Nodules with nonspecific ultrasound pattern according to the 2015 American Thyroid Association malignancy risk stratification system

A comparison to the Thyroid Imaging Reporting and Data System (TIRADS-Na)

Pingping Xiang, MMa, Xiaojiu Chu, MMb, Guofang Chen, MDa, Binbin Liu, BSb, Wenbo Ding, MSC, Zheng Zeng, MMb, Xinping Wu, BSc, Jianhua Wang, MMd, Shuhang Xu, MDe,* Chao Liu, MDa,*

Abstract

New sonographic patterns have been recommended by the 2015 American Thyroid Association (ATA) to stratify nodules in terms of malignancy risk and help guide biopsy decision. This study aimed to compare the ultrasound part of the ATA guidelines and the Thyroid Imaging Reporting and Data System (TIRADS-Na).

In 2013 to 2016, 708 thyroid nodules in 505 patients were confirmed by postoperative histopathology. Hypochoegenicity, solidity, microcalcification, irregular margin, and a taller-than-wide shape were considered features suggesting malignancy. Sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV) were obtained for the TIRADS and ATA guidelines.

Of the 708 nodules, 341 (48.2%) and 367 (51.8%) were benign and malignant, respectively. Based on the ultrasound 2015 ATA guidelines, 62 nodules had nonspecific pattern (both malignant and benign features); malignancy rates of nodules with very low, low, intermediate, and high suspicion, and nonspecific pattern were 0, 17.7%, 67.9%, 90.0%, and 69.4%, respectively (P < .001). Malignancy rates of categories 2/3/4/5 nodules by TIRADS were 0, 8.1%, 67.0%, and 90.1%, respectively (P < .001). Based on pathological results, the AUC, sensitivity, specificity, NPV, and PPV were 0.926, 96.7%, 81.5%, 84.9%, and 96.9% for TIRADS, and 0.920, 93.5%, 82.4%, 85.1%, and 92.1% for ATA patterns, respectively. The TIRADS was generally more efficient than the 2015 ATA guidelines, especially for nodules ≥2 cm in diameter or those with nonspecific pattern.

The TIRADS show a relative superiority over the ultrasound 2015 ATA guidelines, especially for nodules with ≥2 cm diameter or nonspecific pattern.

Abbreviations: AACE = American Association of Clinical Endocrinologists, ATA = American Thyroid Association, AUCs = areas under the curves, FT3 = free triiodothyronine, FT4 = free thyroxine, NPV = negative predictive value, PPV = positive predictive value, ROC = receiver operating curve, TGAb = anti-thyroglobulin antibody, TIRADS = Thyroid Imaging Reporting and Data System, TPOAb = thyroid peroxidase antibody, TSH = serum thyrotropin, TSH = Thyrotropin, US = ultrasound.

Keywords: American Thyroid Association guidelines, sonographic pattern, Thyroid Imaging Reporting and Data System, thyroid nodule, ultrasound

1. Introduction

Thyroid nodules are very common, and an increasing number of people have been recently diagnosed with thyroid cancer, mainly because of the large-scale use of imaging techniques such as computed tomography, magnetic resonance imaging, positron emission tomography, and high-resolution ultrasound (US).[1] In 2013 to 2016, 708 thyroid nodules in 505 patients were confirmed by postoperative histopathology. Hypochoegenicity, solidity, microcalcification, irregular margin, and a taller-than-wide shape were considered features suggesting malignancy. Sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV) were obtained for the TIRADS and ATA guidelines. The authors declare that they have no conflict of interest.

* Correspondence: Shuhang Xu, and Chao Liu, Endocrine and Diabetes Center, Affiliated Hospital of Integrated Traditional Chinese and Western Medicine, Nanjing City, Jiangsu Province, China (e-mail: shuhangxu@163.com, proffichao@163.com).

Copyright © 2019 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

How to cite this article: Xiang P, Chu X, Chen G, Liu B, Ding W, Zeng Z, Wu X, Wang J, Xu S, Liu C. Nodules with nonspecific ultrasound pattern according to the 2015 American Thyroid Association malignancy risk stratification system. Medicine 2019;88(44):e17657.

Received: 17 October 2018 / Accepted: 23 September 2019

http://dx.doi.org/10.1097/MD.0000000000017657
reaching 28.7%.\textsuperscript{12} 50% to 70% of healthy subjects are found to have thyroid nodules by US.\textsuperscript{3,4} Moreover, the rate of thyroid nodule malignancy can reach 5% to 7%.\textsuperscript{5}

As the most effective tool predicting the risk of thyroid nodule malignancy, US assesses nodule composition (solidity, cystic proportion, or spongiform), echogenicity, margin, calcification status, taller-than-wide shape, and vascularity.\textsuperscript{6,7} It is widely accepted that such US features are independent predictors for thyroid nodular malignancy.\textsuperscript{8} However, no individual feature is reliable enough to identify the suspected nodules.\textsuperscript{9} Besides, lack of clinical experience may lead to diagnostic errors if sonographic data are wrongly interpreted. A combination of suspected US characteristics increases the accuracy of malignancy detection.

The Thyroid Imaging Reporting and Data System (TIRADS), a malignancy-risk-stratification tool for thyroid nodule classification initially introduced by Horvath,\textsuperscript{10} quantifies malignancy risks using main sonographic features. Although an updated TIRADS based on the solidity and echogenicity of thyroid nodules has been proposed recently,\textsuperscript{11} no standardized TIRADS risk-stratification system is currently available.\textsuperscript{10,12,13} Therefore, no TIRADS classifications have been widely adopted, particularly in the United States.\textsuperscript{14}

According to the updated management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer, the American Thyroid Association (ATA) has put forward a new US system for malignancy risk stratification based on sonographic features, but its effectiveness has not been validated.\textsuperscript{15} Therefore, the aim of the present study was to compare the US 2015 ATA guidelines and the newly proposed TIRADS-Na for malignancy risk stratification of thyroid tumors.

2. Material and methods

2.1. Study population

In this retrospective study, a total of 946 thyroid nodules from 813 patients were examined from January 2013 to December 2016. Malignant or benign thyroid nodules were diagnosed by surgical pathology at the Affiliated Hospital of Integrated Traditional Chinese and Western Medicine, Nanjing University of Chinese Medicine (a tertiary referral center). The inclusion criteria were:

1. US examination;
2. surgery;
3. thyroid function test; and
4. diagnosis of benign or malignant tumors by postoperative histopathology.

The exclusion criteria were:

1. lack of demographic information;
2. <18 years of age; or
3. lack of US data, including nodule size, composition, echogenicity, margin, shape, and calcification status.

All procedures involving human participants were approved by the ethics committee of the Affiliated Hospital of Integrated Traditional Chinese and Western Medicine, Nanjing University of Traditional Chinese Medicine, in line with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was waived by the committee because of the retrospective nature of the study.

2.2. Image analysis and scoring

Color ultrasound was performed routinely in all patients for thyroid before being included into this study and the images were routinely examined by 5 radiologists. For this study, 2 chief radiologists (5 years of experience in thyroid US) who were not among the 5 radiologists analyzed the ultrasound features. This was to be able to ensure the accuracy of the retrospective analysis of ultrasound features and avoid bias. The 2 radiologists were blind to the final histopathological diagnosis. The images were reviewed independently by the 2 radiologists. Any discordant opinions were solved by discussion.

2.3. US examination and image analysis

The 2 radiologists performed all the classification process. High-resolution US scanning was performed with a 7.5-13-MHz linear-array transducer (HI VISION Preirus, micarcn) by radiologists with >5 years of clinical experience in thyroid US. All US features of thyroid nodules, recorded in our online system, were retrospectively reviewed. The reviewer assessed the following US features of thyroid nodules: composition, echogenicity, calcification status, margin, shape, and comet-tail artefact.\textsuperscript{11,16,17} The comet-tail artifact is characterized by reverberation artifacts within the cystic component\textsuperscript{11} “Partially cystic nodules” with “comet-tail artefact” were considered to be benign. For the solid portion of the partially cystic nodules, the configuration was categorized as concentric, eccentric with an acute angle, eccentric with a blunt angle, and unspecified patterns; eccentric with an acute angle is risk of malignancy. Tumor composition was categorized as solid (no obvious cystic content), predominantly solid (≤50% of solid proportion), predominantly cystic (>50% of cystic proportion), and cystic (no obvious solid content). Compared to normal thyroid tissues and the adjacent thyroid parenchyma, nodule echogenicity was classified as marked hypoechoegenicity, mild hypoechoegenicity, hyperechoegenicity, and isoechogenicity. Spongiform nodules were classified as isoechogenic. Calcification was categorized into microcalcification and macrocalcification (including eggshell calcification\textsuperscript{18}). Calcification large enough to result in posterior acoustic shadowing was considered as macrocalcification.\textsuperscript{14} Combination of macrocalcification with microcalcification was considered to be a malignant microcalcification because the malignancy risk has been shown to be equivalent between the 2 entities.\textsuperscript{17,19} Therefore, a nodule showing both types of calcification (macrocalcification, including rim calcifications, intermingled with microcalcification) was grouped with those showing microcalcification.\textsuperscript{13} Tumor margins were described as regular or irregular (infiltrative, micro-lobulated, or spiculated). Shape was classified as taller-than-wide or wider-than-taller (measured on a transverse view). Comet-tail artifact was defined as intracystic echogenic foci accompanied by reverberation artifacts.\textsuperscript{11} Malignancy was suspected with the following US features: solidity, hypoechoegenicity, irregular margin, microcalcification, a taller-than-wide shape, and vascularity.\textsuperscript{17,19–22}

2.4. TIRADS classification

Based on the criteria proposed by Na et al,\textsuperscript{11} all nodules were classified into 5 categories: TIRADS1, no nodules; TIRADS2, spongiform, purely cystic, or partially cystic nodules with comet-tail artifact (only 1%–3% at the risk of malignancy); TIRADS 3, partially cystic or iso- and hyperechoic nodules without any of the
3 suspicious US features (microcalcification, taller-than-wide shape, and irregular margin) (low suspicion with 3%–15% at risk of malignancy); TIRADS 4, partially cystic or iso- and hyperechoic nodules with any of the 3 suspicious US features, or solid hypoechoic nodules without any of the 3 suspicious US features (intermediate suspicion with 15%–50% at risk of malignancy); TIRADS 5, solid hypoechoic nodules with any of the 3 suspicious US features (high suspicion with >60% at risk of malignancy).

2.5. US pattern by the 2015 ATA

According to the US part of the 2015 ATA guidelines,17 thyroid nodules were categorized as high, intermediate, low, and very low suspicion of malignancy and benignity. Benign nodules were purely cystic, without any solid component. Nodules with very low-suspicion were spongiform or partially cystic, without any sonographic feature of nodules with low, intermediate, or high suspicion. Nodules with low suspicion were isoechoic or hyperechoic, solid or partially cystic, with eccentric solid areas, without microcalcification, irregular margin, extrathyroidal extension, or a taller-than-wide shape. Nodules with intermediate-suspicion were hypoechoic, solid and marginally smooth, without microcalcification, extrathyroidal extension or a taller-than-wide shape. Highly-suspicious nodules were solid hypoechoic or solid hypoechoic component of a partially cystic nodule, with at least 1 of the following features: irregular margin, microcalcification, taller-than-wide shape, rim calcification with or without thyroid function information were excluded. Finally, 708 thyroid nodules from 515 patients were included in this study. Mean patient age was 48.9 ± 16.0 years, ranging from 18 to 82 years. Nodule size was 1.78 ± 1.42 cm, ranging from 0.16 to 7.1 cm. Demographic, clinical, and US features of patients and nodules are summarized in Tables 1 and 2. The patients with benign nodules were aged between 18 and 82 years (54.0 ± 14.0 years), significantly older than those with malignant nodules (ranging from 18 to 81 years; 46.0 ± 16.0 years) (P < .01).

3. Results

3.1. Patients

A total of 4 nodules lacking demographic information, 166 with inadequate US information, 5 being <18 years of age, and 63 without thyroid function information were excluded. Finally, 708 thyroid nodules from 515 patients were included in this study. Mean patient age was 48.9 ± 2.2 years, ranging from 18 to 82 years.

### Table 1

| Parameter              | Benign | Malignant | Total | P value |
|------------------------|--------|-----------|-------|---------|
| Number of nodules      | 341    | 367       | 708   |         |
| Number of patients     | 213    | 302       | 515   |         |
| Age (year)             |        |           |       |         |
| Mean ± SD              | 54.0 ± 14.0 | 46.0 ± 16.0 |       | < .001  |
| Range                  | 18–82  | 18–81     |       |         |
| Sex                    |        |           |       | .053    |
| Male                   | 38     | 76        | 114   |         |
| Female                 | 175    | 226       | 401   |         |
| Diameter (cm)          |        |           |       |         |
| Mean ± SD              | 2.30 ± 2.55 | 0.88 ± 0.70 |       | < .001  |
| Range                  | 0.23–7.1 | 0.16–6.1  |       |         |
| Thyroid function       |        |           |       |         |
| FT3 (pmol/L)           | 4.65 ± 1.03 | 4.78 ± 0.88 |       | .867    |
| FT4 (pmol/L)           | 16.31 ± 3.48 | 16.25 ± 4.18 |       | .591    |
| TSH (μU/ml)            | 1.78 ± 2.04 | 2.09 ± 2.30 |       | .003    |
| TgAb (nmol/L)          | 15.20 ± 24.65 | 17.80 ± 65.30 |       | .004    |
| TPOAb (mIU/ml)         | 36.00 ± 28.05 | 40.6 ± 11.60 |       | .008    |

FT3 = free triiodothyronine, FT4 = free thyroxine, TgAb = anti-thyroglobulin antibody, TPOAb = thyroid peroxidase antibody, TSH = thyrotropin.
Histopathological results were based on the final and standard diagnosis from the surgical specimen, not from biopsies or cytology. After surgery in 315 patients with 708 thyroid nodules, histopathological examination showed 341 benign and 367 malignant cases, including 193 follicular adenomas, 134 goiters, 8 atypical adenomas, 6 oncocyic (Hürthle cell) adenomas, 188 classical papillary carcinomas, 159 papillary thyroid microcarcinomas, 9 follicular variants of papillary carcinomas, 9 follicular carcinomas, 1 medullary carcinoma, and 1 poorly differentiated carcinoma. The diameter of benign nodules was 2.30 ± 2.55 cm (0.26–7.1 cm), significantly larger than that of malignant nodules (0.88 ± 0.70 cm; range, 0.16–6.1 cm) (P < .01). Serum thyrotropin (TSH), anti-thyroid globulin antibody (TGAb), and TPOAb levels in patients with benign nodules were lower compared with those with malignant nodules (1.78 ± 2.04 vs 2.09 ± 2.30 μIU/ml), (15.20 ± 24.65 vs 17.8 ± 65.3 U/ml), and (36.0 ± 28.05 vs 40.6 ± 116.0 U/ml), respectively (P < .05).

There were no significant differences in sex, free triiodothyronine (FT3), and free thyroxine (FT4) between the benign and malignant groups. Compared with benign nodules, malignant ones had significantly higher rates of solid mass, hypoechoic area, irregular margins, microcalcification, and a taller-than-wide shape (P < .01).

### 3.2. Malignancy risk stratification

Based on postoperative histopathology, the malignancy rates of nodules in TIRADS categories 2, 3, 4, and 5 were 0% (0 of 142 nodules), 8.1% (12 of 148 nodules), 67.0% (63 of 94 nodules), and 90.1% (292 of 324 nodules), respectively, with significant differences among categories (P < .01) (Table 3). For nodules <1 cm, the malignancy rates in TIRADS categories 2, 3, 4, and 5 nodules were 0, 23.3%, 69.2%, and 91.6%, with significant differences among categories (P < .01). For nodules >1 to 2 cm, the malignancy rates were 0, 5.3%, 84.2%, and 86.7%, in TIRADS categories 2, 3, 4, and 5, respectively, with significant differences among categories (P < .01). For nodules >2 cm, the malignancy rates were 0, 6.0%, 56.3%, and 92.6% in TIRADS categories 2, 3, 4, and 5, respectively, with significant differences among categories (P < .01) (Table 3). Those groups are based on the definition of microcarcinoma (<1 cm), while 2 cm is the cutoff used by the K-TIRADS and ATA for fine needle aspiration.[11,23]

For nodules with very low, low, intermediate, high suspicion of malignancy, and nonspecific pattern by US ATA guidelines, malignancy rates were 0 (0 of 174 nodules), 17.7% (24 of 136 nodules), 57.9% (11 of 19 nodules), 90.0% (289 of 321 nodules), and 69.4% (43 of 62 nodules), respectively. For nodules <1 cm, the malignancy rates were 0, 29.4%, 79.0%, 92.5%, and 64.3%, respectively, in the very low, low, intermediate, and high suspicion, and nonspecific pattern groups. For nodules 1 to 2 cm, malignancy rates were 0, 13.3%, 50.0%, 87.1%, and 81.3%, respectively. Nodules >2 cm showed malignancy rates of 0, 13.5%, 33.3%, 92.5%, and 61.1%, respectively, with significant differences among patterns (P < .01) (Table 3).

### 3.3. Diagnostic values of TIRADS and ATA guidelines

Of the 708 thyroid nodules, 62 failing to meet the US 2015 ATA guidelines were classified as “nonspecific pattern” (Fig. 1). They were solid or partially solid, isoechoic or hyperechoic, with some suspicious features like irregular margins, taller-than-wide shape or microcalcification. Among them, 19 benign and 43 malignant nodules were confirmed by postoperative histopathology. The diagnostic value of the 2 systems (including the nonspecific pattern nodules or not) is shown in Table 4.

Receiver operating characteristic (ROC) curves revealed the best cutoff for the TIRADS and low suspicion of malignancy for the US 2015 ATA guidelines. For nodules without nonspecific US ATA patterns, AUCs, sensitivities, specificities, PPVs and NPVs were 0.934 (95% CI 0.907–0.948), 96.3% (95% CI 93.6–98.1), 85.4% (95% CI 81.1–89.1), 86.9% (95% CI 83.0–90.2), and 95.8% (95% CI 92.8–97.8), respectively, for the TIRADS, and 0.930 (95% CI 0.912–0.952), 92.6% (95% CI 89.2–95.2), 87.3% (95% CI 83.1–90.7), 88.0% (95% CI 84.0–91.2), and 92.1% (95% CI 88.5–94.9), respectively, for the US ATA system.
For nodules, AUCs, sensitivities, specificities, PPVs and NPVs were 0.926 (95% CI 0.904–0.944), 96.7% (95% CI 94.4–98.3), 81.5% (95% CI 77.0–85.5), 84.9% (95% CI 81.1–88.2), and 95.9% (95% CI 92.9–97.8) for the TIRADS, and 0.920 (95% CI 0.898–0.939), 93.5% (95% CI 90.4–95.8), 82.4% (95% CI 77.9–86.3), 85.1% (95% CI 81.3–88.4), and 92.1% (95% CI 88.5–94.9) for the US ATA system, respectively. The TIRADS had higher AUC, sensitivity, and NPV (P < .05, respectively) (Table 4). For all nodules, AUCs, sensitivities, specificities, PPVs and NPVs were 0.926 (95% CI 0.904–0.944), 96.7% (95% CI 94.4–98.3), 81.5% (95% CI 77.0–85.5), 84.9% (95% CI 81.1–88.2), and 95.9% (95% CI 92.9–97.8) for the TIRADS, and 0.920 (95% CI 0.898–0.939), 93.5% (95% CI 90.4–95.8), 82.4% (95% CI 77.9–86.3), 85.1% (95% CI 81.3–88.4), and 98.0% for the TIRADS, respectively, and 0.859, 88.4%, 75.0%, 91.8%, and 67.1% for US ATA patterns, respectively. The US ATA patterns had a higher NPV compared with the TIRADS (P < .05). There were no differences in AUC, sensitivity, and specificity between the 2 systems. For nodules of 1 to 2 cm, the AUC, sensitivity, specificity, NPV, and PPV were 0.902, 98.7%, 80.3%, 86.4%, and 98.0% for the TIRADS, respectively, and 0.899, 97.4%, 78.7%, 85.2%, and 96.0% for US ATA patterns, indicating a higher NPV in the TIRADS (P < .05). For nodules > 2 cm, the AUC, sensitivity, specificity, NPV, and PPV were 0.941, 87.8%, 91.8%, 72.9%, and 96.7% for the TIRADS, and 0.926, 75.5%, 94.3%, 77.1%, and 93.8% for US ATA patterns. These data indicated that the TIRADS had higher AUC, sensitivity, and NPV (Table 5).

### Table 3

Comparison between the TI-RADS and ATA patterns.

| Scoring system and category | Final diagnosis | P value |
|-----------------------------|-----------------|---------|
|                            | Benign | Malignant | Malignancy rate (%) |
| Overall TIRADS              |        |           |                   |
| 2                           | 142    | 0         | 0                 |
| 3                           | 136    | 12        | 8.1%              |
| 4                           | 31     | 63        | 67.0%             |
| 5                           | 32     | 292       | 90.1%             |
| ATA                         |        |           |                   |
| Very low suspicion          | 170    | 0         | 0                 |
| Low suspicion               | 112    | 24        | 17.7%             |
| Intermediate suspicion      | 6      | 11        | 57.9%             |
| High suspicion              | 32     | 289       | 90.0%             |
| Unspecific                  | 19     | 43        | 69.4%             |
| <1 cm TIRADS                |        |           |                   |
| 2                           | 18     | 0         | 0                 |
| 3                           | 25     | 11        | 23.3%             |
| 4                           | 17     | 30        | 69.2%             |
| 5                           | 16     | 200       | 91.6%             |
| ATA                         |        |           |                   |
| Very low suspicion          | 22     | 0         | 0                 |
| Low suspicion               | 24     | 10        | 29.4%             |
| Intermediate suspicion      | 4      | 15        | 79.0%             |
| High suspicion              | 16     | 198       | 92.5%             |
| Unspecific                  | 10     | 18        | 64.3%             |
| 1–2 cm TIRADS               |        |           |                   |
| 2                           | 31     | 0         | 0                 |
| 3                           | 18     | 1         | 5.3%              |
| 4                           | 3      | 16        | 84.2%             |
| 5                           | 9      | 60        | 86.7%             |
| ATA                         |        |           |                   |
| Very low suspicion          | 35     | 0         | 0                 |
| Low suspension              | 13     | 2         | 13.3%             |
| Intermediate suspension     | 1      | 1         | 50.0%             |
| High suspension             | 9      | 61        | 87.1%             |
| Unspecific                  | 3      | 13        | 81.3%             |
| >2 cm TIRADS                |        |           |                   |
| 2                           | 83     | 0         | 0                 |
| 3                           | 94     | 6         | 6.0%              |
| 4                           | 14     | 18        | 56.3%             |
| 5                           | 2      | 25        | 92.6%             |
| ATA                         |        |           |                   |
| Very low suspicion          | 105    | 0         | 0                 |
| Low suspension              | 77     | 12        | 13.5%             |
| Intermediate suspension     | 2      | 1         | 33.3%             |
| High suspension             | 2      | 25        | 92.6%             |
| Unspecific                  | 7      | 11        | 61.1%             |

P value, Comparison between categories.
Table 4
Diagnostic values of the TI-RADS and ATA 2015 guidelines.

| Scoring system | Cutoff | Accuracy (95%CI) | AUC (95%CI) | Sensitivity (95%CI) | Specificity (95%CI) | PPV (95%CI) | NPV (95%CI) |
|----------------|--------|------------------|-------------|---------------------|---------------------|-------------|-------------|
| Overall, without unspecific nodules | | | | | | | |
| TIRADS 4 | 0.909 | 0.934 | 0.907–0.948 | 96.3 * | 85.4 * | 86.9 | 95.8 |
| ATA Intermediate Suspicion | 0.899 | 0.930 | 0.912–0.952 | 92.6 | 87.3 | 88.0 | 92.1 |
| Overall | | | | | | | |
| TIRADS 4 | 0.894 | 0.926 | 0.904–0.944 * | 96.7 | 81.5 | 84.9 | 95.9 |
| ATA Intermediate Suspicion | 0.847 | 0.920 | 0.898–0.939 | 90.4–95.8 | 77.9–86.3 | 81.3–88.4 | 88.5–94.9 |

Compared to ATA patterns. *P<.05.
Table 5
Subgroup analysis of the TI-RADS and ATA patterns.

| Scoring system | Cutoff | Accuracy | AUC (95%CI) | Sensitivity (95%CI) | Specificity (95%CI) | PPV (95%CI) | NPV (95%CI) |
|----------------|--------|----------|-------------|---------------------|---------------------|-------------|-------------|
| <1 cm          |        |          |             |                     |                     |             |             |
| TRADS          | 4      | 0.864    | 0.853       | [0.809–0.890]       | 83.0                | 80.3        | 93.0        | 59.8        |
| ATA Intermediate Suspicion | 0.852   | 0.859       | [0.816–0.895]       | 88.4                | 75.0                | 91.8        | 67.1        |
| 1–2 cm         |        |          |             |                     |                     |             |             |
| TRADS          | 4      | 0.905    | 0.902       | [0.840–0.946]       | 96.7                | 80.3        | 86.4        | 98.0        |
| ATA Intermediate Suspicion | 0.819   | 0.899       | [0.837–0.944]       | 97.4                | 78.7                | 85.2        | 96.0        |
| >2 cm          |        |          |             |                     |                     |             |             |
| TRADS          | 4      | 0.909    | 0.941       | [0.903–0.967]       | 87.8                | 91.8        | 72.9        | 96.7        |
| ATA Intermediate Suspicion | 0.889   | 0.926       | [0.865–0.955]       | 97.5                | 94.3                | 71.1        | 93.8        |

Compared to ATA patterns. *P < 0.05. AUC = area under the curve, CI = confidence interval, NPV = negative predictive value, PPV = positive predictive value.

4. Discussion

US is used to evaluate the malignancy potential of thyroid nodules and help in biopsy decision. The TI-RADS has been clinically used for a few years. A recent meta-analysis reported pooled sensitivity and specificity of the TI-RADS in differentiated diagnosis of thyroid nodules to be 0.79 and 0.71, respectively. However, it has not yet been adopted by the ATA guidelines, which are much lower than the TI-RADS. In the present study, the malignancy rate obtained for TI-RADS 4 nodules was overtly higher. This elevated malignancy risk in category 4 nodules can be attributed to a selection bias, as for the US part of the 2015 ATA classification.

As shown by ROC curves, for nodules without nonspecific ones, the TI-RADS system had a relatively higher sensitivity, specificity and NPV compared with the US ATA system for malignancy risk stratification. For nodules with nonspecific ones, the TI-RADS system also had a relatively higher AUC, sensitivity and NPV compared with the US ATA system. These findings suggested that the TI-RADS system could be used for nodules with nonspecific patterns according to the US ATA guidelines. Meanwhile, the TI-RADS had a higher diagnostic value compared with US ATA guidelines for all nodules, including those with nonspecific US ATA patterns. In the sub-centimeter group, US ATA patterns had a better NPV (P < 0.05). For nodules of 1 to 2 cm, the TI-RADS had a higher NPV, and slightly but nonsignificantly higher AUC, sensitivity, specificity, and PPV. For those >2 cm in diameter, the TI-RADS had higher AUC, sensitivity, and NPV. Taken together, these findings suggested that the TI-RADS is more efficient than the US part of the ATA guidelines in determining the malignancy of larger nodules.

In the ATA guidelines, a nonspecific group was identified with a high malignancy risk (69.4%) according to US. In the latter group, most nodules were >1 cm in diameter (1.68 ± 1.20 cm), with partially cystic composition, or iso- or hyper-echogenicity. Indeed, only 5-26% of partially cystic nodules are malignant, and iso- and hyperechogenicity are more likely to reflect benignity compared with hypoechogenicity. Recently, Yoon et al also reported that nodules not meeting the criteria for any specific pattern in the US ATA guidelines have a relatively high risk of malignancy (18.2%) as they compared malignancy risk stratification of thyroid nodules by the US ATA guidelines and the TI-RADS. In the present study, the proportion of nodules assigned to the “unspecified” category by the US ATA guidelines was higher (8.7% vs 3.4%), with a dramatically higher malignancy rate in those nodules (69.4% vs 18.2%). All these nonspecific nodules could be classified as TI-RADS 4 with a similar malignancy rate (69.4% vs 67.4%).
US features are very helpful in determining the follow-up procedure for thyroid nodules with benign cytology diagnoses.[10] The TIRADS adopted in this study was proposed by Na et al.[11] combining specific suspicious US features and less specific ones (solidity and echogenicity) in thyroid nodules. It was shown to be useful for risk stratification of thyroid nodules and management decision.[12] Particularly, for cytologically indeterminate thyroid nodules, both the TIRADS and US ATA guidelines allow high-confidence exclusion of malignancy with stringent negativity cutoffs,[29] and high sensitivity may be obtained using the US part of the ATA guidelines.[30]

The limitations of the present study should be addressed. First, the study is retrospective including only patients who had a surgical resection and thus influenced by many other factors to include selection bias and changes in clinical practice over time during data collection. In our study, 51.8% of the nodules was carcinoma which was higher than that of some retrospective studies[31,32] but comparable to those reported studies with similar design.[33] The malignancy rate of our study was higher than that of some retrospective study, such as 37.2% from Han et al [31] and 39% of Xu et al.[32] An important reason is that in the above 2 studies, 25.9% and 90.1% of the nodules were regarded as benign lesions based on cytology and follow-up US, which may cause false negative results. While in another retrospective study with a similar design,[33] 2,544 thyroid nodules in 1,758 patients who underwent thyroidectomy were included. Of all the nodules, 863 (33.9%) were benign, whereas 1,681 (66.1%) were malignant. The malignancy rate was relatively higher than our study. Secondly, only 2 radiologists retrospectively reviewed the US images and classified the nodules according to the US 2015 ATA guidelines and TIRADS, with possible deviations among investigators. Thirdly, The TIRADS-Na used in this study is a relatively new tool.[1] Therefore the universality of the present study was limit for the small application range. Finally, some patients had multiple nodules, which included nodules <1 cm. Nodules >1 cm may be punctured and, when confirmed as PTC, they undergo surgery. The 2 lobes of the thyroid were removed during surgery and contained the nodules <1 cm. Because the study period was from 2013 to 2016 and the ATA guidelines were released in 2015, the understanding of the follow-up observation for thyroid microcarcinoma was not mature at that time, the pathologists suggested that all patients with PTC had to undergo surgery.

In conclusion, both the TIRADS and the US part of the 2015 ATA guidelines have appreciable diagnostic values for thyroid nodules. This study identified thyroid nodules with nonspecific US ATA patterns. The newly proposed TIRADS may be more efficient than the US 2015 ATA guidelines, especially for nodules >2 cm in diameter or those with nonspecific patterns. The identification of nodules with nonspecific US patterns indicates the need for improving the ATA guidelines for risk stratification. The ATA US guideline is used not only for FNA determination, but also for providing a follow-up recommendation for nodules with benign or indeterminate cytology and nodules without FNA. These evaluation and follow-up recommendations remain instructive in the management of nodules with specific patterns. TIRADS is much more used in determining the need of FNA instead of further management. Therefore, a wiser improvement of TIRADS into clinical management of thyroid nodules is required. The best approach is to tailor TIRADS as ATA US guidelines in the aspect of nodules management. The use of US based on these results could be invaluable when combined with clinical features and biopsy results.

Author contributions

Conceptualization: Pingping Xiang, Shuhang Xu, Chao Liu.

Data curation: Pingping Xiang, Xiaoqiu Chu, Guofang Chen, Binbin Liu, Wenbo Ding, Zheng Zeng, Xining Wu, Jianhua Wang.

Formal analysis: Pingping Xiang, Xiaoqiu Chu, Guofang Chen, Binbin Liu, Wenbo Ding, Zheng Zeng, Xining Wu, Jianhua Wang, Shuhang Xu, Chao Liu.

Funding acquisition: Shuhang Xu, Chao Liu.

Project administration: Shuhang Xu, Chao Liu.

Writing – original draft: Pingping Xiang.

Writing – review & editing: Xiaoqiu Chu, Guofang Chen, Binbin Liu, Wenbo Ding, Zheng Zeng, Xining Wu, Jianhua Wang, Shuhang Xu, Chao Liu.

References

[1] Nguyen QT, Lee EJ, Huang MG, et al. Diagnosis and treatment of patients with thyroid cancer. Am Health Drug Benefits 2015;8:30–40.
[2] Wang K, Yang Y, Wu Y, et al. The association between insulin resistance and vasculization of thyroid nodules. J Clin Endocrinol Metab 2015;100:184–92.
[3] Gharib H, Papini E, Garber JR, et al. American Association of Clinical Endocrinologists, American College of Endocrinology, and Associazione Medici Endocrinologi Medical Guidelines for Clinical Practice for the Diagnosis and Management of Thyroid Nodules—2016 Update. Endocr Pract 2016;22:622–39.
[4] Guth S, Theeume U, Aberle J, et al. Very high prevalence of thyroid nodules detected by high frequency (13MHz) ultrasound examination. Eur J Clin Invest 2009;39:699–706.
[5] Burman KD, Wartofsky L. Thyroid nodules. N Engl J Med 2016;374:1294–5.
[6] Yeh MW, Bauer AJ, Bernet VA, et al. American Thyroid Association statement on preoperative imaging for thyroid cancer surgery. Thyroid 2015;25:3–14.
[7] Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid 2016;26:1–33.
[8] Musa A, De Andrea M, Motta M, et al. Predictors of malignancy in children with thyroid nodules. J Pediatr 2015;167:886–92 e881.
[9] Remorini LR, Kramer CK, Leitao CR, et al. Thyroid ultrasound features and risk of carcinoma: a systematic review and meta-analysis of observational studies. Thyroid 2015;25:538–50.
[10] Horvath E, Majloš S, Rossi R, et al. An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. J Clin Endocrinol Metab 2009;94:1748–51.
[11] Na DG, Baek JH, Sung JY, et al. Thyroid imaging reporting and data system risk stratification of thyroid nodules: categorization based on solidity and echogenicity. Thyroid 2016;26:562–72.
[12] Russ G, Royer B, Bigorgne C, et al. Prospective evaluation of thyroid imaging reporting and data system on 4550 nodules with and without elastography. Eur J Endocrinol 2013;168:649–55.
[13] Kwak JI, Han KH, Yoon JH, et al. Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk. Radiology 2011;260:892–9.
[14] Grant EG, Tessler FN, Hoang JK, et al. Thyroid ultrasound reporting lexicon: white paper of the ACR Thyroid Imaging, Reporting and Data System (TIRADS) Committee. J Am Coll Radiol 2013;10:11272–9.
[15] Yoon JH, Lee HS, Kim EK, et al. 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] Moon WJ, Baek JH, Jung SL, et al. Ultrasonography and the ultrasound-based management of thyroid nodules: consensus statement and recommendations. Korean J Radiol 2011;12:1–4.
[17] Moon WJ, Jung SL, Lee JH, et al. Benign and malignant thyroid nodules: US differentiation—multicenter retrospective study. Radiology 2008;247:762–70.
[18] Kim BM, Kim MJ, Kim EK, et al. Sonographic differentiation of thyroid nodules with eggshell calcifications. J Ultrasound Med 2008;27:1425–30.
[19] Brito JP, Gonçalves MR, Al Nofal A, et al. The accuracy of thyroid nodule ultrasound to predict thyroid cancer: systematic review and meta-analysis. J Clin Endocrinol Metab 2014;99:1253–63.
[20] Gul K, Ersoy R, Dirikoc A, et al. Ultrasonographic evaluation of thyroid nodules: comparison of ultrasonographic, cytological, and histopathological findings. Endocrine 2009;36:464–72.
[21] Salmaslioglu A, Erbil Y, Dural C, et al. Predictive value of sonographic features in preoperative evaluation of malignant thyroid nodules in a multinodular goiter. World J Surg 2008;32:1948–54.
[22] Frates MC, Benson CB, Doubilet PM, et al. Prevalence and distribution of carcinoma in patients with solitary and multiple thyroid nodules on sonography. J Clin Endocrinol Metab 2006;91:3411–7.
[23] Cooper DS, et al. American Thyroid Association Guidelines Taskforce on Thyroid N, Differentiated Thyroid CRevised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid 2009;19:1167–214.
[24] Wei X, Li Y, Zhang S, et al. Meta-analysis of thyroid imaging reporting and data system in the ultrasonographic diagnosis of 10,437 thyroid nodules. Head Neck 2016;38:309–15.
[25] Perros P, Boelaert K, Colley S, et al. Guidelines for the management of thyroid cancer. Clin Endocrinol (Oxf) 2014;81(Suppl 1):1–22.
[26] Chan BK, Dessir TS, McDougall IR, et al. Common and uncommon sonographic features of papillary thyroid carcinoma. J Ultrasound Med 2003;22:1083–90.
[27] Watters DA, Ahuja AT, Evans RM, et al. Role of ultrasound in the management of thyroid nodules. Am J Surg 1992;164:654–7.
[28] Lee MJ, Kim EK, Kwak JY, et al. Partially cystic thyroid nodules on ultrasound: probability of malignancy and sonographic differentiation. Thyroid 2009;19:341–6.
[29] Grani G, Lamartina L, Ascoli V, et al. Ultrasoundography scoring systems can rule out malignancy in cytologically indeterminate thyroid nodules. Endocrine 2017;57:236–61.
[30] Trimboli P, Fucini F, Zilioli V, et al. Accuracy of international ultrasound risk stratification systems in thyroid lesions cytologically classified as indeterminate. Diagn Cytopathol 2017;45:113–7.
[31] Ha SM, Ahn HS, Baek JH, et al. Validation of three scoring risk-stratification models for thyroid nodules. Thyroid 2017;27:1550–7.
[32] Xu T, Gu JY, Ye XY, et al. Thyroid nodule sizes influence the diagnostic performance of TIRADS and ultrasound patterns of 2015 ATA guidelines: a multicenter retrospective study. Sci Rep 2017;7:43183.
[33] Gao L, Xi X, Jiang Y, et al. Comparison among TIRADS (ACR TI-RADS and KWAK-TI-RADS) and 2015 ATA Guidelines in the diagnostic efficiency of thyroid nodules. Endocrine 2019.[Epub ahead of print].