Thyroid

DOI: 10.1089/thy.2023.0429

Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study

Xiaoxian Li[1], M.D.* • Chuan Peng[1], M.D.* • Ying Liu[2], M.D. • Yixin Hu[1], M.D. • Liang Yang[1], M.D. • Yiwen Yu[1], M.D., Ph.D. • Hongyan Zeng[3], M.D. • Weijun Huang[4], M.D. • Qian Li[5], M.D. • Nansheng Tao[6], B.D. • Longhui Cao[7], M.D., Ph.D.# • Jianhua Zhou[1], M.D., Ph.D.#

* X.X.L. and C.P. contributed equally to this work.

Author's affiliation:

[1] Department of Ultrasound, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, No. 651 Dongfeng Road East, Guangzhou, 510060, PR China.

[2] Department of Ultrasound, Department of Medical Ultrasonics, The Eighth Affiliated Hospital of Sun Yat-Sen University, No. 3025 Shennan Road, Shenzhen, 518033, PR China.

[3] Department of Ultrasound, Huadu District People's Hospital, No. 48, Xinhua Road, Huadu District, Guangzhou; 510800, P.R. China.

[4] Department of Ultrasound, Foshan First Municipal People's Hospital (The Affiliated Foshan Hospital of Sun Yat-Sen University), No.81 Lingnan North Road, Foshan, 528000, Guangdong Province, P.R. China.

[5] Department of Ultrasound, Affiliated Tumor Hospital of Zhengzhou University, No.127, Dongming Road, Jinshui District, Zhengzhou 450008, P.R. China;

[6] Department of Ultrasound, The Fifth People's Hospital of Nanhai, No.4, Dongyi street, Zhenxing Road, Lishui Town, Nanhai District, Foshan City, Guangdong Province, P.R. China.

[7] Department of Anesthesiology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, No. 651 Dongfeng Road East, Guangzhou, 510060, PR China.

Email address:

Xiaoxian Li, M.D.: lixx43@mail2.sysu.edu.cn

Chuan Peng, M.D.: pengchuan@sysucc.org.cn

Ying Liu, M.D.: liuyzdby@163.com

Yixin Hu, M.D.: huyx@sysucc.org.cn

Liang Yang, M.D.: yangliang@sysucc.org.cn

Yiwen Yu, M.D., Ph.D.: yuyw@sysucc.org.cn

Hongyan Zeng, M.D.: zenghongyan12345@163.com

Weijun Huang, M.D.: huangweijun12345@126.com

Qian Li, M.D.: liqian123452021@yeah.net

Nansheng Tao, B.D.: taonansheng12345@163.com

Longhui Cao, M.D., Ph.D.: caolh@sysucc.org.cn

Jianhua Zhou, M.D., Ph.D.: zhoujh@sysucc.org.cn

[#] **Co-corresponding authors:**

Longhui Cao. Department of Anesthesiology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, No. 651 Dongfeng Road East, Guangzhou, 510060, PR China. Tel: +8620-87343673. E-mail: caolh@sysucc.org.cn

Jianhua Zhou. Department of Ultrasound, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, No. 651 Dongfeng Road East, Guangzhou, 510060, China. Tel: +8620-87343212. E-mail: zhoujh@sysucc.org.cn

Running title: Based on FNA threshold and stratification criteria

Keywords: TI-RADS, thyroid nodules, fine needle aspiration biopsy, multicenter study

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

Abstract

Background: Risk stratification systems for thyroid nodules are limited by low specificity. The fine-needle aspiration (FNA) biopsy size thresholds and stratification criteria are based on evidence from the literature and expert consensus. Our aim was to investigate the optimal FNA biopsy size thresholds in the American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TI-RADS) and artificial intelligence (AI) TI-RADS, and to revise the stratification criteria in AI TI-RADS.

Methods: A total of 2596 thyroid nodules (in 2511 patients) on ultrasound examination with definite pathological diagnoses were retrospectively identified from January 2017 to September 2021 in six participating Chinese hospitals. The modified criteria for ACR TI-RADS were: (1) no FNA for TR3; (2) FNA threshold for TR4 increased to 2.5 cm. The modified criteria for AI TI-RADS were: (1) 6-point nodules upgraded to TR5; (2) no FNA for TR3; (3) FNA threshold for TR4 increased to 2.5 cm. The diagnostic performance and unnecessary FNA rate (UFR) of modified versions were compared with the original ACR TI-RADS.

Results: Compared with original ACR TI-RADS, mACR (modified ACR) TI-RADS yielded higher specificity (73% vs 46%), accuracy (74% vs 51%), area under curve (AUC, 0.80 vs 0.70) and lower UFR (25% vs 48%; all $P < 0.001$) although the sensitivity was slightly decreased (87% vs 93%, $P = 0.057$). Compared with original ACR TI-RADS, mAI (modified AI) TI-RADS yielded higher specificity (73% vs 46%), accuracy (75% vs 51%), AUC (0.81 vs 0.70) and lower UFR (24% vs 48%; all $P < 0.001$), although the sensitivity tended to be slightly decreased (89% vs 93%, $P = 0.13$). There was no significant difference between mACR TI-RADS and mAI TI-RADS in the diagnostic performance and UFR (all $P > 0.05$).

Conclusions: The revised FNA thresholds and stratification criteria of mACR TI-RADS and mAI TI-RADS may be associated with improvements in specificity and accuracy, without significantly sacrificing sensitivity for malignancy detection.

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

INTRODUCTION

The detection rate of thyroid nodules has increased rapidly with the widespread use of ultrasound (US) since the 1990s[1,2]. Fine-needle aspiration cytology (FNAC) remains the gold standard for diagnosis[3]. Although FNA is a relatively safe and cost-effective procedure, performing FNA for all nodules is impractical, inappropriate, and unnecessary because only 10% of patients presenting with thyroid nodules are at risk of malignancy[4]. Overdiagnosis and overtreatment of thyroid nodules are current concerns worldwide, and up to 77% of detected thyroid cancers may be clinically insignificant[2]. It is noteworthy that thyroid cancer-related mortality rates have not increased substantially despite the sharp increase in incidence[5]. Excessive examination and intervention may not only cause anxiety and economic burden for patients but also waste of medical resources for society. Therefore, determining how to reduce unnecessary biopsies while maintaining appropriate sensitivity for malignancy detection, is an issue that requires further investigation.

Various risk stratification systems based on US features have been proposed globally and used to determine which nodules should be subjected to FNA. Size thresholds vary across guidelines, leading to differences in their diagnostic performance and unnecessary biopsy rate. Previous comparative studies showed that Thyroid Imaging Reporting and Data System (TI-RADS) published by the American College of Radiology (ACR) showed the highest specificity and lowest unnecessary biopsy rate as compared to other guidelines[6,7], which was attributed to the larger size thresholds of the ACR guidelines[8]. However, a recent retrospective cohort study reported that 57.4% of biopsied thyroid nodules were benign[9], which indicates that efforts should be taken to improve the diagnostic performance of ACR TI-RADS. Smaller FNA size thresholds may lead to excessive FNAs, while larger thresholds may decrease the sensitivity.

In an effort to achieve higher specificity, Wildman-Tobriner et al[10] applied artificial intelligence (AI) to optimize TI-RADS by assigning new scores for eight ultrasound features in 2019. Our previous study[11] has validated AI TI-RADS significantly improved specificity (70.2% vs 49.2%) despite a slight decrease in sensitivity compared to ACR TI-RADS (82.2% vs 86.7%). AI TI-RADS assigned lower risk levels for 54 malignant nodules, resulting in 29

papillary carcinomas smaller than 1.5 cm were missed diagnosed. Therefore, it is important to investigate how to modify the stratification criteria of AI TI-RADS to compensate for the sacrifice in sensitivity.

The FNA size thresholds and stratification criteria in TI-RADS are based on evidence from the literature and expert consensus, which could be optimized to improve the performance of the system. This study aimed to investigate the optimal FNA size thresholds in ACR TI-RADS and AI TI-RADS, and to revise the stratification criteria in AI TI-RADS.

MATERIALS AND METHODS

This study was approved by the institutional review boards of all participating institutions (approval no. B2021–021-Y01). The requirement for informed consent was waived by the institutional review boards because of the retrospective study design.

Study Patients

Between January 2017 and September 2021, a total of 4001 thyroid nodules from 3517 consecutive patients who underwent thyroid US at six different hospitals in China were retrospectively identified. The eligibility criteria were: (*a*) age ≥ 18, (*b*) the maximum diameter of the nodules was ≥1.0 cm, (*c*) nodules with definitive cytology results (Bethesda category II or VI), definitive core-needle biopsy (CNB) results or surgical resection. US-guided FNA was performed for the thyroid nodules under the recommendation of ACR TI-RADS or before thermal ablation for TR 1 and TR 2 nodules due to compressive or cosmetic symptoms. CNB was usually performed in nodules with prior inconclusive FNA results. The exclusion criteria were: (*a*) nodules with inconclusive final diagnoses (n = 634), (*b*) nodules underwent prior treatments (n = 92), (*c*) nodules with incomplete or poor US images (n = 19). Thus, a total of 3256 thyroid nodules were eligible, including 2336 benign nodules and 920 malignant nodules with a malignancy rate of 28.3%. It was reported that approximately 10% of patients who present with thyroid nodules are at risk of malignancy[4]. To evaluate the diagnostic performance of the modified criteria proposed in this study in the general population, malignant nodules were included using simple random sampling, while benign nodules were included consecutively. Among eligible

malignant nodules, all characteristics were comparable between the exclusion cohort and inclusion cohort (Table S1). A total of 2596 thyroid nodules from 2511 patients were included, including 2336 benign nodules and 260 malignant nodules with a malignancy rate of 10.0% (Fig. 1). We previously reported on 601 of the included in our study evaluating the efficacy of AI TI-RADS[11].

US Examinations and Image Analysis

All nodules underwent US examination within two weeks before biopsy or operation. US examinations were performed using high-frequency linear probes and a real-time US system. The US systems used included GE Logiq 9, Logiq E9, Logiq S8 (GE Medical Systems, Milwaukee, WI, USA); Aixplorer, (Supersonic Imagine, Paris, France); Philips IU22, EPIQ 7 (Philips Medical Systems, Best, the Netherlands); Siemens ACUSON Juniper, Sequoia, S2000 (Siemens Medical Solutions, Mountain View, CA, USA); Toshiba Aplio 400 (Toshiba Medical Systems Corp., Tokyo, Japan); Hitachi Aloka ProSound ALPHA 10 (Hitachi-Aloka Medical, Tokyo, Japan); Esaote MyLab 70 (Esaote, Genoa, Italy); Mindray Resona 7T, DC-8 (Mindray Medical International, Shenzhen, China). All US-guided procedures were performed by radiologists with at least 5-year experience in US.

US image analysis was performed by two experienced radiologists (C.P. and Y.L., with 7 and 8 years of experience, respectively, in thyroid imaging). Blinded to the clinical and pathological data, they independently reviewed all US images and assessed the US features of thyroid nodules according to ACR TI-RADS, including nodule maximum diameter, composition, echogenicity, shape, margin and echogenic foci. Figure S1 shows the scoring system for ACR TI-RADS and AI TI-RADS. When grading a nodule, the reviewer selected one feature from each of the five categories, and the total score determined the nodule's TI-RADS risk level. Recommendations for FNA or US follow-up were based on a nodule's TI-RADS level and its maximum diameter. Images would be reassessed by an expert (J.H.Z., with 22 years of experience in thyroid imaging) when disagreement between two reviewers existed.

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

Exploration of Modified Criteria to ACR TI-RADS and AI TI-RADS

According to our previous study evaluating the efficacy of AI TI-RADS[11], the malignancy rate of 6-point nodules in AI TI-RADS was 43.1%, which was significantly higher than the malignant risk level (5-20%) of TR4 suggested in the ACR TI-RADS. Therefore, we hypothesized that upgrading 6-point nodules from AI TR4 to AI TR5 could improve the sensitivity of AI TI-RADS (hereinafter referred to as “TR4-adjusted AI TI-RADS”).

To explore the optimal nodule size thresholds for FNA recommendation, the thresholds of ACR TI-RADS, AI TI-RADS and TR4-adjusted AI TI-RADS were adjusted, respectively. Five new versions of each guideline were hypothetically established (Table 1). Version 1 simulated FNA size thresholds for TR 4 from 1.5 cm to 2.0 cm. Version 2 simulated FNA size thresholds for TR 3 from 2.5 cm to No FNA. Version 3 simulated FNA size thresholds for TR 3 from 2.5 cm to No FNA and TR 4 from 1.5 cm to 2.0 cm. Version 4 simulated FNA size thresholds for TR 3 from 2.5 cm to No FNA and TR 4 from 1.5 cm to 2.5 cm. Version 5 simulated FNA size thresholds for TR 3 from 2.5 cm to No FNA and TR 4 from 1.5 cm to 3.0 cm. The diagnostic performance and the unnecessary FNA rate (UFR) of recommended FNA in all new versions were calculated and compared with those in the original ACR TI-RADS.

Statistical Analysis

Patient demographics were compared using descriptive statistics. Quantitative data were summarized as the means ± standard deviation and comparing the means using the Mann-Whitney U test. Categorical data were summarized as percentages and compared by the chi-square test. The thyroid nodules were dichotomized into two groups, FNA indicated or not based on the criteria for FNA of each TI-RADS category and their new versions. The diagnostic performance in the detection of thyroid cancer was evaluated by sensitivity, specificity, accuracy, positive predictive value (PPV), negative predictive value (NPV), and the area under the receiver operating characteristic (ROC) curve (AUC) in each guideline, along with the 95% confidence intervals (CI). The UFR was defined as the percentage of the FNA-indicated benign nodules in the total number of nodules included. The McNemar test was used to assess for differences in these measures of diagnostic

performance and the DeLong test was applied to compare AUCs. Statistical analyses were performed using SPSS 22.0 (IBM, Armonk, NY) and R software 4.2.2 (R Foundation, Vienna, Austria). $P < 0.05$ indicated a statistically significant difference.

Results

Study Patients and Nodule Characteristics

Table 2 summarizes the patient demographics and nodule characteristics. Of the 2511 patients included in the study, 1914 (76.2%) were women and 597 (23.8%) were men. Of the total 2596 nodules, 2336 were benign and 260 were malignant. Among the 2336 benign nodules, 1922 were confirmed by cytology, 173 were confirmed by CNB and 241 were confirmed by surgical resection. Among the 260 malignant nodules, 40 were confirmed by cytology, 7 were confirmed by CNB and 213 were confirmed by surgical resection. The detailed pathological diagnoses are summarized in Table S2.

Diagnostic Performance and UFR of Original and New Versions

Table S3-S5 show the diagnostic performance and UFR of original version and new versions of each guideline. Among original versions, the highest sensitivity was observed with the ACR TI-RADS and TR4-adjusted AI TI-RADS (both, 93% [95% CI: 89%-96%]). The highest specificity was observed in the AI TI-RADS (56% [95% CI: 54%-58%]). The lowest UFR was observed in the AI TI-RADS (39% [95% CI: 37%-41%]).

We evaluated the impact on diagnostic performance and UFR by applying higher size thresholds for FNA recommendation to each guideline (Fig. 2). As the nodule size thresholds were raised, the specificity, accuracy, PPV and AUC of each new version gradually increased, while the sensitivity gradually decreased compared to the corresponding original version. Also, the UFR decreased markedly. Without significant difference in the decrease of sensitivity, version 4 of ACR TI-RADS and TR4-adjusted AI TI-RADS performed best relatively, which would be selected as modified ACR TI-RADS (mACR TI-RADS, Fig. 3) and modified AI TI-RADS (mAI TI-RADS, Fig. 4). The final modified criteria for mACR TI-RADS were as follows: (1) TR3 nodules were not recommended for FNA; (2) FNA threshold for TR4 increased to 2.5 cm. The modified criteria for mAI TI-RADS were as

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

follows: (1) 6-point nodules were upgraded from TR4 to TR5; (2) TR3 nodules were not recommended for FNA; (3) FNA threshold for TR4 increased to 2.5 cm. For patients with symptoms or cosmetic issues, FNA was indicated before any treatment.

Compared with the original ACR TI-RADS, mACR TI-RADS yielded higher specificity (73% vs 46%, $P < 0.001$), accuracy (74% vs 51%, $P < 0.001$), AUC (0.80 vs 0.70, $P < 0.001$) and lower UFR (25% vs 48%, $P < 0.001$) although the sensitivity was slightly decreased without a significant difference (87% vs 93%, $P = 0.057$). Compared with the original ACR TI-RADS, mAI TI-RADS yielded higher specificity (73% vs 46%, $P < 0.001$), accuracy (75% vs 51%, $P < 0.001$), AUC (0.81 vs 0.70, $P < 0.001$) and lower UFR (24% vs 48%, $P < 0.001$) although the sensitivity was slightly decreased without significant difference (89% vs 93%, $P = 0.13$). There was no significant difference between the mACR TI-RADS and mAI TI-RADS in the diagnostic performance and UFR (Table 3).

Original Versions vs Modified Versions

Table 4 summarizes the risk stratification and indication of FNA among the four TI-RADS versions. When the ACR TI-RADS and mACR TI-RADS were applied, the malignancy risk of most categories was consistent with those recommended in the ACR TI-RADS white paper; except for TR3 and TR4, where the malignancy risk was slightly lower. A total of 1494 nodules were recommended for FNA according to the ACR TI-RADS, of which 1253 (83.9%) were benign and 241 (16.1%) were malignant. Compared with the original ACR TI-RADS, mACR TI-RADS reduced FNA in 631 nodules, of which 617 (97.8%) were benign. Despite the decrease in the FNA rate of mACR TI-RADS, the malignancy detection rate was higher (26% [227 of 863] vs 16% [241 of 1494]). A total of 1241 nodules were recommended for FNA according to the AI TI-RADS, of which 1019 (82.1%) were benign and 222 (17.9%) were malignant. After modification, mAI TI-RADS reduced FNA in 391 benign nodules and increased 8 more FNA in malignant nodules, compared with the original AI TI-RADS.

Of the 2336 benign nodules, AI TI-RADS and mAI TI-RADS downgraded 113 nodules from ACR TR3 to AI/mAI TR2 and 45 nodules to AI/mAI TR1. Among ACR TR4 nodules, 94 nodules were downgraded to AI/mAI TR3, 22 to AI/mAI TR2, and 32 to AI/mAI TR1. Among

ACR TR5 nodules, 128 nodules were downgraded to AI TR4 (13 to mAI TR4 instead), and 4 to AI/mAI TR3. Ultimately, the new risk level assignments and size thresholds adjustments resulted in 1700 and 1708 benign nodules spared from FNA with the application of mACR TI-RADS and mAI TI-RADS, respectively (Fig. 5).

DISCUSSION

We found that compared with the original ACR TI-RADS, the mACR TI-RADS and mAI TI-RADS had higher specificities (72.8%, 73.1% vs 46.4%), AUCs (0.800, 0.808 vs 0.695) and lower UFRs (24.5%, 24.2% vs 48.3%, all $P < 0.001$) while the sensitivities were slightly but not significantly decreased (87.3% vs 92.7%, $P = 0.057$; 88.5% vs 92.7%, $P = 0.13$).

In the past decade, several associations have issued guidelines based on US features and nodule size to grade the risk of malignancy of thyroid nodules. Previous comparative studies revealed that ACR TI-RADS showed the highest specificity and lowest unnecessary biopsy rate compared with the other guidelines[6,7]. Ha et al[8] proved the main reason lied in the larger size thresholds of the ACR guidelines (mildly suspicious nodules, 2.5 cm; moderately suspicious nodules, 1.5 cm) when compared with the American Thyroid Association (ATA) and Korean Thyroid Association/Korean Society of Thyroid Radiology (KTA/KSThR) guidelines (1.5 cm and 1.0 cm, respectively). As the nodule size thresholds of the ATA and KTA/KSThR guidelines were raised, the diagnostic performance and unnecessary biopsy rates became similar to those seen with the ACR guideline[8]. A recent study[12] proposed the recommended FNA nodule size in Kwak TI-RADS 4b could be raised to 15 mm, 4a could not consider FNA and the ATA guideline intermediate suspicion could be raised to 15 mm or 20 mm, low suspicion and very low suspicion could not consider FNA. As far as we know, this is the first study to explore whether the size thresholds of ACR TI-RADS could be optimized to improve the diagnostic performance and decrease the unnecessary biopsy rates.

In our previous study[11], ACR TR3 nodules accounted for 31.1% of the total nodules, with a malignancy rate of 0.5%. In this study, ACR TR3 nodules accounted for 21.3% of the total nodules, with a malignancy rate of 1.4%. The most recent ATA guideline suggests that observation without FNA is a reasonable option for nodules in the very low suspicion

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

category with a risk of malignancy < 3%[13]. In this study, TR3 nodules were mainly mixed cystic/ solid (28.5%) or solid, hyper/ isoechoic (71.3%). Mixed cystic/ solid nodules account for one-third to one-half of all US-detected thyroid nodules[14-16] whose malignancy rate varies but is generally low (~5%), especially in predominately cystic nodules[14,15]. Malignant nodules always show an eccentric solid component with moderately or highly suspicious characteristics such as decreased echogenicity, lobulation, or punctate echogenic foci[17-19]. Hyperechogenicity and isoechogenicity suggest benign disease[20]. Rosario et al[21] previously reported a rate of malignancy of only 1.5% for solid, iso- or hyperechoic nodules without suspicious US features, which agreed with the rate of < 3% reported by other studies[22-24]. Therefore, Rosario et al[21] suggested FNA was less necessary in the case of iso- or hyperechoic nodules that did not show suspicious US characteristics, provided the patient was closely followed by US. Although ACR TR3 nodules account for a relatively high proportion, their malignancy rate is quite low. Increasing the FNA recommendation threshold could substantially reduce the biopsy of benign nodules.

ACR TR4 nodules accounted for 30.8% of the total nodules, with a malignancy rate of 4.5%. Nguyen et al[25] analyzed 112,128 patients and concluded the risk of local invasion, nodal metastases, or distant metastases was low for DTC tumors < 4 cm, and there was no size threshold associated with a sharp rise in adverse outcomes. Increasing tumor size did not affect survival until a threshold of 2.5 cm. Furthermore, the dimension of nodules on US has been reported to be larger than their size at gross pathology by 5 mm on average[26,27]. These findings suggest that increased FNA size thresholds may not lead to significantly increased risk of morbidity and mortality.

When AI TI-RADS is applied, due to the simplification, only the features that are important in the differential diagnosis are retained, such as solid nodule composition. A large number of benign nodules will be downgraded, widening the gap between their scores and malignant nodules. The malignancy rate of 6-point nodules in AI TR4 was higher than that of 4-point nodules and 5-point nodules, closer to TR5 instead. Therefore, the sensitivity of the mAI TI-RADS became the same as that of the original ACR TI-RADS after the upgrade of 6-point nodules from AI TR4 to AI TR5, which solved the missed diagnosis problem of the original AI TI-RADS.

This study has several limitations. First, it is a retrospective study and, therefore, selection bias may be inevitable. Also, nodules were selected on the basis of the specific risk stratification system (ACR TI-RADS) or clinically significant issues like compressive symptoms. To minimize this limitation, we conducted a multicenter study involving a large sample. Second, nodules with inconclusive final diagnoses were excluded. It's difficult to assess the malignancy rate and TI-RADS performance among this subgroup. Inclusion criteria for future studies will add a follow-up criterion to study nodules that lack pathological diagnoses but remain stable over time (considered benign). Third, the composite reference standard including FNA cytology and CNB histology used in our study may lead to false-negative and false-positive results.

In conclusion, the mACR TI-RADS and mAI TI-RADS based on FNA thresholds and stratification rules adjustments may significantly improve the specificity and accuracy without sacrificing sensitivity compared with the original ACR TI-RADS. Further validation is required in a larger, prospective, longitudinal study.

Author contributions:

Guarantors of integrity of entire study, Xiaoxian Li, Chuan Peng, Jianhua Zhou; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, Xiaoxian Li, Longhui Cao; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, Yixin Hu, Liang Yang, Yiwen Yu; clinical studies, Ying Liu, Hongyan Zeng, Weijun Huang, Qian Li, Nansheng Tao; experimental studies, Ying Liu, Hongyan Zeng, Weijun Huang, Qian Li, Nansheng Tao; statistical analysis, Xiaoxian Li; and manuscript editing, all authors.

Author disclosure statements:

Xiaoxian Li disclosed no relevant relationships. Chuan Peng disclosed no relevant relationships. Ying Liu disclosed no relevant relationships. Yixin Hu disclosed no relevant relationships. Liang Yang disclosed no relevant relationships. Yiwen Yu disclosed no relevant relationships. Hongyan Zeng disclosed no relevant relationships. Weijun Huang

disclosed no relevant relationships. Qian Li disclosed no relevant relationships. Nansheng Tao disclosed no relevant relationships. Longhui Cao disclosed no relevant relationships. Jianhua Zhou disclosed no relevant relationships. Activities related to the present article: disclosed no relevant relationships.

Any funding information:

None.

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

References

1. Wiltshire JJ, Drake TM, Uttley L, et al. Systematic Review of Trends in the Incidence Rates of Thyroid Cancer. Thyroid 2016;26(11):1541-1552, doi:10.1089/thy.2016.0100

2. Vaccarella S, Franceschi S, Bray F, et al. Worldwide Thyroid-Cancer Epidemic? The Increasing Impact of Overdiagnosis. N Engl J Med 2016;375(7):614-7, doi:10.1056/NEJMp1604412

3. Singh Ospina N, Brito JP, Maraka S, et al. Diagnostic accuracy of ultrasound-guided fine needle aspiration biopsy for thyroid malignancy: systematic review and meta-analysis. Endocrine 2016;53(3):651-61, doi:10.1007/s12020-016-0921-x

4. Durante C, Grani G, Lamartina L, et al. The Diagnosis and Management of Thyroid Nodules: A Review. JAMA 2018;319(9):914-924, doi:10.1001/jama.2018.0898

5. La Vecchia C, Malvezzi M, Bosetti C, et al. Thyroid cancer mortality and incidence: a global overview. Int J Cancer 2015;136(9):2187-95, doi:10.1002/ijc.29251

6. Ha EJ, Na DG, Baek JH, et al. US Fine-Needle Aspiration Biopsy for Thyroid Malignancy: Diagnostic Performance of Seven Society Guidelines Applied to 2000 Thyroid Nodules. Radiology 2018;287(3):893-900, doi:10.1148/radiol.2018171074

7. Grani G, Lamartina L, Ascoli V, et al. Reducing the Number of Unnecessary Thyroid Biopsies While Improving Diagnostic Accuracy: Toward the "Right" TIRADS. J Clin Endocrinol Metab 2019;104(1):95-102, doi:10.1210/jc.2018-01674

8. Ha SM, Baek JH, Na DG, et al. Diagnostic Performance of Practice Guidelines for Thyroid Nodules: Thyroid Nodule Size versus Biopsy Rates. Radiology 2019;291(1):92-99, doi:10.1148/radiol.2019181723

9. Pandya A, Caoili EM, Jawad-Makki F, et al. Retrospective Cohort Study of 1947 Thyroid Nodules: A Comparison of the 2017 American College of Radiology TI-RADS and the 2015 American Thyroid Association Classifications. AJR Am J Roentgenol 2020;214(4):900-906, doi:10.2214/AJR.19.21904

10. Wildman-Tobriner B, Buda M, Hoang JK, et al. Using Artificial Intelligence to Revise ACR TI-RADS Risk Stratification of Thyroid Nodules: Diagnostic Accuracy and Utility. Radiology 2019;292(1):112-119, doi:10.1148/radiol.2019182128

11. Liu Y, Li X, Yan C, et al. Comparison of diagnostic accuracy and utility of artificial intelligence-optimized ACR TI-RADS and original ACR TI-RADS: a multi-center validation study based on 2061 thyroid nodules. Eur Radiol 2022;32(11):7733-7742, doi:10.1007/s00330-022-08827-y

12. Qi TY, Chen X, Liu H, et al. Comparison of thyroid nodule FNA rates recommended by ACR TI-RADS, Kwak TI-RADS and ATA guidelines. Eur J Radiol 2022;148(110152, doi:10.1016/j.ejrad.2022.110152

13. 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):1-133, doi:10.1089/thy.2015.0020

14. Lee MJ, Kim EK, Kwak JY, et al. Partially cystic thyroid nodules on ultrasound: probability of malignancy and sonographic differentiation. Thyroid 2009;19(4):341-6, doi:10.1089/thy.2008.0250

15. 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(9):3411-7, doi:10.1210/jc.2006-0690

16. Bellantone R, Lombardi CP, Raffaelli M, et al. Management of cystic or predominantly cystic thyroid nodules: the role of ultrasound-guided fine-needle aspiration biopsy. Thyroid 2004;14(1):43-7, doi:10.1089/105072504322783830

17. Shi YZ, Jin Y, Zheng L. Partially cystic thyroid nodules on ultrasound: The associated factors for malignancy. Clin Hemorheol Microcirc 2020;74(4):373-381, doi:10.3233/CH-190582

18. Kim DW, Park JS, In HS, et al. Ultrasound-based diagnostic classification for solid and partially cystic thyroid nodules. AJNR Am J Neuroradiol 2012;33(6):1144-9, doi:10.3174/ajnr.A2923

19. Kim DW, Lee EJ, In HS, et al. Sonographic differentiation of partially cystic thyroid nodules: a prospective study. AJNR Am J Neuroradiol 2010;31(10):1961-6, doi:10.3174/ajnr.A2204

20. Moon WJ, Jung SL, Lee JH, et al. Benign and malignant thyroid nodules: US differentiation--multicenter retrospective study. Radiology 2008;247(3):762-70, doi:10.1148/radiol.2473070944

21. Rosario PW, Salles DS, Purisch S. Fine-needle biopsy should be performed in solid hypoechoic thyroid nodules greater than one centimeter even in the absence of suspicious ultrasonographic characteristics. Thyroid 2010;20(8):939-40, doi:10.1089/thy.2010.0026

22. Cappelli C, Castellano M, Pirola I, et al. The predictive value of ultrasound findings in the management of thyroid nodules. QJM 2007;100(1):29-35, doi:10.1093/qjmed/hcl121

23. Papini E, Guglielmi R, Bianchini A, et al. Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound and color-Doppler features. J Clin Endocrinol Metab 2002;87(5):1941-6, doi:10.1210/jcem.87.5.8504

24. Leenhardt L, Hejblum G, Franc B, et al. Indications and limits of ultrasound-guided cytology in the management of nonpalpable thyroid nodules. J Clin Endocrinol Metab 1999;84(1):24-8, doi:10.1210/jcem.84.1.5418

25. Nguyen XV, Roy Choudhury K, Tessler FN, et al. Effect of Tumor Size on Risk of Metastatic Disease and Survival for Thyroid Cancer: Implications for Biopsy Guidelines. Thyroid 2018;28(3):295-300, doi:10.1089/thy.2017.0526

26. Bachar G, Buda I, Cohen M, et al. Size discrepancy between sonographic and pathological evaluation of solitary papillary thyroid carcinoma. Eur J Radiol 2013;82(11):1899-903, doi:10.1016/j.ejrad.2013.07.002

27. Bilginer MC, Ozdemir D, Baser H, et al. Is ultrasonographically detected nodule diameter concordant with pathological tumor size? Int J Surg 2017;42(95-102, doi:10.1016/j.ijsu.2017.04.054

Thyroid
ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

Table 1. New Versions of ACR TI-RADS and AI TI-RADS with FNA Size Threshold Adjustment

System	Original Version	Version 1	Version 2	Version 3	Version 4	Version 5
ACR TI-RADS						
TR3 Mildly Suspicious	≥ 2.5 cm	≥ 2.5 cm	No FNA†	No FNA†	No FNA†	No FNA†
TR4 Moderately Suspicious	≥ 1.5 cm	≥ 2.0 cm†	≥ 1.5 cm	≥ 2.0 cm†	≥ 2.5 cm†	≥ 3.0 cm†
TR5 Highly Suspicious	≥ 1.0 cm	≥ 1.0 cm	≥ 1.0 cm	≥ 1.0 cm	≥ 1.0 cm	≥ 1.0 cm
AI TI-RADS						
TR3 Mildly Suspicious	≥ 2.5 cm	≥ 2.5 cm	No FNA†	No FNA†	No FNA†	No FNA†
TR4 Moderately Suspicious	≥ 1.5 cm	≥ 2.0 cm†	≥ 1.5 cm	≥ 2.0 cm†	≥ 2.5 cm†	≥ 3.0 cm†
TR5 Highly Suspicious	≥ 1.0 cm	≥ 1.0 cm	≥ 1.0 cm	≥ 1.0 cm	≥ 1.0 cm	≥ 1.0 cm
TR4-adjusted AI TI-RADS*						
TR3 Mildly Suspicious	≥ 2.5 cm	≥ 2.5 cm	No FNA†	No FNA†	No FNA†	No FNA†
TR4 Moderately Suspicious	≥ 1.5 cm	≥ 2.0 cm†	≥ 1.5 cm	≥ 2.0 cm†	≥ 2.5 cm†	≥ 3.0 cm†
TR5 Highly Suspicious	≥ 1.0 cm	≥ 1.0 cm	≥ 1.0 cm	≥ 1.0 cm	≥ 1.0 cm	≥ 1.0 cm

Note.—ACR = American College of Radiology, TI-RADS = Thyroid Imaging Reporting and Data System, AI = artificial intelligence, FNA = fine-needle aspiration.

* Refers to 6-point nodules being upgraded from AI TR4 to AI TR5, while other rules were the same as the original AI TI-RADS.

† Indicates FNA size threshold adjustment.

Thyroid
R TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

Table 2. Patient Demographics and Nodule Characteristics

Parameter	Total	Benign	Malignant	*P* Value
No. of patients	2511	2255	256	
Sex				< 0.001
Female	1914 (76.2)	1744 (77.3)	170 (66.4)	
Male	597 (23.8)	511 (22.7)	86 (33.6)	
Age, y*	46 ± 13	47 ± 13	43 ± 13	0.001
No. of nodules	2596	2336	260	
Mean nodule size, cm*	2.5 ± 1.3	2.6 ± 1.3	1.9 ± 1.2	< 0.001
Nodule location				0.114
Left	1171 (45.1)	1049 (44.9)	122 (46.9)	
Right	1360 (52.4)	1233 (52.8)	127 (48.8)	
Isthmus	65 (2.5)	54 (2.3)	11 (4.2)	
Composition				< 0.001
Cystic or almost completely cystic	95 (3.7)	95 (4.1)	0	
Spongiform	5 (0.2)	5 (0.2)	0	
Mixed cystic and solid	882 (34.0)	873 (37.4)	9 (3.5)	
Solid or almost completely solid	1597 (61.5)	1347 (57.7)	250 (96.2)	

Can’t classify	17 (0.7)	16 (0.7)	1 (0.4)	
Echogenicity				< 0.001
Anechoic	95 (3.7)	95 (4.1)	0	
Hyperechoic	84 (3.2)	83 (3.6)	1 (0.4)	
Isoechoic	1159 (44.6)	1141 (48.8)	18 (6.9)	
Hypoechoic	1162 (44.8)	952 (40.8)	210 (80.8)	
Very hypoechoic	79 (3.0)	49 (2.1)	30 (11.5)	
Can’t classify	17 (0.7)	16 (0.7)	1 (0.4)	
Shape				< 0.001
Wider-than-tall	2117 (81.5)	2015 (86.3)	102 (39.2)	
Taller-than-wide	479 (18.5)	321 (13.7)	158 (60.8)	
Margin				< 0.001
Smooth	1867 (71.9)	1824 (78.1)	43 (16.5)	
Ill-defined	505 (19.5)	417 (17.9)	88 (33.8)	
Irregular and/or lobulated	204 (7.9)	90 (3.9)	114 (43.8)	
Extrathyroidal extension	19 (0.7)	4 (0.2)	15 (5.8)	
Can’t classify	1 (0.0)	1 (0.0)	0	

Thyroid
R TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

Echogenic foci[†]			< 0.001
No echogenic foci	1815 (69.9)	1741 (74.5)	74 (28.5)
Large comet-tail artifacts	45 (1.7)	45 (1.9)	0
Macrocalcifications	318 (12.2)	275 (11.8)	53 (20.4)
Peripheral calcifications	48 (1.8)	44 (1.9)	4 (1.5)
Punctate echogenic foci	446 (17.2)	288 (12.3)	158 (60.8)

Note.—Unless otherwise specified, data are reported as number of nodules, with percentages in parentheses.

* Data are reported as mean ± standard deviation for continuous variables.

[†] Nodules could have more than one type of echogenic focus.

Thyroid

R TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

Table 3. Comparison of Diagnostic Performance among four TI-RADS versions

	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)	UFR (%)	AUC
ACR TI-RADS	93 (89, 96） [241/260]	46 (44, 48) [1083/2336]	51 (49, 53) [1324/2596]	16 (14, 18) [241/1494]	98 (97, 99) [1083/1102]	48 (46, 50) [1253/2596]	0.70 (0.67, 0.72)
AI TI-RADS	85 (81, 89） [222/260]	56 (54, 58) [1317/2336]	59 (57-61) [1539/2596]	18 (16, 20) [222/1241]	97 (96, 98) [1317/1355]	39 (37, 41) [1019/2596]	0.71 (0.68, 0.74)
P1	0.01*	< 0.001*	< 0.001*	0.21	0.07	< 0.001*	0.16
mACR TI-RADS	87 (83, 91） [227/260]	73 (71, 75) [1700/2336]	74 (73, 76) [1927/2596]	26 (23, 29) [227/863]	98 (97, 99) [1700/1733]	25 (23, 26) [636/2596]	0.80 (0.77, 0.83)
P1	0.057	< 0.001*	< 0.001*	< 0.001*	0.70	< 0.001*	< 0.001*
P2	0.61	< 0.001*	< 0.001*	< 0.001*	0.01	< 0.001*	< 0.001*
mAI TI-RADS	89 (84, 92） [230/260]	73 (71, 75) [1708/2336]	75 (73, 76) [1938/2596]	27 (24, 30) [230/858]	98 (97, 99) [1708/1738]	24 (23, 26) [628/2596]	0.81 (0.78, 0.83)
P1	0.13	< 0.001*	< 0.001*	< 0.001*	> 0.999	< 0.001*	< 0.001*

P2	0.36	< 0.001*	< 0.001*	< 0.001*	0.04*	< 0.001*	< 0.001*
P3	0.79	0.82	0.75	0.81	0.66	0.80	0.24

Note. — Data in parentheses are 95% CIs, with numerators and denominators in brackets. ACR = American College of Radiology, TI-RADS = Thyroid Imaging Reporting and Data System, AI = artificial intelligence, mACR TI-RADS = modified ACR TI-RADS, mAI TI-RADS = modified AI TI-RADS, PPV = positive predictive value, NPV = negative predictive value, UFR = unnecessary FNA rate, AUC = area under the receiver operating characteristic curve.

P1 represents the comparison with ACR TI-RADS. *P2* represents the comparison with AI TI-RADS. *P3* represents the comparison with mACR TI-RADS.

* $P < 0.05$ indicated statistically significant differences.

Thyroid
R TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)
d and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Thyroid
R TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

Table 4. Comparison of Risk Stratification and Indication of FNA among four TI-RADS versions

Systems	Total nodules (n=2596)*	Benign nodules (n=2336)†	Malignant Nodules (n=260)‡	Suggested risk of malignancy (%)	Calculated risk of malignancy (%)	No. of nodules indicated for FNA	No. of benign nodules among nodules indicated for FNA§	No. of malignant nodules among nodules indicated for FNA※
ACR TI-RADS						**1494**	**1253 (83.9)**	**241 (16.1)**
TR1 Benign	97 (3.7)	97 (4.2)	0 (0.0)	< 2	0	0	0	0
TR2 Not	571 (22.0)	570 (24.4)	1 (0.4)	< 2	0.2	0	0	0
TR3 Mildly	554 (21.3)	546 (23.4)	8 (3.1)	5	1.4	355	350 (98.6)	5 (1.4)
TR4 Moderately	799 (30.8)	763 (32.7)	36 (13.8)	5-20	4.5	564	543 (96.3)	21 (3.7)
TR5 Highly	575 (22.1)	360 (15.4)	215 (82.7)	> 20	37.4	575	360 (62.6)	215 (37.4)
mACR TI-RADS						**863**	**636 (73.7)**	**227 (26.3)**
TR1 Benign	97 (3.7)	97 (4.2)	0 (0.0)	< 2	0	0	0	0
TR2 Not	571 (22.0)	570 (24.4)	1 (0.4)	< 2	0.2	0	0	0
TR3 Mildly	554 (21.3)	546 (23.4)	8 (3.1)	5	1.4	0	0	0
TR4 Moderately	799 (30.8)	763 (32.7)	36 (13.8)	5-20	4.5	288	276 (95.8)	12 (4.2)
TR5 Highly	575 (22.1)	360 (15.4)	215 (82.7)	> 20	37.4	575	360 (62.6)	215 (37.4)

AI TI-RADS						**1241**	**1019 (82.1)**	**222 (17.9)**
TR1 Benign	747 (28.8)	744 (31.8)	3 (1.2)		0.4	0	0	0
TR2 Not	136 (5.2)	135 (5.8)	1 (0.4)		0.7	0	0	0
TR3 Mildly	495 (19.1)	486 (20.8)	9 (3.5)	NA	1.8	295	289 (98.0)	6 (2.0)
TR4 Moderately	771 (29.7)	716 (30.7)	55 (21.2)		7.1	499	475 (95.2)	24 (4.8)
TR5 Highly	447 (17.2)	255 (10.9)	192 (73.8)		43.0	447	255 (57.0)	192 (43.0)
mAI TI-RADS						**858**	**628 (73.2)**	**230 (26.8)**
TR1 Benign	747 (28.8)	744 (31.8)	3 (1.2)		0.4	0	0	0
TR2 Not	136 (5.2)	135 (5.8)	1 (0.4)		0.7	0	0	0
TR3 Mildly	495 (19.1)	486 (20.8)	9 (3.5)	NA	1.8	0	0	0
TR4 Moderately	576 (22.2)	550 (23.5)	26 (10.0)		4.5	216	207 (95.8)	9 (4.2)
TR5 Highly	642 (24.7)	421 (18.0)	221 (85.0)		34.4	642	421 (65.6)	221 (34.4)

Note.—Unless otherwise specified, data are reported as number of nodules, with percentages in parentheses. ACR = American College of Radiology, TI-RADS = Thyroid Imaging Reporting and Data System, AI = artificial intelligence, mACR TI-RADS = modified ACR TI-RADS, mAI TI-RADS = modified AI TI-RADS, FNA= fine-needle aspiration.

* Numbers in parentheses represent percentage of total nodules of each risk level (TR1 to TR5) in total 2596 nodules.

† Numbers in parentheses represent percentage of benign nodules of each risk level (TR1 to TR5) in 2336 benign nodules.

‡ Numbers in parentheses represent percentage of malignant nodules of each risk level (TR1 to TR5) in 260 malignant nodules.

Thyroid
R TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)
d and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

§ Numbers in parentheses represent percentage of benign nodules among nodules indicated for FNA within each risk level (TR1 to TR5).

※ Numbers in parentheses represent percentage of malignant nodules among nodules indicated for FNA within each risk level (TR1 to TR5).

NA: Suggested risk of malignancy was not provided according to Benjamin's research[1].

1. Wildman-Tobriner B, Buda M, Hoang JK, et al. Using Artificial Intelligence to Revise ACR TI-RADS Risk Stratification of Thyroid Nodules: Diagnostic Accuracy and Utility. Radiology 2019;292(1):112-119, doi:10.1148/radiol.2019182128

Thyroid
R TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

Figure Legends

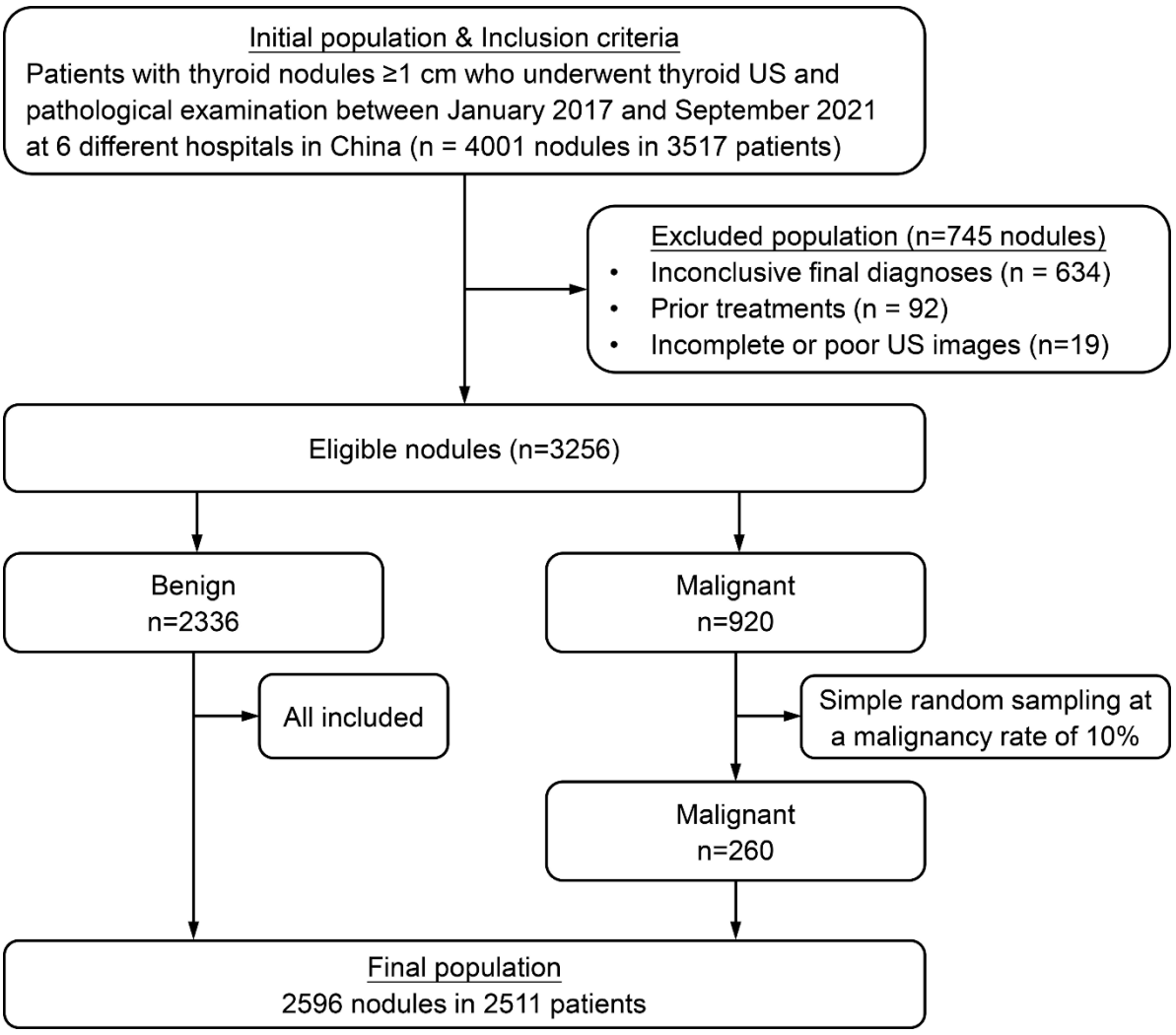

Figure 1. Flowchart of the included patients and number of thyroid nodules. US = ultrasound, n = number of thyroid nodules.

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

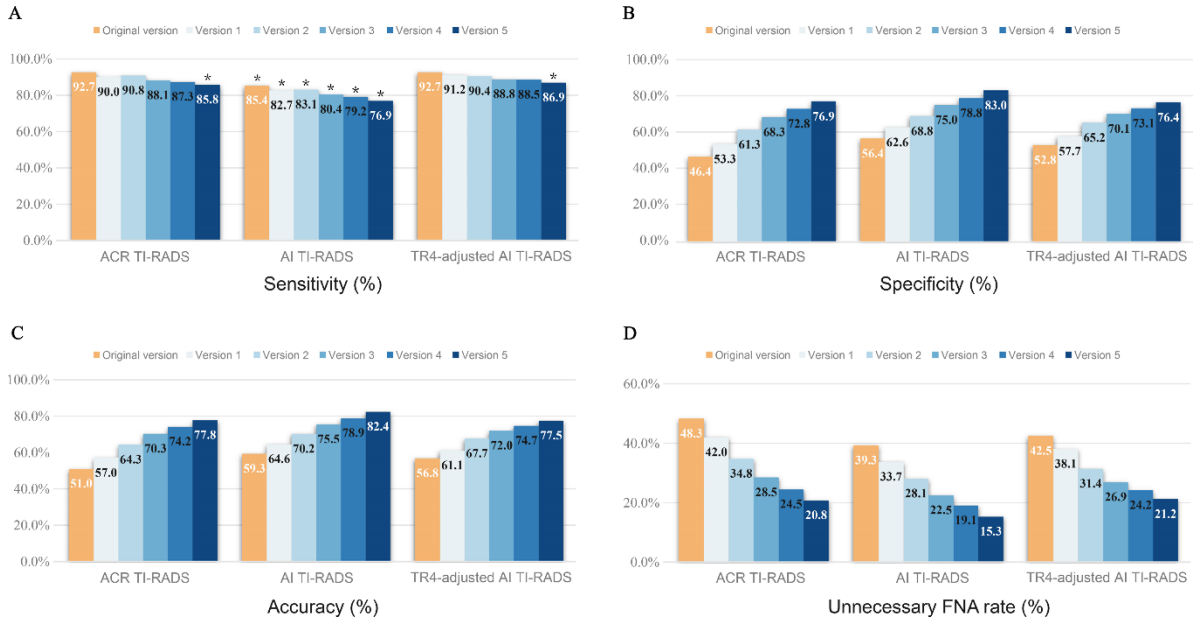

Figure 2. Diagnostic performance and unnecessary FNA rate of new versions of ACR TI-RADS and AI TI-RADS with adjustment (defined in Table 1). TR4-adjusted AI TI-RADS refers to 6-point nodules being upgraded from AI TR4 to AI TR5, while other rules were the same as the original AI TI-RADS. (A) Graph shows the sensitivity gradually decreased as size thresholds were raised. * $P < 0.05$ for the comparison between each new version and the original ACR TI-RADS. (B, C) Graphs show the specificity and accuracy gradually increased with significant difference in all new versions as size thresholds were raised, compared with the original ACR TI-RADS (no specific label). (D) Graph shows the UFR gradually decreased with significant difference in all new versions as size thresholds were raised, compared with the original ACR TI-RADS (no specific label).

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

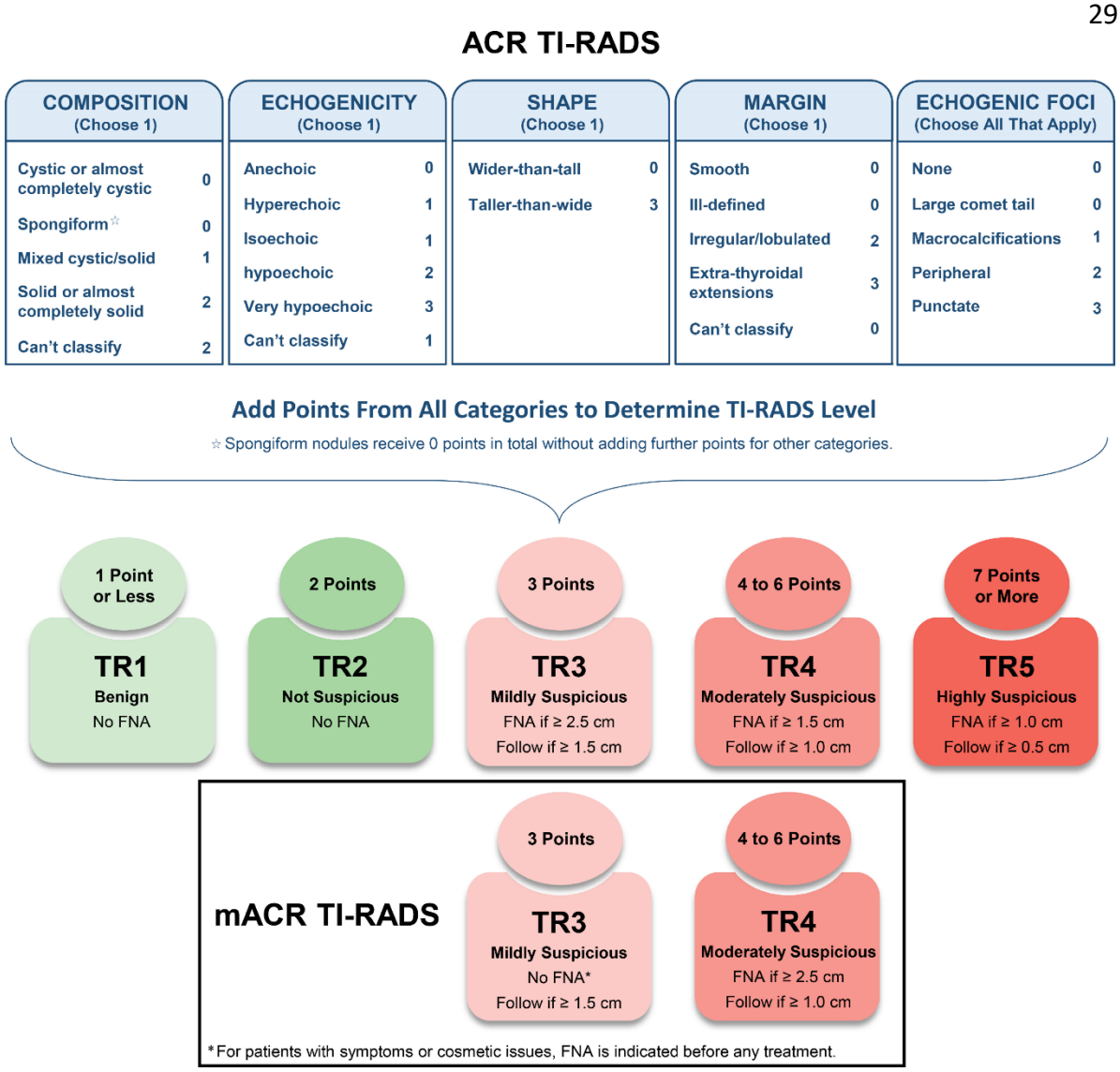

Figure 3. Chart shows comparison of ACR TI-RADS and mACR TI-RADS scheme, including nodule size threshold adjustments in TR3 and TR4. ACR = American College of Radiology, TI-RADS = Thyroid Imaging Reporting and Data System, mACR TI-RADS = modified ACR TI-RADS, FNA = fine-needle aspiration, TR = TI-RADS category.

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

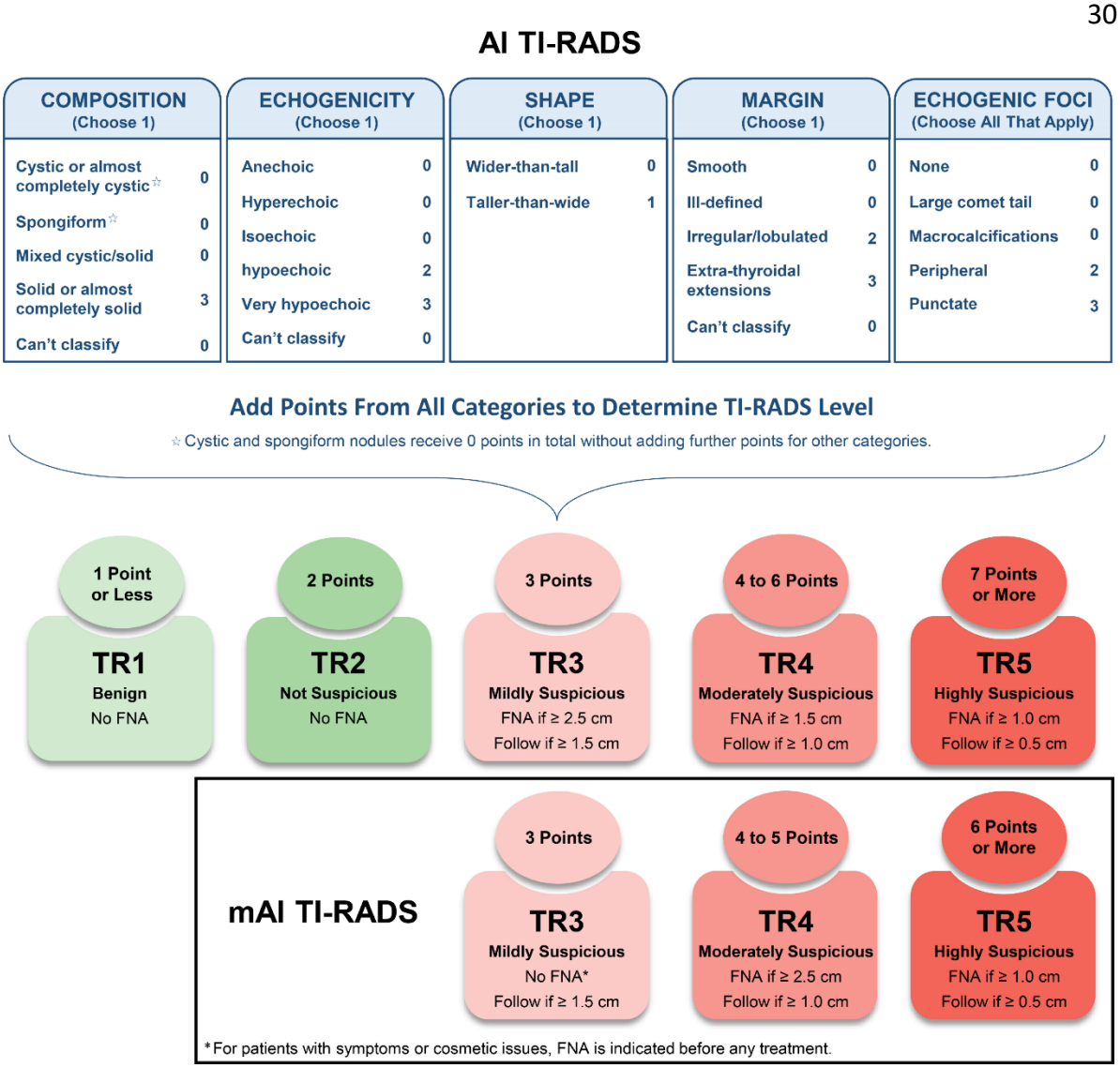

Figure 4. Chart shows comparison of AI TI-RADS and mAI TI-RADS scheme, including nodule size threshold and stratification criteria adjustments. AI = artificial intelligence, TI-RADS = Thyroid Imaging Reporting and Data System, mAI TI-RADS = modified AI TI-RADS, FNA = fine-needle aspiration, TR = TI-RADS category.

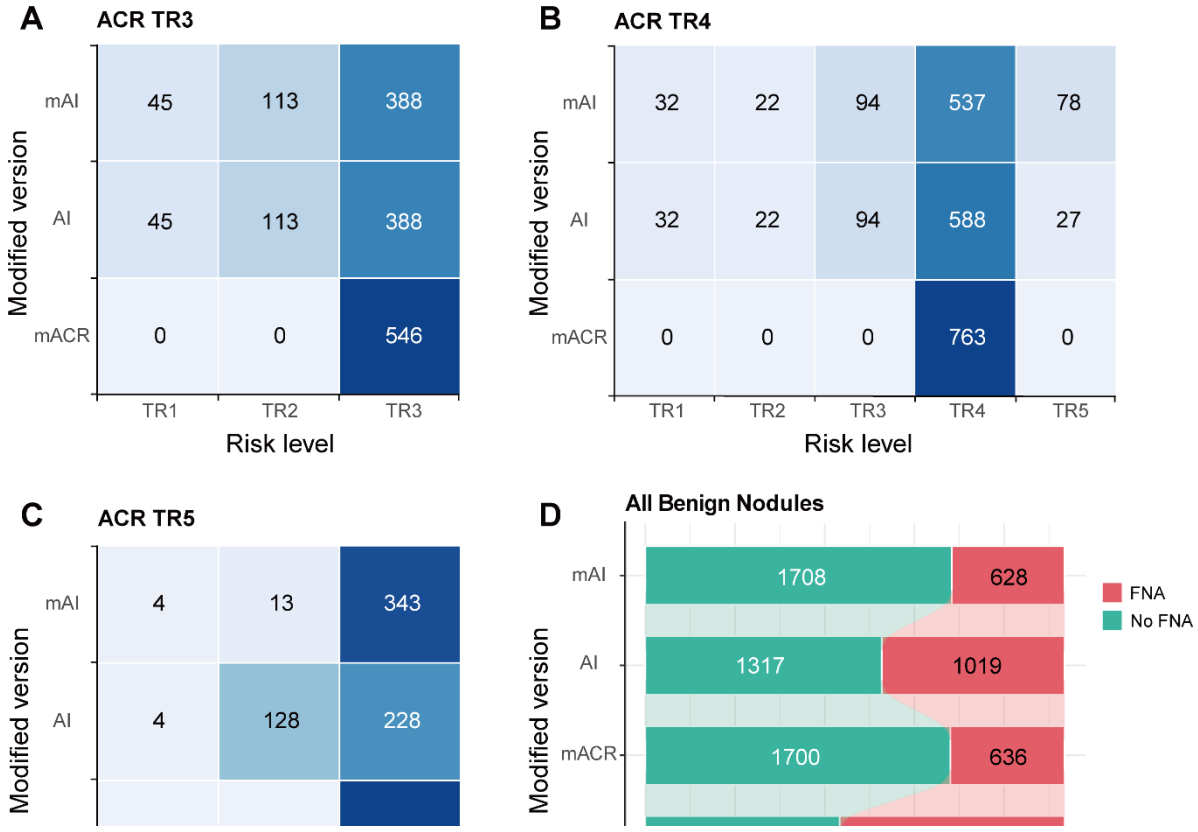

Figure 5. The application value of modified versions in down-grading risk level (A, B, C) and reducing unnecessary FNA (D) of benign nodules. ACR = American College of Radiology, mACR = modified ACR TI-RADS, AI = artificial intelligence, mAI= modified AI TI-RADS, FNA = fine-needle aspiration, TR = TI-RADS category.

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)

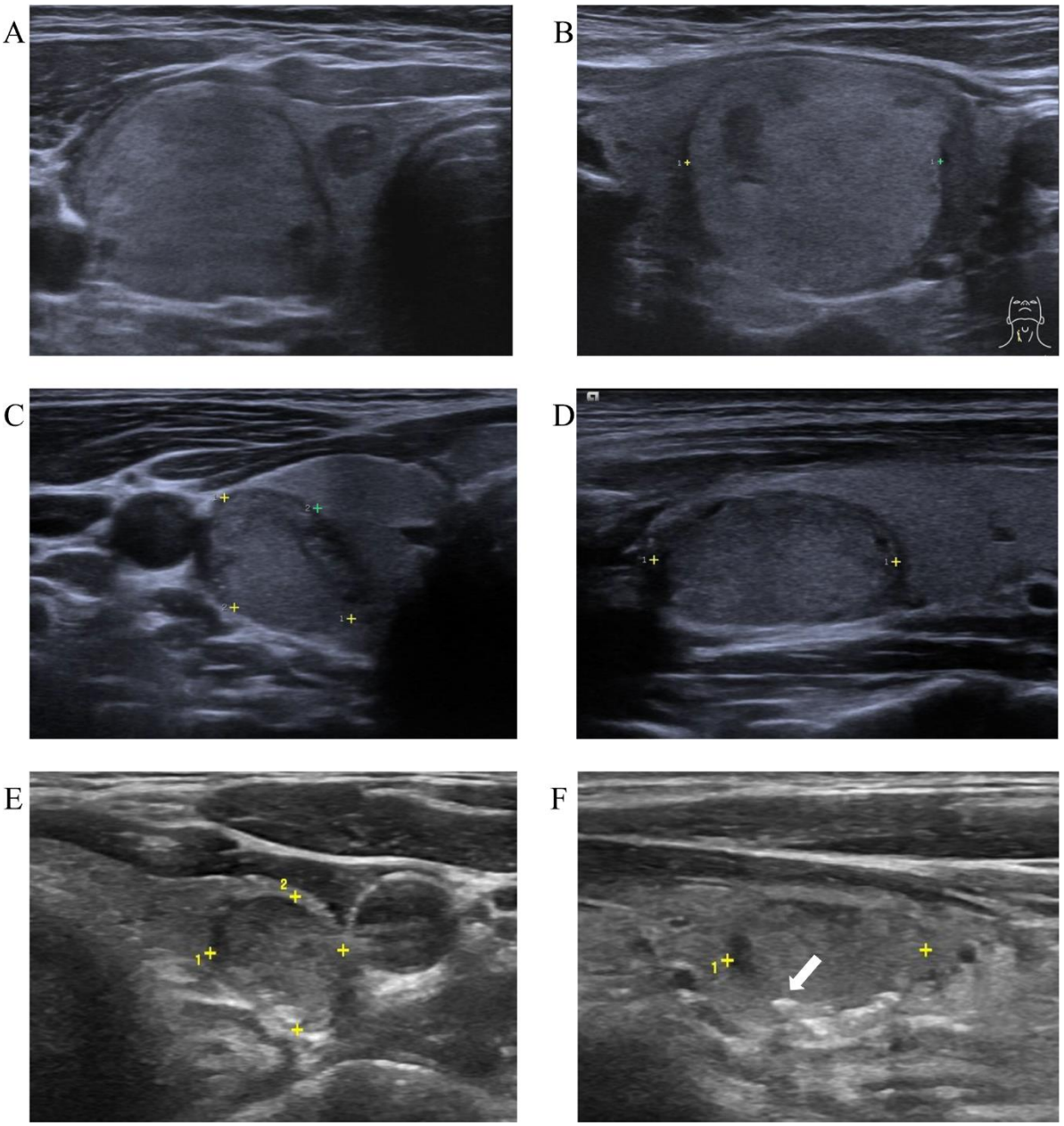

Figure 6. US image of three thyroid nodules. (A, B) Transverse and longitudinal gray-scale US images in a 40-year-old male patient show a 2.7-cm solid, isoechoic, wider-than-tall and smooth thyroid nodule. This nodule scored three points according to ACR TI-RADS or AI TI-RADS, with risk level of TR3 and recommendation for FNA. The same nodule would not be recommended for FNA by mACR TI-RADS or mAI TI-RADS. FNA result suggested that the nodule was benign. (C, D) Transverse and longitudinal gray-scale US images in a 27-year-old male patient show a 2.1-cm solid, hypoechoic, wider-than-tall and smooth thyroid nodule. This nodule scored four points according to ACR TI-RADS and five points according to AI TI-RADS, with risk level of TR4 and recommendation for FNA. The same nodule would

not be recommended for FNA by mACR TI-RADS or mAI TI-RADS. FNA result suggested that the nodule was benign. (E, F) Transverse and longitudinal gray-scale US images in a 50-year-old male patient show a 1.3-cm solid, hypoechoic, taller-than-wide and smooth thyroid nodule with macrocalcification (arrow). This nodule scored eight points according to ACR TI-RADS, with risk level of TR5 and recommendation for FNA. The same nodule was assigned 6 points by AI TI-RADS, with risk level of TR4 and no recommendation for FNA. The same nodule was upgraded to TR5 and recommended for FNA by mAI TI-RADS. Pathologic finding at FNA was papillary carcinoma.

Thyroid
Modified ACR TI-RADS and Modified AI TI-RADS for Thyroid Nodules: A Multicenter Retrospective Study (DOI: 10.1089/thy.2023.0429)