Original Article

# Development and validation of a nomogram based on conventional and contrast-enhanced ultrasound for differentiating malignant from benign thyroid nodules

**Qi-Guo Wang[1#], Mei Li[2#], Guang-Xiu Deng[1], Hai-Qing Huang[1], Qin Qiu[3], Jian-Jun Lin[1^]**

[1]Department of Medical Ultrasound, the First People's Hospital of Qinzhou, Qinzhou, China; [2]Department of Medical Ultrasound, the People's Hospital of Chongzuo, Chongzuo, China; [3]Department of Ultrasound, the People's Hospital of Pubei, Qinzhou, China

*Contributions:* (I) Conception and design: QG Wang, M Li, JJ Lin; (II) Administrative support: JJ Lin; (III) Provision of study materials or patients: QG Wang, GX Deng, HQ Huang, JJ Lin; (IV) Collection and assembly of data: GX Deng, HQ Huang, Q Qiu; (V) Data analysis and interpretation: QG Wang, M Li; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

*Correspondence to:* Jian-Jun Lin, MD. Department of Medical Ultrasound, the First People's Hospital of Qinzhou, No. 8, Mingyang Street, Qinnan District, Qinzhou 535099, China. Email: 2819825313@qq.com.

**Background:** Conventional ultrasound (US) has been routinely used for differential diagnosis of thyroid nodules, but its discriminatory performance remains unsatisfactory. This study aimed to develop and validate a prediction nomogram model based on conventional US and contrast-enhanced ultrasound (CEUS) features for differentiating malignant from benign thyroid nodules.

**Methods:** A total of 815 thyroid nodules with surgical pathology results and complete conventional US and CEUS data were retrospectively collected from the First People's Hospital of Qinzhou between January 2019 and July 2023. The nodules were grouped into a training cohort (n=571) and a validation cohort (n=244) at a 7:3 ratio. Independent risk factors of malignancy were selected by stepwise multivariate logistic regression analysis, and a prediction nomogram model was subsequently constructed. The diagnostic performance of the model was evaluated by the area under the receiver operating characteristic curve (AUC) in both the training and validation cohorts. The unnecessary fine-needle aspiration biopsy (FNAB) rate was calculated.

**Results:** Multivariate logistic regression analysis identified irregular margin, aspect ratio >1, and microcalcification from conventional US images, as well as hypo-enhancement intensity and ring enhancement from CEUS images, as independent predictors for malignancy. The AUC, sensitivity, specificity, and accuracy of the prediction nomogram model were 0.947 [95% confidence interval (CI): 0.928–0.966], 90.4%, 88.8%, and 89.8% in the training cohort, and 0.957 (95% CI: 0.928–0.986), 94.5%, 86.4%, and 91.8% in the validation cohort, respectively. Using the prediction model, the unnecessary FNAB rates reduced from 29.6% to 6.1% in the training cohort and from 29.3% to 6.7% in the validation cohort compared to the Chinese Thyroid Imaging Reporting and Data System. Decision curve analysis demonstrated good clinical utility of the nomogram model.

**Conclusions:** The prediction nomogram model incorporating conventional US and CEUS features could effectively distinguish between malignant and benign thyroid nodules and reduce unnecessary FNAB rates.

**Keywords:** Contrast-enhanced ultrasound (CEUS); thyroid nodules; malignancy; nomogram

^ ORCID: 0009-0009-1279-100X.

Submitted Aug 26, 2024. Accepted for publication Mar 12, 2025. Published online Apr 28, 2025.
doi: 10.21037/qims-24-1796
**View this article at:** https://dx.doi.org/10.21037/qims-24-1796

## Introduction

With the widespread application of conventional ultrasound (US), the incidence of thyroid nodules has gradually increased worldwide (1). The prevalence of thyroid nodules is estimated to be up to 70% in the general population, with malignancies observed in only 5–15% of cases (2,3). Fine-needle aspiration biopsy (FNAB) is routinely used for definitive diagnosis of thyroid malignancies but is limited by its invasive nature and low detection rate (4-6). Therefore, it is critical to develop non-invasive and more accurate methods to differentiate malignant from benign thyroid nodules, aiming to avoid unnecessary FNAB or overtreatment such as surgery.

Conventional US is the first choice for thyroid imaging and the differential diagnosis of thyroid nodules. Malignant nodules exhibit several sonographic characteristics, including a solid component, irregular margins, hypo-echogenicity, microcalcification, and a taller-than-wide shape (7). However, no single sonographic feature has sufficient sensitivity and specificity to detect malignant thyroid nodules accurately (8). Based on these conventional US features, the American College of Radiology (ACR) established a risk stratification system, the Thyroid Imaging Reporting and Data System (ACR-TIRADS), to provide an easy-to-use risk evaluation method and to facilitate appropriate clinical management of thyroid nodules (9). Similar classification systems have also been proposed by other scientific communities, such as the American Thyroid Association (ATA) guideline (10), the Korean Society of Thyroid Radiology (K-TIRADS) (11), the European Thyroid Association (EU-TIRADS) (12), and the Chinese Medical Association (C-TIRADS) (13). FNAB is subsequently recommended for definitive diagnosis according to the nodule's TIRADS risk level and maximum diameter.

Nevertheless, the discriminatory performance of conventional US-based models remains unsatisfactory. TIRADS 4–5 nodules have a broad spectrum of potential malignancy rate greater than 5%, which results in a favorable sensitivity of 89% but an inferior specificity of 70% (14). Under these risk stratification systems, the rate of unnecessary FNAB, ranging from 21.2% to 58.3%, remains considerably high (15,16). Therefore, new techniques that improve diagnostic accuracy with high sensitivity and specificity are warranted for the management of thyroid nodules.

Contrast-enhanced US (CEUS) enables real-time evaluation of microvascular perfusion in thyroid lesions following the administration of microbubble contrast media (17). In recent years, CEUS has increasingly been used to complement conventional US for the differential diagnosis of thyroid nodules (18). Several meta-analyses have demonstrated that CEUS achieves a sensitivity of 84–87% and a specificity of 82–84% in distinguishing malignant from benign thyroid nodules (19-22). We assumed that integrating conventional US and CEUS images would further improve the diagnostic accuracy of malignant thyroid nodules. Therefore, this study aimed to develop a predictive nomogram combining conventional US and CEUS features to facilitate the risk stratification and clinical management of thyroid nodules. We present this article in accordance with the STARD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-24-1796/rc).

## Methods

### *Study design and patient selection*

This retrospective, single-center study was approved by the Ethics Committee of the First People's Hospital of Qinzhou (No. A-20240111) and conducted in accordance with the Declaration of Helsinki and its subsequent amendments. The requirement for informed consent was waived due to the study's retrospective nature. All participants had provided written informed consent for preoperative CEUS examinations. A total of 997 nodules from consecutive patients who underwent surgery for thyroid nodules at this hospital between January 2019 and July 2023 were enrolled. The inclusion criteria for thyroid nodules were as follows: (I) definitive surgical pathology results of thyroid nodules; (II) maximum nodule diameter ≥0.5 cm; (III) complete clinical information; and (IV) complete conventional US and CEUS data within 1 month before surgery. The exclusion criteria were as follows: (I) age younger than 18 years; (II) previously treated thyroid nodules; and (III) lack of surrounding normal parenchyma as a reference.

Ultimately, a total 815 nodules from 689 patients (mean

 

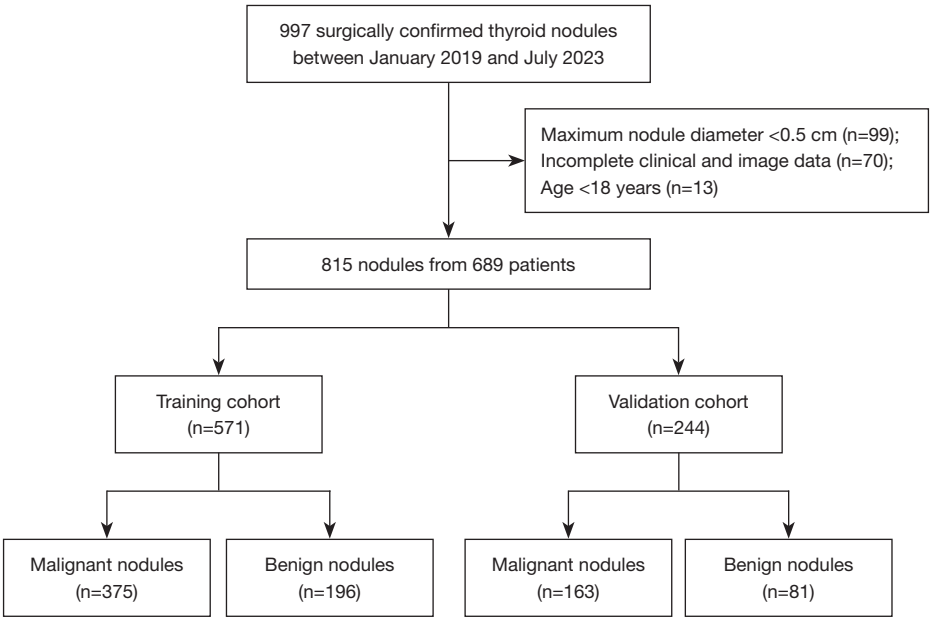


**Figure 1** Flowchart of patient selection.

age: 43.6±12.6 years, range, 18–83 years; 129 males and 560 females) were included in this study. There were 538 malignant nodules and 277 benign nodules. Malignant nodules included 532 papillary thyroid carcinomas, 4 medullary thyroid carcinomas, 1 thyroid metastatic carcinoma, and 1 thyroid squamous cell carcinoma. Benign nodules included 248 nodular goiters and 29 thyroid adenomas. All nodules were randomly divided into a training cohort (n=571, 375 malignant and 196 benign nodules) and a validation cohort (n=244, 163 malignant and 81 benign nodules) at a 7:3 ratio. The patient selection procedure is illustrated in *Figure 1*.

### *Image acquisition*

Both conventional US and CEUS images were acquired with the Canon Aplio i800 system (Canon Medical System Corporation, Tochigi, Japan) equipped with a linear array probe (i8LX5, frequency range, 5–18 MHz). All patients were examined in a supine position with their necks fully exposed and were instructed to avoid swallowing and talking during the examination. After the conventional US examination, the mode was switched to CEUS. SonoVue (Bracco, Milan, Italy) was injected as contrast agent with a volume of 1 mL via the median cubital vein, followed by a bolus of 5 mL normal saline. The mechanical index of CEUS was 0.06–0.08. Dynamic images were observed for 90 seconds and stored on a hard disk for further analysis.

### *Image analysis*

All conventional US and CEUS images were retrospectively reviewed by 2 independent radiologists with over 10 years of experience in thyroid imaging, who were blinded to the clinical information and pathological outcomes of all patients (*Figure 2*). For conventional US (*Figure 2A*,*2C*), the following features of each nodule were recorded: nodule size (maximum diameter), location (left lobe, right lobe, isthmus), echogenicity (hyper-echoic, iso-echoic, hypo-echoic relative to surrounding thyroid parenchyma), homogeneity (heterogeneous, homogeneous), aspect ratio (>1, ≤1), margin (regular, irregular), and microcalcification (present, absent). For the qualitative parameters of CEUS (*Figure 2B*,*2D*), the enhancement intensity (hyper-enhancement, iso-enhancement, hypo-enhancement relative to adjacent thyroid parenchyma at peak), enhancement pattern (heterogeneous, homogeneous), and ring enhancement (present, absent) of each nodule were recorded. In case of disagreement between the radiologists, a consensus was reached through discussion.

### *Statistical analysis*

Continuous variables were presented as mean ± standard

 

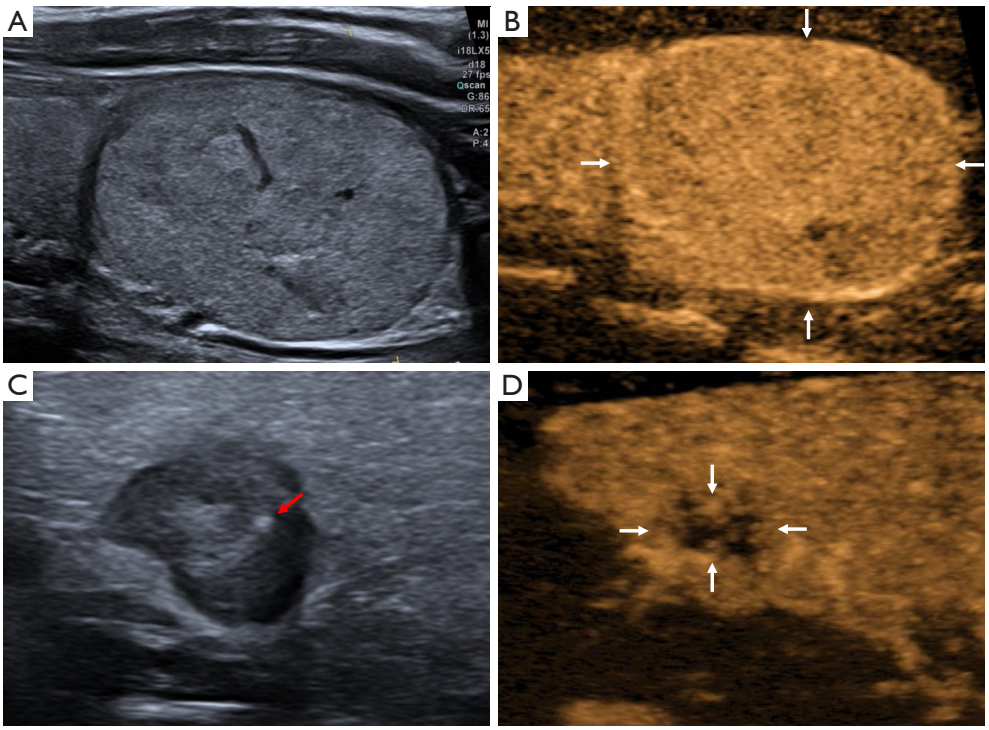


**Figure 2** Ultrasound images of benign and malignant nodules. For a nodular goiter, conventional US showed iso-echoic echogenicity, aspect ratio <1, and regular margin (A), and CEUS showed iso-enhancement and presence of ring enhancement (arrows) (B). For a papillary thyroid carcinoma, conventional US showed hypo-echoic echogenicity, aspect ratio >1, microcalcification (red arrow), and irregular margin (C), and CEUS showed hypo-enhancement and centrifugal enhancement direction (arrows) (D). CEUS, contrast-enhanced ultrasound; US, ultrasound.

deviation or median with interquartile range (IQR), and were compared by using Student's *t*-test or Wilcoxon rank-sum test, respectively. Categorical variables were expressed as frequency with percentage and compared using the chi-square test or Fisher's exact test when necessary. Independent predictive factors were identified by multivariate logistic regression analysis employing a stepwise backward method. A nomogram predicting the risk of malignant thyroid nodules was constructed based on the independent factors in the training cohort. The predictive performance of the constructed prediction model was evaluated by receiver operating characteristic (ROC) curve analysis and area under the curve (AUC) in the training and validation cohorts. The optimal cut-off value of the prediction model was determined by maximizing the Youden index, followed by the calculation of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy. DeLong test was used to compare the diagnostic performance between different models. The unnecessary FNAB rates, evaluated as the percentage of benign nodules among the recommended biopsied nodules, were compared between the constructed nomogram model and the established C-TIRADS (13). A calibration curve was plotted and the Hosmer-Lemeshow test was performed to assess the calibration effect of the prediction model. Decision curve analysis (DCA) was used to evaluate the clinical usefulness of the prediction model by quantifying the net benefits at different risk thresholds. All the statistical analyses were conducted using Stata 16.0 (StataCorp, College Station, TX, USA). A P value <0.05 was considered statistically significant.

## Results

### *Clinical, conventional US, and CEUS characteristics*

The clinical, conventional US, and CEUS characteristics of thyroid nodules in the training and validation cohorts are summarized in *Table 1*. No significant differences were observed in sex, nodule size, location, and pathology

**Table 1** Comparison of characteristics between the training and validation cohorts

| Characteristics | Training cohort (n=571) | Validation cohort (n=244) | P value |
|---|---|---|---|
| Sex | | | 0.28 |
| Male | 101 (17.7) | 51 (20.9) | |
| Female | 470 (82.3) | 193 (79.1) | |
| Age, years | 42.5±12.2 | 44.5±13.1 | 0.04 |
| >45 | 244 (42.7) | 118 (48.4) | 0.14 |
| ≤45 | 327 (57.3) | 126 (51.6) | |
| Thyroid nodule | | | 0.76 |
| Malignant | 375 (65.7) | 163 (66.8) | |
| Benign | 196 (34.3) | 81 (33.2) | |
| Nodule size, cm | 1.4 (0.89, 2.58) | 1.5 (0.83, 2.90) | 0.69 |
| >1.0 | 387 (67.8) | 159 (65.2) | 0.47 |
| ≤1.0 | 184 (32.2) | 85 (34.8) | |
| Location | | | 0.60 |
| Left lobe | 241 (42.2) | 94 (38.5) | |
| Right lobe | 295 (51.7) | 133 (54.5) | |
| Isthmus | 35 (6.1) | 17 (7.0) | |
| US features | | | |
| Echogenicity | | | 0.57 |
| Hyper-/iso-echoic | 108 (18.9) | 42 (17.2) | |
| Hypo-echoic | 463 (81.1) | 202 (82.8) | |
| Homogeneity | | | 0.64 |
| Homogeneous | 20 (3.5) | 7 (2.9) | |
| Heterogeneous | 551 (96.5) | 237 (97.1) | |
| Aspect ratio | | | 0.12 |
| >1 | 208 (36.4) | 75 (30.7) | |
| ≤1 | 363 (63.6) | 169 (69.3) | |
| Margin | | | 0.69 |
| Regular | 219 (38.4) | 90 (36.9) | |
| Irregular | 352 (61.7) | 154 (63.1) | |
| Microcalcification | | | 0.82 |
| Present | 113 (19.8) | 50 (20.5) | |
| Absent | 458 (80.2) | 194 (79.5) | |
| CEUS features | | | |
| Enhancement intensity | | | 0.19 |
| Hyper-/iso-enhancement | 267 (46.8) | 102 (41.8) | |
| Hypo-enhancement | 304 (53.2) | 142 (58.2) | |

**Table 1** (*continued*)

**Table 1** (*continued*)

| Characteristics | Training cohort (n=571) | Validation cohort (n=244) | P value |
|---|---|---|---|
| Enhancement pattern | | | 0.88 |
| Homogeneous | 96 (16.8) | 40 (16.4) | |
| Heterogeneous | 475 (83.2) | 204 (83.6) | |
| Ring enhancement | | | 0.62 |
| Present | 104 (18.2) | 48 (19.7) | |
| Absent | 467 (81.8) | 196 (80.3) | |

Categorical variables are presented as n (%) and continuous variables are presented as mean ± standard deviation or median (interquartile range). CEUS, contrast-enhanced ultrasound; US, ultrasound.

outcomes between these 2 cohorts. Additionally, the conventional US features and CEUS features were comparable between the training and validation cohorts.

### *Correlation analysis of conventional US and CEUS features with nodule pathology*

Univariate analyses of correlations between imaging features and nodule pathology in the training and validation cohorts are presented in *Table 2*. There were more male patients in the malignant group than there were in the benign group of the training cohort, whereas the sex distribution was similar between both groups in the validation cohort. Malignant nodules were significantly smaller than benign nodules (median size: 1.20 *vs.* 2.54 cm in the training cohort, P<0.001; 1.17 *vs.* 2.90 cm in the validation cohort, P<0.001). In the training cohort, patients with malignant nodules were significantly younger than those with benign nodules (P=0.03), but no significant difference was found in the validation cohort (P=0.22). Malignant nodules exhibited higher proportions of hypo-echoic feature, aspect ratio >1, irregular margin, and microcalcification compared to benign nodules in both cohorts. Regarding CEUS features, hypo-enhancement intensity and heterogeneous enhancement patterns were more prevalent in malignant nodules than in benign nodules (P<0.001). Ring enhancement was more frequently found in benign nodules compared to malignant nodules.

### *Prediction model construction*

Multivariate logistic regression analysis in the training cohort identified irregular margin, aspect ratio (>1), presence of microcalcification, hypo-enhancement intensity, and absence of ring enhancement as independent predictive factors for malignant thyroid nodules (*Table 3*). These 5 factors were incorporated into a prediction model, which was presented as a nomogram predicting the malignancy risk of thyroid nodules (*Figure 3*).

### *Performance of prediction model*

The ROC curves of the prediction model and individual feature (margin, aspect ratio, microcalcification, enhancement intensity, ring enhancement) were plotted to illustrate the performance in differentiating malignant from benign thyroid nodules in both the training cohort (*Figure 4A*) and validation cohort (*Figure 4B*). The AUC, sensitivity, specificity, PPV, NPV, accuracy, and cut-off value of prediction model were 0.947 [95% confidence interval (CI): 0.928–0.966], 90.4%, 88.8%, 93.9%, 82.9%, 89.8%, and 0.683 in the training cohort, and 0.957 (95% CI: 0.928–0.986), 94.5%, 86.4%, 93.3%, 0.88.6%, 91.8%, and 0.615 in the validation cohort, respectively. The DeLong test demonstrated that the AUC of the prediction model was significantly higher than that of any single conventional US and CEUS feature (all P<0.001, *Table 4*), indicating superior performance in discriminating malignant nodules from benign nodules. Using the prediction model, the unnecessary FNAB rates reduced from 29.6% to 6.1% in the training cohort and from 29.3% to 6.7% in the validation cohort compared to C-TIRADS (*Table 5*).

Calibration plots and Hosmer-Lemeshow test showed good agreement between the observed probability and the probability predicted by the nomogram in the training cohort (P=0.37, *Figure 5A*) and validation cohort (P=0.59, *Figure 5B*). The DCA curves of the prediction model and individual conventional US and CEUS features in the

 

**Table 2** Comparison of characteristics between malignant and benign thyroid nodules

| Characteristics | Training cohort | | | Validation cohort | | |
|---|---|---|---|---|---|---|
| | Malignant (n=375) | Benign (n=196) | P value | Malignant (n=163) | Benign (n=81) | P value |
| Sex | | | 0.007 | | | 0.49 |
| Male | 78 (20.8) | 23 (11.7) | | 32 (19.6) | 19 (23.5) | |
| Female | 297 (79.2) | 173 (88.3) | | 131 (80.4) | 62 (76.5) | |
| Age, years | 41.7±11.5 | 44.1±13.2 | 0.023 | 43.7±12.7 | 45.9±13.9 | 0.22 |
| >45 | 147 (39.2) | 97 (49.5) | 0.012 | 74 (45.4) | 44 (54.3) | 0.19 |
| ≤45 | 228 (60.8) | 99 (50.5) | | 89 (54.6) | 37 (45.7) | |
| Nodule size, cm | 1.2 (0.79, 1.86) | 2.54 (1.40, 3.73) | <0.001 | 1.17 (0.76, 2.10) | 2.9 (1.3, 4.2) | <0.001 |
| >1.0 | 223 (59.5) | 164 (83.7) | <0.001 | 91 (55.8) | 68 (84.0) | <0.001 |
| ≤1.0 | 152 (40.5) | 32 (16.3) | | 72 (44.2) | 13 (16.0) | |
| Location | | | 0.53 | | | 0.008 |
| Left lobe | 156 (41.6) | 85 (43.4) | | 52 (31.9) | 42 (51.9) | |
| Right lobe | 193 (51.5) | 102 (52.0) | | 97 (59.5) | 36 (44.4) | |
| Isthmus | 26 (6.9) | 9 (4.6) | | 14 (8.6) | 3 (3.7) | |
| US features | | | | | | |
| Echogenicity | | | <0.001 | | | <0.001 |
| Hyper-/iso-echoic | 48 (12.8) | 60 (30.6) | | 18 (11.0) | 24 (29.6) | |
| Hypo-echoic | 327 (87.2) | 136 (69.4) | | 145 (89.0) | 57 (70.4) | |
| Homogeneity | | | 0.047 | | | 0.17 |
| Homogeneous | 9 (2.4) | 11 (5.6) | | 3 (1.8) | 4 (4.9) | |
| Heterogeneous | 366 (97.6) | 185 (94.4) | | 160 (98.2) | 77 (95.1) | |
| Aspect ratio | | | <0.001 | | | <0.001 |
| >1 | 196 (52.3) | 12 (6.1) | | 70 (42.9) | 5 (6.2) | |
| ≤1 | 179 (47.7) | 184 (93.9) | | 93 (57.1) | 76 (93.8) | |
| Margin | | | <0.001 | | | <0.001 |
| Regular | 50 (13.3) | 169 (86.2) | | 20 (12.3) | 70 (86.42) | |
| Irregular | 325 (86.7) | 27 (13.8) | | 143 (87.7) | 11 (13.6) | |
| Microcalcification | | | <0.001 | | | <0.001 |
| Present | 106 (28.3) | 7 (3.6) | | 49 (30.1) | 1 (1,2) | |
| Absent | 269 (71.7) | 189 (96.4) | | 114 (69.9) | 80 (98.8) | |
| CEUS features | | | | | | |
| Enhancement intensity | | | <0.001 | | | <0.001 |
| Hyper-/iso-enhancement | 109 (29.1) | 158 (80.6) | | 33 (20.3) | 69 (85.2) | |
| Hypo-enhancement | 266 (70.9) | 38 (19.4) | | 130 (79.7) | 12 (14.8) | |
| Enhancement pattern | | | <0.001 | | | <0.001 |
| Homogeneous | 38 (10.1) | 58 (29.6) | | 16 (9.8) | 24 (29.6) | |
| Heterogeneous | 337 (89.9) | 138 (70.4) | | 147 (90.2) | 57 (70.4) | |
| Ring enhancement | | | <0.001 | | | <0.001 |
| Present | 4 (1.1) | 100 (51.0) | | 1 (0.6) | 47 (58.0) | |
| Absent | 371 (98.9) | 96 (49.0) | | 162 (99.4) | 34 (42.0) | |

Categorical variables are presented as n (%) and continuous variables are presented as mean ± standard deviation or median (interquartile range). CEUS, contrast-enhanced ultrasound; US, ultrasound.

**Table 3** Multivariate logistic regression analysis for risk factors of malignant thyroid nodules in the training cohort

| Variables | β | Odds ratio | 95% CI | P value |
|---|---|---|---|---|
| Margin (irregular) | 2.477 | 11.9 | 6.517–21.730 | <0.001 |
| Aspect ratio (>1) | 1.702 | 5.487 | 2.486–12.109 | <0.001 |
| Microcalcification (present) | 1.494 | 4.454 | 1.745–11.370 | 0.002 |
| Enhancement intensity (hypo) | 1.016 | 2.762 | 1.486–5.135 | 0.001 |
| Ring enhancement (absent) | 2.984 | 19.761 | 6.263–62.350 | <0.001 |
| Constant | –4.218 | 0.015 | 0.005–0.046 | <0.001 |

CI, confidence interval.

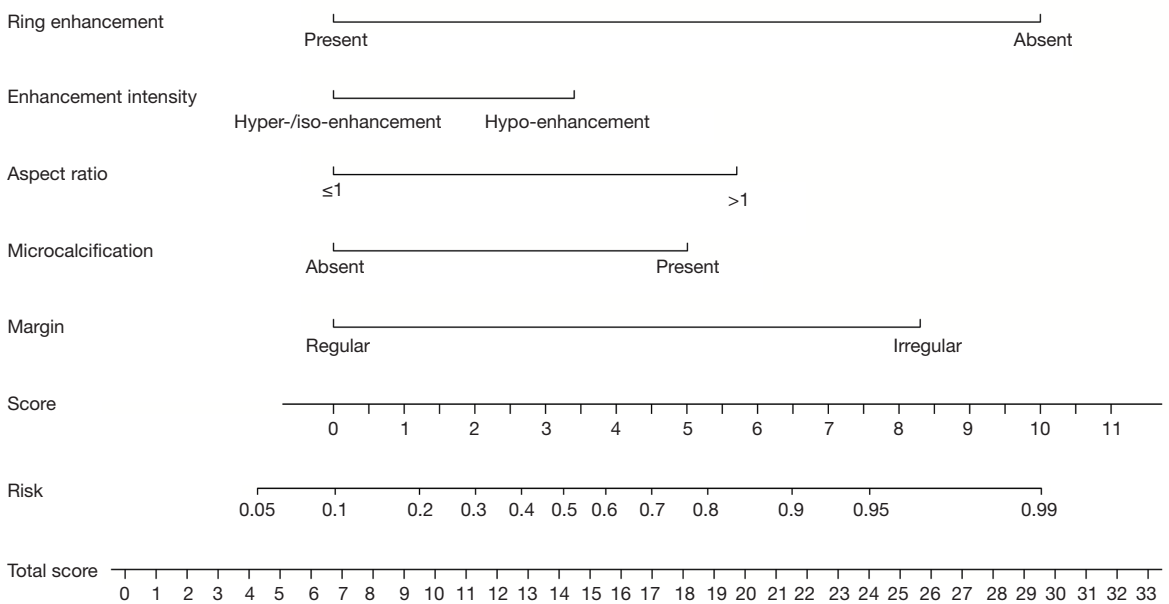


**Figure 3** Predictive nomogram to assess the malignancy risk of thyroid nodules.

training and validation cohorts are plotted in *Figure 6*. The prediction model conferred more clinical net benefits than any single feature in predicting malignant nodules at all threshold probabilities.

## Discussion

In the present study, we developed and validated a prediction model incorporating features from conventional and CEUS images for the differential diagnosis of thyroid nodules. These features encompassed aspect ratio, margins, and microcalcification from conventional US modality, and enhancement intensity and ring enhancement from CEUS modality. The prediction model demonstrated excellent diagnostic performance in both the training and validation cohorts, outperforming any single conventional US and CEUS feature, as well as reducing the unnecessary FNAB rates. An easy-to-use nomogram was subsequently constructed to facilitate the individualized prediction of malignant thyroid nodules in clinical practice.

The incidence of thyroid cancer has rapidly increased in recent years, but the mortality rate has remained relatively low and unchanged (23,24). In China, the extensive utilization of imaging modalities, mainly conventional US examinations, and the confusing use of various TIRADS classification systems have led to overdiagnosis and overtreatment of thyroid nodules (13,25). Recently, a modified TIRADS, named as C-TIRADS, has been established to fit China's medical conditions for malignancy risk stratification of thyroid nodules (13,26). C-TIRADS

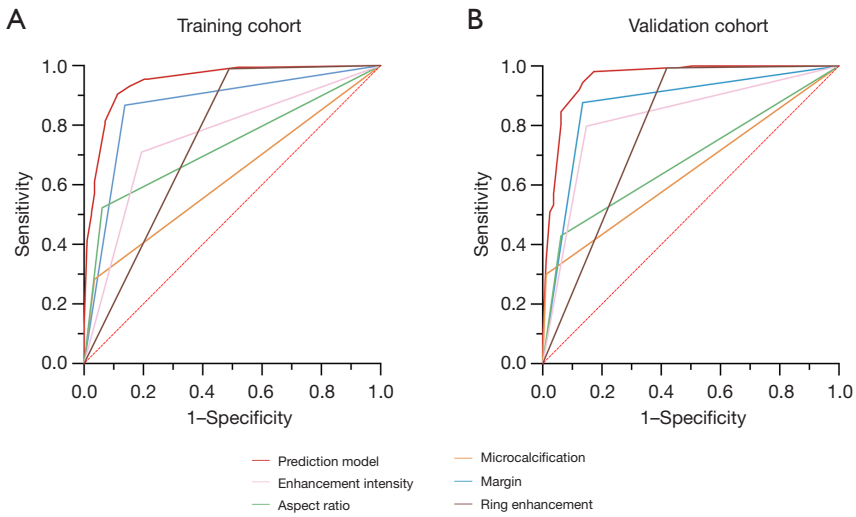


**Figure 4** ROC curves of the prediction model and single ultrasound feature in the training cohort (A) and validation cohort (B). ROC, receiver operating characteristic.

**Table 4** Predictive performance of the prediction model in the training and validation cohorts

| Model | AUC (95% CI) | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|
| Training cohort | | | | |
| Prediction model | 0.947 (0.928–0.966) | 89.8% | 90.4% | 88.8% |
| Margin | 0.865 (0.835–0.894) | 86.5% | 86.7% | 86.2% |
| Aspect ratio | 0.731 (0.700–0.761) | 66.6% | 52.3% | 93.9% |
| Microcalcification | 0.624 (0.597–0.650) | 51.7% | 28.3% | 96.4% |
| Enhancement intensity | 0.758 (0.722–0.794) | 74.3% | 70.9% | 80.6% |
| Ring enhancement | 0.750 (0.714–0.785) | 82.5% | 98.9% | 51.0% |
| Validation cohort | | | | |
| Prediction model | 0.957 (0.928–0.986) | 91.8% | 94.5% | 86.4% |
| Margin | 0.871 (0.826–0.915) | 87.3% | 87.7% | 86.4% |
| Aspect ratio | 0.684 (0.638–0.730) | 59.8% | 42.9% | 93.8% |
| Microcalcification | 0.644 (0.607–0.681) | 52.9% | 30.1% | 98.8% |
| Enhancement intensity | 0.825 (0.775–0.874) | 81.6% | 79.8% | 85.2% |
| Ring enhancement | 0.787 (0.733–0.841) | 85.7% | 99.4% | 58.0% |

AUC, area under the curve; CI, confidence interval.

has shown good performance in differentiating between malignant and benign nodules, especially with the C-TIRADS 4B classification, which achieved a sensitivity of 94% and a specificity of 70% (27). Comparisons between C-TIRADS and other classification systems such as ACR-TIRADS, K-TIRADS, EU-TIRADS, and the ATA guidelines have been conducted in the Chinese population (16,28-30). Despite comparable discrimination efficacy among these systems, C-TIRADS has shown some advantages in reducing the unnecessary FNAB rate (16,28,30). Nevertheless, the unnecessary FNAB rate of C-TIRADS remains high at 20–60% (16,28,30). In our

**Table 5** Comparison of unnecessary FNAB rates of the nomogram model and C-TIRADS

| Models | No. of recommended FNABs | No. of malignant nodules | No. of benign nodules | Unnecessary FNAB rates, % | P value |
|---|---|---|---|---|---|
| Nomogram model | | | | | |
| Training cohort | 361 | 339 | 22 | 6.1 (22/361) | – |
| Validation cohort | 165 | 154 | 11 | 6.7 (11/165) | – |
| All | 526 | 493 | 33 | 6.3 (33/526) | – |
| C-TIRADS | | | | | |
| Training cohort | 382 | 269 | 113 | 29.6 (113/382) | <0.001 |
| Validation cohort | 147 | 104 | 43 | 29.3 (43/147) | <0.001 |
| All | 529 | 373 | 156 | 29.5 (156/529) | <0.001 |

C-TIRADS, Chinese-Thyroid Imaging Reporting and Data System; FNAB, fine-needle aspiration biopsy.

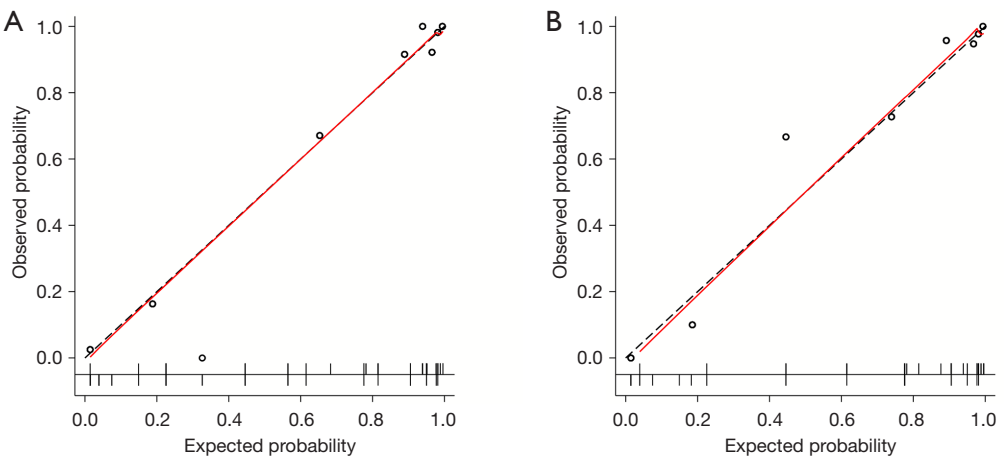


**Figure 5** Calibration curves of the prediction model in the training cohort (A) and validation cohort (B).

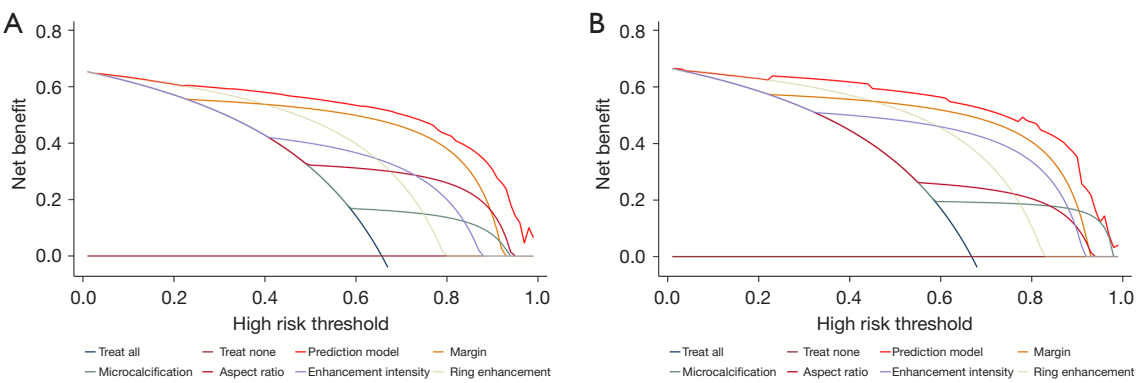


**Figure 6** DCA of the prediction model and single ultrasound feature in the training cohort (A) and validation cohort (B). DCA, decision curve analysis.

study, the unnecessary FNAB rates of C-TIRADS were 29.6% in the training cohort and 29.3% in the validation cohort. When applying the constructed nomogram model, the unnecessary FNAB rates decreased to 6.1% and 6.7%, respectively.

TIRADS classification systems are established based on conventional US features, whereas the emerging CEUS technology provides additional information on microvascular perfusion. The combined use of CEUS and TIRADS has gained much attention, where nodules of specific TIRADS categories were re-graded by CEUS schedule and assigned a new CEUS-TIRADS level or score (31-36). These studies demonstrated that the combination of CEUS and TIRADS significantly improved the diagnostic accuracy of malignant thyroid nodules compared to TIRADS alone and reduced unnecessary FNAB rate (31-36). Despite better performance, these CEUS + TIRADS re-grading systems are more complex than traditional TIRADS and may be challenging in clinical implementation.

Recent studies have tried to construct effective and easy-to-use nomogram models for predicting the risk of malignant thyroid nodules by incorporating multi-model US features. For instance, Yi *et al.* developed a nomogram using conventional US, elastography (ES), and CEUS data from 447 participants, incorporating features such as shape, ES score, ring enhancement, and margin after enhancement (37). This nomogram model achieved an AUC of 0.936 in the training dataset and 0.902 in the validation dataset. Similarly, Pang *et al.* created a nomogram that included the ACR-TIRADS category, ES score, and enhancement pattern in ACR-TIRADS 3–5 nodules, yielding an AUC of 0.93 in both training and validation cohorts (38). However, the malignancy of some nodules in these studies was confirmed by cytological evaluations after FNAB, which potentially introduced false-negative results (37,39).

In our study, all thyroid nodules were surgically confirmed. We identified several high-risk conventional US features for malignancy, consistent with previous studies assessing the discriminative ability of sonographic features in thyroid nodules (40). Features such as irregular margins, aspect ratio >1, and presence of microcalcification were identified as independent predictive factors for malignancy and included in the final prediction model. In CEUS, the hypo-enhancement intensity has been recognized as the strongest and best-established indicator of malignancy. Hypo-enhancement is often associated with inadequate neovascularization and insufficient blood supply in malignant nodules, especially those with a maximum diameter ≤1.0 cm (41,42). Our study found that malignant nodules were significantly smaller than benign ones [1.2 (IQR: 0.78–1.9) *vs.* 2.7 (IQR: 1.4–3.9) cm, P<0.001]. In the entire cohort, a markedly higher proportion of malignant nodules had a size ≤1.0 cm compared to benign nodules (41.6% *vs.* 16.2%, P<0.001). Among the 269 small nodules (diameter ≤1.0 cm) in our study, hypo-enhancement was present in 77.2% of malignant nodules but only 40.0% of benign nodules. The existence of calcification, focal necrosis, and fibrosis in thyroid lesions may cause a heterogenous enhancement pattern, another CEUS risk feature for malignancy (43). In our study, 90.0% of malignant nodules and 70.4% of benign nodules exhibited a heterogeneous pattern (P<0.001). However, this feature was excluded from the prediction model after multivariate analysis. Therefore, our final prediction nomogram model was constructed using 5 features—aspect ratio, margin, microcalcification, enhancement intensity, and ring enhancement—yielding an AUC of 0.947 in the training cohort and 0.957 in the validation cohort.

Our study integrated key conventional US and CEUS features into a single, easy-to-use nomogram. This nomogram not only enhanced diagnostic accuracy but also maintains simplicity for clinical application. The high AUC values in both training (0.947) and validation (0.957) cohorts indicated robust performance. Moreover, our prediction model, based on conventional US and CEUS features, showed a significantly better diagnostic performance in the validation cohort than those established by other studies incorporating additional ES scores (37,39).

There were several limitations in our study. Firstly, the retrospective nature of the study may have introduced selection bias. Secondly, all thyroid nodules were collected from a single institution. The reproducibility of our prediction model needs to be investigated through prospective, multi-center studies in the future. Thirdly, the features of conventional US and CEUS are subjective, leading to unavoidable inconsistencies between radiologists. With the progress in artificial intelligence, radiomics-based nomogram models have shown promising outcomes in more effectively and objectively differentiating malignant from benign nodules (44,45).

## Conclusions

Our study presents a validated prediction nomogram

model that combines conventional US and CEUS features, offering an accurate and practical tool for the differential diagnosis of thyroid nodules. This model has the potential to reduce unnecessary FNAB and improve the clinical management of thyroid nodules.

## Acknowledgments

None.

## Footnote

*Reporting Checklist*: The authors have completed the STARD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-24-1796/rc

*Funding:* None.

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-24-1796/coif). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments. This study was approved by the Ethic Committee of the First People's Hospital of Qinzhou (No. A-20240111). The requirement for informed consent was waived due to the study's retrospective nature.



## References

1. Dean DS, Gharib H. Epidemiology of thyroid nodules. Best Pract Res Clin Endocrinol Metab 2008;22:901-11.
2. Durante C, Grani G, Lamartina L, Filetti S, Mandel SJ, Cooper DS. The Diagnosis and Management of Thyroid Nodules: A Review. JAMA 2018;319:914-24.
3. Wolinski K, Stangierski A, Ruchala M. Comparison of diagnostic yield of core-needle and fine-needle aspiration biopsies of thyroid lesions: Systematic review and meta-analysis. Eur Radiol 2017;27:431-6.
4. Baier ND, Hahn PF, Gervais DA, Samir A, Halpern EF, Mueller PR, Harisinghani MG. Fine-needle aspiration biopsy of thyroid nodules: experience in a cohort of 944 patients. AJR Am J Roentgenol 2009;193:1175-9.
5. Raab SS, Vrbin CM, Grzybicki DM, Sudilovsky D, Balassanian R, Zarbo RJ, Meier FA. Errors in thyroid gland fine-needle aspiration. Am J Clin Pathol 2006;125:873-82.
6. Seningen JL, Nassar A, Henry MR. Correlation of thyroid nodule fine-needle aspiration cytology with corresponding histology at Mayo Clinic, 2001-2007: an institutional experience of 1,945 cases. Diagn Cytopathol 2012;40 Suppl 1:E27-32.
7. Tae HJ, Lim DJ, Baek KH, Park WC, Lee YS, Choi JE, Lee JM, Kang MI, Cha BY, Son HY, Lee KW, Kang SK. Diagnostic value of ultrasonography to distinguish between benign and malignant lesions in the management of thyroid nodules. Thyroid 2007;17:461-6.
8. Melany M, Chen S. Thyroid Cancer: Ultrasound Imaging and Fine-Needle Aspiration Biopsy. Endocrinol Metab Clin North Am 2017;46:691-711.
9. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, Cronan JJ, Beland MD, Desser TS, Frates MC, Hammers LW, Hamper UM, Langer JE, Reading CC, Scoutt LM, Stavros AT. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. J Am Coll Radiol 2017;14:587-95.
10. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L. 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;26:1-133.
11. Shin JH, Baek JH, Chung J, Ha EJ, Kim JH, Lee YH, et al. Ultrasonography Diagnosis and Imaging-Based Management of Thyroid Nodules: Revised Korean Society of Thyroid Radiology Consensus Statement and Recommendations. Korean J Radiol 2016;17:370-95.
12. 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;6:225-37.
13. 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;70:256-79.
14. Li W, Wang Y, Wen J, Zhang L, Sun Y. Diagnostic Performance of American College of Radiology TI-RADS: A Systematic Review and Meta-Analysis. AJR Am J Roentgenol 2021;216:38-47.
15. Huang BL, Ebner SA, Makkar JS, Bentley-Hibbert S, McConnell RJ, Lee JA, Hecht EM, Kuo JH. A Multidisciplinary Head-to-Head Comparison of American College of Radiology Thyroid Imaging and Reporting Data System and American Thyroid Association Ultrasound Risk Stratification Systems. Oncologist 2020;25:398-403.
16. Jin Z, Pei S, Shen H, Ouyang L, Zhang L, Mo X, Chen Q, You J, Zhang S, Zhang B. Comparative Study of C-TIRADS, ACR-TIRADS, and EU-TIRADS for Diagnosis and Management of Thyroid Nodules. Acad Radiol 2023;30:2181-91.
17. Zhan J, Ding H. Application of contrast-enhanced ultrasound for evaluation of thyroid nodules. Ultrasonography 2018;37:288-97.
18. Cantisani V, Bertolotto M, Weskott HP, Romanini L, Grazhdani H, Passamonti M, Drudi FM, Malpassini F, Isidori A, Meloni FM, Calliada F, D'Ambrosio F. Growing indications for CEUS: The kidney, testis, lymph nodes, thyroid, prostate, and small bowel. Eur J Radiol 2015;84:1675-84.
19. Ding Y, Peng Y, Zhang J, Pan X, Huang X, Zhang CQ. Diagnostic value of contrast-enhanced ultrasound in the diagnosis of papillary thyroid microcarcinoma: A systematic review and meta-analysis. Medicine (Baltimore) 2024;103:e37768.
20. Trimboli P, Castellana M, Virili C, Havre RF, Bini F, Marinozzi F, D'Ambrosio F, Giorgino F, Giovanella L, Prosch H, Grani G, Radzina M, Cantisani V. Performance of contrast-enhanced ultrasound (CEUS) in assessing thyroid nodules: a systematic review and meta-analysis using histological standard of reference. Radiol Med 2020;125:406-15.
21. 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;11:1067-77.
22. 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:e0231775.
23. La Vecchia C, Malvezzi M, Bosetti C, Garavello W, Bertuccio P, Levi F, Negri E. Thyroid cancer mortality and incidence: a global overview. Int J Cancer 2015;136:2187-95.
24. La Vecchia C, Negri E. Thyroid cancer: The thyroid cancer epidemic - overdiagnosis or a real increase? Nat Rev Endocrinol 2017;13:318-9.
25. Vaccarella S, Franceschi S, Bray F, Wild CP, Plummer M, Dal Maso L. Worldwide Thyroid-Cancer Epidemic? The Increasing Impact of Overdiagnosis. N Engl J Med 2016;375:614-7.
26. Zhou J, Song Y, Zhan W, Wei X, Zhang S, Zhang R, et al. Thyroid imaging reporting and data system (TIRADS) for ultrasound features of nodules: multicentric retrospective study in China. Endocrine 2021;72:157-70.
27. Hu Y, Xu S, Zhan W. Diagnostic performance of C-TIRADS in malignancy risk stratification of thyroid nodules: A systematic review and meta-analysis. Front Endocrinol (Lausanne) 2022;13:938961.
28. Cai Y, Yang R, Yang S, Lu L, Ma R, Xiao Z, Lin N, Huang Y, Chen L. Comparison of the C-TIRADS, ACR-TIRADS, and ATA guidelines in malignancy risk stratification of thyroid nodules. Quant Imaging Med Surg 2023;13:4514-25.
29. Huang H, Zhu MJ, Gao Q, Huang YL, Li WM. Comparison of Diagnostic Values of ACR TI-RADS versus C-TIRADS Scoring and Classification Systems for the Elderly Thyroid Cancers. Int J Gen Med 2023;16:4441-51.
30. Zhu H, Yang Y, Wu S, Chen K, Luo H, Huang J. Diagnostic performance of US-based FNAB criteria of the 2020 Chinese guideline for malignant thyroid nodules: comparison with the 2017 American College of Radiology guideline, the 2015 American Thyroid Association guideline, and the 2016 Korean Thyroid Association guideline. Quant Imaging Med Surg 2021;11:3604-18.
31. Huang Y, Hong Y, Xu W, Song K, Huang P. Contrast-Enhanced Ultrasound Improves the Accuracy of the ACR TI-RADS in the Diagnosis of Thyroid Nodules Located in the Isthmus. Ultraschall Med 2022;43:599-607.
32. Li J, Dou J, Li H, Xiao F, Yu J, Xie M, Zhou P, Liang L, Zhou G, Che Y, Liu C, Cong Z, Liu F, Han Z, Liang P. Contrast Enhancement Ultrasound Improves Diagnostic

 

Accuracy for Thyroid Nodules: A Prospective Multicenter Study. J Endocr Soc 2023;8:bvad145.
33. 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.
34. Zhu T, Chen J, Zhou Z, Ma X, Huang Y. Differentiation of Thyroid Nodules (C-TIRADS 4) by Combining Contrast-Enhanced Ultrasound Diagnosis Model With Chinese Thyroid Imaging Reporting and Data System. Front Oncol 2022;12:840819.
35. Si CF, Yu J, Cui YY, Huang YJ, Cui KF, Fu C. Comparison of diagnostic performance of the current score-based ultrasound risk stratification systems according to thyroid nodule size. Quant Imaging Med Surg 2024;14:9234-45.
36. Yang YP, Zhang GL, Zhou HL, Dai HX, Huang X, Liu LJ, Xie J, Wang JX, Li HJ, Liang X, Yuan Q, Zeng YH, Xu XH. Diagnostic efficacy of the contrast-enhanced ultrasound thyroid imaging reporting and data system classification for benign and malignant thyroid nodules. Quant Imaging Med Surg 2024;14:5721-36.
37. Yi D, Fan L, Zhu J, Yao J, Peng C, Xu D. The diagnostic value of a nomogram based on multimodal ultrasonography for thyroid-nodule differentiation: A multicenter study. Front Oncol 2022;12:970758.
38. Pang L, Yang X, Zhang P, Ding L, Yuan J, Liu H, Liu J, Gong X, Yu M, Luo W. Development and Validation of a Nomogram Based on Multimodality Ultrasonography Images for Differentiating Malignant from Benign American College of Radiology Thyroid Imaging, Reporting and Data System (TI-RADS) 3-5 Thyroid Nodules. Ultrasound Med Biol 2024;50:557-63.
39. Guo BL, Ouyang FS, Ouyang LZ, Liu ZW, Lin SJ, Meng W, Huang XY, Chen HX, Yang SM, Hu QG. Development and validation of an ultrasound-based nomogram to improve the diagnostic accuracy for malignant thyroid nodules. Eur Radiol 2019;29:1518-26.
40. Remonti LR, Kramer CK, Leitão CB, Pinto LC, Gross JL. Thyroid ultrasound features and risk of carcinoma: a systematic review and meta-analysis of observational studies. Thyroid 2015;25:538-50.
41. Bartolotta TV, Midiri M, Galia M, Runza G, Attard M, Savoia G, Lagalla R, Cardinale AE. Qualitative and quantitative evaluation of solitary thyroid nodules with contrast-enhanced ultrasound: initial results. Eur Radiol 2006;16:2234-41.
42. Zhang Y, Luo YK, Zhang MB, Li J, Li J, Tang J. Diagnostic Accuracy of Contrast-Enhanced Ultrasound Enhancement Patterns for Thyroid Nodules. Med Sci Monit 2016;22:4755-64.
43. Zhang B, Jiang YX, Liu JB, Yang M, Dai Q, Zhu QL, Gao P. Utility of contrast-enhanced ultrasound for evaluation of thyroid nodules. Thyroid 2010;20:51-7.
44. Ren JY, Lin JJ, Lv WZ, Zhang XY, Li XQ, Xu T, Peng YX, Wang Y, Cui XW. A Comparative Study of Two Radiomics-Based Blood Flow Modes with Thyroid Imaging Reporting and Data System in Predicting Malignancy of Thyroid Nodules and Reducing Unnecessary Fine-Needle Aspiration Rate. Acad Radiol 2024;31:2739-52.
45. Ren JY, Lv WZ, Wang L, Zhang W, Ma YY, Huang YZ, Peng YX, Lin JJ, Cui XW. Dual-modal radiomics nomogram based on contrast-enhanced ultrasound to improve differential diagnostic accuracy and reduce unnecessary biopsy rate in ACR TI-RADS 4-5 thyroid nodules. Cancer Imaging 2024;24:17.

**Cite this article as:** Wang QG, Li M, Deng GX, Huang HQ, Qiu Q, Lin JJ. Development and validation of a nomogram based on conventional and contrast-enhanced ultrasound for differentiating malignant from benign thyroid nodules. Quant Imaging Med Surg 2025;15(5):4641-4654. doi: 10.21037/qims-24-1796