Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Clinical-Ultrasound Nomogram for Malignancy Prediction in Thyroid Nodules with Follicular Neoplasm-Like Sonographic Features: Comparison with Four International TIRADS Systems

Zehui Fu
the First People's Hospital of Changzhou
Mengyuan Li
the First People's Hospital of Changzhou
Hui Chen
the First People's Hospital of Changzhou
Feng Qian
the First People's Hospital of Changzhou
Xiao Xie
the First People's Hospital of Changzhou
Ying He
the First People's Hospital of Changzhou
Jian Zhang
the First People's Hospital of Changzhou
Yejun Lu
the First People's Hospital of Changzhou
Siqi Wang
siqiwangg@sina.com

the First People's Hospital of Changzhou

Research Article

Keywords: Ultrasound, Thyroid nodule, Follicular neoplasm, Nomogram, TIRADS

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11398320/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Clinical-Ultrasound Nomogram for Malignancy Prediction in Thyroid Nodules with Follicular Neoplasm-Like Sonographic Features: Comparison with Four International TIRADS Systems

Zehui Fu, Department of Ultrasound, the First People's Hospital of Changzhou, Changzhou, China
Mengyuan Li, Department of Ultrasound, the First People's Hospital of Changzhou, Changzhou, China.
Hui Chen, Department of Ultrasound, the First People's Hospital of Changzhou, Changzhou, China.
Feng Qian, Department of Ultrasound, the First People's Hospital of Changzhou, Changzhou, China.
Xiao Xie, Department of Ultrasound, the First People's Hospital of Changzhou, Changzhou, China.
Ying He, Department of Ultrasound, the First People's Hospital of Changzhou, Changzhou, China.
Jian Zhang, Department of Ultrasound, the First People's Hospital of Changzhou, Changzhou,
Yejun Lu, Department of Ultrasound, the First People's Hospital of Changzhou, Changzhou, China.
Siqi Wang, Department of Radiology, the First People's Hospital of Changzhou, Changzhou, China.
Corresponding author: Siqi Wang, E-mail: siqiwangg@sina.com

Abstract

Background: Thyroid nodules with follicular neoplasm-like ultrasound features (FNLUF-TNs) are frequently misdiagnosed due to their inconspicuous malignant manifestations. We aimed to develop and validate a clinical-ultrasound nomogram for predicting malignancy risk in FNLUF-TNs, and to compare its diagnostic performance with conventional TIRADS.

Methods: This retrospective study included patients with 458 FNLUF-TNs (defined as oval, solid/predominantly solid, well-circumscribed nodules) who underwent preoperative ultrasound and surgical resection. Univariate and multivariate analyses were conducted to screen out independent predictors of benign and malignant nodules, and a nomogram model based on clinical and ultrasonic features was constructed. The diagnostic performance of the nomogram was assessed by calculating sensitivity, specificity, accuracy, and AUC, and further compared with the corresponding performance of four TIRADS. Calibration, clinical utility, and interobserver agreement were also evaluated.

Results: Malignancy was confirmed in 127 of 458 nodules (27.7%). Independent predictors of malignancy included younger age, male gender, smaller transverse diameter, hypoechogenicity, calcification, trabecular formation, and irregular interrupted halo (all $p < 0.05$). The nomogram achieved an AUC of 0.824 (95% CI 0.786–0.858), accuracy of 87.55%, specificity of 93.96%, significantly outperforming all four TIRADS systems (all $p < 0.001$). Calibration curves showed good consistency between predicted and actual malignancy risk. Decision curve analysis demonstrated

favorable net clinical benefit. Interobserver agreement for key ultrasound features was excellent (all $\kappa > 0.81$).

Conclusion: This clinical ultrasound nomogram provides accurate and individualized malignancy risk stratification for FNLUF-TNs, outperforming conventional TIRADS in diagnostic performance while showing good reproducibility and clinical value, which may help optimize management and minimize unnecessary interventions for this challenging subgroup.

Trial registration: Not applicable.

Keywords Ultrasound, Thyroid nodule, Follicular neoplasm, Nomogram, TIRADS

Background

Thyroid carcinoma is the most prevalent malignant tumor of the endocrine system, accounting for approximately 90% of all endocrine malignancies(1, 2). Ultrasound stands as the preferred imaging modality for differentiating benign from malignant thyroid nodules, and several ultrasound risk stratification systems (RSSs) have been developed to further enhance their diagnostic performance in accurately assessing the malignant risk of such lesions(3). Among these, the American College of Radiology Thyroid Imaging Reporting and Data System (ACR-TIRADS)(4), the Korean Society of Thyroid Radiology Thyroid Imaging Reporting and Data System (KSThR-TIRADS)(5), and the Chinese Thyroid Imaging Reporting and Data System (C-TIRADS)(6) are representative thyroid reporting systems from different geographical

regions that have proven highly valuable in the diagnosis of papillary thyroid carcinoma.

However, preoperative differentiation of follicular thyroid neoplasm (FTN) — encompassing follicular thyroid carcinoma (FTC) and follicular thyroid adenoma (FTA)—remains challenging, as these lesions share overlapping clinical manifestations and sonographic features(7, 8). Notably, even fine needle aspiration cytology (FNAC) fails to reliably distinguish benign from malignant entities within FTN(1). Histopathological examination is the only gold standard for distinguishing an FTA from FTC(9). The advent of Follicular Thyroid Imaging Reporting and Data System (F-TIRADS) has, to a certain extent, facilitated the differentiation between FTC and FTA and outperformed existing RSSs(10).

Nevertheless, F-TIRADS was established based on a retrospective analysis of the features of nodules pathologically diagnosed as FTA or FTC postoperatively. Moreover, most previous studies have focused primarily on differentiating FTA from FTC(1, 11). In clinical practice, a substantial number of nodules with similar sonographic features are actually of other pathological types, such as nodular goiter, papillary thyroid carcinoma, and medullary thyroid carcinoma. It remains a major challenge to make an accurate preoperative diagnosis of such nodules(12).

In this study, we innovatively defined nodules presenting with an oval shape, solid or solid-predominant composition, and well-defined borders as thyroid nodules with

follicular neoplasm-like sonographic features (FNLUF-TNs). We then analyzed their sonographic features and compared the diagnostic performance of four TIRADS for these lesions, with the aim of improving preoperative diagnostic efficacy for this subset of thyroid nodules.

Methods

This retrospective study was approved by the ethics committee. Informed consent was waived due to retrospective design.

Patients

Thyroid sonographic images were consecutively collected at the First People's Hospital of Changzhou between January 2023 and August 2025. The inclusion criteria were (1) histopathological results were obtained, (2) thyroid nodules demonstrated features of FNLUF-TNs, (3) no history of thyroid surgery or preoperative treatment like thermal ablation therapy. The exclusion criteria were (1) duplicate examinations, (2) absence of definitive diagnosis, (3) unclear ultrasound images. Based on the above criteria, a total of 442 patients with 458 thyroid nodules were finally enrolled, including 331 benign and 127 malignant surgically confirmed nodules. The flowchart in **Fig. 1** illustrates the detailed process of patient inclusion and exclusion.

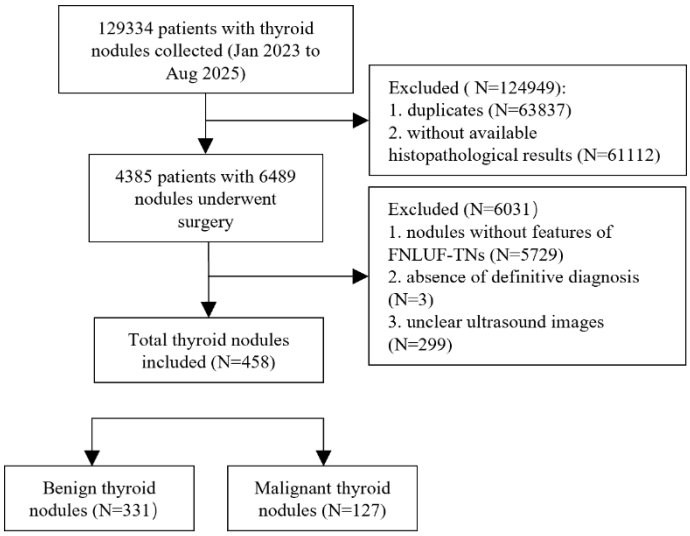


Fig. 1 Flowchart of the included patients

Image acquisition and evaluation

Conventional ultrasound examinations were performed using various models of color Doppler ultrasound systems (Philips Epiq7, Philips Epiq5, GE Logiq E9, Mindray Resona E9 Elite, SUMSUN XR85 et al) equipped with 5 – 12-MHz linear array transducers. Two radiologists with 5 and 12 years of experience in thyroid, independently analyzed the images, recorded the grayscale features of nodules, and classified them according to ACR, K, C, and F TIRADS criteria. They were all blinded to the patients' clinical information and pathological diagnosis.

Ultrasound features of the nodules included shape (wider-than-tall or taller-than-wide), composition (predominantly solid or solid), echogenicity (hyperechoic, isoechoic or hypoechoic), calcification (none, macrocalcification, microcalcification or rim calcification), trabecular formation (absent or present), halo (absent, regular complete or irregular interrupted) and diffuse lesions (absent or present). Microcalcification was

defined as a punctate echogenic focus measuring ≤1 mm, with or without posterior acoustic shadowing within the solid component. Macrocalcification was defined as an echogenic focus >1 mm accompanied by posterior acoustic shadowing. Rim calcification was defined as a peripheral curvilinear hyperechoic line (complete or incomplete) (13). The trabecular formation refers to a radial central scar or reticularly extending hypoechoic bands within the nodule on ultrasound images(14).

Blood flow within nodules was evaluated using color Doppler flow imaging and graded semi-quantitatively according to the Adler criteria(15): grade 0 (no flow), grade 1 (1–2 dot like signals), grade 2 (3–4 dot like or one long vessel), and grade 3 (≥5 dot like or ≥2 long vessels).

Statistical analysis

Statistical analyses were performed using SPSS 22.0 software (IBM Corp, Chicago, Illinois, USA), MedCalc 18.2.1 software (MedCalc Software, Ostend, Belgium), and R software (Version 4.3.3; R Foundation for Statistical Computing, Vienna, Austria). Cohen's Kappa was used interobserver agreement of ultrasound features with values 0.01–0.20 slight agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement and 0.81–0.99 almost perfect agreement(16, 17). Quantitative data with a normal distribution were presented as mean ± standard deviation and compared using the independent-samples *t* - test. Non-normally distributed continuous data were expressed as median (interquartile range) and

analyzed with nonparametric tests. Qualitative data were reported as frequencies (percentages) and compared using the chi-squared test or Fisher's exact test. Variables with statistically significant results in univariate analysis were further included in multivariate logistic regression analysis, and a clinical-ultrasound nomogram model was constructed based on the identified independent predictive variables. For each TIRADS system, the optimal cut-off value was determined via receiver operating characteristic curve (ROC) analysis at the maximum Youden index. Sensitivity, specificity and area under the ROC curve (AUC) were calculated for all models. ROC curves were plotted to evaluate the diagnostic discriminatory ability of the nomogram model, and the Delong test was used to compare AUC differences between the nomogram model and the four distinct TIRADS. The Hosmer-Lemeshow test and calibration curves were employed to assess the calibration performance of the models. The nomogram underwent internal validation with 1,000 bootstrap resamples; calibration curve was used to evaluate model consistency. Additionally, decision curve analysis (DCA) was conducted to evaluate the clinical utility of the models at different threshold probabilities by quantifying the net benefit across the entire study cohort. A two-sided P value < 0.05 was considered statistically significant.

Results

Demographic and pathology data

In this study, a total of 442 patients were included, comprising 322 benign and 120 malignant cases. Significant differences were observed between the two groups in terms

of gender distribution and age. The proportion of male patients was higher in the malignant group than in the benign group (36.7% vs. 22.7%, $p = 0.003$). The mean age of patients in the malignant group was significantly lower than that in the benign group (44.79 ± 16.02 years vs. 52.58 ± 14.39 years, $p < 0.001$).

A total of 458 thyroid nodules were histopathologically examined. Among them, 331 nodules were benign, including 132 nodular goiters, 153 follicular adenomas, 2 oncocytic adenomas, 4 parathyroid adenomas, and 36 adenomatous goiters. The remaining 127 nodules were malignant, consisting of 98 papillary thyroid carcinomas, 22 follicular carcinomas, 6 medullary thyroid carcinomas, 2 anaplastic thyroid carcinomas, 2 lymphomas, and 1 squamous cell carcinoma.

Ultrasound features and model construction

The ultrasound features of the FNLUF-TNs in the malignant and benign groups are compared in **Table 1**. Univariate analysis showed that maller transverse diameter, isoechoic and hypoechoic echogenicity, microcalcification, macrocalcification, rim calcification, trabecular formation, absence of a halo, and irregularly interrupted halo were more frequently observed in the malignant group (all $p < 0.05$).There were no significant differences in nodule location or Alder blood flow grade between the two groups (all $p > 0.05$).

Table 1 Ultrasound features of nodules

	Benign	Malignant	t/χ^2	p

		nodules (n=331)	nodules (n=127)		value
Size, cm (mean±SD)					
	Super-inferior diameter	3.05±1.20	2.44±1.14	4.95	0.000
	Transverse diameter	3.99±1.65	3.07±1.44	5.51	0.000
	Anter-posterior diameter	2.32±0.98	2.00±0.90	3.22	0.001
	Maximum diameter	3.99±1.65	3.09±1.43	5.41	0.000
	Mean diameter	3.13±1.23	2.50±1.13	4.96	0.000
Location				3.45	0.178
	Right lobe	165（49.8%）	75（59.1%）		
	Isthmus	9（2.7%）	4（3.1%）		
	Left lobe	157（47.4%）	48（37.8%）		
Shape				10.12	0.002
	Wider-than-tall	321（97.0%）	114（89.8%）		
	Taller-than-wide	10（3.0%）	13（10.2%）		
Composition				16.82	0.000
	Predominantly solid	117（35.3%）	20（15.7%）		
	Solid	214（64.7%）	107（84.3%）		
Echogenicity				62.27	0.000
	Hyperechoic	15（4.5%）	1（0.8%）		
	Isoechoic	283（85.5%）	73（57.5%）		
	Hypoechoic	33（10.0%）	53（41.7%）		
Calcification				97.14	0.000
	None	292（88.2%）	59（46.5%）		
	Macrocalcification	16（4.8%）	14（11.0%）		
	Microcalcification	19（5.7%）	50（39.4%）		
	Rim calcification	4（1.2%）	4（3.1%）		
Trabecular formation				36.21	0.000
	Absent	251（75.8%）	59（46.5%）		
	Present	80（24.2%）	68（53.5%）		
Halo				59.07	0.000
	Absent	239（72.2%）	70（55.1%）		
	Regular complete	71（21.5%）	14（11.0%）		
	Irregular interrupted	21（6.3%）	43（33.9%）		
Diffuse lesions				17.359	0.000
	Absent	322（97.3%）	111（87.4%）		
	Present	9（2.7%）	16（12.6%）		
Blood flow				2.524	0.471
	Grade 0	16（4.8%）	9（7.1%）		
	Grade 1	69（20.8%）	24（18.9%）		
	Grade 2	115（34.7%）	37（29.1%）		
	Grade 3	131（39.6%）	57（27.7%）		

Multivariate logistic regression analysis was performed to identify independent predictors of malignant FNLUF-TNs. Older age (OR = 0.973, 95%CI: 0.955–0.991, $p < 0.05$), female gender (OR = 0.420, 95%CI: 0.232–0.760, $p < 0.05$), and larger transverse diameter (OR = 0.722, 95%CI: 0.598–0.872, $p < 0.05$) were associated with a decreased risk of malignancy. Conversely, hypoechogenicity (OR = 4.978, 95%CI: 2.672–9.274, $p < 0.05$), presence of calcification (OR = 1.466, 95%CI: 1.277–1.683, $p < 0.05$), trabecular formation (OR = 5.116, 95%CI: 2.821–9.277, $p < 0.05$), and presence of a irregular interrupted halo (OR = 2.156, 95%CI: 1.509–3.080, $p < 0.05$) were identified as independent risk factors for malignant nodules(**Fig. 2 and 3**). Among these, trabecular formation and hypoechogenicity exhibited the strongest predictive effects, conferring a more than 5-fold and 4.9-fold increased risk of malignancy, respectively.

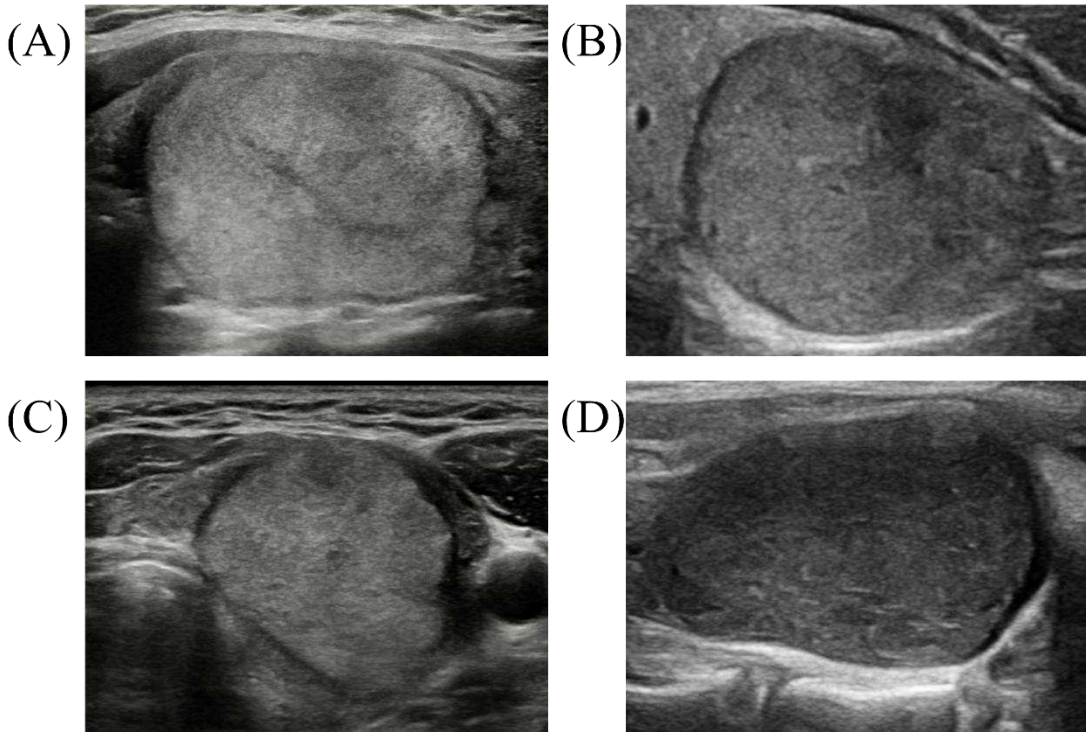


Fig. 2 Grayscale ultrasound image of a benign thyroid neoplasm. **A** and **B** Oncocytic adenoma, **C** Follicular thyroid adenoma, **D** Nodular goiter

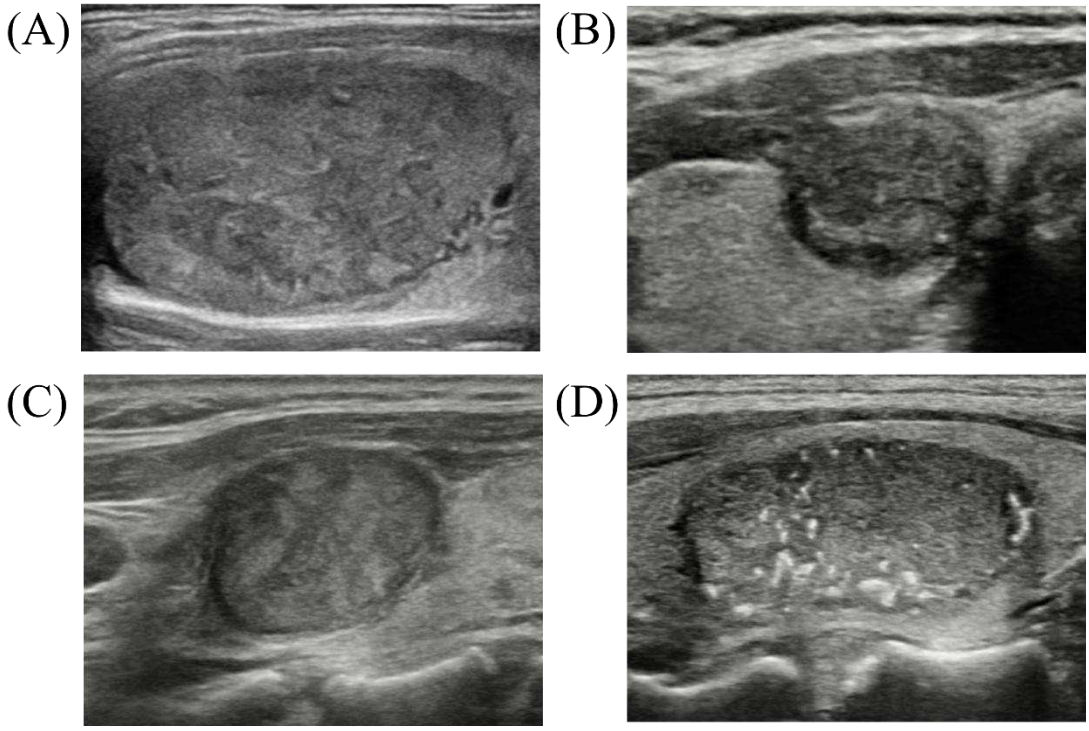


Fig. 3 Grayscale ultrasound image of a malignant thyroid neoplasm. **A** and **B** Papillary thyroid carcinoma, **C** Follicular thyroid carcinoma, **D** Medullary thyroid carcinoma

Based on the above independent factors, a clinical-ultrasound nomogram model was established (**Fig. 4**). Internal validation was performed using the bootstrap method with 1000 resamples. The calibration curve of the nomogram showed good consistency between the predicted probability and the actual probability, suggesting favorable calibration. The DCA demonstrated that the model had a favorable net clinical benefit within a wide range of threshold probabilities, confirming its clinical practicability (**Fig. 5**).

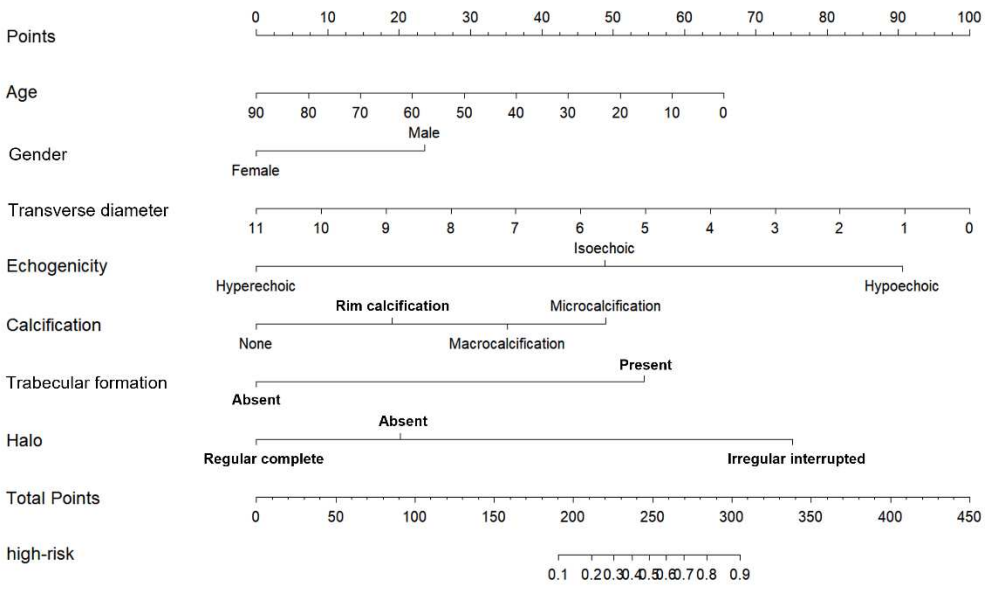


FIGURE. 4 Nomogram prediction model for follicular neoplasm-like thyroid nodules on ultrasonography

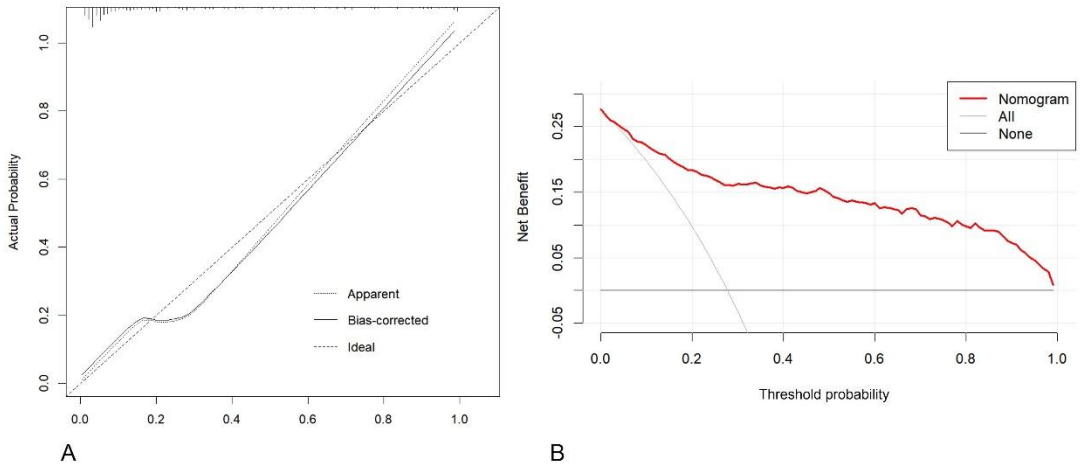


FIGURE. 5 Evaluation indicators of the nomogram model. A. Calibration curve of the nomogram model B. Decision curve of the nomogram model

Comparison of the diagnostic performance

The nomogram model demonstrated superior diagnostic performance compared with the four TIRADS systems. It achieved the highest AUC of 0.824 (95%CI: 0.786–0.858), accuracy of 87.55%, specificity of 93.96%, and Youden index of 0.648, significantly outperforming ACR-TIRADS, KSThR-TIRADS, C-TIRADS, and F-TIRADS (Z = 3.288, 3.389, 6.073 and 6.406, respectively; all $p < 0.001$). Among the TIRADS systems, ACR-TIRADS demonstrated the best overall diagnostic performance (AUC = 0.748, 95%CI: 0.705–0.787), followed by KSThR-TIRADS (AUC = 0.738, 95%CI: 0.695–0.778). C-TIRADS and F-TIRADS showed lower diagnostic efficacy, with AUCs of 0.677 and 0.624, respectively (**Figure. 6, Table 2**). These findings indicate that the clinical-ultrasound nomogram model is a more reliable and accurate tool for predicting malignant follicular neoplasm-like thyroid nodules, providing better individualized risk

stratification than conventional TIRADS systems.

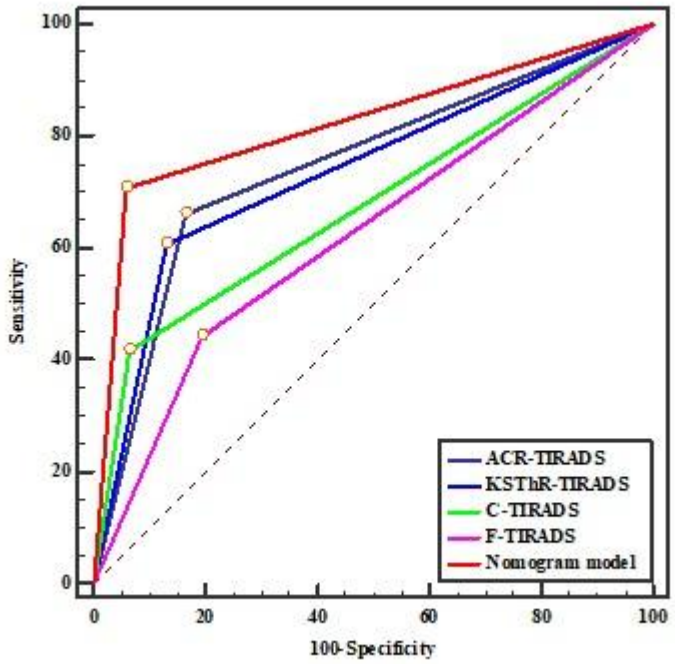


FIGURE. 6 The ROC curves comparison of four TIRADS and nomogram model. TIRADS Thyroid Imaging Reporting and Data System, ACR-TIRADS American College of Radiology TIRADS, KSThR-TIRADS Korean Society of Thyroid Radiology TI-RADS, C-TIRADS Chinese TI-RADS, F-TIRADS Follicular TI-RADS

Table 2 Diagnostic performance of the four TIRADS and nomogram model

	Cut off value	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC(95% CI)	Youden index	Z*	P*
ACR-TIRADS	⩾4	66.14	83.38	78.60	0.748(0.705-0.787)	0.495(0.405-0.579)	3.288	0.001
KSThR-TIRADS	⩾4a	60.63	87.01	79.69	0.738(0.695-0.778)	0.476(0.379-0.566)	3.389	0.000
C-TIRADS	⩾4b	41.73	93.66	79.26	0.677(0.632-0.720)	0.354(0.273-0.449)	6.073	0.000

F-TIRADS	⩾6	55.07	72.35	70.52	0.624(0.578-0.668)	0.274(0.144-0.397)	6.406	0.000
Nomogram model	> 0.476	70.87	93.96	87.55	0.824(0.786-0.858)	0.648(0.566-0.719)	—	—

TIRADS Thyroid Imaging Reporting and Data System, ACR-TIRADS American College of Radiology TIRADS, KSThR-TIRADS Korean Society of Thyroid Radiology TI-RADS, C-TIRADS Chinese TI-RADS, F-TIRADS Follicular TI-RADS, AUC the area under the receiver operating characteristic curve, *Comparison with the nomogram model by DeLong test

Observer agreement of ultrasound features

Table 3 presents the interobserver agreement between the two observers regarding the ultrasound features of thyroid nodules. Composition, echogenicity, calcification, trabecular formation, halo, diffuse lesions, and blood flow all showed almost perfect interobserver agreement (κ = 0.813, 0.813, 0.909, 0.859, 0.913, 0.933, and 0.929, respectively; all $p < 0.001$).

Table 3 Interobserver variation in the evaluation of the main ultrasound features of thyroid nodules

Features	Composition	Echogenicity	Calcification	Trabecular formation	Halo	Diffuse lesions	Blood flow
κ value	0.813	0.813	0.909	0.859	0.913	0.933	0.929
P value	＜0.001	＜0.001	＜0.001	＜0.001	＜0.001	＜0.001	＜0.001

Discussion

Preoperative differentiation of FNLUF-TNs remains challenging, mainly due to overlapping morphological features between benign and malignant lesions and the unsatisfactory diagnostic performance of FNAC for follicular neoplasms(18). In

clinical practice, conventional TIRADS were primarily developed for papillary thyroid carcinoma, which exhibits typical ultrasound features such as taller-than-wide, microcalcifications, and irregular margins (19, 20). However, these systems yield relatively low accuracy for the diagnosis of follicular neoplasms. Moreover, several other pathological types of thyroid nodules—including nodular goiter and medullary thyroid carcinoma—can exhibit ultrasound features resembling those of follicular neoplasms, as illustrated by our defined FNLUF-TNs, further complicating preoperative diagnosis.

The present study established a clinical ultrasound nomogram model for the preoperative individualized prediction of malignancy risk in FNLUF-TNs. Our results demonstrated that the nomogram model achieved significantly higher diagnostic performance than ACR-TIRADS, KSThR-TIRADS, C-TIRADS, and F-TIRADS, with favorable calibration and clinical net benefit, supporting its clinical utility in risk stratification and management optimization. Multivariate logistic regression identified seven independent predictors for malignant FNLUF-TNs, including younger age, male gender, smaller transverse diameter, hypoechogenicity, presence of calcification, trabecular formation, and irregular interrupted halo sign. Consistent with previous finding(21), patients with malignant nodules were younger than those with benign nodules. FNLUF-TNs thyroid nodules in female patients tended to be benign. According to the study by Li et al(10), a larger mean diameter of follicular neoplasms correlates with a higher malignant potential, and $\geqslant 5$ cm is highly suggestive of

malignancy (assigned 8 points). In the present study, nodules with a smaller transverse diameter were more likely to be malignant, in line with the conclusions of Yang et al(22).

Hypoechogenicity is considered to be attributable to nodule compaction, and its marked high cellularity strongly suggests malignancy(23), which is consistent with our findings. Although isoechogenicity and hyperechogenicity usually indicate benign lesions, the present study found that isoechoic FNLUF TNs may also be malignant. Calcification is an important feature of thyroid nodules, as microcalcification is strongly associated with papillary thyroid carcinoma but has limited diagnostic value for other pathological types of malignant nodules(24, 25). The diagnostic significance of macrocalcification and rim calcification remains controversial(26-28). Among them, a study by Li et al(10) regarded rim calcification as a strong indicator of FTC (assigning a score of 6), while the positive predictive value of macrocalcification was 13.2% – 64.8% for malignancy(24). Hwa et al(13) reported that among all thyroid nodules, the presence of macrocalcification was significantly associated with an elevated risk of papillary thyroid carcinoma (PTC), whereas rim calcification conferred a significantly higher risk of follicular thyroid carcinoma (FTC). By contrast, macrocalcification did not significantly increase the risk of FTC, nor did rim calcification raise the risk of PTC. In our study, the presence of any calcification indicated an elevated malignant risk of FNLUF TNs, with microcalcification showing the strongest predictive effect, followed by macrocalcification. This may be attributed to the relatively high proportion of papillary thyroid carcinoma in our sample. Jeh et al(23) found that about 83% of the

FTC had no calcifications suggesting that calcification should not be overemphasized in the diagnosis of follicular neoplasms. Trabecular formation is a recently recognized feature of follicular neoplasms, characterized by trabecular cell arrangement on microscopic examination(14). Li et al(10) first incorporated this feature into their stratification guideline for follicular neoplasms, considering it a predictor of malignant risk with an assigned score of 2. In the present study, this feature indicated a moderate malignant risk. A halo sign is highly prevalent in FNLUF TNs. The prevailing view of the halo composition is that it consists of compressed vessels around thyroid nodules(29, 30). Previous studies have regarded a halo as a feature of malignant nodules, particularly an irregularly thickened halo(31). However, the present study found that a regular and complete halo is characteristic of benign FNLUF TNs, whereas an interrupted or irregular halo indicates malignant FNLUF TNs. These findings confirm that comprehensive evaluation of clinical and ultrasound features can improve the identification of high-risk FNLUF-TNs.

Notably, blood flow grade was not identified as an independent risk factor. This may be explained by the generally rich vascularity seen in FNLUF-TNs, implying that evaluating blood flow adds little to diagnostic performance and can be excluded from predictive models.

The nomogram model achieved an AUC of 0.824, accuracy of 87.55%, specificity of 93.96%, and Youden index of 0.648, all of which were significantly higher than those

of the four TIRADS systems. Conventional TIRADS rely heavily on macro-calcification, taller-than-wide shape, and marked hypoechogenicity, which are less common in FNLUF-TNs. In contrast, our model was specifically developed and validated for FNLUF-TNs, incorporating unique features such as trabecular formation and halo regularity, thus filling the diagnostic gap for this specific subgroup. Calibration curves showed excellent consistency between predicted and actual probabilities, and decision curve analysis confirmed obvious net clinical benefit within a wide range of threshold probabilities, indicating that the model can help clinicians avoid unnecessary biopsies and surgeries while reducing missed diagnoses. Interobserver agreement for key ultrasound features was excellent ($\kappa > 0.81$), suggesting high reproducibility in routine clinical settings.

This study has several limitations. First, this is a retrospective single-center study, which may lead to potential selection bias. Second, external validation in multi-institutional cohorts is needed to verify the generalizability of the nomogram. Third, in clinical practice, most follicular nodules are only followed up with observation rather than undergoing surgery, resulting in a disproportionately high proportion of suspicious malignant lesions in the study. Despite these limitations, our model still provides a reliable and practical tool for FNLUF-TNs.

Conclusions

In conclusion, the clinical-ultrasound nomogram based on age, gender, transverse

diameter, echogenicity, calcification, trabecular formation, and halo sign enables accurate and individualized malignancy risk stratification for FNLUF-TNs. The model outperforms conventional TIRADS with high diagnostic performance, good reproducibility, and favorable clinical utility, and is expected to be widely used in preoperative evaluation to optimize clinical decision-making and reduce unnecessary interventions.

List of abbreviations

FNLUF-TNs	Thyroid nodules with follicular neoplasm-like ultrasound features
TIRADS	Thyroid imaging reporting and data system
ACR-TIRADS	American College of Radiology Thyroid Imaging Reporting and Data System
KSThR-TIRADS	Korean Society of Thyroid Radiology Thyroid Imaging Reporting and Data System
C-TIRADS	Chinese Thyroid Imaging Reporting and Data System
F-TIRADS	Follicular Thyroid Imaging Reporting and Data System
RSSs	Risk stratification systems
FTN	Follicular thyroid neoplasm
FTC	Follicular thyroid carcinoma
FTA	Follicular thyroid adenoma
ROC	Receiver operating characteristic curve
AUC	Area under the ROC curve

DCA	Decision curve analysis

Declarations

Ethics approval and consent to participate

This retrospective study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the First People's Hospital of Changzhou. Informed consent was waived due to the retrospective design.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare no competing interests.

Funding

This study has received funding by the Health Commission of Changzhou (Grant No. QN202501).

Authors' contributions

All authors made substantial contributions to this paper: (1) Substantial contributions to the conception or design of the work (Zehui Fu and Siqi Wang) or the acquisition (Mengyuan Li, Feng Qian and Yejun LU), analysis (Zehui Fu and Xiao Xie), or interpretation of the data (Mengyuan Li and Jian Zhang); (2) Drafting the work (Zehui Fu) or revising it critically for important intellectual content (Siqi Wang and Hui Chen); (3) Final approval of the completed version (all of the authors). Each author approved the submitted version and agreed to be personally accountable for the author's own contributions.

Acknowledgements

The authors thank all study participants for contributing.

References

1. Wang Y , Lu W , Kong D. Deep learning-based ultrasound diagnostic model for follicular thyroid carcinoma. *Eur Radiol* (2025) 36: 357-366. doi:10.1007/s00330-025-11840-6
2. Feng JW, Ye J, Qi GF, Hong LZ, Wang F, Liu SY, et al. A comparative analysis of eight machine learning models for the prediction of lateral lymph node metastasis in patients with papillary thyroid carcinoma. *Front Endocrinol (Lausanne)* (2022),13(0), 1004913. doi:10.3389/fendo.2022.1004913
3. Almeida AMLD , Delfim RLC , Vidal APA, Maria CACMC, Ana CLS, Marcos FG, et al. Combining the American Thyroid Association's Ultrasound Classification with Cytological Subcategorization Improves the Assessment of Malignancy Risk in Indeterminate Thyroid Nodules. *Thyroid*(2020) 31: 922-932. doi:10.1089/thy.2019.0575
4. Franklin NT, William DM, Edward GG, Jenny KH, Lincoln LB, Sharlene A, 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: 587-595. doi:10.1016/j.jacr.2017.01.046
5. Ji YL, Jung HB, Eun a, Jin YS, Jung HS, Ji HK, et al. 2020 Imaging Guidelines for Thyroid Nodules and Differentiated Thyroid Cancer: Korean Society of Thyroid Radiology. *Korean J Radiol*, (2021), 22: 840-860. doi:10.3348/kjr.2020.0578
6. Zhou JQ, Yin LX, Wei X, Zhang YF, Wang Y, Li X, et al. 2020 Chinese guidelines

for ultrasound malignancy risk stratification of thyroid nodules: the C-TIRADS. *Endocrine* (2020) 70(2):256-279. doi:10.1007/s12020-020-02441-y

7. Hosseini Sarkhosh SM, Shirzad N, Taghvaei M, Soltani A, Khosravi F, Ghasemi M, et al. Prediction of thyroid malignancy risk using clinical and ultrasonography features and a machine learning approach. Eur Radiol (2025) 35(7):5157-5167. doi:10.1007/s00330-024-11321-2
8. Goundan PN, Mamou J, Rohrbach D, Feleppa E, Machi J, Kato H, et al. A Preliminary Study of Quantitative Ultrasound for Cancer-Risk Assessment of Thyroid Nodules. *Front Endocrinol (Lausanne)* (2021) 12:627698. doi:10.3389/fendo.2021.627698
9. Hou TK, Nie F. Ultrasound features of thyroid follicular adenoma and follicular carcinoma: A narrative review. *J Int Med Res* (2026) 54(3):3000605261418838. doi:10.1177/03000605261418838
10. Li JM, Li C, Zhou XH, Huang JX, Yang PC, Cang YD, et al. US Risk Stratification System for Follicular Thyroid Neoplasms. *Radiology* (2023) 309(2):e230949. doi:10.1148/radiol.230949
11. Kim MK, Park H, Oh YL, Song DE, Kim TY, Kim WB, et al. Association of Ultrasonography Features of Follicular Thyroid Carcinoma With Tumor Invasiveness and Prognosis Based on WHO Classification and TERT Promoter Mutation. *Korean J Radiol* (2024) 25(1):103-112. doi:10.3348/kjr.2023.0812
12. Matrone A, Gambale C, Pieroni E, Ferrari SM, Elisei R, Antonelli A, et al. Ultrasound features and risk stratification system in NIFT-P and other follicular-

patterned thyroid tumors. *Eur J Endocrinol* (2023) 189(1):175-182. doi:10.1093/ejendo/lvad090

13. Shin HS, Na DG, Paik W, Kim DS, Kim JH, Lee JH, et al. Malignancy Risk Stratification of Thyroid Nodules with Macrocalcification and Rim Calcification Based on Ultrasound Patterns. *Korean J Radiol* (2021) 22(4):663-671. doi:10.3348/kjr.2020.0834
14. Kuo TC, Wu MH, Chen KY, Hsieh MS, Chen A, Chen CN. Ultrasonographic features for differentiating follicular thyroid carcinoma and follicular adenoma. *Asian J Surg* (2020) 43(1):339-346. doi:10.1016/j.asjsur.2019.04.016
15. 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):553-559. doi:10.1016/0301-5629(90)90020-d
16. Li WY, Gao LY, Wang YH, Wang Y, Zhang XH, Chen Q, et al. The role of microflow patterns combined with greyscale ultrasound in enhancing diagnostic validity and reducing unnecessary biopsy rate of thyroid nodules. *Eur Radiol* (2025) 36(3):1771-1782. doi:10.1007/s00330-024-11378-z
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* (1977) 33(1):159-174. doi:10.2307/2529310
18. Zhu XX, Liu F, Liu JY, Li ZH, Ma Y. Thyroglobulin-to-tumor volume ratio combined with ultrasound features for diagnosing thyroid follicular neoplasms. *Front Endocrinol (Lausanne)* (2025) 16:1626766. doi:10.3389/fendo.2025.1626766

19. 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):225-237. doi:10.1159/000478927

20. Na HY, Moon JH, Choi JY, Kim JH, Lee JJ, Lee HJ, et al. Preoperative diagnostic categories of fine needle aspiration cytology for histologically proven thyroid follicular adenoma and carcinoma, and Hurthle cell adenoma and carcinoma: Analysis of cause of under- or misdiagnoses. *PLoS ONE* (2020) 15(11):e0241597. doi:10.1371/journal.pone.0241597

21. Moon HJ, Sung JM, Kim EK, Yoon JH, Youk JH, Kwak JY. Diagnostic performance of gray-scale US and elastography in solid thyroid nodules. *Radiology* (2012) 262(3):1002-1013. doi:10.1148/radiol.11110839

22. 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.16.03041

23. Jeh SK, Jung SL, Kim BS, Ahn KJ, Choi HS, Ahn MI, et al. Evaluating the degree of conformity of papillary carcinoma and follicular carcinoma to the reported ultrasonographic findings of malignant thyroid tumor. *Korean J Radiol* (2007) 8(3):192-197. doi:10.3348/kjr.2007.8.3.192

24. Durante C, Hegedüs L, Na DG, Moon JH, Papini E, Zhanna S, et al. International Expert Consensus on US Lexicon for Thyroid Nodules. *Radiology* (2023)

309(3):e231481. doi:10.1148/radiol.231481

25. Wei PY, Jiang ND, Ding JW, Yang Z, Gong YW, Zhang T, et al. The Diagnostic Role of Computed Tomography for ACR TI-RADS 4-5 Thyroid Nodules With Coarse Calcifications. *Front Oncol* (2020) 10:911. doi:10.3389/fonc.2020.00911

26. 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(3):762-770. doi:10.1148/radiol.2473070944

27. Na DG, Baek JH, Sung JY, Kim JH, 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(4):562-572. doi:10.1089/thy.2015.0460

28. Bai S, Wu LH, Su YH, Wang B, Li L, Zhang Y, et al. Ultrasound-Based Prediction Model for Distinguishing Malignant from Benign Thyroid Nodules with Peripheral Calcification. *Acad Radiol* (2025) 32(10):4486-4499. doi:10.1016/j.acra.2025.01.043

29. Yao XY, Li X, Mei F, Wang XX, Wu YY, Zhang L, et al. Prospective Validation of an Ultrasound Capsule-Based Model for Predicting Follicular Thyroid Carcinoma Without High-Risk Features. *Cancer Med* (2025) 14(13):e71190. doi:10.1002/cam4.71190

30. Chen XJ, Huang LJ, Mao F, Chen Q, Li H, Yang J, et al. Value of CEUS features in diagnosing thyroid nodules with halo sign on B-mode ultrasound. *BMC Med Imaging* (2023) 23:11. doi:10.1186/s12880-023-00967-x

31. Shi WZ, Zhang M, Tang WY, Chen L, Liu Y, Wang Z, et al. The Association Between the Thickness of the Hypoechoic Halo of Thyroid Nodules and Thyroid Cancer: A Retrospective Study. *Acad Radiol* (2025) 32(4):1906-1917. doi:10.1016/j.acra.2024.08.037