

20. Ahn HS, Lee JB, Seo M, Park SH, Choi BI. Distinguishing benign from malignant thyroid nodules using thyroid ultrasonography: Utility of adding superb microvascular imaging and elastography. *Radiol Med* (2018) 123(4):260–70. doi: 10.1007/s11547-017-0839-2

21. Pei S, Cong S, Zhang B, Liang C, Zhang L, Liu J, et al. Diagnostic value of multimodal ultrasound imaging in differentiating benign and malignant TI-RADS category 4 nodules. *Int J Clin Oncol* (2019) 24(6):632–9. doi: 10.1007/s10147-019-01397-y

22. Wong R, Farrell SG, Grossmann M. Thyroid nodules: diagnosis and management. *Med J Aust* (2018) 209(2):92–8. doi: 10.5694/mja17.01204

23. Kovacevic B, Vilmann P. EUS tissue acquisition: From a to b. *Endosc Ultrasound* (2020) 9(4):225–31. doi: 10.4103/eus.eus_21_20

24. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda system for reporting thyroid cytopathology: A meta-analysis. *Acta cytol* (2012) 56(4):333–9. doi: 10.1159/000339959

25. Alexander LF, Patel NJ, Caserta MP, Robbin ML. Thyroid ultrasound: Diffuse and nodular disease. *Radiol Clin North Am* (2020) 58(6):1041–57. doi: 10.1016/j.rcl.2020.07.003

26. Shaha AR. Comment on "The diagnosis and management of thyroid nodules: A review". *Gland Surg* (2018) 7(5):424–5. doi: 10.21037/gs.2018.06.01

27. Kim DW, Jung SJ, Eom JW, Kang T. Color Doppler features of solid, round, isoechoic thyroid nodules without malignant sonographic features: a prospective cytopathological study. *Thyroid* (2013) 23(4):472–6. doi: 10.1089/thy.2012.0238

28. Lu R, Meng Y, Zhang Y, Zhao W, Wang X, Jin M, et al. Superb microvascular imaging (SMI) compared with conventional ultrasound for evaluating thyroid nodules. *BMC Med Imaging* (2017) 17(1):65. doi: 10.1186/s12880-017-0241-5

# Frontiers in Oncology

**Advances knowledge of carcinogenesis and tumor progression for better treatment and management**

The third most-cited oncology journal, which highlights research in carcinogenesis and tumor progression, bridging the gap between basic research and applications to imrpove diagnosis, therapeutics and management strategies.

---

## Discover the latest Research Topics

See more →


**Frontiers**

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

**Contact us**

+41 (0)21 510 17 00
frontiersin.org/about/contact