

## Comparison of diagnostic performance and FNA management of the ACR-TIRADS and Chinese-TIRADS based on surgical histological evidence

## Weilu Dong, Yiyun Wu, Ting Cai, Xiangqian Wang

Department of Ultrasound Medicine, The Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Province Hospital of Chinese Medicine), Nanjing, China

*Contributions:* (I) Conception and design: W Yun, W Dong; (II) Administrative support: X Wang; (III) Provision of study materials or patients: W Dong, T Cai; (IV) Collection and assembly of data: W Dong, T Cai, X Wang; (V) Data analysis and interpretation: W Dong; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*Correspondence to:* Yiyun Wu. Department of Ultrasound Medicine, The Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Province Hospital of Chinese Medicine), 155 Hanzhong Road, Qinhuai District, Nanjing, China. Email: yfy0043@njucm.edu.cn.

**Background:** The sonography-based risk stratification of thyroid nodules is based on the different sonographic features of benign and malignant thyroid nodules. The American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TIRADS) and the Chinese-TIRADS define different risks for malignancy categories and the size criteria for fine-needle aspiration (FNA). Few studies have compared their diagnostic performance and FNA management approaches. Thus, we sought to compare the diagnostic performance and FNA management approaches of the ACR-TIRADS and Chinese-TIRADS based on surgical histological evidence.

**Methods:** This retrospective study included patients with complete thyroid ultrasound images and histologic evidence who were consecutively selected from The Affiliated Hospital of Nanjing University of Chinese Medicine. A total of 333 nodules from 252 patients with definitive surgical histological findings were examined. Ultrasonography categories and FNA management proposals were assigned according to the ACR-TIRADS and Chinese-TIRADS. The thyroid nodules were divided into 2 groups based on a cut-off size of 1 cm. The diagnostic performance and recommended and unnecessary FNA rates for each group were compared for both systems.

**Results:** Overall, 280 malignant thyroid nodules (84.1%) and 53 benign nodules (15.9%) were analyzed. Across all groups, the ACR-TIRADS had higher sensitivity, specificity, positive and negative predictive values, and accuracy, and a higher area under the curve (AUC) than the Chinese-TIRADS. However, there was a significant difference in the negative predictive value between the ACR-TIRADS and Chinese-TIRADS of <1 and  $\geq$ 1 cm, and the overall AUCs differed significantly (39.5% *vs.* 35.1%, P=0.007 for <1 cm; 70.0% *vs.* 62.8%, P=0.014 for  $\geq$ 1 cm; 0.843 *vs.* 0.806, P=0.037 for all). The rate for recommending the FNA of the nodules in the ACR-TIRADS was lower than that in the Chinese-TIRADS (25.2% *vs.* 56.5%, P<0.0001), but there was no significant difference between the ACR-TIRADS and Chinese-TIRADS in terms of the unnecessary FNA rate (14.3% *vs.* 13.8%, P=0.931). The kappa statistics for the consistency of the ACR-TIRADS and Chinese-TIRADS FNA recommendations for nodules sized  $\geq$ 1.0 to <1.5,  $\geq$ 1.5 to <2.5, and  $\geq$ 2.5 cm were 0.084, 0.635, and 0.909, respectively.

**Conclusions:** The ACR-TIRADS had slightly better diagnostic performance and a lower recommended FNA rate than the Chinese-TIRADS for thyroid nodules. For thyroid nodules  $\geq 1.5$  cm in size, the FNA recommendations of the 2 guidelines had good consistency.

**Keywords:** Thyroid Imaging Reporting and Data System (TIRADS); risk; guidelines; thyroid nodule; fine-needle aspiration (FNA)

Submitted Jun 27, 2022. Accepted for publication Jan 12, 2023. Published online Feb 23, 2023. doi: 10.21037/qims-22-685

View this article at: https://dx.doi.org/10.21037/qims-22-685

## Introduction

The prevalence of thyroid nodules in China is approximately 20–47% (1-4). Only 5–15% of thyroid nodules are malignant (5,6). Currently, the detection of thyroid nodules largely depends on ultrasonography, which has a detection rate of 19–68% worldwide (6-8). Since the absolute prevalence of malignancy is low, previous research has focused on strategies to stratify the risk of cancer and optimize the use of fine-needle aspiration (FNA) or surgery.

Sonologists conduct the risk stratification of thyroid nodules based on the different sonographic features of benign and malignant thyroid nodules. In 2009, inspired by the Breast Imaging Reporting and Data System, Horvath et al. (9) proposed the Thyroid Imaging Reporting and Data System (TIRADS). Since then, several thyroid risk stratification systems have been designed, including the Korean (K)-TIRADS (10), European (Eu)-TIRADS (11), and American College of Radiology (ACR)-TIRADS (12). Several of these TIRADSs are used in China, which may lead to confusion when different doctors read the same thyroid ultrasound report. Apart from the confusing application of TIRADS in China, the Chinese experience may differ to that of other countries due to the higher incidence of thyroid cancer in China. Thus, in 2020, the Chinese Medical Association published the Chinese (C)-TIRADS based on medical conditions in China (13).

Currently, both the ACR-TIRADS and C-TIRADS are used in China. The ACR-TIRADS uses a weighted method to calculate the score of each ultrasound sign to determine the nodule's risk level, while the C-TIRADS counts the number of positive signs. The ACR-TIRADS differs greatly from the C-TIRADS. Notably, the ACR-TIRADS 1 defines nodules based on whether they are benign, while the C-TIRADS 1 defines nodules based on whether they exist or are lacking. The C-TIRADS has 4A, 4B, and 4C subgroups, but the ACR-TIRADS does not. The malignancy risks for the ACR-TIRADS categories 1–5 are 0%, 0–2%, 2–5%, 5–20%, and >20%, respectively. Conversely, the malignancy risks for the C-TIRADS categories 3–5 are <2%, 2–10%, 10–50%, 50–90%, and >90%, respectively. These different methods for assessing the same thyroid nodule may produce different classification results. For instance, a solid isoechoic nodule with smooth margins would be classified as ACR-TIRADS 3 or C-TIRADS 4A. These guidelines report completely different risks for malignancy categories and size criteria for FNA, and few studies have compared their diagnostic performance and use in FNA management. This study sought to examine the differences between these 2 systems. We present the following article in accordance with the STARD reporting checklist (available at https://qims. amegroups.com/article/view/10.21037/qims-22-685/rc).

## Methods

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The Ethics Committee of the Affiliated Hospital of Nanjing University of Chinese Medicine approved this retrospective study that reviewed the thyroid ultrasound images and medical records of patients (No. 2015NL-023-02), and waived the requirement for informed consent.

## Study cobort

We retrospectively analyzed the ultrasound images of 333 nodules from 252 participants who underwent thyroid surgery between January 2020 and April 2021 (*Figure 1*). All the patients were enrolled in this study consecutively. Nodules without complete ultrasound images, histologic evidence, or definite histologic results were excluded. The mean age of participants was  $45.1\pm12.7$  years (range, 17-81 years). The median size of the thyroid nodules was  $0.8\pm0.9$  cm (range, 0.3-6.1 cm), and the distribution of nodule size was skewed. Among the 252 participants, 55 (21.8%) were male, and 197 (78.2%) were female; 177 participants had single nodules, 70 participant had 2 nodules (*Table 1*). The indications for the surgeries

#### Quantitative Imaging in Medicine and Surgery, Vol 13, No 3 March 2023



Figure 1 A CONSORT diagram of the study cohort selection. ACR, American College Radiology; TIRADS, Thyroid Imaging Reporting and Data System; C, Chinese.

were as follows: benign nodules >4 cm causing compressive symptoms, malignant nodules that caused patient anxiety, or other surgery indications.

## Ultrasound analysis

All the thyroid ultrasound images were reviewed by 2 doctors who had >5 years of work experience. The doctors were blinded to the surgical pathology findings, and every TIRADS classification was determined via consensus; in cases of disagreement, a third highly experienced sonologist joined the discussion.

Both the ACR-TIRADS and C-TIRADS were employed to classify the thyroid nodules according to the assigned points of the nodule's structure, echogenicity, margin, shape, and echogenic foci. However, there are some differences between the lexicons of ACR-TIRADS and C-TIRADS. For example, the ACR-TIRADS defines a spongiform nodule as "predominantly (>50%) small cystic spaces", while the C-TIRADS defines it as "multiple tiny cystic spaces occupying the entire nodule without aggregated solid tissues". The difference in this definition may be due to the different classification rules. The ACR-TIRADS and C-TIRADS also have completely different scoring rules (12-15). Under the ACR-TIRADS, 0–3 points are assigned to each ultrasound feature and summed to determine the category (*Figure 2*). Under the C-TIRADS, the number of positive ultrasound features is counted, and 1 point is then subtracted if a comet tail sign (negative feature) is present. The total points are used to decide the final category (*Figure 2*).

The ACR-TIRADS and C-TIRADS have different criteria for FNA management. Under the ACR-TIRADS, FNA is recommended for nodules classed as category 3 (nodules  $\geq$ 2.5 cm), category 4 (nodules  $\geq$ 1.5 cm), and category 5 (nodules  $\geq$ 1.0 cm). Under the C-TIRADS,

Table 1 Patient and nodule characteristics

| Variables                                   | Value         |
|---|---------------|
| Patient characteristics (n=252)             |               |
| Age in year (mean $\pm$ SD)                 | 45.1±12.7     |
| Sex, n (%)                                  |               |
| Male  | 55 (21.8)     |
| Female                                      | 197 (78.2)    |
| Nodule characteristics (n=333)              |               |
| Size in cm [median (IQR)]                   | 0.8 (0.6–1.2) |
| Echogenicity, n (%)                         |               |
| Hypoechoic                                  | 176 (52.9)    |
| Very hypoechoic                             | 112 (33.6)    |
| Hyperechoic                                 | 5 (1.5)       |
| Isoechoic                                   | 40 (12.0)     |
| Margin, n (%)                               |               |
| Smooth                                      | 48 (14.4)     |
| III-defined                                 | 59 (17.7)     |
| Lobulated or irregular                      | 217 (65.2)    |
| ETE   | 1 (0.3)       |
| Irregular and ETE                           | 8 (2.4)       |
| Structure, n (%)                            |               |
| Solid                                       | 321 (96.4)    |
| Mixed cystic and solid                      | 12 (3.6)      |
| Shape, n (%)                                |               |
| Taller-than-wide                            | 164 (49.2)    |
| Wider-than-tall                             | 169 (50.8)    |
| Echogenic foci, n (%)                       |               |
| Macrocalcifications                         | 16 (4.8)      |
| Macrocalcifications and microcalcifications | 25 (7.5)      |
| Peripheral (rim)                            | 1 (0.3)       |
| Peripheral (rim) and microcalcifications    | 1 (0.3)       |
| Microcalcifications                         | 113 (33.9)    |
| None  | 177 (53.2)    |
| Surgical histology, n (%)                   |               |
| Benign                                      | 53 (15.9)     |
| Malignant                                   | 280 (84.1)    |

Isoechoic means the echogenicity is similar to the surrounding thyroid parenchyma. Hypoechoic means the echogenicity is lower than the surrounding thyroid parenchyma. Very hypoechoic means the echogenicity is lower than that of the strap muscles of the neck. SD, standard deviation; IQR, interquartile range; ETE, extra-thyroidal extension.

## Dong et al. ACR-TIRADS & C-TIRADS for thyroid nodules

FNA is not recommended for category 3. For C-TIRADS 4A nodules, FNA is recommended when the nodules are  $\geq 1.5$  cm or for multiple 4A nodules  $\geq 1.0$  cm found immediately adjacent to the capsule, trachea, or recurrent laryngeal nerve. For nodules scored C-TIRADS 4B and above, FNA is recommended when the nodules are  $\geq 1.0$  cm, multiple, or found immediately adjacent to the capsule, trachea, or recurrent laryngeal nerve with a diameter  $\geq 0.5$  cm (*Figure 2*).

## Data and statistical analysis

MedCalc 15.8 (MedCalc, Mariakerke, Belgium) and SPSS 26 (IBM, Armonk, NY, USA) were used for the statistical analysis. The age of the patients was normally distributed and is presented as the mean ± standard deviation. The size distribution of the nodules was skewed and is presented as the median (interquartile range). The Pearson chi-square test was used to compare the proportions. The cut-off points of the ACR-TIRADS and C-TIRADS were determined using receiver operating characteristic (ROC) curves, and the risks of malignancy were compared to the cut-off values. A cross-comparison between the ACR-TIRADS and C-TIRADS categories was performed, and the malignancy rate of each category was calculated. Adopting the surgical-pathological results as the gold standard, the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy of the ACR-TIRADS and C-TIRADS for nodules <1.0 cm,  $\geq 1.0$  cm, and all nodules were calculated. The areas under the curve (AUCs) were calculated using the Delong test. The unnecessary FNA rate was defined as the proportion of recommended FNA nodules (RFNs) among the pathological benign nodules (16). The use of FNA before surgery was also compared between the ACR-TIRADS and C-TIRADS groups. The weighted kappa statistic was used to evaluate the level of agreement between the recommendations for FNAs according to the ACR-TIRADS and C-TIRADS. Histograms were generated using GraphPad Prism (version 9; GraphPad Software, San Diego, CA). Statistical significance was set at P<0.05.

## **Results**

## Pathological characteristics of participants and nodules

Of the 252 patients, 35 (13.9%) had benign nodules and 217 (86.1%) had malignant nodules. The mean age of

#### Quantitative Imaging in Medicine and Surgery, Vol 13, No 3 March 2023



Figure 2 The ACR-TIRADS and C-TIRADS. ACR, American College Radiology; TIRADS, Thyroid Imaging Reporting and Data System; C, Chinese; FNA, fine-needle aspiration; TR, TIRADS.

patients in the malignant group (44.4 $\pm$ 12.3 years; range, 20–76 years) was lower than that of patients in the benign group (50.0 $\pm$ 14.0 years; range, 17–81 years; P=0.025). In this study, there were 55 (21.8%) men and 197 (78.2%) women, and the difference in terms of gender was

significant (P<0.0001). Among the 55 men, 9 (16.4%) and 46 (83.6%) had benign and malignant nodules, respectively. Among the 197 women, 26 (13.2%) and 171 (86.8%) had benign and malignant nodules, respectively. The was no significant difference in the distribution of the malignant

| 0 ,            | 0,        |            |               |                        |
|----------------|-----------|------------|---------------|------------------------|
| Classification | Total (n) | Benign (n) | Malignant (n) | Rate of malignancy (%) |
| ACR-TIRADS     |           |            |               |                        |
| 2              | 10        | 10         | 0             | 0.0                    |
| 3              | 16        | 14         | 2             | 12.5                   |
| 4              | 52        | 19         | 33            | 63.5                   |
| 5              | 255       | 10         | 245           | 96.1                   |
| C-TIRADS       |           |            |               |                        |
| 3              | 13        | 13         | 0             | 0.0                    |
| 4A             | 23        | 16         | 7             | 30.4                   |
| 4B             | 44        | 11         | 33            | 75.0                   |
| 4C             | 229       | 13         | 216           | 94.3                   |
| 5              | 24        | 0          | 24            | 100.0                  |
| Total          | 333       | 53         | 280           | 84.1                   |

Table 2 Malignancy rate within each category of the ACR-TIRADS and C-TIRADS

ACR, American College Radiology; TIRADS, Thyroid Imaging Reporting and Data System; C, Chinese.

 Table 3 Cross-comparison between the ACR-TIRADS and C-TIRADS classifications

| C-TIRADS - |    | Total |    |     |       |
|------------|----|-------|----|-----|-------|
|            | 2  | 3     | 4  | 5   | Iotai |
| 3          | 10 | 2     | 1  | 0   | 13    |
| 4A         | 0  | 12    | 10 | 1   | 23    |
| 4B         | 0  | 2     | 31 | 11  | 44    |
| 4C         | 0  | 0     | 10 | 219 | 229   |
| 5          | 0  | 0     | 0  | 24  | 24    |
| Total      | 10 | 16    | 52 | 255 | 333   |

ACR, American College Radiology; TIRADS, Thyroid Imaging Reporting and Data System; C, Chinese.

and benign nodules in terms of gender (P=0.548).

Among the 333 thyroid nodules analyzed, 53 (15.9%) were benign and 280 (84.1%) were malignant. The median maximum diameter of the nodules was 0.8 cm (0.6–1.2 cm). The malignant nodules were significantly smaller than the benign nodules ( $0.9\pm1.6~vs.~1.9\pm1.6~cm;~P<0.0001$ ). The pathological findings for the nodules showed 1 (0.3%) follicular carcinoma, 279 (83.8%) papillary thyroid carcinomas (PTC), 34 (10.2%) nodular goiters, 1 (0.3%) atypical thyroid adenoma, 12 (3.6%) follicular adenomas, and 2 (0.6%) of each of oncocytomas, granulomatous thyroiditis, and lymphocytic thyroiditis.

## Malignancy rate and cross-comparison of the ACR-TIRADS and C-TIRADS classifications

The ACR-TIRADS and C-TIRADS were applied to all the thyroid nodules. The risks of malignancy in ACR-TIRADS 2, 3, 4, and 5 were 0.0%, 12.5%, 63.5%, and 96.1%, respectively, and those in C-TIRADS 3, 4A, 4B, 4C, and 5 were 0.0%, 30.4%, 75.0%, 94.3%, and 100.0%, respectively (*Table 2*).

One difference between the ACR-TIRADS and C-TIRADS is that unlike the C-TIRADS, the ACR-TIRADS has no subclassifications (*Table 3*). The cross-

#### Quantitative Imaging in Medicine and Surgery, Vol 13, No 3 March 2023

|                 | Size <1.0 cm           |                        |          | Size ≥1.0 cm           |                        |        | All                    |                        |        |
|-----------------|------------------------|------------------------|----------|------------------------|------------------------|--------|------------------------|------------------------|--------|
| Method          | ACR-TIRADS             | C-TIRADS               | <b>D</b> | ACR-TIRADS             | C-TIRADS               | Ρ      | ACR-TIRADS             | C-TIRADS               | Ρ      |
|                 | 5                      | 4C                     | Р        | 5                      | 4C                     |        | 5                      | 4C                     |        |
| Sensitivity (%) | 88.1<br>(82.7–91.9)    | 87.5<br>(82.1–91.5)    | 0.876    | 86.4<br>(77.7–92.0)    | 81.8<br>(72.5–88.5)    | 0.410  | 87.5<br>(83.1–90.9)    | 85.7<br>(81.1–89.3)    | 0.535  |
| Specificity (%) | 62.5<br>(42.7–78.8)    | 54.2<br>(35.1–72.1)    | 0.558    | 96.6<br>(82.8–99.4)    | 93.1<br>(78.0–98.1)    | 1.000  | 81.1<br>(68.6–89.4)    | 75.5<br>(62.4–85.1)    | 0.480  |
| PPV (%)         | 94.9<br>(90.7–97.3)    | 93.9<br>(89.3–96.5)    | 0.655    | 98.7<br>(93.0–99.8)    | 97.3<br>(90.7–99.3)    | 0.972  | 96.1<br>(92.9–97.9)    | 94.9<br>(91.4–97.0)    | 0.510  |
| NPV (%)         | 39.5<br>(25.6–55.3)    | 35.1<br>(21.8–51.3)    | 0.007*   | 70.0<br>(54.6–81.9)    | 62.8<br>(47.9–75.6)    | 0.014* | 55.1<br>(44.1–65.7)    | 50.0<br>(39.3–60.7)    | 0.519  |
| Accuracy (%)    | 85.2<br>(79.8–89.3)    | 83.8<br>(78.3–88.1)    | 0.690    | 88.9<br>(81.9–93.4)    | 84.6<br>(77.0–90.1)    | 0.335  | 86.5<br>(82.4–89.8)    | 84.1<br>(79.8–87.6)    | 0.382  |
| AUC             | 0.753<br>(0.690–0.809) | 0.708<br>(0.643–0.768) | 0.140    | 0.915<br>(0.801–0.929) | 0.875<br>(0.848–0.958) | 0.070  | 0.843<br>(0.800–0.881) | 0.806<br>(0.759–0.847) | 0.037* |

Table 4 Diagnostic performances of the ACR-TIRADS and C-TIRADS

95% confidence interval in parentheses; \*, P values <0.05 indicated a significant difference. FNA, fine-needle aspiration; ACR, American College Radiology; TIRADS, Thyroid Imaging Reporting and Data System; C, Chinese; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the curve.

comparison showed that ACR-TIRADS 2 corresponded to C-TIRADS 3, ACR-TIRADS 3 to C-TIRADS 4A, ACR-TIRADS 4 to C-TIRADS 4B, and ACR-TIRADS 5 to C-TIRADS 4C and 5 (*Table 3*).

# Diagnostic performance of the ACR-TIRADS and C-TIRADS

The diagnostic efficacy and ROC curves of the ACR-TIRADS and C-TIRADS for all the nodules and subgroups based on size are presented in *Table 4*. The ROC analysis of the nodule points showed that the best diagnostic cutoff values for ACR-TIRADS and C-TIRADS were 5 and 4C, respectively. The sensitivity, specificity, NPV, PPV, and accuracy of the ACR-TIRADS in the subgroups and all nodules were slightly higher than those of the C-TIRADS, but the differences were not significant (P>0.05). Conversely, the NPV in <1 and  $\geq$ 1 cm groups differed significantly (P=0.007 and P=0.014; *Table 4*).

The difference in the AUCs between the ACR-TIRADS and C-TIRADS was not statistically significant in the <1-cm and  $\geq$ 1-cm groups (0.753 vs. 0.708, P=0.140; 0.915 vs. 0.875, P=0.070; *Figure 3*). The AUC of the ACR-TIRADS for all nodules was significantly higher than that of the C-TIRADS (0.843 vs. 0.806, P=0.037; *Figure 3*).

## FNA recommendation results for ACR-TIRADS and C-TIRADS

Of the total 333 thyroid nodules, 84 (25.2%) were categorized as RFNs by the ACR-TIRADS, while 188 (56.5%) were categorized as RFNs by the C-TIRADS (P<0.001; *Table 5*). The malignancy risks of the RFNs did not differ significantly between the ACR-TIRADS and C-TIRADS (85.7% vs. 86.2%, P=0.801). The RFN rates for the benign and malignant nodules under the ACR-TIRADS were lower than those under the C-TIRADS (22.6% vs. 49.1%, P=0.009; 25.7% vs. 57.9%, P<0.0001; *Figure 4*). The unnecessary FNA rate of ACR-TIRADS did not differ significantly to that of the C-TIRADS (14.3% vs. 13.8%, P=0.931).

For nodules that were  $\geq 1.0$  to < 1.5 cm in diameter, the kappa statistic for the consistency between the ACR-TIRADS and C-TIRADS was 0.084 (P=0.260), and 0.635 (P<0.001), and 0.909 (P<0.0001) for nodules  $\geq 1.5$  to < 2.5and  $\geq 2.5$  cm, respectively (*Table 6*), indicating a good agreement in terms of the FNA recommendations between the ACR-TIRADS and C-TIRADS for nodules  $\geq 1.5$  cm.

## Discussion

We compared the diagnostic performance and FNA

## Dong et al. ACR-TIRADS & C-TIRADS for thyroid nodules



Figure 3 Diagnostic performance of the ACR-TIRADS and C-TIRADS. (A) Diagnostic efficacy among nodules <1.0 cm in size. (B) Diagnostic efficacy among nodules  $\geq$ 1.0 cm in size. (C) Diagnostic efficacy among all nodules. (D) AUCs for <1.0 cm,  $\geq$ 1.0 cm, and all nodules. \*, P $\leq$ 0.05. ACR, American College Radiology; TIRADS, Thyroid Imaging Reporting and Data System; C, Chinese; AUC, area under the curve; ACC, accuracy; SEN, sensitivity; SPE, specificity; PPV, positive predictive value; NPV, negative predictive value; ns, not statistically significant.

| Table 5 Recommendations for the FNA of | of nodules based on the ACR | -TIRADS and C-TIRADS classifications |
|--|-----------------------------|--------------------------------------|
|--|-----------------------------|--------------------------------------|

| System     | No. of RFN<br>(n) | Rate of RFN<br>(%) | Malignancy risk of<br>RFN (%) | Rate of RFN among benign nodules (%) | Rate of RFN among malignant nodules (%) | Unnecessary FNA<br>rate (%) |
|------------|-------------------|--------------------|-------------------------------|--------------------------------------|---|-----------------------------|
| ACR-TIRADS | 84                | 25.2 (20.9–30.2)   | 85.7 (76.7–91.8)              | 22.6 (13.5–35.5)                     | 25.7 (20.9–31.1)                        | 14.3 (8.4–23.3)             |
| C-TIRADS   | 188               | 56.5 (51.1–61.7)   | 86.2 (80.5–90.4)              | 49.1 (36.1–62.1)                     | 57.9 (52.0–63.5)                        | 13.8 (9.6–19.5)             |
| Р          | -                 | <0.001*            | 0.801                         | 0.009*                               | <0.001*                                 | 0.931                       |

95% confidence interval in parentheses. \*, P values <0.05 indicated a significant difference. RFN, recommending FNA nodules; FNA, fineneedle aspiration; ACR, American College Radiology; TIRADS, Thyroid Imaging Reporting and Data System; C, Chinese.

management approaches of the ACR-TIRADS and C-TIRADS based on surgical histological evidence. Our cross-comparison of the ACR-TIRADS and C-TIRADS may serve as a reference for other physicians. Contrary to several recent studies (17,18), we found that the sensitivity, specificity, PPV, NPV, accuracy, and AUC of the ACR-TIRADS were higher than those of the C-TIRADS, regardless of the nodule size. The RFN rate under the ACR-TIRADS was significantly lower than that under the C-TIRADS, which is in line with a recent meta-analysis (19). The ACR-TIRADS and C-TIRADS showed good agreement in terms of the FNA recommendations for nodules  $\geq$ 1.5 cm.

In contrast to the malignancy risks reported previously for the ACR-TIRADS and C-TIRADS (12,13), the malignancy risks of ACR-TIRADS categories 2–5 in this study were 0.0%, 12.5%, 63.5%, and 96.1%, respectively. Those of C-TIRADS categories 3, 4A, 4B, 4C, and 5 were 0.0%, 30.4%, 75.0%, 94.3%, and 100.0%, respectively. The malignancy risks based on the surgical study cohort were higher than those reported by the 2 guidelines, except for ACR-TIRADS 2 and C-TIRADS 3.

The AUC was used as a parameter to assess the global



Figure 4 Recommendation for the FNA of nodules based on the ACR-TIRADS and C-TIRADS classifications. \*\*, P $\leq$ 0.01; \*\*\*, P $\leq$ 0.001. ACR, American College Radiology; TIRADS, Thyroid Imaging Reporting and Data System; C, Chinese; FNA, fine-needle aspiration; RFN, recommending FNA nodules; ns, not statistically significant.

diagnostic efficacy of the ACR-TIRADS and C-TIRADS. A cut-off size of 1 cm was chosen because it is a pathologic diagnostic criterion for microcarcinoma (20), the incidence of which is rapidly growing in China (6,21). Our results showed that both the ACR-TIRADS and C-TIRADS showed a good diagnostic performance in predicting malignancy in thyroid nodules, especially in larger nodules. Nevertheless, all the diagnostic parameters of the ACR-TIRADS were higher than those of the C-TIRADS, and the NPV differed significantly in both the <1-cm and ≥1-cm subgroups (39.5% and 35.1%, P=0.007; 70.0% and 62.8%, P=0.014). The AUCs of the ACR-TIRADS were significantly higher than those of the C-TIRADS, particularly for all nodules, (0.843 vs. 0.806, P=0.037). Thus, we concluded that the ACR-TIRADS had slightly better diagnostic efficacy for thyroid nodules than the C-TIRADS.

According to the ACR-TIRADS, FNA is not recommended for nodules <1 cm. However, the C-TIRADS guidelines indicate that if a 4A nodule >1 cm or a 4B nodule >0.5 cm is detected multiple times, or if the nodule is immediately adjacent to the capsule, trachea, or recurrent laryngeal nerve, FNA should be considered. A decision to conduct FNA should take into account both the doctor's skills and the anxiety level of the patient if the nodule is <5 mm. Zhu *et al.* (18) reported that the ACR-TIRADS had the lowest rate of FNAs compared to other guidelines. Similarly, we also found that the overall RFN rates of the ACR-TIRADS were lower than those of the C-TIRADS

| Thrashold of size (cm) | Total (n)  |                   | C-TIRADS (n) |     |               | 16    | р        |
|------------------------|------------|-------------------|--------------|-----|---------------|-------|----------|
|                        | iotai (II) | ACR-TIRADS (II) - | Follow-up    | FNA | Consideration | ĸ     | Г        |
| <0.5                   | 42         | Follow-up         | 30           | 0   | 12            | NA    | NA       |
|                        |            | FNA               | 0            | 0   | 0             |       |          |
| [0.5, 1.0)             | 174        | Follow-up         | 85           | 84  | 5             | NA    | NA       |
|                        |            | FNA               | 0            | 0   | 0             |       |          |
| [1.0, 1.5)             | 58         | Follow-up         | 2            | 19  | 0             | 0.084 | 0.260    |
|                        |            | FNA               | 1            | 36  | 0             |       |          |
| [1.5, 2.5)             | 31         | Follow-up         | 2            | 2   | 0             | 0.635 | <0.001*  |
|                        |            | FNA               | 0            | 27  | 0             |       |          |
| ≥2.5                   | 28         | Follow-up         | 7            | 0   | 0             | 0.909 | <0.0001* |
|                        |            | FNA               | 1            | 20  | 0             |       |          |

Table 6 Consistency of the ACR-TIRADS and C-TIRADS in the follow-up and FNA recommendations

\*, P values <0.05 indicated a significant difference. NA, not available; FNA, fine-needle aspiration; ACR, American College Radiology; TIRADS, Thyroid Imaging Reporting and Data System; C, Chinese.

## Dong et al. ACR-TIRADS & C-TIRADS for thyroid nodules

for all nodules and benign and malignant nodules. However, the unnecessary FNA rate of the ACR-TIRADS did not differ significantly to that of the C-TIRADS (14.3% vs. 13.8%, P=0.931). Some studies have compared several TIRADSs using a large number of thyroid nodules (17,18,22,23); however, few studies have examined the consistency of FNA recommendations among different TIRADSs. The consistency test in this study showed that the ACR-TIRADS and C-TIRADS FNA recommendations had good agreement for nodules  $\geq 1.5$  cm.

The K-TIRADS and Eu-TIRADS are 2 other wellestablished risk stratification systems for thyroid nodules. Unlike the C-TIRADS and ACR-TIRADS, which assign points to ultrasound features, both the K-TIRADS and Eu-TIRADS classify thyroid nodules directly based on thyroid ultrasound images, and neither guideline uses subclassifications. The Eu-TIRADS classifies mildly hypoechoic nodules into 4 categories, and nodules with at least 1 malignant sign are classified into 5 categories. The K-TIRADS defines solid hypoechoic nodules without any suspicious signs or partially cystic isoechoic nodules with any suspicious signs as 1 of 4 categories, and nodules with any suspicious signs. The Eu-TIRADS has more classification categories, which may cause anxiety in patients.

The present study had several limitations. First, the study cohort included patients who underwent thyroid surgery, which could have led to selection bias. Second, the radiologists retrospectively analyzed all the thyroid nodules using previous ultrasound images; thus, the radiologists might have confused microcalcifications with ambiguous comet tail signs. Third, interobserver variability and consistency should be considered when reassigning the TIRADS classifications. Fourth, while all the patients were consecutively recruited, the pathology types of the study samples were biased. Most malignant thyroid nodules were PTC, and only 1 was follicular carcinoma. This may be related to the higher incidence of PTC. Fifth, as this was a single-center study, more multicenter studies need to be conducted to validate our results.

In conclusion, the ACR-TIRADS appears to have slightly better diagnostic performance and a lower recommended FNA rate than the C-TIRADS for thyroid nodules. For thyroid nodules  $\geq 1.5$  cm, both systems have good consistency in their FNA recommendations. In daily clinical work, it is necessary to pay attention to the differences in the recommendations of different TIRADSs. Each patient's situation, including their anxiety or financial situation, should be considered when deciding whether to perform FNA or actively observe a thyroid nodule.

## **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-22-685/rc

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at https://qims. amegroups.com/article/view/10.21037/qims-22-685/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 Ethics Committee of the Affiliated Hospital of Nanjing University of Chinese Medicine approved this retrospective study that reviewed the thyroid ultrasound images and medical records of patients (No. 2015NL-023-02), and waived the requirement for informed consent.

*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

 Xu W, Chen Z, Li N, Liu H, Huo L, Huang Y, Jin X, Deng J, Zhu S, Zhang S, Yu Y. Relationship of anthropometric measurements to thyroid nodules in a

1721

Chinese population. BMJ Open 2015;5:e008452.

- Jiang H, Tian Y, Yan W, Kong Y, Wang H, Wang A, Dou J, Liang P, Mu Y. The Prevalence of Thyroid Nodules and an Analysis of Related Lifestyle Factors in Beijing Communities. Int J Environ Res Public Health 2016;13:442.
- Liu Y, Lin Z, Sheng C, Zhu Y, Huang Y, Zhong N, Jia Z, Qu S. The prevalence of thyroid nodules in northwest China and its correlation with metabolic parameters and uric acid. Oncotarget 2017;8:41555-62.
- Li Y, Teng D, Ba J, Chen B, Du J, He L, et al. Efficacy and Safety of Long-Term Universal Salt Iodization on Thyroid Disorders: Epidemiological Evidence from 31 Provinces of Mainland China. Thyroid 2020;30:568-79.
- 5. Kaderli RM, Trepp R. From thyroid nodules to thyroid cancer. Ther Umsch 2020;77:419-25.
- 6. Zamora EA, Khare S, Cassaro S. Thyroid nodule. Treasure Island (FL): StatPearls Publishing; 2022 Jan.
- Durante C, Grani G, Lamartina L, Filetti S, Mandel SJ, Cooper DS. The Diagnosis and Management of Thyroid Nodules: A Review. JAMA 2018;319:914-24.
- 8. 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.
- 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.
- Shin JH, Baek JH, Chung J, Ha EJ, Kim JH, Lee YH, et al. Ultrasonography Diagnosis and Imaging-Based Management of Thyroid Nodules: Revised Korean Society of Thyroid Radiology Consensus Statement and Recommendations. Korean J Radiol 2016;17:370-95.
- 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.
- 12. 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.

- Zhou J, Yin L, Wei X, Zhang S, Song Y, Luo B, et al. 2020 Chinese guidelines for ultrasound malignancy risk stratification of thyroid nodules: the C-TIRADS. Endocrine 2020;70:256-79.
- 14. 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.
- Tessler FN, Middleton WD, Grant EG. Thyroid Imaging Reporting and Data System (TI-RADS): A User's Guide. Radiology 2018;287:29-36.
- Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. Thyroid 2017;27:1341-6.
- 17. Qi Q, Zhou A, Guo S, Huang X, Chen S, Li Y, Xu P. 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) 2021;12:763897.
- 18. 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.
- Castellana M, Castellana C, Treglia G, Giorgino F, Giovanella L, Russ G, Trimboli P. Performance of Five Ultrasound Risk Stratification Systems in Selecting Thyroid Nodules for FNA. J Clin Endocrinol Metab 2020;105:dgz170.
- 20. Pacini F. Thyroid microcarcinoma. Best Pract Res Clin Endocrinol Metab 2012;26:421-9.
- Du L, Wang Y, Sun X, Li H, Geng X, Ge M, Zhu Y. Thyroid cancer: trends in incidence, mortality and clinicalpathological patterns in Zhejiang Province, Southeast China. BMC Cancer 2018;18:291.
- 22. Gao L, Xi X, Jiang Y, Yang X, Wang Y, Zhu S, Lai X, Zhang X, Zhao R, Zhang B. Comparison among TIRADS (ACR TI-RADS and KWAK- TI-RADS) and 2015 ATA

## Dong et al. ACR-TIRADS & C-TIRADS for thyroid nodules

Guidelines in the diagnostic efficiency of thyroid nodules. Endocrine 2019;64:90-6.

23. Koc AM, Adıbelli ZH, Erkul Z, Sahin Y, Dilek I. Comparison of diagnostic accuracy of ACR-TIRADS,

**Cite this article as:** Dong W, Wu Y, Cai T, Wang X. Comparison of diagnostic performance and FNA management of the ACR-TIRADS and Chinese-TIRADS based on surgical histological evidence. Quant Imaging Med Surg 2023;13(3):1711-1722. doi: 10.21037/qims-22-685 American Thyroid Association (ATA), and EU-TIRADS guidelines in detecting thyroid malignancy. Eur J Radiol 2020;133:109390.