



Comparison of the diagnostic performance and clinical role of different ultrasound-based thyroid malignancy risk stratification systems for medullary thyroid carcinoma

Liu Jiang^{1#^}, Hai-Bin Zhu^{2#}, Zhen-Wei Liang¹, Lei Chen¹, Xiu-Ming Sun¹, Yu-Hong Shao¹, Lu-Zeng Chen¹

¹Department of Ultrasonography, Peking University First Hospital, Beijing, China; ²Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Radiology, Peking University Cancer Hospital and Institute, Beijing, China

Contributions: (I) Conception and design: L Jiang, HB Zhu, LZ Chen; (II) Administrative support: None; (III) Provision of study materials or patients: LZ Chen; (IV) Collection and assembly of data: L Jiang, HB Zhu, ZW Liang, L Chen, XM Sun; (V) Data analysis and interpretation: L Jiang, HB Zhu, ZW Liang, L Chen, YH Shao, LZ Chen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

#These authors contributed equally to this work.

Correspondence to: Lu-Zeng Chen, MD. Department of Ultrasonography, Peking University First Hospital, 8 Xi Shi Ku Street, Xi Cheng District, Beijing 100034, China. Email: chenluzeng@bjmu.edu.cn.

Background: This study sought to investigate the applicability of different ultrasound (US) thyroid risk stratification systems in diagnosing medullary thyroid carcinoma (MTC) and determining the need for biopsy.

Methods: In total, 34 MTC nodules, 54 papillary thyroid carcinoma (PTC) nodules, and 62 benign thyroid nodules were examined in this study. All the diagnoses were histopathologically confirmed postoperatively. All the thyroid nodule sonographic features were recorded and categorized by 2 independent reviewers according to the Thyroid Imaging Reporting and Data System (TIRADS) of the American College of Radiology (ACR), the American Thyroid Association (ATA) guidelines, the European Thyroid Association (EU) TIRADS, the Kwak-TIRADS, and the Chinese TIRADS (C-TIRADS). The sonographic differences and risk stratifications of the MTCs, PTCs, and benign thyroid nodules were analyzed. The diagnostic performance and recommended biopsy rates for each classification system were evaluated.

Results: The risk stratifications of MTCs were all higher than the benign thyroid nodules ($P < 0.01$) and lower than PTCs ($P < 0.01$) with each classification system. Hypoechoogenicity and malignant marginal features were independent risk factors for identifying malignant thyroid nodules, and the area under the receiver operating characteristic curve (AUC) for identifying MTCs was lower than that for identifying PTCs (0.873 vs. 0.954, respectively). The AUCs, sensitivity, specificity, positive predictive values, negative predictive values, and accuracy values of the 5 systems for MTC were all lower than those for PTC. The best cut-off values for diagnosing MTC were TIRADS (TR) 4 in the ACR-TIRADS, intermediate suspicion in the ATA guidelines, TR 4 in EU-TIRADS, and TR 4b in both the Kwak-TIRADS and the C-TIRADS. The Kwak-TIRADS had the highest recommended biopsy rate for MTCs (97.1%), followed by the ATA guidelines, the EU-TIRADS (88.2%), the C-TIRADS (85.3%), and the ACR-TIRADS (79.4%).

Conclusions: The US-based thyroid malignancy risk stratification systems analyzed in this study were able to satisfactorily identify MTC and recommend biopsy, but the diagnostic performance of these systems for MTC was not as good as that for PTC.

[^] ORCID: 0000-0001-7352-2975.

Keywords: Medullary thyroid carcinoma (MTC); thyroid cancer; Thyroid Imaging Reporting and Data System (TIRADS); ultrasonography; biopsy

Submitted Sep 15, 2022. Accepted for publication Apr 07, 2023. Published online May 04, 2023.

doi: 10.21037/qims-22-968

View this article at: <https://dx.doi.org/10.21037/qims-22-968>

Introduction

Medullary thyroid carcinoma (MTC) accounts for 3–5% of all malignant thyroid tumors and is a type of neuroendocrine tumor that originates from the parafollicular cells of the thyroid gland and secretes calcitonin (1-3). Compared with papillary thyroid carcinoma (PTC), which accounts for >80% of primary thyroid carcinomas and has a good prognosis, MTC is more malignant and has a poorer prognosis (4). The early diagnosis and timely treatment of MTC play important roles in the prognosis of affected patients. Most previous studies on the diagnosis of MTC have sought to compare the sonographic features of MTC with those of benign thyroid nodules or PTC and have reported that the sonographic findings of MTC appear to be similar to those of benign thyroid nodules (5-7). The ultrasonic diagnosis of MTC remains a challenge, especially for inexperienced clinicians. Thus, identifying the most effective ultrasonic assessment methods for MTC has important clinical significance.

Currently, ultrasound (US) is still the first choice for identifying and diagnosing thyroid nodules; however, a challenge arises in determining how to distinguish between benign and malignant thyroid nodules to minimize unnecessary biopsies and surgeries as much as possible. Many scholars and clinical societies have proposed thyroid risk stratification systems based on US findings, such as the Thyroid Imaging Reporting and Data System (TIRADS), to improve the accuracy of diagnosis and provide effective guidance for clinicians (8-10). The classification methods of different thyroid risk stratification systems differ and include regression equations as well as weighting, counting, and pattern-based methods (11-15). To date, several widely recognized thyroid classification systems have been developed, including the 2017 American College of Radiology (ACR) TIRADS (11), the 2015 American Thyroid Association (ATA) management guidelines (12), the European Thyroid Association (EU) TIRADS (13), the Kwak-TIRADS (14), and most recently, the 2020 Chinese TIRADS (C-TIRADS) (15). These different classification

systems have shown reliable diagnostic performance and effectively reduced unnecessary biopsies and the overtreatment of benign thyroid nodules to a certain extent. However, most previous studies on thyroid classification systems have focused on PTC, and very few studies have evaluated the efficacy of routinely used and newly developed thyroid classification systems in assessing MTC.

In this study, we used the 2017 ACR-TIRADS, the 2015 ATA guidelines, the EU-TIRADS, the Kwak-TIRADS, and the latest C-TIRADS to compare the risk stratifications and sonographic features of thyroid nodules, including MTC, PTC, and benign thyroid nodules, and then evaluated the diagnostic accuracy of the 5 thyroid risk stratification systems for MTC. This study sought to investigate the diagnostic performance of the currently used classification systems for MTC and evaluate the reliability of these guidelines in informing decisions regarding the need for fine-needle aspiration biopsy (FNAB). We present the following article in accordance with the STARD reporting checklist (available at <https://qims.amegroups.com/article/view/10.21037/qims-22-968/rc>).

Methods

The Medical Ethics Committee of Peking University First Hospital approved this study, and the requirement for informed consent was waived due to the monocentric and retrospective nature of this study. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Patients

From January 2009 to December 2020, 51 MTC patients confirmed by operation and biopsy at Peking University First Hospital were enrolled in this study. Patients were excluded from the study if they met any of the following exclusion criteria: (I) had not undergone a preoperative thyroid sonographic examination; (II) had US images of thyroid nodules that did not meet the requirements for

analysis; and/or (III) had incomplete clinical information. Ultimately, 27 patients with 34 MTC lesions were enrolled in this study. Additionally, 50 patients with PTC (with a total of 54 nodules) and 59 patients with benign thyroid nodules (with a total of 62 nodules), who underwent surgery during the same period as the patients with MTC, were also enrolled in this study as the control groups. The 62 benign thyroid nodules included follicular adenoma (n=50) and nodular goiter (n=12).

Sonographic examinations

All the ultrasonic examinations were performed using 7–14 MHz high-frequency linear transducers (i.e., Philips EPIQ7, Philips Medical Systems, Amsterdam, The Netherlands; GE Volume E8, GE Healthcare, Chicago, IL, USA; Siemens ACUSON S2000, Munich, Germany; and ABVS, Siemens Medical Solutions, Erlangen, Germany). The ultrasonic images of the thyroid nodules, including the grayscale and color Doppler images, were saved in a picture archiving and communication system (PACS).

US feature analysis

The preoperative ultrasonic images were retrospectively reviewed in a PACS reporting workstation by 2 physicians with 10 and 6 years of thyroid US experience, respectively. As the image quality displayed by different reporting workstations may differ slightly, the 2 reviewers used the same workstation. All the US images of the thyroid nodules were evaluated independently by the 2 reviewers who were blinded to the clinical information and surgical pathology results. The results of the reviewer who achieved better accuracy were used for the further statistical analysis. Before the evaluation started, the reviewers reviewed the literature and instructions on the 5 thyroid classification systems in detail, and then analyzed 80 thyroid nodules together to ensure interobserver reliability.

Evaluation criteria of US features

The following characteristics of the nodules were recorded: (I) location (right lobe, left lobe, or isthmus); (II) size (anteroposterior, transverse, and longitudinal diameters); (III) orientation (wider-than-tall shape or taller-than-wide shape); (IV) margins (smooth, ill-defined, irregular/microlobulated, or extrathyroidal extension); (V) composition (cystic, solid, mixed predominantly solid/

solid component $\geq 50\%$, mixed predominantly cystic/solid component $< 50\%$, spongiform, or unidentified); (VI) echogenicity (hyperechoic, isoechoic, hypoechoic, markedly hypoechoic, or anechoic); (VII) presence of hyperechoic foci (none, comet-tail artifacts, microcalcification/punctate hyperechoic foci ≤ 1 mm, macrocalcification/hyperechoic foci > 1 mm, or peripheral/rim calcification); (VIII) blood flow (absent, some, or abundant); and (IX) suspected cervical lymph node metastasis (present or absent). Wider-than-tall shape was defined as an anteroposterior diameter shorter than the transverse diameter, and taller-than-wide shape was defined as an anteroposterior diameter longer than the transverse diameter. A nodule was defined as spongiform if the interior of the nodule was occupied by multiple tiny cystic spaces without solid tissues. A nodule was classified as unidentified if it had unevaluable nodule features due to heavy calcifications. When both macrocalcification and microcalcification existed in a nodule, it was classified as a microcalcification. The abnormal appearance of suspected lymph node metastasis included a round shape, an obscure boundary, heterogeneity with cystic components, loss of normal echogenicity of the central lymphatic hilum, punctate hyperechoic foci, and hypervascularity in the peripheral portion of lymph node (16,17).

Categorization of the thyroid nodules according to the thyroid risk stratification systems

Each thyroid nodule was categorized according to the following 5 thyroid risk stratification systems: 2017 ACR-TIRADS, 2015 ATA guidelines, EU-TIRADS, Kwak-TIRADS, and C-TIRADS. Each of the 5 thyroid risk stratification systems defines the categories of risk according to a set of sonographic features; the higher the category, the greater the malignant risk of the thyroid nodules.

The 2017 ACR-TIRADS assesses thyroid nodules based on an evaluation of the following 5 key features: structure, echogenicity, shape, margin, and echogenic foci. Each of the 5 US features is described in detail and weighted by set points. When evaluating the thyroid nodules, the US radiologists calculate the total number of points according to the set points of each US feature and then classify the risk stratification according to the following 5 levels: TIRADS (TR) 1 (benign), TR 2 (not suspicious), TR 3 (mildly suspicious), TR 4 (moderately suspicious), and TR 5 (highly suspicious) (11).

Both the ATA guidelines and EU-TIRADS use pattern-based approaches that aim to apply ultrasonic evaluation

criteria to the least extent possible, making them simpler than other classification systems. Based on the US features described in each classification level pattern, the risk categories in the ATA guidelines are benign, very low suspicion, low suspicion, intermediate suspicion, and high suspicion (12), and those of EU-TIRADS are TR 1 (normal), TR 2 (benign), TR 3 (low risk), TR 4 (intermediate risk), and TR 5 (high risk) (13).

The Kwak-TIRADS and C-TIRADS classify thyroid nodules by counting the number of suspicious US features, making them relatively easy to apply clinically. Additionally, category 4 in both the Kwak-TIRADS and C-TIRADS are subdivided into TR 4a, 4b, and 4c. According to Kwak-TIRADS, suspicious US features include a solid or almost solid composition, hypoechoic intensity, irregular margins, microcalcifications, and a taller-than-wide shape; these features are used to calculate the score for classification as TR 3 (probably benign, no suspicious US features), TR 4a (low suspicion, 1 suspicious feature), TR 4b (intermediate suspicious, 2 suspicious features), TR 4c (moderate concern but not classic, 3 or 4 suspicious features), and TR 5 (highly suspicious, 5 suspicious features) (14). Among the 5 positive US features included in the C-TIRADS, 3 of the features differ somewhat to those included in the Kwak-TIRADS, and include a solid composition, markedly hypoechoic nodules, ill-defined/irregular margins, or extrathyroidal extension. Conversely, the C-TIRADS uses a negative feature (comet-tail artifact) and several positive US features to calculate the risk stratification score, from which 1 point is subtracted if a negative feature is present (15).

Additionally, each risk grade has corresponding recommended strategies for clinical management, including the recommendation to undergo FNAB or plans for long-term follow-up. In the high-risk categories of all 5 classifications, FNAB is usually recommended when the maximum diameter of the thyroid nodule is ≥ 1 cm; however, in the low-risk categories, the recommended size thresholds differ (Table 1). Thus, according to the clinical decisions recommended by each thyroid risk stratification system, we divided the MTC nodules into the following 2 groups: (I) the “FNAB required” group; and (II) the “FNAB not required” group.

Statistical analysis

The statistical analysis was performed using SPSS Version 22.0 (IBM, Armonk, NY, USA). Using the pathology results as the gold standard, the accuracy of the 2 reviewers in

diagnosing the thyroid nodules was calculated respectively. The data of the reviewer who achieved better accuracy were used for the further statistical analysis. The variables are described as the mean \pm standard deviation (SD) or the number (percentage). The distributions among the disease groups were compared using a 1-way analysis of variance or the chi-square test. Receiver operating characteristic (ROC) curves were used to assess the diagnostic performance of the 5 classification systems. The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy and the areas under the ROC curves (AUCs) were calculated with 95% confidence intervals (CIs), and the cut-off values were determined using the maximum Youden index. Kappa coefficients were calculated to assess the interobserver agreement; 0.0–0.20, 0.21–0.40, 0.41–0.60, 0.61–0.80, and 0.81–1.00 indicated poor, fair, moderate, substantial, and perfect agreement, respectively. Two-sided P values <0.05 were considered statistically significant.

Results

Clinical characteristics

The baseline characteristics of the patients are summarized in Table 2. The mean sizes of the nodules in the MTC, PTC, and benign groups were 1.74 ± 1.23 , 1.11 ± 0.69 , and 1.82 ± 0.81 cm, respectively (MTC *vs.* PTC, $P=0.002$; MTC *vs.* benign group, $P=0.173$; and PTC *vs.* benign group, $P<0.001$). The lymph node metastases were confirmed by surgical pathology in 4 MTC patients and 9 PTC patients.

Comparison of the 5 thyroid risk stratification systems and US characteristics among the MTCs, PTCs, and benign thyroid nodules

Using the pathology results as the gold standard, the accuracy of the 2 reviewers with 10 and 6 years of thyroid US experience were 85.3% and 84.6%, respectively. Thus, we used the data of the reviewer with 10 years of experience for the later analysis. The evaluations of the 5 risk stratification systems and US characteristics among the MTCs, PTCs, and benign thyroid nodules are set out in Table 2. Statistically significant differences in the 5 stratification systems were detected among the MTCs, PTCs, and benign thyroid nodules. In terms of the orientation, margins, composition, echogenicity, calcification, and blood flow, the differences between the benign and malignant thyroid nodules were statistically significant. Significant

Table 1 Summary of the fine-needle aspiration biopsy recommendations of the 5 thyroid risk stratification systems for medullary thyroid carcinomas

Guidelines	FNAB size cut-off (largest dimension)	MTC nodules, n	Recommend FNAB, n (%)
ACR-TIRADS		34	27 (79.4%)
TR1 (benign)	No biopsy	0	0
TR2 (not suspicious)	No biopsy	1	0
TR3 (mildly suspicious)	Recommend FNAB at ≥ 25 mm; follow-up at ≥ 15 mm	1	0
TR4 (moderately suspicious)	Recommend FNAB at ≥ 15 mm; follow-up at ≥ 10 mm	14	11
TR5 (highly suspicious)	Recommend FNAB at ≥ 10 mm; follow-up at ≥ 5 mm	18	16
ATA guidelines		34	30 (88.2%)
Benign	No biopsy	0	0
Very low suspicion	Consider FNAB at ≥ 2 cm, observation is also a reasonable option	1	1
Low suspicion	Recommend FNAB at ≥ 15 mm	1	1
Intermediate suspicion	Recommend FNAB at ≥ 10 mm	10	9
High suspicion	Recommend FNAB at ≥ 10 mm	21	18
Non-ATA pattern	Recommend FNAB at ≥ 15 mm	1	1
EU-TIRADS		34	30 (88.2%)
TR1 (normal)	No biopsy	0	0
TR2 (benign)	No biopsy	0	0
TR3 (low risk)	Recommend FNAB at ≥ 20 mm	3	3
TR4 (intermediate risk)	Recommend FNAB at ≥ 15 mm	11	10
TR5 (high risk)	Recommend FNAB at ≥ 10 mm	20	17
Kwak-TIRADS		34	33 (97.1%)
TR2 (benign)	No biopsy	0	0
TR3 (probably benign)	No biopsy	1	0
TR4a (low suspicion for malignancy)	Recommend FNAB	1	1
TR4b (intermediate suspicion for malignancy)	Recommend FNAB	14	14
TR4c (moderate concern but not classic for malignancy)	Recommend FNAB	15	15
TR5 (highly suggestive of malignancy)	Recommend FNAB	3	3
C-TIRADS		34	29 (85.3%)
TR2 (benign)	No biopsy	0	0
TR3 (probably benign)	No biopsy	1	0
TR4a (low suspicion)	Recommend FNAB at ≥ 15 mm; at ≥ 10 mm when A* occurs	14	13
TR4b (moderate suspicion)	Recommend FNAB at ≥ 10 mm; at ≥ 5 mm when A* occurs	7	6
TR4c (high suspicion)	Similar to TR 4b	11	9
TR5 (highly suggestive of malignancy)	Similar to TR 4b; recommend FNAB of any size if B* occur	1	1

A*, category 4a–5 nodules are multiple, or immediately adjacent to the trachea or recurrent laryngeal nerve; B*, there are typical cervical metastatic lymph nodes of thyroid cancer. ACR-TIRADS, the 2017 American College of Radiology TIRADS (11); ATA guidelines, the 2015 American Thyroid Association management guidelines (12); EU-TIRADS, European Thyroid Association TIRADS (13); C-TIRADS, the latest 2020 Chinese TIRADS (15). MTC, medullary thyroid carcinoma; TIRADS, Thyroid Imaging Reporting and Data System; FNAB, fine-needle aspiration biopsy.

Table 2 Comparison of the baseline characteristics, the 5 thyroid risk stratification systems, and ultrasonic features among the medullary thyroid carcinoma, papillary thyroid carcinoma, and benign groups

Category	Parameter	MTCs (n=34)	PTCs (n=54)	Benign group (n=62)	P value (malignant vs. benign group)	P value (MTC vs. PTC)	P value (MTC vs. benign group)
Characteristic	Sex (male/female)	13/14	20/30	13/46	0.01	0.491	0.01
	Age (years, mean \pm SD)	51.59 \pm 13.68	43.18 \pm 10.64	50.98 \pm 15.64	0.075	0.086	0.441
Guidelines	ACR-TIRADS (categories, mean \pm SD)	4.24 \pm 0.89	4.86 \pm 0.36	3.26 \pm 0.81	<0.001	<0.001	<0.001
	ATA guidelines (categories, mean \pm SD)	4.27 \pm 0.91	4.91 \pm 0.35	3.21 \pm 0.87	<0.001	<0.001	<0.001
	EU-TIRADS (categories, mean \pm SD)	4.29 \pm 0.84	4.93 \pm 0.26	3.29 \pm 0.80	<0.001	<0.001	<0.001
	Kwak-TIRADS (categories, mean \pm SD)	4.21 \pm 1.18	5.11 \pm 0.66	4.06 \pm 1.01	<0.001	<0.001	<0.001
	C-TIRADS (categories, mean \pm SD)	4.71 \pm 1.19	5.80 \pm 0.79	3.84 \pm 0.66	<0.001	<0.001	<0.001
Orientation	Taller-than-wide	8	27	1	<0.001	0.078	0.009
	Wider-than-tall	26	27	61			
Margins	Well-defined/halo	20	9	59	<0.001	<0.001	<0.001
	Ill-defined	1	1	0			
	Irregular/microlobulated	13	42	3			
	Extrathyroidal extension	0	2	0			
Composition	Cystic, spongiform	0	0	0	0.004	0.600	0.351
	Mixed predominately solid	4	2	15			
	Mixed predominately cystic	0	0	2			
	Solid	30	52	45			
Echogenicity	Anechoic	0	0	0	<0.001	0.24	<0.001
	Hyperechoic	1	0	0			
	Isoechoic	2	0	45			
	Hypoechoic	31	54	17			
Echotexture	Homogeneous	6	13	21	0.197	>0.99	0.747
	Heterogeneous	28	41	41			
Calcifications	Absent/comet-tail	12	12	47	<0.001	0.210	<0.001
	Micro	14	34	4			
	Macro/peripheral	6	8	7			
	Foci indeterminate	2	0	4			
Posterior features	Enhancement	1	0	0	0.033	>0.99	0.122
	Shadowing	5	6	1			
	No	28	48	61			
Blood flow	Yes	31	12	61	<0.001	<0.001	0.378
	No	3	42	1			

ACR-TIRADS, the 2017 American College of Radiology TIRADS (11); ATA guidelines, the 2015 American Thyroid Association management guidelines (12); EU-TIRADS, European Thyroid Association TIRADS (13); C-TIRADS, the latest 2020 Chinese TIRADS (15). SD, standard deviation; MTC, medullary thyroid carcinoma; PTC, papillary thyroid carcinoma; TIRADS, Thyroid Imaging Reporting and Data System.

Table 3 Diagnostic accuracy of the 5 ultrasound-based thyroid risk stratification systems

Groups	Parameters	ACR-TIRADS	ATA guidelines	EU-TIRADS	Kwak-TIRADS	C-TIRADS
Benign/malignant	AUC (95% CI)	0.897 (0.841–0.953)	0.885 (0.824–0.946)	0.908 (0.854–0.963)	0.911 (0.862–0.960)	0.891 (0.841–0.942)
	Cut-off	1–4 ^a , 5 ^b	1–4 ^a , 5 ^b	1–4 ^a , 5 ^b	1–4b ^a , 4c–5 ^b	1–4a ^a , 4b–5 ^b
	Sensitivity (%)	73.9	79.5	80.7	77.3	78.4
	Specificity (%)	90.3	88.7	88.7	88.7	88.7
	PPV (%)	91.5	90.9	91.0	90.7	90.8
	NPV (%)	70.9	75.3	76.4	73.3	74.3
	Accuracy (%)	80.7	83.3	84.0	82.0	76.0
MTC/benign	AUC (95% CI)	0.843 (0.760–0.927)	0.836 (0.743–0.928)	0.868 (0.794–0.942)	0.856 (0.778–0.934)	0.790 (0.694–0.886)
	Cut-off	1–3 ^a , 4–5 ^c	1–3 ^a , 4–5 ^c	1–3 ^a , 4–5 ^c	1–4a ^a , 4b–5 ^c	1–4a ^a , 4b–5 ^c
	Sensitivity (%)	52.9	61.8	61.8	52.9	55.9
	Specificity (%)	90.3	88.7	88.7	88.7	88.7
	PPV (%)	75.0	75.0	75.0	72.0	73.1
	NPV (%)	77.8	88.9	80.9	77.5	78.6
	Accuracy (%)	77.1	79.2	79.2	76.0	77.1
PTC/benign	AUC (95% CI)	0.931 (0.882–0.980)	0.916 (0.857–0.975)	0.933 (0.884–0.983)	0.946 (0.906–0.985)	0.956 (0.917–0.991)
	Cut-off	1–4 ^a , 5 ^d	1–4 ^a , 5 ^d	1–4 ^a , 5 ^d	1–4b ^a , 4c–5 ^d	1–4b ^a , 4c–5 ^d
	Sensitivity (%)	87.4	90.7	92.6	92.6	92.6
	Specificity (%)	90.3	88.7	88.7	88.7	88.7
	PPV (%)	88.7	87.5	87.7	87.7	87.7
	NPV (%)	88.9	91.7	93.2	93.2	93.2
	Accuracy (%)	88.8	89.7	90.5	90.5	90.5

^a, benign group; ^b, malignant group, including MTC and PTC lesions; ^c, MTC group; ^d, PTC group. TIRADS, Thyroid Imaging Reporting and Data System; ACR-TIRADS, the 2017 American College of Radiology TIRADS (11); ATA guidelines, the 2015 American Thyroid Association management guidelines (12); EU-TIRADS, European Thyroid Association TIRADS (13); C-TIRADS, the latest 2020 Chinese TIRADS (15). MTC, medullary thyroid carcinoma; PTC, papillary thyroid carcinoma; AUC, area under the receiver operating characteristic curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

differences in the margins, echogenicity, and calcification were also observed between the MTCs and benign thyroid nodules, among which, there was a significant difference in the margins between the MTCs and PTCs ($P < 0.001$). Additionally, there was a significant difference in the blood flow between the MTCs and PTCs ($P < 0.001$).

Comparison of the diagnostic efficacy of the 5 thyroid risk stratification systems

The AUCs, cut-off values, sensitivity, specificity, PPVs, NPVs, and accuracy of the 5 risk stratification systems are listed and compared in *Table 3*. The AUCs for identifying

malignant thyroid nodules using the ACR-TIRADS, ATA guidelines, EU-TIRADS, Kwak-TIRADS, and C-TIRADS were 0.897, 0.885, 0.908, 0.911, and 0.891, respectively. The 5 classification systems showed similar diagnostic performances.

The ROC curves of the ACR-TIRADS, ATA guidelines, EU-TIRADS, Kwak-TIRADS, and C-TIRADS had AUCs of 0.843, 0.836, 0.868, 0.856, and 0.790, respectively, for differentiating between the MTCs and benign thyroid nodules and AUCs of 0.931, 0.916, 0.933, 0.946, and 0.956, respectively, for differentiating between the PTCs and benign thyroid nodules. For identifying MTCs, the EU-TIRADS showed the best diagnostic efficiency, with the

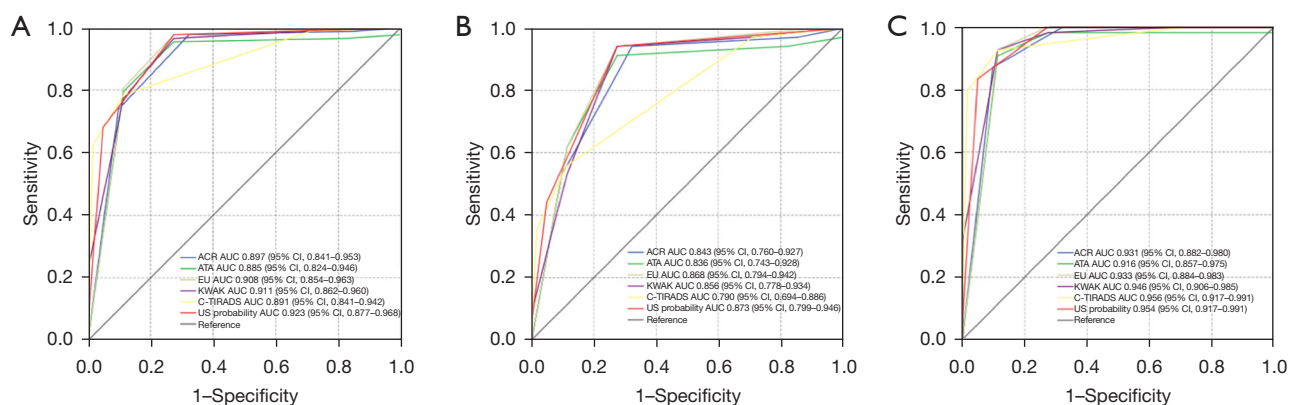


Figure 1 ROC curves for comparison between the 5 US-based thyroid risk stratification systems and US probability. (A) Differentiation between benign thyroid nodules and malignant thyroid nodules; (B) differentiation between MTCs and benign thyroid nodules; (C) differentiation between PTCs and benign thyroid nodules. ACR-TIRADS, the 2017 American College of Radiology TIRADS (11); ATA guidelines, the 2015 American Thyroid Association management guidelines (12); EU-TIRADS, European Thyroid Association TIRADS (13); C-TIRADS, the latest 2020 Chinese TIRADS (15). ROC, receiver operating characteristic; US, ultrasound; MTC, medullary thyroid carcinoma; PTC, papillary thyroid carcinoma; TIRADS, Thyroid Imaging Reporting and Data System.

highest AUC, sensitivity, PPV, and accuracy (0.868, 61.8%, 75.0%, and 79.2%, respectively), and the ACR-TIRADS showed the best specificity (90.3%). For identifying PTCs, the C-TIRADS showed the best diagnostic efficiency, with the highest AUC, sensitivity, NPV, and accuracy (0.956, 92.6%, 93.2%, and 90.5%, respectively), and the ACR-TIRADS showed the best specificity (90.3%). The AUCs, sensitivity, PPVs, NPVs, and accuracy for identifying the MTCs using the 5 classification systems were all lower than those for the PTCs; however, the specificity for identifying the MTCs and PTCs was similar. The best cut-off values for diagnosing the MTCs and PTCs among the 5 classification systems were as follows: TR 4 and 5 in the ACR-TIRADS; intermediate suspicion and high suspicion in the ATA guidelines; TR 4 and 5 in the EU-TIRADS; and 4b and 4c in both the Kwak- and C-TIRADS.

Notably, 1 MTC nodule (2.94%, 1/34) and 1 PTC nodule (1.85%, 1/54) were considered unclassifiable under the ATA guidelines. All the thyroid nodules were considered classifiable under the other 4 systems.

In addition, according to the clinical recommendations of each guideline, the number of “FNAB required” cases among MTC nodules were 27 (79.4%) according to the ACR-TIRADS, 30 (88.2%) according to the ATA guidelines, 30 (88.2%) according to the EU-TIRADS, 33 (97.1%) according to the Kwak-TIRADS, and 29 (85.3%) according to the C-TIRADS (Table 1). No significant differences were observed in the clinical recommendations

for FNAB across the 5 systems.

Diagnostic efficacy of US characteristics for distinguishing between benign and malignant thyroid nodules

The multivariate logistic regression results showed that echogenicity and margins were independent factors for differentiating between benign and malignant thyroid nodules. The ultrasonic probability of the malignant nodules was calculated accordingly. The AUC for the ultrasonic probability of malignancy was 0.923 (95% CI, 0.877–0.968). The AUC for the ultrasonic probability of differentiation between the MTCs and benign thyroid nodules was 0.873 (95% CI, 0.799–0.946). The AUC for the ultrasonic probability of differentiation between the PTCs and benign thyroid nodules was 0.954 (95% CI, 0.917–0.991). ROC curves were used to compare the 5 guidelines and the ultrasonic probability (Figure 1).

Interobserver agreement of the 5 TI-RADS classifications

The kappa coefficients for the ACR-TIRADS, ATA guidelines, EU-TIRADS, Kwak-TIRADS, and C-TIRADS were 0.970, 0.969, 0.968, 0.939, and 0.963, respectively. The kappa coefficients for diagnosing the MTCs by the 5 TIRADS classifications were 0.969, 0.913, 0.957, 0.947, and 0.955, respectively. No significant differences were observed among the 5 classification systems.

Discussion

Many studies have proven that malignant thyroid nodules display certain sonographic features, such as a solid composition, hypoechoic or markedly hypoechoic nodules, ill-defined or irregular margins, a taller-than-wide shape, microcalcification, and extrathyroidal extension (15,18). These malignant US features are mainly consistent with PTC due to the high prevalence of PTC, which accounts for >80% of all malignant thyroid tumors. Most thyroid risk stratification systems, including the 5 classifications examined in our study, are based on these malignant sonographic features. Because the malignant sonographic characteristics of MTC are not typical of PTC, the application of these classification systems to MTC remains controversial. We compared the 5 thyroid guidelines for MTCs, PTCs, and benign thyroid nodules and found that the risk stratification of MTC was higher than that of benign thyroid nodules and lower than that of PTC, and there were statistically significant differences among the MTCs, PTCs, and benign thyroid nodules. Thus, we observed that the MTC lesions tended to have some malignant sonographic features, but these malignant features were not typical of PTCs. This finding is consistent with the findings of some previous studies (19). Thus, we further compared the sonographic features of MTCs with those of benign thyroid nodules and PTCs.

In our study, the sizes of the MTCs and benign thyroid nodules were both significantly larger than those of the PTCs, which is consistent with the results of other studies (3,20). Our results showed that the MTCs had some malignant sonographic characteristics, including taller-than-wide orientation, malignant marginal features, and hypoechoic or markedly hypoechoic nodules and microcalcification. In our study, the MTC lesions were mostly solid (31/34) and hypoechoic or markedly hypoechoic (31/34). These features are attributable to the abundance of collagen fibers and blood vessels, which are not prone to necrosis and liquefaction (21). A total of 41.2% of the MTC lesions showed malignant signs in terms of tumor margins, and exhibited ill-defined, irregular or microlobulated margins, but this proportion was significantly lower than that of the PTCs (83.3%). Zhou *et al.* reported that 34% of MTCs showed spiculated margins, and MTCs with a diameter >1 cm tended to have smooth margins (22). This is in line with our results; however, another study showed that 60% of MTC lesions had ill-defined or spiculated margins (23). We presume that differences in the resolution

of the ultrasonic instruments and the experience levels of the observers affected the results. Microcalcification was observed in 41.2% of all the MTCs, and taller-than-wide shape was observed in 23.5% of all the MTCs. These proportions were lower than those observed in the PTCs; however, in line with the findings of previous studies (20,23), the difference was not significant. In addition, the MTC lesions had a more abundant blood supply than the PTCs (91.2% *vs.* 22.2%, $P < 0.05$), which is also consistent with the findings of previous studies (24,25).

We are of the view that the increased vascularity of MTC is related to its neuroendocrine properties. The abundant blood supply of MTC lesions leads to rapid tumor growth and compression of the surrounding tissues to form pseudocapsules and thus these tumors are usually larger than PTCs, and their well-defined margins are relatively more visible. Further, the multivariate logistic regression results showed that the hypoechoogenicity and malignant marginal features were independent risk factors for identifying malignant thyroid nodules, and the ultrasonic probability AUCs for identifying MTCs and PTCs were 0.873 (95% CI, 0.799–0.946) and 0.954 (95% CI, 0.917–0.991), respectively. Thus, the possibility of malignancy should be considered when any of the above suspicious malignant US characteristics are observed, especially hypoechoogenicity and malignant marginal features. The lymph node metastasis rate of MTC in our research was 14.8% (4/27), and 2 of these 4 patients had multifocal lesions. Notably, the lymph node metastasis rate of PTC was slightly lower than that of PTC in this study (18.0%, 9/50) and significantly lower than that reported in other studies (23,26). This might be due to the small sample size used in our study and the early detection of these cases.

To confirm the utility of the thyroid classification systems, we compared the efficacy of these 5 classification systems for the diagnosis of malignant thyroid tumors, including MTCs and PTCs. Our results showed that all 5 classification systems proved to be effective risk stratification methods in diagnosing malignant thyroid nodules, and had similar and good diagnostic performance. Among them, the Kwak-TIRADS had the best sensitivity, and the ACR-TIRADS had the best specificity. Our results also revealed the excellent diagnostic efficiency of the 5 guidelines in identifying PTCs. Notably, the C-TIRADS had higher AUC than the other 4 systems (0.956 *vs.* 0.916–0.946).

The 5 classification systems also had good diagnostic efficiency for MTCs; however, the AUCs for the MTCs were all lower than those for the PTCs (0.843 for the

ACR-TIRADS, 0.836 for the ATA guidelines, 0.868 for the EU-TIRADS, 0.856 for the Kwak-TIRADS, and 0.790 for the C-TIRADS). In addition, the best cut-off values for the PTCs according to the ACR-TIRADS, ATA guidelines, EU-, Kwak-, and C-TIRADS were TR5, high suspicion, TR5, TR 4c, and TR 4c, respectively; however, the corresponding best cut-off values for the MTCs were TR4, intermediate suspicion, TR4, TR 4b, and TR 4b. Thus, the risk stratifications for diagnosing MTCs were all 1 level lower than those for PTCs. Previous studies also found that MTC nodules were more likely to be classified as intermediate risk than PTC nodules due to their less suspicious malignant features (12,23).

In diagnosing MTCs, the EU-TIRADS showed the best diagnostic efficiency, with the highest AUC, sensitivity, and accuracy (0.868, 95% CI, 0.794–0.942; 80.7%, and 84%, respectively), and the ACR-TIRADS showed the best specificity (90.3%). Apart from the C-TIRADS, some studies examining the diagnostic efficacy of the other 4 guidelines reported similar results (27,28). The AUC of the C-TIRADS was lower than that of the other 4 guidelines in evaluating the MTCs (0.790, 95% CI, 0.694–0.886), and as such, the diagnostic efficiency of the C-TIRADS should be considered relatively poor. Our results showed that 94.1% of the MTCs (32/34) were stratified as intermediate to high risk in the ACR-, EU-, and Kwak-TIRADS, while 91.2% (31/34) of the MTCs in the ATA guidelines and only 55.9% (19/34) in the C-TIRADS, and 41.2% (14/34) were classified as low risk in C-TIRADS. We observed that 1 of the suspicious malignant characteristics in the C-TIRADS is a markedly hypoechoic nodule, which differs to that of the other 4 classifications, in which the set of suspicious echo characteristics are all hypoechoic. Only 2 MTC nodules were markedly hypoechoic among our MTC lesions, which might have led to a reduction in the risk classification performance of the C-TIRADS for MTCs.

The ultimate purpose of classifying thyroid nodules using each system is to make reasonable treatment plans. However, the sonographic characteristics and size cut-off values of the thyroid nodules recommended for FNAB vary across the different guidelines (29,30). In relation to the clinical recommendations of each guideline, the percentages of MTC nodules recommended for FNAB were 79.4% in the ACR-TIRADS, 88.2% in the ATA guidelines, 88.2% in the EU-TIRADS, 97.1% in the Kwak-TIRADS, and 85.3% in the C-TIRADS. The Kwak-TIRADS had the highest percentage of recommended FNAB cases, as FNAB was suggested for the nodules classified as TR4a and above.

Only 1 MTC nodule was classified as TR3 under the Kwak-TIRADS and not recommended for FNAB; however, this MTC nodule was classified as TR2 according to the ACR-TIRADS, as very low suspicion according to the ATA guidelines, and as TR3 according to the EU-TIRADS and C-TIRADS. We reviewed the sonograms and found that this MTC nodule was a solid-cystic mass with the characteristics of a benign thyroid nodule; however, as the maximum diameter of this nodule was >2.0 cm, FNAB was recommended under the ATA guidelines and EU-TIRADS but not under the other 3 guidelines. Moreover, the ATA guidelines and EU- and C-TIRADS had similar percentages of MTC nodules recommended for FNAB, while the ACR-TIRADS had the lowest percentage, which can be explained by the ACR-TIRADS having a larger size criteria for FNAB. Two MTC nodules classified as TR4 under the ACR-TIRADS were not recommended for FNAB due to the maximum diameters being smaller than the size cut-off of 1.5 cm. Conversely, these 2 MTC nodules were recommended for FNAB under the ATA guidelines and the EU- and C-TIRADS, as the size cut-off values for the similar categories were all 1.0 cm. Among all the MTC nodules not recommended for FNA, 3 nodules with an average diameter of 0.56–0.61 cm did not show obvious blood flow signals, which may be because the nodules were too small to be easily detected. Conversely, the other 4 MTC nodules had an abundant blood supply. If a thyroid nodule with an abundant blood supply is classified as TR4/4a or higher under the guidelines, we are of the view that FNAB should be performed to avoid any delays in the diagnosis and treatment of MTC. Thus, while the risk stratification of a small portion of MTC lesions may be underestimated due to atypical malignant sonographic manifestations, the adoption of the recommendations at different classification levels could improve the diagnostic efficiency and avoid delays in the diagnosis and treatment of MTC.

Our study compared the diagnostic performance of 5 thyroid risk stratification systems for MTC, including the most recently published C-TIRADS. To date, many studies have examined the applicability of different guidelines in thyroid carcinoma; however, these studies have mostly focused on PTC, and few studies have focused on MTC using the most commonly used and latest guidelines. The results of this study indicated that the 5 current thyroid risk stratification systems satisfactorily stratify MTCs, providing a foundation for further management. Additionally, the overall interobserver agreement was excellent between

the 2 US physicians with 10 and 6 years of experience, respectively, in diagnosing all the thyroid nodules and MTCs using the 5 guidelines.

Our studies had some limitations. First, it was a single-center retrospective study with a low incidence of MTC. The proportion of patients with MTC was low, and the patients who had incomplete clinical information and US images at our center were not included in this study, which might have caused inherent selection bias. A comprehensive multicenter prospective study needs to be conducted to confirm these findings in the future. Second, 2 US doctors retrospectively read the preoperative US images on a reporting workstation, which might have decreased the image quality and resulted in incomplete data. The 2 US doctors reviewed the 5 guidelines together, but their understandings of these classifications might nonetheless have differed. Thus, we tested the consistency of their classification results, and both doctors showed good agreement with each other.

Conclusions

In conclusion, according to the clinical decisions recommended by the thyroid risk stratification systems, the diagnostic efficacy of the 5 guidelines in detecting MTC, including the ACR-TIRADS, ATA guidelines, EU-TIRADS, Kwak-TIRADS, and C-TIRADS, is satisfactory, but it is not as good as that for PTC.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at <https://qims.amegroups.com/article/view/10.21037/qims-22-968/rc>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://qims.amegroups.com/article/view/10.21037/qims-22-968/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as

revised in 2013). The study was approved by the Medical Ethics Committee of Peking University First Hospital, and the requirement of individual consent for this retrospective analysis was waived.

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

References

1. Trimboli P, Giovanella L, Crescenzi A, Romanelli F, Valabrega S, Spriano G, Cremonini N, Guglielmi R, Papini E. Medullary thyroid cancer diagnosis: an appraisal. *Head Neck* 2014;36:1216-23.
2. Gorman B, Charboneau JW, James EM, Reading CC, Wold LE, Grant CS, Gharib H, Hay ID. Medullary thyroid carcinoma: role of high-resolution US. *Radiology* 1987;162:147-50.
3. Lee S, Shin JH, Han BK, Ko EY. Medullary thyroid carcinoma: comparison with papillary thyroid carcinoma and application of current sonographic criteria. *AJR Am J Roentgenol* 2010;194:1090-4.
4. Valderrabano P, Klippenstein DL, Tourtelot JB, Ma Z, Thompson ZJ, Lilienfeld HS, McIver B. New American Thyroid Association Sonographic Patterns for Thyroid Nodules Perform Well in Medullary Thyroid Carcinoma: Institutional Experience, Systematic Review, and Meta-Analysis. *Thyroid* 2016;26:1093-100.
5. Gharib H, Papini E, Paschke R, Duick DS, Valcavi R, Hegedüs L, Vitti P; AACE/AME/ETA Task Force on Thyroid Nodules. American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and European Thyroid Association medical guidelines for clinical practice for the diagnosis and management of thyroid nodules: Executive Summary of recommendations. *J Endocrinol Invest* 2010;33:287-91.
6. Brito JP, Gionfriddo MR, Al Nofal A, Boehmer KR, Leppin AL, Reading C, Callstrom M, Elraiyah TA, Prokop LJ, Stan MN, Murad MH, Morris JC, Montori VM. The accuracy of thyroid nodule ultrasound to predict thyroid cancer: systematic review and meta-analysis. *J Clin*

- Endocrinol Metab 2014;99:1253-63.
7. Kim SH, Kim BS, Jung SL, Lee JW, Yang PS, Kang BJ, Lim HW, Kim JY, Whang IY, Kwon HS, Jung CK. Ultrasonographic findings of medullary thyroid carcinoma: a comparison with papillary thyroid carcinoma. *Korean J Radiol* 2009;10:101-5.
 8. Horvath E, Majlis S, Rossi R, Franco C, Niedmann JP, Castro A, Dominguez M. An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. *J Clin Endocrinol Metab* 2009;94:1748-51.
 9. Park JY, Lee HJ, Jang HW, Kim HK, Yi JH, Lee W, Kim SH. A proposal for a thyroid imaging reporting and data system for ultrasound features of thyroid carcinoma. *Thyroid* 2009;19:1257-64.
 10. Xu SY, Zhan WW, Wang WH. Evaluation of Thyroid Nodules by a Scoring and Categorizing Method Based on Sonographic Features. *J Ultrasound Med* 2015;34:2179-85.
 11. 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.
 12. 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.
 13. 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.
 14. Kwak JY, Han KH, Yoon JH, Moon HJ, Son EJ, Park SH, Jung HK, Choi JS, Kim BM, Kim EK. Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk. *Radiology* 2011;260:892-9.
 15. 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.
 16. McCook TA, Putman CE, Dale JK, Wells SA. Review: Medullary carcinoma of the thyroid: radiographic features of a unique tumor. *AJR Am J Roentgenol* 1982;139:149-55.
 17. Pombo F, Rodríguez E, Cao JI, Martínez-Isla C. Cervical lymph node metastases of medullary thyroid carcinoma: CT findings. *Eur Radiol* 1997;7:99-101.
 18. Tessler FN, Middleton WD, Grant EG. Thyroid Imaging Reporting and Data System (TI-RADS): A User's Guide. *Radiology* 2018;287:1082.
 19. Woliński K, Rewaj-Łosyk M, Ruchała M. Sonographic features of medullary thyroid carcinomas--a systematic review and meta-analysis. *Endokrynol Pol* 2014;65:314-8.
 20. Trimboli P, Nasrollah N, Amendola S, Rossi F, Ramacciato G, Romanelli F, Aurello P, Crescenzi A, Laurenti O, Condorelli E, Ventura C, Valabrega S. Should we use ultrasound features associated with papillary thyroid cancer in diagnosing medullary thyroid cancer? *Endocr J* 2012;59:503-8.
 21. Wang H, Liu M, Yang J, Song Y. High frequency ultrasound features and pathological characteristics of medullary thyroid carcinoma. *Pak J Pharm Sci* 2016;29:2269-71.
 22. Zhou L, Chen B, Zhao M, Zhang H, Liang B. Sonographic features of medullary thyroid carcinomas according to tumor size: comparison with papillary thyroid carcinomas. *J Ultrasound Med* 2015;34:1003-9.
 23. Yun G, Kim YK, Choi SI, Kim JH. Medullary thyroid carcinoma: Application of Thyroid Imaging Reporting and Data System (TI-RADS) Classification. *Endocrine* 2018;61:285-92.
 24. Zhu J, Li X, Wei X, Yang X, Zhao J, Zhang S, Guo Z. The application value of modified thyroid imaging report and data system in diagnosing medullary thyroid carcinoma. *Cancer Med* 2019;8:3389-400.
 25. Zhang P, Liu H, Yang X, Pang L, Gu F, Yuan J, Ding L, Zhang J, Luo W. Comparison of contrast-enhanced ultrasound characteristics of inflammatory thyroid nodules and papillary thyroid carcinomas using a quantitative time-intensity curve: a propensity score matching analysis. *Quant Imaging Med Surg* 2022;12:5209-21.
 26. Oh HS, Kwon H, Song E, Jeon MJ, Song DE, Kim TY, Lee JH, Hong SJ, Kim WB, Shong YK, Baek JH, Kim WG. Preoperative Clinical and Sonographic Predictors for Lateral Cervical Lymph Node Metastases in Sporadic Medullary Thyroid Carcinoma. *Thyroid* 2018;28:362-8.
 27. 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;287:893-900.
28. Ha EJ, Na DG, Moon WJ, Lee YH, Choi N. Diagnostic Performance of Ultrasound-Based Risk-Stratification Systems for Thyroid Nodules: Comparison of the 2015 American Thyroid Association Guidelines with the 2016 Korean Thyroid Association/Korean Society of Thyroid Radiology and 2017 American College of Radiology Guidelines. *Thyroid* 2018;28:1532-7. Erratum in: *Thyroid* 2019;29:159.
29. 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.
30. Trimboli P, Treglia G, Guidobaldi L, Romanelli F, Nigri G, Valabrega S, Sadeghi R, Crescenzi A, Faquin WC, Bongiovanni M, Giovanella L. Detection rate of FNA cytology in medullary thyroid carcinoma: a meta-analysis. *Clin Endocrinol (Oxf)* 2015;82:280-5.

Cite this article as: Jiang L, Zhu HB, Liang ZW, Chen L, Sun XM, Shao YH, Chen LZ. Comparison of the diagnostic performance and clinical role of different ultrasound-based thyroid malignancy risk stratification systems for medullary thyroid carcinoma. *Quant Imaging Med Surg* 2023;13(6):3776-3788. doi: 10.21037/qims-22-968