Original Article

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

Hui Zhu, Yan Yang, Senmin Wu, Kai Chen, Hongxia Luo, Jianle Huang

Department of Ultrasound, the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, China

Contributions: (I) Conception and design: J Huang, H Luo, H Zhu; (II) Administrative support: J Huang, H Luo, H Zhu; (III) Provision of study materials or patients: S Wu, K Chen; (IV) Collection and assembly of data: J Huang, Y Yang; (V) Data analysis and interpretation: H Zhu, Y Yang, J Huang, H Luo; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Hongxia Luo, Jianle Huang. 109 Xueyuan Western Road, Wenzhou, China. Email: luohongxia0127@163.com; wenzhoufey@126.com.

Background: To investigate the diagnostic performance of the ultrasonography-based fine-needle aspiration biopsy criteria of the Chinese Thyroid Imaging Reporting and Data System (C-TIRADS) for malignant nodules compared to 3 other guidelines.

Methods: This study included 2,309 thyroid nodules in 1,697 patients with histopathological and cytopathological diagnoses of benign and malignant nodules from January 2018 to August 2020. The clinical and ultrasonographic features of the nodules were retrospectively reviewed and classified according to the Chinese guideline (C-TIRADS), the American College of Radiology guideline (ACR-TIRADS), the American Thyroid Association guideline (ATA guideline), and the Korean Thyroid Association guideline (K-TIRADS). The diagnostic performance of the guidelines and their unnecessary fine-needle aspiration biopsy rates were calculated using randomized, blinded trials.

Results: Of the 2,309 nodules, 1,418 (61.4%) were benign and 891 (38.6%) were malignant, with 884 (99.21%) papillary carcinomas. The accuracy of C-TIRADS was 84.71%, followed by the guidelines of ACR-TIRADS (82.11%), K-TIRADS (81.64%), and the ATA guideline (78.56%). Furthermore, the area under the receiver operating characteristic curve (AUC) was the highest for the C-TIRADS (0.905). Similar results were revealed for both the diagnostic performance and AUC of nodules smaller and larger than 10 mm. The ACR-TIRADS showed the lowest unnecessary biopsy rate (17.54%), followed by the C-TIRADS (22.61%), ATA guideline (27.90%), and the K-TIRADS (28.67%).

Conclusions: The C-TIRADS demonstrated high diagnostic performance and a relatively low unnecessary biopsy rate in detecting thyroid cancer compared to the 3 other guidelines. However, further understanding of the ultrasonography-based fine-needle aspiration biopsy criteria of the C-TIRADS should be gained in the future.

Keywords: Thyroid nodule; Chinese Thyroid Imaging Reporting and Data System (C-TIRADS); American College of Radiology guideline (ACR-TIRADS); American Thyroid Association (ATA); Korean Thyroid Association guideline (K-TIRADS)

Submitted Dec 15, 2020. Accepted for publication Apr 19, 2021.
doi: 10.21037/qims-20-1365
View this article at: http://dx.doi.org/10.21037/qims-20-1365

Introduction

Thyroid nodules are commonly seen in clinical practice, and their detection mainly depends on ultrasound (US) (1,2). The prevalence of thyroid nodules detected by US is 19–68% worldwide and 20–35% in China (3). Of all the nodules, only 10–15% are malignant, and distinguishing these nodules is essential to subsequent therapy (4). High-resolution US is recommended as the best tool for the diagnosis of thyroid nodules and for the identification of features such as echo, shape, margin, and calcification, among others. After Horvath converted the Thyroid Imaging Reporting and Data System (TIRADS) from the Breast Imaging Reporting and Data System (BI-RADS) in 2009, different international societies, including the American College of Radiology (ACR), American Thyroid Association (ATA), and the Korean Thyroid Association/Korean Society of Thyroid Radiology (KTA/KSThR) among others, published guidelines to aid in thyroid nodule diagnosis and fine-needle aspiration biopsy (FNAB) (5). However, the accuracy of US and US-based guidelines is low in differentiating benign and malignant thyroid nodules, ranging from 41.8–84.0% (6-8). Meanwhile, the criteria for FNAB differs among the various guidelines, with the ACR demonstrating the lowest unnecessary FNAB rate of 25.3% (7,8). In 2020, the Superficial Organ and Vascular Ultrasound Group of the Society of Ultrasound in Medicine of the Chinese Medical Association published the Chinese-TIRADS (C-TIRADS), which consisted of counting methods and proved to be more convenient and practical than weighting methods for clinical application (3). One study showed that the sensitivity of C-TIRADS was 92.7%, which was much higher than the other guidelines (9). However, little is known about its diagnostic performance and the unnecessary rate of US-based FNAB.

The purpose of this study was to investigate the diagnostic performance and unnecessary US-based FNAB rate of C-TIRADS for malignant nodules compared to the ACR, ATA, and KTA/KSThR guidelines.

Methods

This is a retrospective study. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the institutional board of the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University (No. LCKY2020-415), and individual consent for this retrospective analysis was waived.

Patients

From January 2018 to August 2020, 1782 consecutive patients with 2417 thyroid nodules who underwent high-resolution US examination and US-guided FNAB or surgery afterward at the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University were involved in this study. The inclusion criteria were as follows: (I) preoperative or pre-FNAB US examinations with US reports were performed, and the images were saved as JPEG files; (II) the target nodules had undergone initial surgical resection or US-guided FNAB within 2 weeks of US examination; (III) surgery or FNAB was carried out within 2 weeks of US examination. The exclusion criteria were as follows: (I) the target nodules lacked a final pathological diagnosis after surgical resection; (II) cytopathological diagnosis of atypia or follicular lesion of undetermined significance and suspicion of malignancy by US-guided FNAB. Finally, 108 nodules were excluded due to a lack of US characteristics (n=37), lack of final pathological diagnosis after surgical resection (n=51), cytopathological diagnosis of atypia or follicular lesion of undetermined significance (n=16), and suspicion of malignancy (n=4) after FNAB. A total of 2,309 thyroid nodules in 1,697 patients were included in this study, including nodules that were diagnosed as benign or malignant on the basis of surgery (n=1,933), and those with initial definitive benign or malignant results at US-guided FNAB (n=376; *Figure 1*).

US examination and image analysis

All US examinations and FNABs were performed with high-frequency linear probes ranging in frequency from 5 to 14 MHz, and a variety of commercially available US systems: Philips EPIQ7C, EPIQ7 (Philips Medical Systems, Best, the Netherlands), GE Volume E10, E8 (GE Medical Systems, Milwaukee, WI, USA), Siemens ACUSON OXANA 2, S2000 (Siemens Medical Solutions, Mountain View, CA, USA), Esaote MyLab Class C, MyLab65, MyLabTwice eHD (Esaote, Genoa, Italy), Hitachi HI VISION Preirus (Hitachi-Aloka Medical, Tokyo, Japan), and Mindray Resona 7T (Mindray Medical International, Shenzhen, China). The US images and clips included both the transverse and longitudinal real-time imaging of thyroid nodules. The position, maximum diameter, composition,

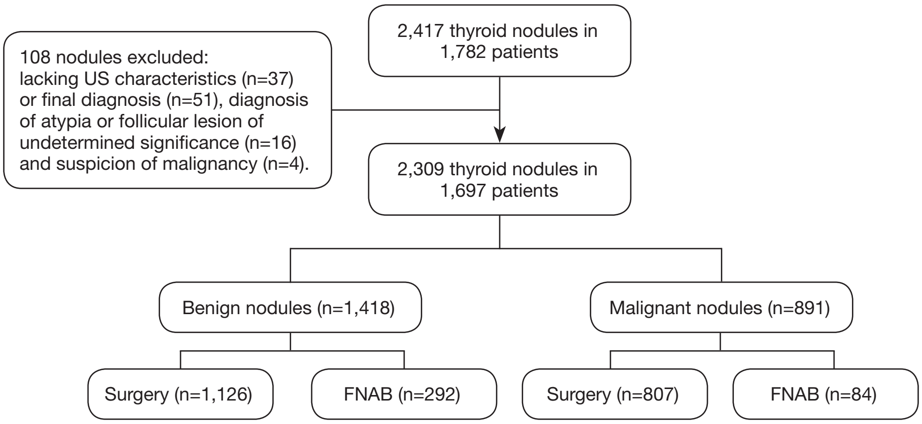

Figure 1 Flowchart showing the study participants. US, ultrasound; FNAB, fine-needle aspiration biopsy.

echogenicity, shape, margin, and echogenic foci were evaluated for each nodule. The position of nodules was categorized into the left lobe, right lobe, or isthmus. The left and right lobes were divided into upper, middle, and lower zones. Two US experts (YY, HL) who had 23 and 15 years of experience in performing thyroid US examination and interventional procedures retrospectively reviewed the images and clips. They were blinded to the final diagnosis of nodules and the identification of patients. Before the study started, several clinical sessions of lecture-based and hands-on instructions that explained the C-TIRADS 2020, ACR-TIRADS 2017, ATA guideline 2016, and KTA/KSThR TIRADS 2016 (K-TIRADS 2016) in detail were provided to the US experts. Subsequently, they analyzed 100 nodules jointly to establish a standard. Finally, YY and HL reviewed 1154 and 1155 nodules which were divided at random.

The descriptions of composition, echogenicity, shape, margin, and echogenic foci for nodules differ among guidelines. In the C-TIRADS, the composition is categorized as solid (entirely composed of solid tissue, without any cystic components), predominately solid (solid components accounting for more than 50%), predominately cystic (solid components accounting for <50%), cystic (completely or almost completely cystic), and spongiform (multiple tiny cystic spaces occupying the entire nodules without aggregated solid tissues). The echogenicity is defined as hyperechoic, isoechoic, hypoechoic, markedly hypoechoic, and anechoic. The echogenicity of hyperechoic, isoechoic, and hypoechoic can be an echo higher, similar, or lower than the surrounding thyroid parenchyma. The echogenicity of markedly hypoechoic is defined as an echo lower than that of the strap muscles of the neck. Anechoic was seen in cystic nodules, and was usually accompanied by posterior enhancement. The shape is defined as orientation in the C-TIRADS, and is categorized as vertical or horizontal. Vertical nodules are considered those with an anteroposterior diameter larger than the longitudinal diameter on transverse or longitudinal sections. Meanwhile, horizontal nodules are considered those with an anteroposterior diameter less than or equal to the longitudinal diameter on both transverse and longitudinal sections. The margin of nodules is defined as circumscribed (clear, smooth, and complete curve), irregular (spiculated, angular, or microlobulated), ill-defined (difficult to distinguish from the surrounding thyroid parenchyma), and extrathyroidal extension (the nodule spread to the thyroid capsule). The echogenic foci in the C-TIRADS are categorized as microcalcifications, comet-tail artifacts, punctate echogenic foci of undetermined significance, macrocalcifications, peripheral calcifications, and no echogenic foci. Microcalcifications and punctate echogenic foci of undetermined significance refer to echogenic foci less than 1 mm and macrocalcifications larger than 1 mm. Punctate echogenic foci of undetermined significance refer to those echogenic foci which may be microcalcifications, dense colloid materials, or other components on gray scale US without shadowing (the echo behind the echogenic foci is lower than that of the surrounding tissue at the same depth) and comet-tail artifact (punctate echogenic foci with a dense tapering trail of echoes in the rear). Peripheral calcifications appear as a continuous or discontinuous ring or arc involving more than a third of the margin. Positive features in the C-TIRADS are solid composition, markedly hypoechoic vertical orientation, ill-defined/irregular margin/extrathyroidal extension, and microcalcifications. One point was added when any of these were present. Negative features include comet-tail artifacts, and we subtracted 1 point when they appeared (3). Unlike C-TIRAD, in the ACR-TIRADS, the composition is categorized as cystic or almost completely cystic, spongiform, mixed cystic or solid, and solid or almost completely solid. The shapes of nodules are measured on transverse sections. Taller-than-wide shape, similar to vertical in the C-TIRADS, is defined as measurements higher than the width of nodules. Similar to C-TIRADS, the margin of nodules in the ACR-TIRADS is defined as smooth, ill-defined, lobulated or irregular, and extrathyroidal extension (frank invasion of adjacent soft tissue and/or vascular structures). However, an ill-defined margin does not affect the calculated points in the ACR-TIRADS. The echogenic foci are categorized as none or large comet-tail artifacts, macrocalcifications (caused acoustic shadowing), peripheral calcifications (complete or incomplete along margin), and punctate echogenic foci (smaller than macrocalcifications with non-shadowing). The definitions of echogenicity in the ACR-TIRADS are similar to those in C-TIRADS. Very hypoechoic in the ACR-TIRADS equates to markedly hypoechoic in C-TIRADS (10). The ATA guideline focuses on irregular margins, microcalcifications, taller-than-wide shape, disrupted rim calcifications with small extrusive hypoechoic soft tissue component, and extrathyroidal extension. Similar to ACR-TIRADS, the taller-than-wide shape in the ATA guideline is measured on transverse sections. However, the echogenicity of nodules does not include markedly hypoechoic as in the 2 aforementioned guidelines. Extrathyroidal extension is divided into minor extrathyroidal extension and gross extrathyroidal invasion, which are included in the ATA intermediate risk and high risk, respectively (11). The K-TIRADS mainly focuses on microcalcification, nonparallel orientation, and spiculated/microlobulated margin. The shape of nonparallel orientation (equal to vertical in the C-TIRADS) is measured on transverse or longitudinal sections (5).

The frequency of malignancy and diagnostic performance based on classified categories of each guideline were calculated according to the guidelines of the C-TIRADS (3), ACR-TIRADS (10), ATA guidelines (11), and the K-TIRADS (5). Furthermore, the unnecessary FNAB rate was determined with retrospective analysis based on the biopsy criteria of each guideline. The criteria of FNAB recommendations for each guideline are listed in *Table 1*.

US-guided FNAB and surgery

US-guided FNAB was performed by experienced radiologist on nodules suspicious for malignancy based on their US features. In brief, a 25-gauge needle attached to a 2-mL disposable syringe was used to aspirate the thyroid nodule. After aspiration was performed at least twice, the aspirated material was expelled on 2 glass slides, followed by 95% alcohol for cell block processing or saline for Papanicolaou staining. One of 8 experienced cytopathologists reviewed the slides. Cytopathological reports at our institution were categorized according to the criteria recommended by the Bethesda classification (12). Total or near total thyroidectomy or hemithyroidectomy was performed in patients whose nodules were highly suspicious for malignancy and who had multiple tumors or lymph node (LN) metastasis according to the ATA guideline. Surgery was performed by 4 experienced surgeons at our institution.

Histopathological and cytopathological results from surgery or US-guided FNAB were considered as the reference standard. Nodules with benign or malignant results at surgery or initial US-guided FNAB were considered benign or malignant, respectively.

Statistical analysis

Statistical analysis was performed at the patient level as well as the nodule level. Patient and nodule characteristics are presented as mean ± standard deviation (SD), while the calculated malignancy risk of each guideline is presented as a percentage. The formulas for calculating the malignancy rates of different categories for different guidelines were as follows: malignancy rate for a particular category = (nodules proven as malignant pathologically or cytopathologically in this category)/(all the nodules in this category). The diagnostic performance of different guidelines for malignancy prediction was compared in terms of sensitivity, specificity, positive predictive value, negative predictive value, and accuracy using the following generalized estimating formulae: sensitivity = true-positive/(true-positive + false-negative) (TP/TP + FN); specificity = true − negative/(true-negative + false-positive) (TN/TN + FP); positive predictive value = TP/(TP + FP); negative predictive value = TN/(TN + FN); accuracy = (TP + TN)/(TP + TN + FP + FN). To further assess the

Table 1 Summary of US FNAB recommendations of the 4 guidelines for thyroid nodules

Guidelines	FNAB size cutoff
ACR-TIRADS 2017	
TR1 (benign)	No biopsy
TR2 (not suspicious)	No biopsy
TR3 (mildly suspicious)	Recommend FNAB if ≥25 mm; follow if ≥15 mm
TR4 (moderately suspicious)	Recommend FNAB if ≥15 mm; follow if ≥10 mm
TR5 (highly suspicious)	Recommend FNAB if ≥10 mm; follow if ≥5 mm
ATA 2016	
Benign	No biopsy
Very low suspicion	Consider FNAB at ≥20 mm; observation without FNAB is also a reasonable option
Low suspicion	Recommend FNAB if ≥15 mm
Intermediate suspicion	Recommend FNAB if ≥10 mm
High suspicion	Recommend FNAB if ≥10 mm
K-TIRADS 2016	
K-TIRADS 2 (benign)	No biopsy
(I) Spongiform	Recommend FNAB if ≥20 mm
(II) Partially cystic nodule with comet-tail artifact	No biopsy
(III) Pure cyst	No biopsy
K-TIRADS 3 (low suspicion)	Recommend FNAB if ≥15 mm
K-TIRADS 4 (intermediate suspicion)	Recommend FNAB if ≥10 mm
K-TIRADS 5 (high suspicion)	Recommend FNAB if ≥10 mm; selective FNAB if >0.5
C-TIRADS 2020	
C-TR 2	No biopsy
C-TR 3	No biopsy
C-TR 4A	Recommend FNAB if ≥15 mm; recommend FNAB if ≥10 mm when 1 occurs
C-TR 4B	Recommend FNAB if ≥10 mm; recommend FNAB if ≥5 mm when 1 occurs
C-TR 4C	Similar to C-TR 4B
C-TR 5	Similar to C-TR 4B; recommend FNAB of any size if 2 occur

1: multifocality, subcapsular nodule, trachea, and recurrent laryngeal nerve invasion. 2: there are typical cervical metastatic lymph nodes of thyroid cancer. US, ultrasound; FNAB, fine-needle aspiration biopsy; ACR-TIRADS 2017, the 2017 Thyroid Imaging Reporting and Data System of the American College of Radiology (10); ATA 2016, the 2016 American Thyroid Association guideline (11); K-TIRADS 2016, the 2016 Thyroid Imaging Reporting and Data System of the Korean Thyroid Association/Korean Society of Thyroid Radiology (5); C-TIRADS 2020, the 2020 Chinese Thyroid Imaging Reporting and Data System (3); TR, TIRADS.

diagnostic value of the guidelines, the areas under the receiver operating characteristic (ROC) curves (AUC) were compared. In malignant nodule diagnosis, the potentially unnecessary FNAB rates and false-positive rates were calculated using the following equations: unnecessary FNAB rate = the number of benign nodules among FNAB-required nodules/total nodules; false-positive rate = FP/(FP + TN). These rates were compared among guidelines. The association between characteristics or categories and final diagnosis was evaluated with Student's *t*-test, the Mann-

Whitney U test (for continuous variables), and the chi-square test (for categorical variables). Statistical analyses were performed using SPSS software 19.0 (IBM Corp., RRID:SCR_002865). A P value less than 0.05 was defined as a significant difference.

Results

Clinical and US characteristics

Of the 2,309 thyroid nodules in the 1,697 patients included, the number of benign nodules was 1,418 (61.4%) and the number of malignant nodules was 891 (38.6%). There were 1336 females and 361 males, and the mean age was 49.7±12.2 years (age range, 9–87 years; P<0.01). Of all the nodules, the mean size was 13.1±10.6 mm (range, 0.2–69 mm; P<0.01). The numbers of benign and malignant nodules in the upper zone, middle zone, lower zone, and isthmus were 248 and 235, 548 and 336, 594 and 283, and 28 and 37, respectively (P<0.01; *Table 2*). Their US features are also shown in *Table 2*. Among the 891 malignant nodules, surgery was performed on 807 nodules. The majority of them were papillary thyroid carcinomas (PTC; n=800), and included 20 follicular variant thyroid carcinomas. The remaining 7 nodules included 4 follicular carcinomas, 2 medullar carcinomas, and 1 poorly differentiated carcinoma. Among the 1,126 benign nodules diagnosed via surgical resection, 967 were nodular hyperplasias, 30 were follicular adenomas, 74 were thyroiditis, 2 were eosinophilic adenomas, 7 were thyroid fibrosis, and 46 were cysts. The number of nodules in each size category is shown in *Table 3*.

Malignancy risk according to the categories in the 4 guidelines

The results of malignancy risk according to the 4 guidelines (ACR-TIRADS, ATA guidelines, K-TIRADS, and C-TIRADS) are listed and compared in *Table 4*. With the suggested risk of malignancy proposed by each guideline, we found that the calculated malignancy risk in each category matched well within the suggested range, except for low suspicion, intermediate suspicion, and high suspicion categories in the ATA guideline, and low suspicion and intermediate suspicion categories in the K-TIRADS. A total of 108 isoechoic nodules (4.9%) with suspicious US features could not be classified as any specific pattern with the ATA guideline.

Diagnostic performance of the 4 guidelines

Table 5 shows the diagnostic performance of the 4 guidelines with different cutoffs. Based on the results, we set the TIRADS 5 (TR5, highly suspicious), high suspicion, the K-TIRADS 5 (high suspicion), and the C-TIRADS 4C (C-TR 4C) as the cutoffs of malignancy for the ACR-TIRADS, ATA guideline, the K-TIRADS, and the C-TIRADS, respectively. Among the guidelines, specificity, positive predictive value, and accuracy were highest with the C-TIRADS (90.55%, 83.37%, and 84.71%, respectively) followed by the ACR-TIRADS (74.82%, 70.05%, and 82.11%, respectively) and the K-TIRADS (76.02%, 70.36%, and 81.64%, respectively). The ATA guideline had higher negative predictive value and sensitivity (96.77% and 96.51%, respectively) compared to the ACR-TIRADS (94.99% and 93.71%, respectively), the K-TIRADS (92.77% and 90.57%, respectively), and the C-TIRADS (85.43% and 75.42%, respectively). For further evaluation, ROC curves of each guideline were plotted; the AUCs were 0.854, 0.824, 0.849, and 0.905 for the ACR-TIRADS, ATA guideline, K-TIRADS, and C-TIRADS, respectively (all P<0.01; *Table 6*). The results indicated that the AUC of the C-TIRADS was significantly higher than the other guidelines.

In order to determine the diagnostic performance of guidelines in terms of different nodule sizes, we chose a cutoff diameter of 10 mm, which has been used as a criterion for microcarcinoma for pathologic diagnosis (13). The sensitivity, specificity, positive predictive value, negative predictive value, and accuracy of guidelines for nodules smaller or larger than 10 mm are shown in *Table* 7. The C-TIRADS presented with the highest specificity, positive predictive value, and accuracy in nodules smaller and larger than 10 mm. And the ATA guideline revealed the highest negative predictive value and sensitivity in both groups. The ROC curves of the ACR-TIRADS, ATA guideline, the K-TIRADS, and the C-TIRADS yielded AUCs of 0.793, 0.733, 0.791, and 0.858, respectively, for nodules smaller than 10 mm (*Table 6*; all P<0.01), and 0.901, 0.862. 0.885, and 0.940, respectively, for nodules larger than 10 mm (10 mm included; *Table 6*; all P<0.01). Thus, we could conclude that the diagnostic performance of the C-TIRADS was significantly higher than that of other guidelines for nodules regardless of their size.

Comparisons of unnecessary FNAB rates of the 4 guidelines

Among the 4 guidelines, the ACR-TIRADS had the

Table 2 Clinical characteristics and US features

Parameter	Total	No. of benign lesions	No. of malignant lesions	P value
Patients	1,697	957	740	
Mean age (years)§	49.7±12.2	50.3±12.1	45.5±11.9	<0.01
Males§, n (%)	361 (21.3)	172 (18.0)	189 (25.5)	<0.01
Females§, n (%)	1,336 (78.7)	785 (82.0)	551 (74.5)	<0.01
No. of nodules	2,309	1,418	891	
Mean size of mass (mm)	13.1±10.6	15.7±12.0	9.1±5.9	<0.01
Distribution location (total)	2,309	1,418	891	<0.01
Upper zone	483	248	235	
Middle zone	884	548	336	
Lower zone	887	594	283	
Isthmus	65	28	37	
Composition, n (%)	2,309	1,418	891	<0.01
Solid	1,817 (78.7)	934 (65.9)	883 (99.1)	
Predominately solid/cystic	390 (16.9)	382 (26.9)	8 (0.9)	
Cystic	93 (4.0)	93 (6.6)	0 (0.0)	
Spongiform	9 (0.4)	9 (0.6)	0 (0.0)	
Echogenicity, n (%)	2,309	1,418	891	<0.01
Hyperechoic/isoechoic	486 (21.1)	449 (31.7)	37 (4.2)	
Hypoechoic	1,524 (66.0)	819 (57.8)	705 (79.1)	
Markedly hypoechoic	190 (8.2)	41 (2.9)	149 (16.7)	
Anechoic	109 (4.7)	109 (7.7)	0 (0.0)	
Margin, n (%)	2,506	1,458	1,048	<0.01
Microlobulated/irregular margin	389 (15.5)	67 (4.6)	322 (30.7)	
Extrathyroidal extension	117 (4.7)	8 (0.6)	109 (10.4)	
Ill-defined	197 (7.9)	40 (2.7)	157 (15.0)	
Circumscribed	1,803 (72.0)	1,343 (92.1)	460 (43.9)	
Calcification, n (%)	2,244	1,472	1,038	<0.01
Microcalcifications	777 (34.6)	340 (23.1)	570 (54.9)	
Comet-tail artifacts	37 (1.7)	37 (2.5)	0 (0.0)	
Macrocalcifications	230 (10.3)	184 (12.5)	179 (17.2)	
Peripheral calcifications	52 (2.3)	32 (2.2)	20 (1.9)	
No echogenic foci	1,148 (51.2)	879 (59.7)	269 (25.9)	
Shape, n (%)	2,309	1,418	891	<0.01
Horizontal (wider-than-tall)	1,561 (67.6)	1,227 (86.5)	334 (37.5)	
Vertical (taller-than-wide)	748 (32.4)	191 (13.5)	557 (62.5)	

Data in parentheses are percentages unless otherwise indicated. §, analysis based on the number of patients. US, ultrasound.

Table 3 Number of nodules in different size categories

Size (mm)	Nodules, n (%)	Benign lesions, n (%)	Malignant lesions, n (%)	Mean size ± SD
≤10	1,229 (53.2)	625 (44.1)	604 (67.8)	6.2±1.8
10–19.9	621 (26.9)	386 (27.2)	235 (26.4)	13.3±2.8
20–29.9	215 (9.3)	179 (12.6)	36 (4.0)	24.4±2.9
30–39.9	177 (7.7)	153 (10.8)	14 (1.6)	34.4±3.1
≥40	67 (2.9)	75 (5.3)	2 (0.2)	46.9±6.4
Total	2,309	1,418	891	13.1±10.6

Data in parentheses are percentages. SD, standard deviation.

lowest unnecessary FNAB rate (17.54%), followed by the C-TIRADS, ATA guideline, and the K-TIRADS (22.61%, 27.90%, and 28.67%, respectively), and the same order was seen in terms of the false-positive rates for FNAB (28.56%, 36.81%, 45.75%, and 46.69%, respectively; *Table 8*). Furthermore, the number of malignant nodules not recommended for FNAB was 618, 522, 594, and 607 for the ACR-TIRADS, C-TIRADS, ATA guideline, and the K-TIRADS, respectively.

Discussion

Guidelines such as the ACR-TIRADS, C-TIRADS, ATA guideline, and the K-TIRADS are risk stratification systems that use risk category criteria. The main purpose of these guidelines is to distinguish malignant nodules from benign ones in a quick and relatively accurate way. They classify nodules into different categories, calculate the malignancy risk for each category, and recommend FNAB based on the category and size of each nodule. Then, the FNAB results confirm whether the nodules are malignant or benign, and surgery is taken into account if necessary. In this study, we evaluated the clinical and US features of 2,309 nodules in 1,697 patients, compared the diagnostic performance of the ACR-TIRADS, C-TIRADS, ATA guideline, and the K-TIRADS for these nodules using the cutoff of risk categories, and reviewed the unnecessary FNAB rates of each guideline retrospectively. Nodules with initial surgical resection or US-guided FNAB were included. As this was a retrospective study, we lacked a follow-up of nodules diagnosed with atypia or follicular lesion of undetermined significance and suspicion of malignancy by FNAB. Thus, they were excluded, and the pathological or cytopathological diagnoses of nodules were used as the gold standard. Then, we calculated the diagnostic performance using different cutoffs and chose the best one as the cutoff of each guideline. Finally, our results showed a variable range of sensitivity and specificity, from 75.42% to 96.51%, and 67.06% to 90.55%, respectively. The C-TIRADS had the highest specificity, positive predictive value, and accuracy; however, it also had the lowest sensitivity, regardless of nodule size. The AUC for the ROC curve of the C-TIRADS ranked first among the 4 guidelines, both in nodules smaller than 10 mm and in those larger than 10 mm. In nodules larger than 10 mm, the C-TIRADS showed higher sensitivity, specificity, accuracy, and AUC than those in subcentimeter nodules. Although the AUC was one of the parameters assessing global diagnostic accuracy, we believed that the high AUC of the C-TIRADS in different nodule sizes was not only due to its high specificity, but also due to having a relatively high negative predictive value and the highest positive predictive value among the guidelines. This meant that although the C-TIRADS could classify some malignant nodules into lower categories, the nodules included in the C-TR 4C and C-TR 5 were most likely malignant nodules. This helped in distinguishing malignant nodules from benign ones accurately. Furthermore, the higher sensitivity, specificity, accuracy, and AUC in nodules larger than 10 mm indicated its better diagnostic performance in these nodules. In addition, we found that the calculated malignancy risks in the ACR-TIRADS and C-TIRADS were generally well correlated within the range of the suggested malignancy risk. However, the calculated malignancy risk of low suspicion, intermediate suspicion, and high suspicion of the ATA guideline, and low suspicion and intermediate suspicion of the K-TIRADS, were lower than that of the suggested malignancy risk.

With the development of ultrasonography, the detection of thyroid nodules, especially those smaller than 10 mm or even 5 mm, has increased rapidly. Meanwhile,

Table 4 Comparison of malignancy risk according to categories in various guidelines

Guideline	Benign nodules (n=1,418)	Malignant nodules (n=891)	Total (n=2,309)	Suggested risk of malignancy (%)	Calculated risk of malignancy (%)	P value
ACR-TIRADS 2017						<0.01
TR1 (benign)	98	0	98	≤2	0	
TR2 (not suspicious)	141	0	141	≤2	0	
TR3 (mildly suspicious)	287	1	288	<5	0.35	
TR4 (moderately suspicious)	535	55	590	5–20	9.32	
TR5 (highly suspicious)	357	835	1,192	>20	70.05	
ATA 2016						<0.01
Benign	94	0	94	<1	0	
Very low suspicion	13	0	13	<3	0	
Low suspicion	451	4	455	5–10	0.88	
Intermediate suspicion	342	26	368	10–20	7.07	
High suspicion	442	829	1,271	>70–90	65.22	
Not specified§	76	32	108	–	29.63	
K-TIRADS 2016						<0.01
K-TIRADS 2 (benign)	112	0	112	<3	0	
K-TIRADS 3 (low suspicion)	458	7	465	3–15	1.51	
K-TIRADS 4 (intermediate suspicion)	508	77	585	15–50	13.16	
K-TIRADS 5 (high suspicion)	340	807	1,147	>60	70.36	
C-TIRADS 2020						<0.01
C-TR 2	34	0	34	0	0	
C-TR 3	362	0	362	<2	0	
C-TR 4A	579	31	610	2–10	5.08	
C-TR 4B	309	188	497	10–50	37.83	
C-TR 4C	130	626	756	50–90	82.80	
C-TR 5	4	46	50	<90	92.00	

Data in parentheses are the raw data used to calculate the percentages. §, nodules that did not meet the criteria of the ATA guideline for any pattern (isoechoic or hyperechoic solid nodule with 1 or more of the following features: microcalcifications, taller-than-wide shape, extrathyroidal extension). ACR-TIRADS 2017, the 2017 Thyroid Imaging Reporting and Data System of the American College of Radiology (10); ATA 2016, the 2016 American Thyroid Association guideline (11); K-TIRADS 2016, the 2016 Thyroid Imaging Reporting and Data System of the Korean Thyroid Association/Korean Society of Thyroid Radiology (5); the 2020 C-TIRADS, the 2020 Chinese Thyroid Imaging Reporting and Data System (3); TR, TIRADS.

the description of US features for each nodule is more precise than before. Studies have proven that US features including solid composition; markedly hypoechoic, ill-defined, or irregular margins; extrathyroidal extension; vertical orientation; and microcalcifications are independent features of malignancy (9,14). The 4 guidelines we chose were risk stratification systems determined by these features. The 2017 ACR-TIRADS estimated the malignancy of nodules by total scores, adding the score of each US feature. Different categories, ranging from TR 1 to TR 5, were determined by different scores. However, the score of each US feature was set mainly by expert opinion

Table 5 Comparison of diagnostic performance for malignant thyroid nodules in various guidelines

Guideline	Cutoff	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Accuracy (%)
ACR TI-RADS 2017	TR4 (moderately suspicious)	99.89	37.09	49.94	99.81	61.33
	TR5 (highly suspicious)	93.71	74.82	70.05	94.99	82.11
ATA 2016	Intermediate suspicion	99.53	41.58	52.17	99.29	64.20
	High suspicion	96.51	67.06	65.22	96.77	78.56
K-TIRADS 2016	K-TIRADS 4 (intermediate suspicion)	99.21	40.20	51.04	98.79	62.97
	K-TIRADS 5 (high suspicion)	90.57	76.02	70.36	92.77	81.64
C-TIRADS 2020	C-TR 4B	96.52	68.76	66.00	96.92	79.47
	C-TR 4C	75.42	90.55	83.37	85.43	84.71
	C-TR 5	5.16	99.65	90.20	62.59	63.20

Data in parentheses are the raw data used to calculate the percentages. ACR-TIRADS 2017, the 2017 Thyroid Imaging Reporting and Data System of the American College of Radiology (10); ATA 2016, the 2016 American Thyroid Association guideline (11); K-TIRADS 2016, the 2016 Thyroid Imaging Reporting and Data System of the Korean Thyroid Association/Korean Society of Thyroid Radiology (5); C-TIRADS 2020, the 2020 Chinese Thyroid Imaging Reporting and Data System (3); TR, TIRADS.

Table 6 The AUCs for ROC curves of different guidelines

Guideline	AUCs for ROC curves	P value
ACR-TIRADS 2017		
Total	0.854	<0.01
Nodule size <10 mm	0.793	<0.01
Nodule size ≥10 mm	0.901	<0.01
ATA 2016		
Total	0.824	<0.01
Nodule size <10 mm	0.733	<0.01
Nodule size ≥10 mm	0.862	<0.01
K-TIRADS 2016		
Total	0.849	<0.01
Nodule size <10 mm	0.791	<0.01
Nodule size ≥10 mm	0.885	<0.01
C-TIRADS 2020		
Total	0.905	<0.01
Nodule size <10 mm	0.858	<0.01
Nodule size ≥10 mm	0.940	<0.01

ROC, receiver operating characteristic; AUC, the area under the ROC curves; ACR-TIRADS 2017, the 2017 Thyroid Imaging Reporting and Data System of the American College of Radiology (10); ATA 2016, the 2016 American Thyroid Association guideline (11); K-TIRADS 2016, the 2016 Thyroid Imaging Reporting and Data System of the Korean Thyroid Association/Korean Society of Thyroid Radiology (5); C-TIRADS 2020, the 2020 Chinese Thyroid Imaging Reporting and Data System (3).

Table 7 Comparison of Diagnostic Performance for Different Nodule Sizes in Various Guidelines

Guideline	Nodule size	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Accuracy (%)
ACR TI-RADS 2017	<10 mm	94.37	63.04	71.16	92.06	78.44
	≥10 mm	92.33	84.11	67.77	96.81	86.30
ATA 2016	<10 mm	96.80	57.14	69.24	94.71	76.94
	≥10 mm	95.86	74.97	57.69	98.07	80.45
K-TIRADS 2016	<10 mm	94.37	62.72	70.98	92.02	78.28
	≥10 mm	82.58	86.51	68.90	93.21	85.46
C-TIRADS 2020	<10 mm	74.17	84.48	82.20	77.19	79.41
	≥10 mm	78.05	95.33	85.82	92.31	90.74

Data in parentheses are the raw data used to calculate the percentages. ACR-TIRADS 2017, the 2017 Thyroid Imaging Reporting and Data System of the American College of Radiology (10); ATA 2016, the 2016 American Thyroid Association guideline (11); K-TIRADS 2016, the 2016 Thyroid Imaging Reporting and Data System of the Korean Thyroid Association/Korean Society of Thyroid Radiology (5); C-TIRADS 2020, the 2020 Chinese Thyroid Imaging Reporting and Data System (3).

Table 8 Comparison of unnecessary FNAB rates for the diagnosis of thyroid cancer in all nodules

Guidelines	FNABs, n (%)	Malignant nodules among FNAB nodules, n (%)	Benign nodules among FNAB nodules (A), n (%)	Unnecessary FNAB rate (%)	False-positive rate[§] (%)
ACR-TIRADS 2017	679 (29.41)	274 (40.35)	405 (59.65)	17.54 (A/2,309)	28.56 (405/1,418)
ATA 2016	879 (39.94)	265 (30.14)	614 (69.86)	27.90 (A/2,201)	45.75 (614/1,342)
K-TIRADS 2016	946 (40.97)	284 (30.02)	662 (69.98)	28.67 (A/2,309)	46.69 (662/1,418)
C-TIRADS 2020	891 (38.59)	369 (41.41)	522 (58.59)	22.61 (A/2,309)	36.81 (522/1,418)

Data in parentheses are percentages for no. of FNABs, no. of malignant nodules among FNAB nodules, and no. of benign nodules among FNAB nodules. [§], data in parentheses are the raw data from which the false-positive rate was calculated. FNAB, fine-needle aspiration biopsy; ACR-TIRADS 2017, the Thyroid Imaging Reporting and Data System of 2017 American College of Radiology (10); ATA 2016, the 2016 American Thyroid Association guideline (11); K-TIRADS 2016, the Thyroid Imaging Reporting and Data System of 2016 Korean Thyroid Association/Korean Society of Thyroid Radiology (5); C-TIRADS 2020, the 2020 Chinese Thyroid Imaging Reporting and Data System (3).

instead of statistical analysis (14). Hence, the accuracy of ACR-TIRADS in the malignancy prediction of thyroid nodules remains questionable. The 2016 ATA guideline and the 2016 K-TIRADS 2016 are risk stratification systems that do not use mathematical calculations. Studies revealed that about 3.4–13.9% of nodules did not meet the criteria for any pattern in the ATA guideline (i.e., isoechoic or hyperechoic solid nodule with 1 or more of the following features: microcalcifications, taller-than-wide shape, extrathyroidal extension) (7,15,16), it may be thus impractical to use the ATA guideline in clinical practice. In our study, we also found 108 (4.7%) nodules which did not meet the criteria. With a malignancy risk of 29.63%, they were hard to categorize, as their calculated malignancy risk was lower than that of high suspicion but higher than that of intermediate suspicion. With regard to the FNAB criteria of the ATA guideline, we consider that FNAB should be selectively indicated in these nodules when the size is larger than 10 mm (the FNAB cutoff of intermediate suspicion and high suspicion in the ATA guideline). Similar to the ATA guideline, the K-TIRADS is a pattern-based system, and its categories are determined by weighting different US features, such as microcalcifications or solid composition. Considering the fact that the same US features exhibit different weights in different studies, it was hard to assume that the K-TIRADS weighted the

US features in an appropriate manner (17-19). This might explain the reason why the calculated malignancy risks of some categories in the K-TIRADS and the ATA guideline were lower than those suggested. Moreover, the diagnosis of atypia or follicular lesion of undetermined significance and suspicion of malignancy based on FNAB were excluded in our study, which might have led to a high proportion of PTC in malignant nodules. As US features like echogenic foci are more common in PTC than in follicular thyroid carcinomas, and follicular thyroid carcinomas are often classified as low suspicion in the guidelines, the number of malignant nodules might be underestimated as low suspicion in the K-TIRADS and ATA guideline (11). The C-TIRADS was established based on the counting method, which is more convenient and practical in clinical practice than weighting methods (9). The counting method was thought to be the simplest as compared to previously published risk stratification systems, and the score of each US feature is determined by statistical analysis (9). In one study, all examined nodules met the criteria of the C-TIRADS regardless of the size or US features, and its diagnostic performance was satisfactory compared to the 6 guidelines compared, including the ACR-TIRADS, ATA guideline, and the K-TIRADS. This was consistent with our study. The C-TIRADS had the highest specificity, accuracy, and AUC for nodules smaller or larger than 10 mm. This could be explained by the lower size criterion of nodules for FNAB. Meanwhile, the high sensitivity of 96.51% of the ATA guideline might probably have been due to the higher prevalence of PTC in this study. As for the diagnostic performance of the 4 guidelines apart from the C-TIRADS, several studies have shown similar results to our own (7,20).

Research has been published regarding the relationship between nodule location and nodule malignancy. However, the results have been conflicting. Ramundo's report showed that nodules in the middle zone of the thyroid had a higher risk of malignancy, while Jasim reported that those in the isthmus had the highest risk of malignancy (21,22). Meanwhile, Zhang *et al.* concluded that nodules in the upper zone of the thyroid had the highest malignancy risk (23). In this study, we found that nodule location in the isthmus had the highest malignancy risk (56.92%), followed by the upper zone (48.65%). Furthermore, we found that the sex and age of patients were significantly different between benign and malignant masses, and the female-to-male ratio was 4.56 (785:172) and 2.92 (551:189), respectively. As the malignancy risk of nodules can be influenced by nonthyroid nodular variables, such as gender, ethnicity, age, family history, body mass index, and radiation exposure (9), these results appear reasonable.

In considering whether thyroid nodules should be subjected to FNAB, a US-based risk stratification system plays an essential role (11-13,24). However, US features of each category and size cutoffs for FNAB differ among guidelines. A deep understanding of FNAB is required to optimize patient management (7). The criteria of FNAB for each guideline are listed in *Table 1* in this study. The percentage of nodules that underwent FNAB ranged from 29.41% to 40.97% for the guidelines. The K-TIRADS had the highest percentage of FNAB nodules and the highest unnecessary FNAB rate. In contrast, the percentage of FNAB nodules and unnecessary FNAB rate of the ACR-TIRADS was the lowest. In the ATA guideline, the percentage of FNAB nodules and the unnecessary FNAB rate were relatively high. Furthermore, when the "not specified" pattern was considered to undergo FNAB with a size cutoff of 10 mm, the unnecessary FNAB rate of the ATA guideline increased to 28.58%. This could have occurred because the wrong cutoff of FNAB was set for those nodules. The C-TIRADS had a relatively low percentage of FNAB nodules and a low unnecessary FNAB rate. Together with its high specificity, accuracy, and AUC for the ROC curve, the diagnostic performance of the US-based FNAB criteria for the C-TIRADS indicated that this guideline was better than that of the ACR-TIRADS, ATA guideline, and the K-TIRADS. Although thyroid cancers are slow-growing tumors and have a better prognosis than other malignant carcinomas, delayed diagnosis may influence the outcome of patients (22,25). However, Nam *et al.* reported that PTC without suspicious US features exhibited better prognosis than those that met the criteria (23). Furthermore, unnecessary FNAB may cause considerable anxiety to patients, and also involves a substantial financial and medical burden (7). Thus, the use of FNAB should be considered in an appropriate indication after a shared decision is made with patients. As can be seen, US practitioners require a deeper understanding of the US-based FNAB criteria of the different guidelines.

Our study had some limitations. First, this study only included nodules from 1 hospital, which does not fully represent the Chinese population. Second, the majority of diagnostic criteria we used were pathological results, with the remaining being from definite cytopathological results, which might have led to selection bias. Third, the malignancy rate and PTC proportion were high (38.6% and 99.1%, respectively). This might have been due to the low or intermediate suspicion of US features of follicular carcinomas, which would have led to follow-ups instead of surgery or FNAB for these nodules. Furthermore, our hospital is a tertiary referral center that treats patients with more serious diseases. Our hospital is also located in an urban city in Zhejiang province, where the incidence of thyroid cancer was reported to be higher than other places (26). Taken together, these factors might have influenced the diagnostic performance of the 4 guidelines. Fourth, most guidelines do not recommend the management of nodules smaller than 10 mm, especially those smaller than 5 mm. However, the C-TIRADS, as an exception, recommends FNAB for nodules smaller than 5 mm in certain conditions. With 53.2% of nodules smaller than 10 mm in this study, it limited the comparison of the C-TIRADS with other guidelines, and the comparison of our results with those of other previous studies, to some degree. Additionally, we did not evaluate the variation in interpretation among the readers. However, we explained the guidelines and had the radiologists analyze nodules for standardization purposes before the study started to avoid reading bias.

Conclusions

In conclusion, the 2020 C-TIRADS guideline demonstrated high diagnostic performance and a relatively low unnecessary FNAB rate in detecting thyroid cancer compared to the 2017 ACR-TIRADS, 2015 ATA guideline, and the 2016 K-TIRADS. US practitioners should acquire a greater understanding of the US-based FNAB criteria of the C-TIRADS.

Acknowledgments

Funding: None.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/qims-20-1365). 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 institutional board of the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University (No. LCKY2020-415), and individual consent for this retrospective analysis was waived.

References

1. Hegedüs L. Clinical practice. The thyroid nodule. New Engl J Med 2004;351:1764-71.
2. Vanderpump MP. The epidemiology of thyroid disease. Br Med Bull 2011;99:39-51.
3. Zhou J, Yin L, Wei X, Zhang S, Song Y, Luo B, Li J, Qian L, Cui L, Chen W, Wen C, Peng Y, Chen Q, Lu M, Chen M, Wu R, Zhou W, Xue E, Li Y, Yang L, Mi C, Zhang R, Wu G, Du G, Huang D, Zhan W; Superficial Organ and Vascular Ultrasound Group of the Society of Ultrasound in Medicine of the Chinese Medical Association; Chinese Artificial Intelligence Alliance for Thyroid and Breast Ultrasound. 2020 Chinese guidelines for ultrasound malignancy risk stratification of thyroid nodules: the C-TIRADS. Endocrine 2020;70:256-79.
4. Reiners C, Wegscheider K, Schicha H, Theissen P, Vaupel R, Wrbitzky R, Schumm-Draeger PM. Prevalence of Thyroid Disorders in the Working Population of Germany Ultrasonography Screening in 96,278 Unselected Employees. Thyroid 2004;14:926-32.
5. Shin JH, Baek JH, Chung J, Ha EJ, Kim JH, Lee YH, Lim HK, Moon WJ, Na DG, Park JS, Choi YJ, Hahn SY, Jeon SJ, Jung SL, Kim DW, Kim EK, Kwak JY, Lee CY, Lee HJ, Lee JH, Lee JH, Lee KH, Park SW, Sung JY; Korean Society of Thyroid Radiology (KSThR) and Korean Society of Radiology. Ultrasonography Diagnosis and Imaging-Based Management of Thyroid Nodules: Revised Korean Society of Thyroid Radiology Consensus Statement and Recommendations. Korean J Radiol 2016;17:370-95.
6. Iannuccilli JD, Cronan JJ, Monchik JM. Risk for

Malignancy of Thyroid Nodules as Assessed by Sonographic Criteria The Need for Biopsy. J Ultrasound Med 2004;23:1455-64.
7. 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.
8. Yoon JH, Han K, Kim EK, Moon HJ, Kwak JY. Diagnosis and Management of Small Thyroid Nodules: A Comparative Study with Six Guidelines for Thyroid Nodules. Radiology 2017;283:560-9.
9. Zhou J, Song Y, Zhan W, Wei X, Zhang S, Zhang R, Gu Y, Chen X, Shi L, Luo X, Yang L, Li Q, Bai B, Ye X, Zhai H, Zhang H, Jia X, Dong Y, Zhang J, Yang Z, Zhang H, Zheng Y, Xu W, Lai L, Yin L; Superficial Organ and Vascular Ultrasound Group of the Society of Ultrasound in Medicine of Chinese Medical Association; Chinese Artificial Intelligence Alliance for Thyroid and Breast Ultrasound. Thyroid imaging reporting and data system (TIRADS) for ultrasound features of nodules: multicentric retrospective study in China. Endocrine 2020. [Epub ahead of print].
10. 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.
11. 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.
12. Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. Thyroid 2017;27:1341-6.
13. Lang W, Borrusch H, Bauer L. Occult Carcinomas of the Thyroid Evaluation of 1,020 Sequential Autopsies. Am J Clin Pathol 1988;90:72-6.
14. Li J, Ma X, Cui K. Re ACR Thyroid Imaging, Reporting and Data System (TI-RADS) White Paper of the ACR TI-RADS Committee. J Am Coll Radiol 2018;15:380-1.
15. Yoon JH, Lee HS, Kim EK, Moon HJ, Kwak JY. Malignancy risk stratification of Thyroid nodules Comparison between the Thyroid Imaging Reporting and Data System and the 2014 American Thyroid Association Management Guidelines. Radiology 2016;278:917-24.
16. Middleton WD, Teefey SA, Reading CC, Langer JE, Beland MD, Szabunio MM, Desser TS. Comparison of Performance Characteristics of American College of Radiology TI-RADS, Korean Society of Thyroid Radiology TIRADS, and American Thyroid Association Guidelines. AJR Am J Roentgenol 2018;210:1148-54.
17. Delfim RLC, Veiga LCGD, Vidal APA, Lopes FPPL, Vaisman M, Teixeira PFDS. Likelihood of malignancy in thyroid nodules according to a proposed Thyroid Imaging Reporting and Data System (TI-RADS) classification merging suspicious and benign ultrasound features. Arch Endocrinol Metab 2017;61:211-21.
18. Zhang WB, Xu HX, Zhang YF, Guo LH, Xu SH, Zhao CK, Liu BJ. Comparisons of ACR TI-RADS, ATA guidelines, Kwak TI-RADS, and KTA/KSThR guidelines in malignancy risk stratification of thyroid nodules. Clin Hemorheol Microcirc 2020;75:219-32.
19. 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.
20. 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.
21. Ramundo V, Lamartina L, Falcone R, Ciotti L, Lomonaco C, Biffoni M, Giacomelli L, Maranghi M, Durante C, Grani G. Is thyroid nodule location associated with malignancy risk? Ultrasonography 2019;38(3):231-235.
22. Jasim S, Baranski TJ, Teefey SA, Middleton WD. Investigating the Effect of Thyroid Nodule Location on the Risk of Thyroid Cancer. Thyroid 2020;30:401-7.
23. Zhang F, Oluwo O, Castillo FB, Gangula P, Castillo M, Farag F, Zakaria S, Zahedi T. Thyroid nodule location on ultrasonography as a predictor of malignancy. Endocr Pract 2019;25:131-7.
24. Nam SY, Shin JH, Han BK, Ko EY, Ko ES, Hahn SY, Chung JH. Preoperative Ultrasonographic Features of Papillary Thyroid Carcinoma Predict Biological Behavior. J Clin Endocrinol Metab 2013;98:1476-82.
25. Ramundo V, Lamartina L, Falcone R, Ciotti L, Lomonaco

C, Biffoni M, Giacomelli L, Maranghi M, Durante C, Grani G. Is thyroid nodule location associated with malignancy risk? Ultrasonography 2019;38:231-5.

26. Du L, Zhao Z, Zheng R, Li H, Zhang S, Li R, Wei W, He J. Epidemiology of Thyroid Cancer: Incidence and Mortality in China, 2015. Front Oncol 2020;10:1702.

Cite this article as: 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(8):3604-3618. doi:10.21037/qims-20-1365

