



# Explore the diagnostic performance of 2020 Chinese Thyroid Imaging Reporting and Data Systems by comparing with the 2017 ACR-TIRADS guidelines: a single-center study

Miaomiao Cai<sup>1</sup> · Libo Chen<sup>1</sup> · Limin Shui<sup>1</sup> · Xuan Lv<sup>1</sup> · Hui Wang<sup>1</sup>

Received: 10 October 2022 / Accepted: 8 January 2023 / Published online: 17 March 2023  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

**Objective** To compare the diagnostic efficacy of the Chinese Thyroid Imaging Reporting and Data Systems (C-TIRADS) with the well-accepted ACR-TIRADS guidelines in identifying benign from malignant thyroid nodules.

**Methods** A total of 2064 nodules were collected from 1627 patients undergoing thyroid ultrasonography in our center between October 2019 and November 2021. Nodules were divided into two groups: “≥1 cm” and “<1 cm”. Ultrasound features of each nodule were observed and recorded by two physicians with more than 15 years of experience and classified according to the ACR-TIRADS and C-TIRADS guidelines, respectively.

**Results** The area under the curve of the ACR-TIRADS guideline was higher than that of the C-TIRADS guideline (0.922,  $P = 0.017$ ), the specificity and positive predictive value of the C-TIRADS guideline were higher (81.64%, 88.72%, all  $P < 0.05$ ), which was more significant in the subgroup of nodules <1 cm ( $P = 0.001$ ). In addition, there was no statistical difference between the two guidelines in the diagnostic efficacy indicators for nodules ≥1 cm. The ACR-TIRADS effectively reduced unnecessary biopsies compared with the C-TIRADS ( $P < 0.05$ ).

**Conclusions** There was high agreement between the two guidelines for the diagnosis of thyroid nodules, C-TIRADS guidelines had a higher specificity and simplicity while were inferior to the ACR-TIRADS guidelines in terms of reducing the number of biopsies.

**Keywords** Thyroid nodules · Ultrasound · Fine-needle aspiration biopsy · TIRADS

## Introduction

Thyroid nodules are a relatively common disease of the endocrine system, and in recent years, studies have shown that the detection rate of malignant thyroid nodules is increasing yearly and the age of onset is gradually becoming younger. The reported prevalence ranges from 10 to 50% due to differences in age, race, and gender composition distribution, as well as relatively small sample sizes in previous studies [1, 2]. The number of patients with overdiagnosed low-risk nodules is predicted to reach 5.1 million from 2019 to 2030 [3]. Accurate assessment of the malignant risk of thyroid

nodules remains a key clinical concern as those overdiagnosed will receive extensive and unnecessary treatment. Current guidelines for the evaluation of thyroid nodules recommend high-resolution ultrasound as the core method for deciding whether to schedule FNAB and subsequent treatment options for patients [4–7].

To better assess the ultrasound characteristics of thyroid nodules and to produce standardized reports, researchers have recommended the TIRADS as the standard risk stratification scheme, which has been in existence for more than a decade and different versions have been developed. The most widely investigated is the thyroid nodule risk stratification system published by the American College of Radiology in 2017. ACR-TIRADS [8] is a score-based system that allows for the classification of all nodules. Thyroid nodules are classified into five categories based on ultrasound characteristics of the composition, echogenicity, shape, margins, and strongly echogenic foci. The decision to perform FNAB or further

✉ Libo Chen  
lbchen@jlu.edu.cn

<sup>1</sup> China-Japan Union Hospital, Jilin University, Changchun, China

examinations is based on the classification of the nodule as well as its size.

The ACR TIRADS Committee estimated that the malignant risk of TR5 nodules was only greater than 20% [9]. FNAB results are required prior to surgery according to internationally accepted principles for the diagnosis and management of thyroid nodules. FNAB has not been widely implemented in a few local primary hospitals and there is no uniform application management guideline in China [10]. Based on the above situation, the Chinese Society of Ultrasound Medicine has developed the “2020 Chinese guideline for ultrasound risk stratification of thyroid nodules for malignancy: C-TIRADS” [11], which uses the counting method to classify thyroid nodules, following the well-established BI-RADS for breast masses, and the rate of suspected malignancy for each category is highly similar to that of BI-RADS.

Currently, fewer studies have been reported on the C-TIRADS, which gradually became popular in China although the most widely used was ACR-TIRADS before its publication. The aim of this study is to investigate the diagnostic performance and the unnecessary FNAB efficacy by comparing it with the highly recognized ACR-TIRADS guideline to identify their strengths and weaknesses.

## Materials and methods

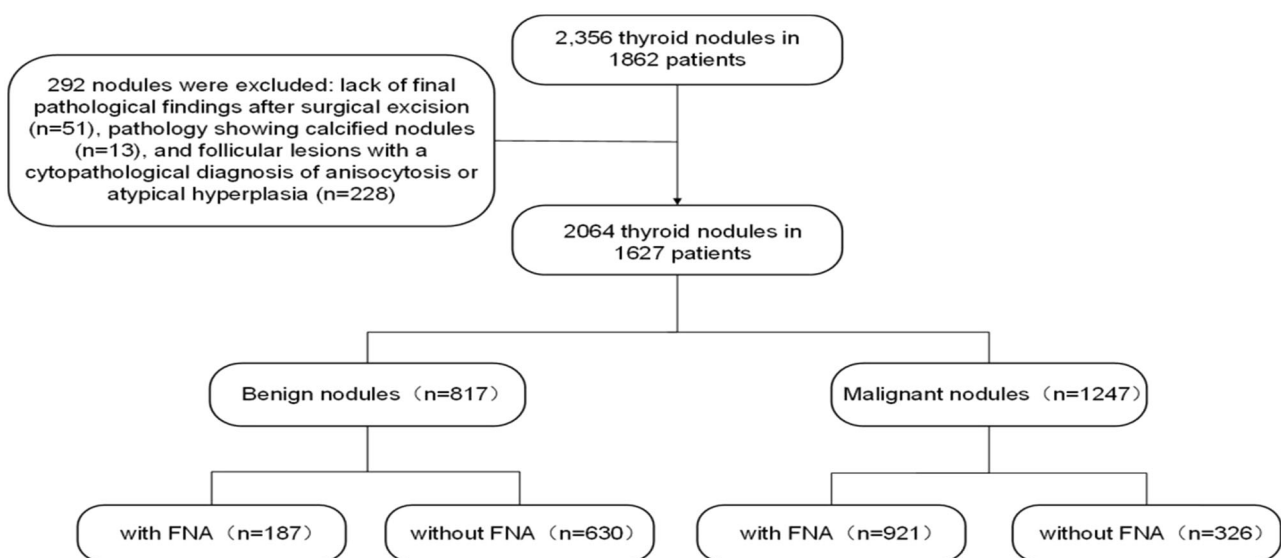
This study was approved by the Institutional Review Board and Ethics Committee of Sino-Japanese Friendship Hospital Affiliated with Jilin University (code: 20221124001).

## Patients

The study population was 2356 thyroid nodules in 1862 consecutive patients who underwent thyroid ultrasonography with FNA biopsy or had postoperative pathological findings at our hospital from October 2019 to November 2021. Inclusion criteria were as follows: (1) nodules with definite pathological findings; (2) FNA biopsy or preoperative ultrasound with full report and saving of images as JPEG files. Exclusion criteria were as follows: (1) absence of complete preoperative ultrasound images of thyroid nodules; (2) “zombie” thyroid nodules; (3) lack of definitive pathological diagnosis after surgical resection; (4) atypical or few heterogeneous cells diagnosed by FNAB. Of these, 292 nodules were excluded due to lack of final pathological findings after surgical resection ( $n = 51$ ), Pathological findings of calcified nodules ( $n = 13$ ), cytopathological diagnosis of follicular lesions with heterogeneous cells or atypical hyperplasia ( $n = 228$ ). Ultimately, a total of 2064 thyroid nodules from 1627 patients were included in this study (Fig. 1).

## Ultrasonography

All conventional examinations were performed with PHILIP EPIQ7 ultrasound diagnostic instrument equipped with a broadband line array probe ranging from 5 to 18 MHz. Ultrasound images of thyroid nodules were recorded and stored in JPEG file format by a US specialist with over 15 years of experience. The location of the nodule (right lobe, left lobe, and isthmus) should be recorded, Repeatedly measure the upper and lower



**Fig. 1** Flowchart of the selection of patients with 2064 thyroid nodules

**Table 1** Number of thyroid nodules in different size categories

Groups	Number of nodules (%)	Pathologically benign (%)	Pathological malignancy (%)	Nodule size (cm)	Z	P
<0.5	374 (18.1)	119 (31.8)	255 (68.2)	0.30 (0.30, 0.40), 0.34 ± 0.07	-12.004	<0.001
0.5–1.0	675 (32.7)	215 (31.9)	460 (68.1)	0.70 (0.60, 0.80), 0.68 ± 0.14		
1.0–1.5	508 (24.6)	133 (26.2)	375 (73.8)	1.20 (1.10, 1.30), 1.18 ± 0.13		
1.5–2.0	178 (8.6)	69 (38.8)	109 (61.2)	1.70 (1.60, 1.80), 1.67 ± 0.13		
2.0–2.5	74 (3.6)	54 (73.0)	20 (27.0)	2.20 (2.10, 2.30), 2.19 ± 0.13		
≥2.5	255 (12.4)	227 (89.0)	28 (11.0)	3.80 (3.10, 4.80), 4.00 ± 1.14		

Data in parentheses are percentages

diameters, right and left diameters, anterior and posterior diameters of the nodules three times and record the average values.

### Nodules analysis

All selected thyroid nodules were evaluated by two experts with more than 15 years of experience in ultrasound diagnosis without providing pathological findings. They did not analyze all nodules first and then discuss the results to reach a consensus. Instead, each expert simultaneously analyzed each nodule according to the C-TIRADS and ACR-TIRADS guidelines, immediately discussed the nodule to reach a consensus and then proceeded to the next nodule analysis. When two experts disagree, they discuss it with a third expert who has 20 years of experience in diagnosis. After reaching a consensus on the previous nodules, this protocol serves as a standard for subsequent analysis. The classification was based on the ultrasound images of nodules in five aspects: composition, echogenicity, morphology, margins and strongly echogenic foci.

### Statistics

SPSS 26.0 and MedCalc 20.0.22 were used for statistical analysis. Quantitative data were presented as median and quartile, while qualitative data were presented as composition ratio. The distribution of nodule size between the groups was analyzed by the Mann–Whitney test. The generation of the receiver operating characteristic (ROC) curve was based on the pathological findings to determine the optimal cut-off value for benign and malignant thyroid nodules by the ACR-TIRADS and C-TIRADS guidelines. The DeLong test was used to compare the differences in the area under the curve (AUC) between the two guidelines. The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (+LR), negative likelihood ratio (–LR) and accuracy of the two guidelines were compared by the  $\chi^2$  test. Agreement between the two guidelines and with pathology was compared by Cohen’s kappa coefficient. The rate of unnecessary

biopsies was further calculated for both according to the guidelines.  $P < 0.05$  was considered a significant difference.

## Results

### Basic characteristics of nodules

A total of 1247 nodules were malignant and 817 were benign of all thyroid nodules. In all, 944 nodules were located in the left lobe, 1063 nodules were located in the right lobe and 57 nodules were located in the isthmus of the thyroid. There were 1049 nodules less than 1 cm in diameter, of which 334 were benign and 715 were malignant. The number of nodules in different categories is shown in Table 1.

### Risk of malignancy for different categories of nodules in both guidelines

The actual malignant rate for nodules graded TR3 and above in the ACR-TIRADS guideline is higher than the recommended malignant rate, which is more pronounced in nodules less than 1 cm in diameter. The actual malignant rate of nodules within the TR3 to TR4C according to the C-TIRADS guideline was also higher than the recommended, but the difference was smaller than that of the ACR-TIRADS guideline. In addition, the C-TIRADS guideline assessed accurately that the malignant rate of nodules in TR5 was greater than 90% (Table 2). Evaluation of the various ultrasound signs in the C-TIRADS guideline revealed that the malignant rate of nodules with very hypoechoic, vertical and microcalcified features was above 85% (Table 3).

### Comparison of diagnostic efficacy in different guidelines

All 2064 nodules could be classified according to both guidelines, as shown in the ROC curve (Fig. 2), the optimal diagnostic cut-off value for the ACR-TIRADS guideline was 3.5, which means the thyroid nodules are diagnosed as

**Table 2** Comparison of malignancy risk stratification results between C-TIRADS and ACR-TIRADS guidelines

Guidelines	Overall ( <i>n</i> = 2064)						<1 cm ( <i>n</i> = 1049)					
	Grade	Total number	Malignant	Benign	Calculated malignant rate	Recommended malignant rate	Total number	Malignant	Benign	Calculated malignant rate		
ACR-TIRADS	TR1	67 (3.2)	1	66	1.5	≤2	47 (4.5)	1	46	2.1		
	TR2	356 (17.2)	7	349	2.0	≤2	108 (10.3)	6	102	5.6		
	TR3	245 (11.9)	39	206	15.9	<5	107 (10.2)	29	78	27.1		
	TR4	352 (17.1)	225	127	63.9	5–20	194 (18.5)	123	71	63.4		
	TR5	1044 (50.6)	975	69	93.4	>20	593 (56.5)	556	37	93.8		
C-TIRADS	TR2	25 (1.2)	0	25	0	0	13 (1.2)	0	13	0		
	TR3	428 (20.7)	23	405	5.4	<2	153 (14.6)	18	135	11.8		
	TR4A	281 (13.6)	44	237	15.7	2–10	151 (14.4)	35	116	23.2		
	TR4B	283 (13.7)	214	69	75.6	10–50	171 (16.3)	137	34	80.1		
	TR4C	940 (45.5)	866	74	92.1	50–90	523 (49.9)	490	33	93.7		
	TR5	107 (5.2)	100	7	93.5	>90	38 (3.6)	35	3	92.1		

Data in parentheses are percentages

malignant when  $TR \geq 4$ , and diagnosed as benign when  $TR < 4$ . C-TIRADS guideline could distinguish between benign and malignant nodules with a cut-off value of 3.5, meaning that a nodule is diagnosed as malignant when it is rated as moderately suspicious for malignancy or higher, and as benign when it is rated as low suspicion for malignancy or lower. A comparison of the overall diagnostic agreement between the two guidelines for thyroid nodules by the above assessment criteria showed high agreement ( $Kappa = 0.86 > 0.75$ ,  $P < 0.001$ ). The specific value is shown in Table 4. The results showed higher specificity and PPV of C-TIRADS guideline among all nodules (81.64%, 88.72%, both  $P < 0.05$ ). The sensitivity and NPV of ACR-TIRADS guideline were higher (96.23%, 96.26%, both  $P > 0.05$ ). The diagnostic effectiveness of the two guidelines for nodules of different sizes is shown in Table 5. The difference in AUC between nodules  $< 1$  cm and  $\geq 1$  cm groups was not statistically significant (all  $P > 0.05$ ).

### Comparison of the unnecessary FNAB rates between C-TIRADS and ACR-TIRADS guidelines

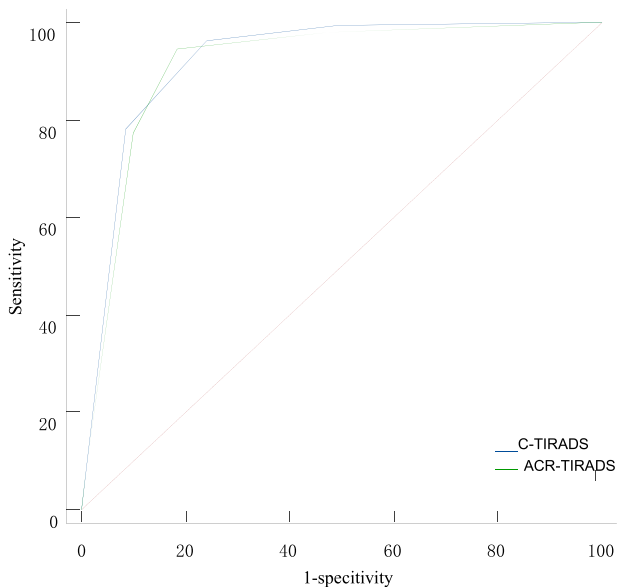
FNAB was performed in 1108 of all thyroid nodules. 695 of these nodules were recommended for biopsy according to the C-TIRADS guideline, and 630 of these nodules were recommended for biopsy according to the ACR-TIRADS guideline. The unnecessary biopsy rate for the C-TIRADS guideline was higher than the ACR-TIRADS guideline (24.9%, 20.2%), and the same result was seen in terms of the false-positive rate for FNAB (21.2%, 15.5%). The difference was statistically significant (all  $P < 0.05$ ). Furthermore, There is almost no difference in the rate of missed biopsies between the two guidelines (Table 6).

### Discussion

The optimal diagnostic threshold of the C-TIRADS guideline in this retrospective study was 4B, whereas Chen and Wu [12] showed that the diagnostic threshold of this guideline was 4C, which may be due to the bias of the data source; the optimal diagnostic threshold of the ACR-TIRADS guideline was TR4, which is consistent with the findings of Mao et al. [13]. The variability in optimal diagnostic thresholds between the two guidelines is due to the different scoring criteria and malignant risk for each category. The C-TIRADS has a slightly higher accuracy rate (89.49%) than the ACR-TIRADS (88.23%), which improves the accuracy of classification while taking into account clinical usability. The AUC of the ACR-TIRADS guideline (0.922) was slightly higher than that of the C-TIRADS guideline (0.913),  $P < 0.05$ , indicating a better overall diagnostic performance.

**Table 3** Frequency of ultrasound findings in all and <1 cm thyroid nodules according to the C-TIRADS

Ultrasound signs		Nodules		Benign		Malignant		Risk of malignancy	
		All (n = 2064)	<1 cm (n = 1049)	All (n = 817)	<1 cm (n = 334)	All (n = 1247)	<1 cm (n = 715)	All (%)	<1 cm (%)
Solidity	Yes	1557	874	341	179	1216	695	78.1	79.5
	No	507	175	476	155	31	20	6.1	11.4
Very hypoechoic	Yes	326	194	23	13	303	181	92.9	93.3
	No	1738	855	794	321	944	534	54.3	62.5
Vertical position	Yes	869	513	71	34	798	479	91.8	93.4
	No	1195	536	746	300	449	236	37.6	44.0
Blurred, irregular, outwardly invasive margins	Yes	1310	708	167	71	1143	637	87.3	90.0
	No	754	341	650	263	104	78	13.8	22.9
Microcalcification	Yes	526	195	60	17	466	178	88.6	91.3
	No	1538	854	757	317	781	537	50.8	62.9
Comet tail artifact	Yes	30	15	30	15	0	0	0	0
	No	2034	1034	787	319	1247	715	61.3	69.1



**Fig. 2** The ROC curve of the C-TIRADS and ACR guidelines

Thyroid nodules are evaluated in five different ultrasound characteristics based on the ACR-TIRADS guidelines. Note that all ultrasound features are scored individually, except for the strong echogenic foci, which are scored superimposed. Thyroid nodules are classified into five risk categories, ranging from TR1 (benign) to TR5 (highly suspicious of malignancy). The C-TIRADS guidelines show that solid, microcalcified, very hypoechoic, blurred margins, irregular margins or extrathyroidal invasion, and vertical position (aspect ratio >1) are ultrasound features of nodules that are suspicious of malignancy (Fig. 3). Microcalcifications are one of the most specific features of malignancy with a specificity of 85.8–95% [14–17]. In contrast to the

ACR-TIRADS guidelines, which score different types of calcifications, the C-TIRADS guidelines only score microcalcifications for risk stratification, whereas previous studies found that eggshell discontinuity coarse calcifications and peripheral calcifications were highly associated with malignancy [18–20]. In addition, the C-TIRADS guidelines score only very hypoechoic, whereas the study found 78% of papillary thyroid carcinomas are hypoechoic, and since 30.6–55% of benign nodules are also hypoechoic, it is a sensitive sign but not specific [21–23]. ACR-TIRADS scores hypoechoic as 2, which may explain its higher sensitivity than the C-TIRADS guidelines. Focal strong echogenicity within thyroid nodules is classified into three types according to the C-TIRADS, including microcalcifications, comet tail artifacts, and punctate strong echogenicity of uncertain significance, one or more of which may be present in the same nodule. The C-TIRADS guideline is innovative in not scoring focal strong echoes of the uncertain significance of <1 mm. "Comet tail" artifact is considered a benign feature and given a score of –1 according to the C-TIRADS while the ACR-TIRADS does not. These may explain its higher specificity.

The C-TIRADS guidelines have higher PPV, +LR, and –LR in both groups of nodules, which may be due to their more rigorous control of malignant signs, and the five malignant signs were derived from multivariate logistic regression analysis with large multicenter data, which is relatively realistic and reliable, and the interpretation of each ultrasound feature as only yes and no. By counting the number of five positive and one negative ultrasound features to determine the category. This simplified classification method was easy to improve the interobserver agreement. In contrast, the ACR-TIRADS guidelines

**Table 4** Comparison of diagnostic consistency between ACR-TIRADS and C-TIRADS guidelines

Groups		Pathological malignant	Pathological benign	Kappa <sup>a</sup>	P
ACR-TIRADS	Malignant	1200	196	0.75	<0.001
	Benign	47	621		
C-TIRADS	Malignant	1180	150	0.78	<0.001
	Benign	67	667		
Kappa <sup>b</sup>		0.86			
P value		<0.001			

<sup>a</sup>Kappa was calculated for the ACR-TIRADS and C-TIRADS guidelines separately compared with pathology

<sup>b</sup>Kappa was calculated for the overall comparison of the two guidelines

**Table 5** Comparison of the diagnostic efficacy of the two guidelines for thyroid nodules

Groups	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	AUC	Accuracy (%)	+LR	−LR
All nodules								
ACR-TIRADS	96.23 [94.98, 97.19]	76.01 [72.90, 78.87]	85.96 [84.00, 87.72]	92.96 [90.69, 94.73]	0.922 [0.908, 0.935]	88.23 [86.77, 89.55]	4.01 [3.55, 4.53]	0.05 [0.04, 0.07]
C-TIRADS	94.63 [93.19, 95.78]	81.64 [78.78, 84.20]	88.72 [86.87, 90.35]	90.87 [88.50, 92.81]	0.913 [0.899, 0.928]	89.49 [88.09, 90.74]	5.15 [4.46, 5.96]	0.07 [0.05, 0.08]
P	0.055	0.005	0.030	0.152	0.017	0.198	–	–
<1 cm nodule								
ACR-TIRADS	94.97 [93.03, 96.40]	67.66 [62.32, 72.60]	86.28 [83.63, 88.56]	86.26 [81.35, 90.07]	0.894 [0.871, 0.918]	86.27 [84.05, 88.22]	2.94 [2.51, 3.43]	0.07 [0.05, 0.10]
C-TIRADS	92.58 [90.35, 94.35]	79.04 [74.20, 83.20]	90.44 [88.02, 92.42]	83.28 [78.61, 87.13]	0.891 [0.867, 0.916]	88.27 [86.18, 90.08]	4.42 [3.58, 5.45]	0.09 [0.07, 0.12]
P	0.063	0.001	0.012	0.323	0.524	0.169	–	–
≥1 cm nodule								
ACR-TIRADS	97.93 [96.22, 98.91]	81.78 [77.98, 85.06]	85.55 [82.45, 88.19]	97.29 [95.06, 98.57]	0.943 [0.927, 0.958]	90.25 [88.27, 91.93]	5.38 [4.45, 6.50]	0.03 [0.02, 0.05]
C-TIRADS	97.37 [95.52, 98.50]	83.44 [79.75, 86.58]	86.62 [83.57, 89.19]	96.64 [94.30, 98.08]	0.933 [0.916, 0.951]	90.74 [88.80, 92.37]	5.88 [4.81, 7.18]	0.03 [0.02, 0.06]
P	0.544	0.497	0.591	0.588	0.053	0.705	–	–

Data in parentheses are 95% confidence intervals

require multiple interpretations of each ultrasound feature, which may overlap between categories, and the risk of malignancy of some ultrasound features may vary depending on other features. In addition, the scores of each ultrasound feature are mainly determined by expert opinion rather than statistical analysis, which may explain the lower accuracy of ACR-TIRADS in predicting the malignant risk of thyroid nodules compared with C-TIRADS. Our study takes interobserver variability into account to some extent, and it has been suggested that artificial intelligence can effectively limit interobserver variability through standardized mathematical algorithms [24].

This study demonstrated that in nodules ≥1 cm, C-TIRADS showed higher sensitivity, specificity, PPV, accuracy, and AUC than nodules in the <1 cm subgroup, indicating better diagnostic performance in these nodules. In nodules ≥1 cm in diameter, there was no significant difference between the diagnostic efficacy of the ACR-TIRADS

guideline and the C-TIRADS guideline based on ultrasound modality, and in the nodule diameter <1 cm group, the specificity of the C-TIRADS guideline (79.04%) was significantly higher than that of the ACR-TIRADS guideline (67.66%), while there was no significant difference in sensitivity. The findings of Zhou et al. [25] were similar to the present study. In contrast, the results of Zhu et al. [26] showed that the C-TIRADS guideline had a high specificity regardless of whether the nodule diameter was smaller or larger than 1 cm, and its sensitivity was significantly lower than that of the ACR-TIRADS guideline. Currently, studies on the diagnostic ability of each guideline for subcentimeter nodules are less reported and controversial, but it is clear that in our study, the diagnostic efficacy of both guidelines for nodules <1 cm in diameter was not inferior to that of nodules ≥1 cm.

Recommendations for FNA or ultrasound follow-up were based on the grade and its maximum diameter of the

**Table 6** Comparison of unnecessary biopsy rates between C-TIRADS and ACR-TIRADS guidelines

Guidelines	Grade	Nodule size	No. of FNABs (A)	Malignant nodules among A (B)	Benign nodules among A (C)	Unnecessary biopsy rate (C/A)	Positive rate (B/A)	False-positive rate (C/817)	Missed biopsy rate (D)
ACR-TIRADS	-	-	630	503	127	20.2 [17.2, 23.5]	79.8 [76.5, 82.8]	15.5 [13.2, 18.2]	59.7 [56.9, 62.4]
	TR3	≥25 mm	64 (26.1)	3	61	95.3 [87.1, 98.4]	4.7 [1.6, 12.9]	7.5 [5.9, 9.5]	92.3 [79.7, 97.4]
	TR4	≥15 mm	115 (32.7)	81	34	29.6 [22.0, 38.5]	70.4 [61.5, 78.0]	4.2 [2.9, 5.8]	64.0 [57.5, 69.9]
	TR5	≥10 m	451 (43.2)	419	32	7.1 [5.1, 9.9]	92.9 [90.2, 94.9]	3.9 [2.8, 5.5]	57.0 [53.9, 60.1]
	-	-	695	522	173	24.9 [21.8, 28.2]	75.1 [71.8, 78.2]	21.2 [18.5, 24.1]	58.1 [55.4, 60.9]
C-TIRADS	TR4A	≥15 mm	98 (34.9)	5	93	94.9 [88.6, 97.8]	5.1 [2.2, 11.4]	11.4 [9.4, 13.7]	88.6 [76.0, 95.1]
	TR4B	≥10 mm	112 (39.6)	77	35	31.3 [23.4, 40.3]	68.7 [59.7, 76.6]	4.3 [3.1, 5.9]	64.0 [57.4, 70.2]
	TR4C	≥10 mm	416 (44.3)	375	41	9.9 [7.4, 13.1]	90.1 [86.9, 92.7]	5.0 [3.7, 6.7]	56.7 [53.4, 59.9]
	TR5	≥10 mm	69 (64.5)	65	4	5.8 [2.3, 13.9]	94.2 [86.0, 97.7]	0.5 [0.2, 1.3]	35.0 [26.4, 44.8]
	P	-	-	-	-	0.040	0.040	0.003	0.439

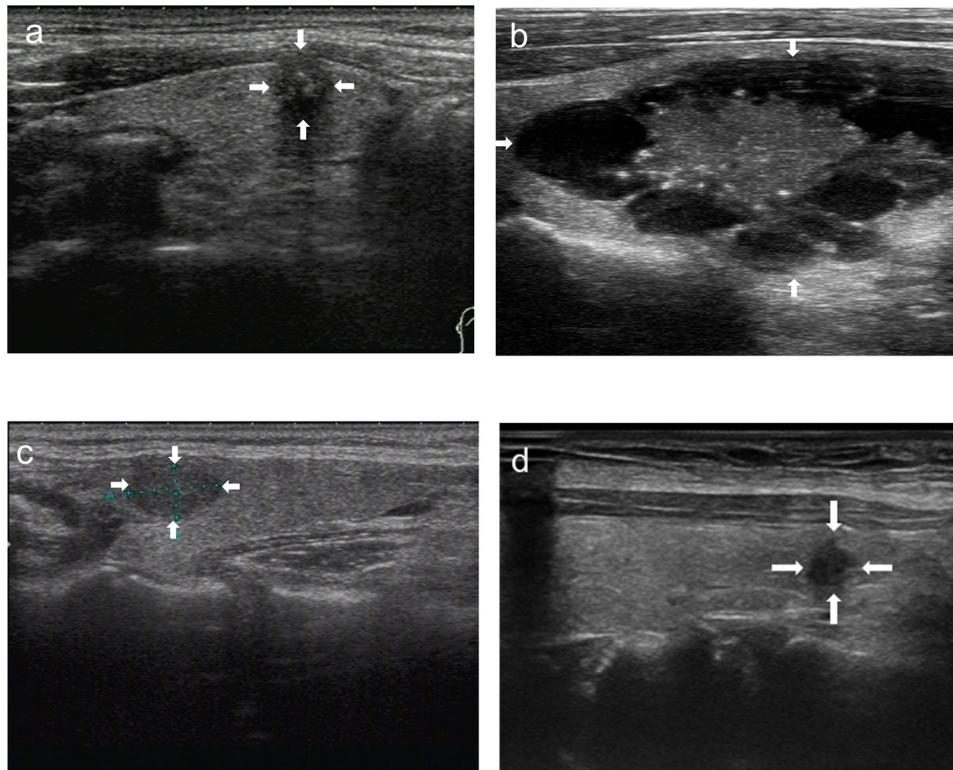
Data in parentheses are percentages; the numbers in middle brackets are 95% confidence intervals; calculation formula for D is (number of pathological malignancies-B)/number of pathological malignancies

nodule. A biopsy is recommended for nodules in risk grades TR3 to TR5 according to ACR-TIRADS, which have a size threshold of 2.5 cm for FNA for nodules of mild suspicion of malignancy. A biopsy is recommended for nodules in grades TR4A and above according to the C-TIRADS, with a size threshold of 1.5 cm for nodules of TR4A. if the nodule is multifocal, or immediately adjacent to the perineum, trachea, or involves the recurrent laryngeal nerve, the diameter for biopsy is reduced to 1 cm. Both guidelines, while reducing the rate of unnecessary biopsies, are expected to result in a higher percentage of missed malignant nodules. This is unavoidable because some malignant tumors have benign ultrasound features. In the present study, the rate of unnecessary biopsies was lower in ACR-TIRADS (20.2%) than in C-TIRADS (24.9%), *P* < 0.05. This may be due to the higher biopsy threshold for nodules of equal malignant risk compared with C-TIRADS. The nodal leakage rate was slightly higher in ACR-TIRADS (59.7%) than in C-TIRADS guidelines (58.1%), and the difference was not statistically significant (*P* > 0.05).

This study also has some limitations. First, most nodules are diagnosed based on pathological findings and the rest based on definite cytopathological findings, while the medical level in China is lower than that in developed countries, FNA has not been widely carried out in Chinese hospitals at all levels and its value has not been widely recognized in China, which may lead to selection bias. Second, this was a single-center retrospective study. Although the consistency of diagnostic results for all thyroid nodules was guaranteed, the heterogeneity of the patient population was less than that of a multicenter study. Third, our hospital is a tertiary referral center for patients with more severe diseases, which may have led to a sample bias that increased the proportion of malignant nodules and decreased the number of low-grade nodules, thus affecting the diagnostic efficacy of the guidelines, which explains why the calculated malignancy rate for certain grades of nodules classified according to the guidelines is much greater than the recommended. Finally, the ACR-TIRADS guidelines do not recommend FNAB for nodules smaller than 1 cm, however, C-TIRADS recommends FNAB for nodules smaller than 0.5 cm under certain conditions. in this study, 50.8% of nodules were smaller than 1 cm, which somewhat limits the comparison of C-TIRADS with other guidelines.

### Conclusion

Both guidelines showed excellent diagnostic efficacy and overall diagnostic consistency for thyroid nodules. The ability to diagnose nodules in the <1 cm subgroup was not inferior to nodules ≥1 cm. Compared with ACR-TIRADS,



**Fig. 3** **a** It is a solid, hypoechoic, vertically located nodule that invaded the thyroid peritoneum. The pathological finding showed a malignant nodule. ACR guidelines classified it as TR5, meaning a malignant nodule. C-TIRADS guidelines classified it as TR4C, meaning a malignant nodule. The diagnostic results are the same for both guidelines, indicating a good consistency of diagnosis. **b** It is a mixed cystic-solid, microcalcification, irregularly shaped nodule. The pathological finding showed a malignant nodule. Classified by the ACR guidelines as TR5, meaning a malignant nodule. The C-TIRADS guidelines evaluate multiple predominantly solid and/or predominantly cystic nodules with similar ultrasonic manifestations in the thyroid as TR3, meaning benign nodules. The ACR guidelines are correct,

indicating that the sensitivity of the ACR guidelines is relatively high. **c** It is a solid, hypoechoic, smooth-edged nodule. The pathological finding showed a benign nodule. Classified by ACR guidelines as TR4, meaning a malignant nodule. And classified by C-TIRADS guidelines as TR4A, meaning a benign nodule. The diagnosis was correct by C-TIRADS guidelines, indicating its higher specificity. **d** It is a solid, hypoechoic, marginal lobulated nodule. The pathological finding showed a malignant nodule. Classified as TR4 by the ACR guidelines, meaning a malignant nodule. Classified as TR4B by the C-TIRADS guidelines, meaning a malignant nodule. The diagnosis is correct in both guidelines, indicating a high degree of agreement between the two guidelines

although the C-TIRADS guideline has a slightly higher rate of unnecessary biopsies, it is simple and significantly improves the stratification efficiency of thyroid nodules, which is more suitable for the Chinese situation.

### Compliance with ethical standards

**Conflict of interest** The authors declare no competing interests.

**Ethical approval** All procedures performed in research involving human participants comply with institutional and/or national research council ethical standards.

### References

1. L. Xu, F. Zeng, Y. Wang, Y. Bai, X. Shan, L. Kong, Prevalence and associated metabolic factors for thyroid nodules: a cross-sectional study in Southwest of China with more than 120 thousand populations. *BMC Endocr. Disord.* **21**(1), 175 (2021)
2. B. Song, Z. Zuo, J. Tan, J. Guo, W. Teng, Y. Lu, C. Liu, Association of thyroid nodules with adiposity: a community-based cross-sectional study in China. *BMC Endocr. Disord.* **18**(1), 3 (2018)
3. M. Li, R. Zheng, L. Dal Maso, S. Zhang, W. Wei, S. Vaccarella, Mapping overdiagnosis of thyroid cancer in China. *Lancet Diabetes Endocrinol.* **9**(6), 330–332 (2021)
4. S. Guth, U. Theune, J. Aberle, A. Galach, C.M. Bamberger, Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. *Eur. J. Clin. Invest.* **39**(8), 699–706 (2009)
5. S.H. Choi, E.-K. Kim, J.Y. Kwak, M.J. Kim, E.J. Son, Interobserver and intraobserver variations in ultrasound assessment of thyroid nodules. *Thyroid* **20**, 167–172 (2010)
6. T.L. Henrichsen, C.C. Reading, Thyroid ultrasonography. Part 2: nodules. *Radio. Clin. North Am.* **49**, 417–424 (2011)
7. W.J. Moon, J.H. Baek, S.L. Jung et al. Ultrasonography and the ultrasound-based management of thyroid nodules. *Korean J. Radio.* **12**, 1–14 (2011)



8. F.N. Tessler, W.D. Middleton, E.G. Grant, J.K. Hoang, L.L. Berland, S.A. Teefey, J.J. Cronan, M.D. Beland, T.S. Desser, M.C. Frates, L.W. Hammers, U.M. Hamper, J.E. Langer, C.C. Reading, L.M. Scoutt, A.T. Stavros, ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. *J. Am. Coll. Radio.* **14**(5), 587–595 (2017)
9. W.D. Middleton, S.A. Teefey, C.C. Reading et al. Multi-institutional analysis of thyroid nodule risk stratification using the American College of Radiology Thyroid Imaging Reporting and Data System. *AJR Am. J. Roentgenol.* **208**(6), 1331–1341 (2017)
10. B.R. Haugen, E.K. Alexander, K.C. Bible et al. 2015 American Thyroid Association Management Guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* **26**(1), 1–133 (2016)
11. Superficial Organs and Vasculature Group of Ultrasound Medicine Branch of Chinese Medical Association, Chinese Thyroid and Breast Ultrasound Artificial Intelligence Consortium. 2020 Chinese guidelines for ultrasound malignancy risk stratification of thyroid nodules: C-TIRADS. *Chin J Ultrasound Imaging* **30**(3), 185–200 (2021)
12. Q.F. Chen, S.Z. Wu, Comparative study on the diagnostic efficacy of C-TIRADS and ACR-TIRADS for stratifying the risk of malignancy in thyroid nodules. *Chin. J. Ultrasound Imaging* **30**(10), 861–867 (2021)
13. S. Mao, L.P. Zhao, X.H. Li, Y.F. Sun, H. Su, Y. Zhang, K.L. Li, D.C. Fan, M.Y. Zhang, Z.G. Sun, S.C. Wang, The diagnostic performance of 2020 Chinese Ultrasound Thyroid Imaging Reporting and Data System in thyroid nodules. *Zhonghua Yi Xue Za Zhi* **101**(45), 3748–3753 (2021)
14. J.D. Iannuccilli, J.J. Cronan, J.M. Monchik, Risk for malignancy of thyroid nodules as assessed by sonographic criteria: the need for biopsy. *J. Ultrasound Med.* **23**(11), 1455–1464 (2004)
15. E. Koike, S. Noguchi, H. Yamashita et al. Ultrasonographic characteristics of thyroid nodules: prediction of malignancy. *Arch. Surg.* **136**(3), 334–337 (2001)
16. B.K. Chan, T.S. Desser, I.R. McDougall, R.J. Weigel, R.B. Jeffrey Jr., Common and uncommon sonographic features of papillary thyroid carcinoma. *J. Ultrasound Med.* **22**(10), 1083–1090 (2003)
17. G.H. Klinck, T. Winship, Psammoma bodies and thyroid cancer. *Cancer* **12**(4), 656–662 (1959)
18. M.L. Khoo, J.L. Freeman, I.J. Witterick et al. Underexpression of p27/Kip in thyroid papillary microcarcinomas with gross metastatic disease. *Arch. Otolaryngol. Head. Neck Surg.* **128**(3), 253–257 (2002)
19. L.B. Ferreira, E. Gimba, J. Vinagre, M. Sobrinho-Simões, P. Soares, Molecular aspects of thyroid calcification. *Int J. Mol. Sci.* **21**(20), 7718 (2020)
20. L. Yin, W. Zhang, W. Bai, W. He, Relationship between morphologic characteristics of ultrasonic calcification in thyroid nodules and thyroid carcinoma. *Ultrasound Med. Biol.* **46**(1), 20–25 (2020)
21. P. Jun, L.C. Chow, R.B. Jeffrey, The sonographic features of papillary thyroid carcinomas: pictorial essay. *Ultrasound Q.* **21**, 39–45 (2005)
22. D.Y. Yoon, J.W. Lee, S.K. Chang, C.S. Choi, E.J. Yun, Y.L. Seo, K.H. Kim, H.S. Hwang, Peripheral calcification in thyroid nodules: ultrasonographic features and prediction of malignancy. *J. Ultrasound Med.* **26**(10), 1349–1355 (2007)
23. J.R. Wienke, W.K. Chong, J.R. Fielding et al. Sonographic features of benign thyroid nodules: interobserver reliability and overlap with malignancy. *J. Ultrasound Med.* **22**, 1027–1031 (2003)
24. S. Sorrenti, V. Dolcetti, M. Radzina et al. Artificial intelligence for thyroid nodule characterization: where are we standing? *Cancers (Basel)* **14**(14), 3357 (2022)
25. Q. Qi, A. Zhou, S. Guo, X. Huang, S. Chen, Y. Li, P. Xu, Explore the diagnostic efficiency of Chinese Thyroid Imaging Reporting and Data Systems by comparing with the other four systems (ACR TI-RADS, Kwak-TIRADS, KSThR-TIRADS, and EU-TIRADS): a single-center study. *Front Endocrinol. (Lausanne)* **12**, 763897 (2021)
26. H. Zhu, Y. Yang, S. Wu, K. Chen, H. Luo, J. Huang, 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. *Quant. Imaging Med. Surg.* **11**(8), 3604–3618 (2021)

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.