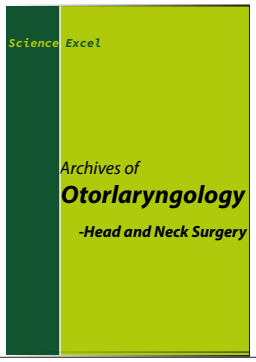
Archives of Otolaryngology-Head and Neck Surgery

Original



Comparison of ACR-TIRADS to the ATA Guidelines for Thyroid Nodules: A Neck to Neck Comparison

Judy K Qiang[1], Carrie Betel[2], Kalesha Hack[2], Zeina Ghorab[3], Julie Gilmour[1,4], Manijeh Mohammadi[5], Kirsteen Burton[2], Kevin M Higgins[6], Ilana Halperin[1,7]

[1]*Department of Medicine, Division of Endocrinology and Metabolism, University of Toronto; Toronto, Ontario, Canada*
[2]*Department of Medical Imaging, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada*
[3]*Department of Pathology, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada*
[4]*Department of Medicine, St. Michael's Hospital, Toronto, Ontario, Canada*
[5]*Department of Medicine, Division of Endocrinology and Metabolism, McMaster University, Hamilton, Ontario, Canada*
[6]*Department of Surgery, Division of Otolaryngology, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada*
[7]*Department of Medicine Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada*

Correspondence

Ilana Halperin

Sunnybrook Health Sciences Centre Room H121, 2075 Bayview Avenue, Toronto, ON M4N 3N5, Canada

E-mail: ilana.halperin@sunnybrook.ca

Phone: (416) 480-6056

- Received Date: 16 Nov 2021
- Accepted Date: 27 Nov 2021
- Publication Date: 02 Jan 2022

Keywords

Thyroid nodules, cancer risk, ultrasound, ACR-TIRADS, ATA

Abstract

Introduction: The goal of this study was to compare the performance characteristics of the American College of Radiology Thyroid Imaging Reporting and Data System (ACR-TIRADS) and the American Thyroid Association (ATA) systems in identifying malignant thyroid nodules.

Methods: In a retrospective chart review, ultrasound images of all thyroid nodules biopsied in 2014-2015 at a Canadian academic centre were reviewed by two radiologists. The ultrasound characteristics of thyroid nodules were compared with cytologic or pathologic results to determine the positive predictive value (PPV), negative predictive value (NPV), sensitivity and specificity for TIRADS and ATA in predicting cancer risk. Clinical course of nodules not requiring follow up or intervention according to ACR-TIRADS was described. Vascularity was added to ACR-TIRADS to determine whether sensitivity of TIRADS improves.

Results: A total of 417 thyroid nodules were reviewed, 82% were benign (Bethesda II). The sensitivity, specificity, PPV, and NPV were 97%, 11%, 9%, 98%, and 70%, 29%, 18%, and 81% for ATA and TIRADS, respectively. Of the 10 nodules that did not need ultrasound follow up based on TIRADS criteria, 2 were malignant, the rest were FLUS. If vascularity was added to TIRADS (TIRADS-Vasc), the number of malignant cases missed could have been reduced by 43% (from 7 to 4 cases).

Conclusions: TIRADS is more specific but less sensitive than ATA, and misses a small number of malignant nodules. Clinicians need to use their judgement to decide which nodules require biopsy as some malignant cases will be missed using TIRADS alone.

Introduction

Thyroid nodules are common in the general population [1-4]. Most clinically encountered nodules are discovered incidentally on ultrasound [5,6]. Given that the majority of these thyroid incidentalomas are benign [7], it is important to avoid unnecessary interventions that offer minimal benefit and may cause potential harm.

In order to risk stratify thyroid nodules, several tools have been developed. In 2015, the American Thyroid Association (ATA) published guidelines outlining high risk ultrasound features [8]. In 2017, the American College of Radiology devised a point based system (ACR-TIRADS) in which higher scores are assigned for the presence of high risk ultrasound features [9].

Previous studies have compared the efficacy of the ATA and the ACR-TIRADS systems in predicting thyroid nodule malignancy risk, and have found that ACR-TIRADS is more specific but may be less sensitive than ATA [10-15]. This has raised the concern that some thyroid cancer cases may be missed by ACR-TIRADS. However, TIRADS does recommend ongoing surveillance ultrasound for nodules not meeting size threshold for FNA, which may reduce the number of malignant nodules lost to follow up [9]. Vascularity on ultrasound is another potential marker of malignancy. This has been reported to correlate with malignancy risk by some [16-18], but not all studies [19].

To our knowledge, the diagnostic accuracy of the ATA and ACR-TIRADS system have not been compared using real world data in a Canadian context. Furthermore, the outcomes of nodules not requiring follow up or intervention as per ACR-TIRADS have not been thoroughly elucidated in the literature. We aimed to to compare the performance characteristics of the ATA and the ACR-TIRADS scoring systems and to determine

Citation: Qiang JK, Betel C, Hack K, et al. Comparison of ACR-TIRADS to the ATA Guidelines for Thyroid Nodules: A Neck to Neck Comparison. Arch Otolaryngol Head Neck Surg 2022; 1(1):1-8.

if unnecessary FNAs could be minimized without missing non-benign or malignant thyroid nodules. We describe the the percentage of malignant nodules missed by TIRADS and their clinical outcomes. We also explore the utility of adding vascularity to TIRADS (TIRADS-Vasc) to determine if the sensitivity of ACR-TIRADS could be improved.

Materials and methods

This was a retrospective cohort study. We generated a database of all biopsied and/or surgically resected thyroid nodules between Janaury 2014 and December 2015 at a large academic hospital in Toronto, Ontario, Canada. Nodules were selected for biopsy based on physician discretion.

Patient and nodule characteristics are described. For the continuous variables, we determined the mean and standard deviations for each group and we compared between group differences using independent sample t-tests. Univariable logistic regression was used to determine the likelihood of having a malignant nodule comparing men and women.

Two radiologists (CB and KB) blinded to the original radiologic report and histologic/cytologic data independently reviewed the ultrasonographic features of these thyroid nodules in a retrospective fashion. The first 35 nodules were reviewed in duplicate, and inter-rater agreement was determined. Subsequent nodules were split between the two radiologists. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the ATA system were determined. The radiologists recommended the need for FNA based on the 2015 ATA guidelines. We used the histology report instead of the cytologic result as the "gold standard" diagnostic test when it was available. In cases where there were multiple nodules in the same thyroid gland, each nodule was examined independently.

Nodules were divided into three groups, which included benign (Bethesda category II), malignant or suspicious for malignancy (Bethesda V and VI), and non-benign (atypia of undetermined significance (AUS)/follicular lesion of undetermined significance (FLUS), Bethesda III, and follicular neoplasm/suspicious for follicular neoplasm, Bethesda IV) [20]. Nodules with a non-diagnostic or unsatisfactory cytology (Bethesda category I) were excluded from our database.

Author JQ (final year Endocrinology resident physician) assigned an ACR-TIRADS score and risk category to each nodule based on the chracteristics reported by the radiologists (Supplementary Table 1), and decided whether a nodule required biopsy according to malignancy risk and size of lesion. We then compared these decisions with the gold standard tests (either cytologic or histologic data, where available) to determine the sensitivity, specificity, PPV, and NPV of ACR-TIRADS. The number of avoidable FNAs, as well as the number of non-benign and malignant cases that were missed by each system were reported in percentages.

We then explored the utility of adding vascularity to determine if we could improve the diagnostic accuracy of the ACR-TIRADS system (TIRADS-Vasc) as vascularity has been reported to be possibly predictive of malignancy risk in some studies. Vascularity was rated as a binary value of present or absent based on retrospective review of still ultrasound images using color Doppler flow and rarely with pulsed Doppler flow. We assigned vascular nodules one TIRADS category higher than their original risk category to assess the number of missed non-benign and malignant cases that could be reduced, using descriptive statistics.

We also calculated both inter and intra-rater reliability for reporting vascularity. A third radiologist KH and author CB, independently re-analyzed 25 randomly selected thyroid nodules, including those where TIRADS-Vasc changed initial management, with emphasis on nodular vascularity. We compared the inter-rater agreement between KH and CB as well as intra-rater agreement for author CB overtime by calculating Cohen's kappa.

Follow up data on the malignant or non-benign thyroid nodules that were classified as not requiring FNA initially using ACR-TIRADS criteria were obtained from the hospital electronic medical records system to determine the proportion of cases that had subsequent radiologic follow up. Ouctomes of nodules not requiring follow up as per ACR-TIRADS are determined by chart review.

Analyses were performed using Microsoft Excel version 15.38 and SAS University Edition. This study was approved by the Sunnybrook Hospital research ethics board. All patient data were de-identified and stored in an encrypted fashion.

Results

Performance characteristics of ATA and ACR-TIRADS

A total of 438 thyroid nodules underwent biopsy and/or surgical resection at our institution between January, 2014 and December, 2015. Approximately 5% of the nodules were excluded due to one of the following reasons: 1. unsatisfactory/non-diagnostic cytology and no surgical histology, 2. incomplete ultrasound report precluding us from assigning a ACR-TIRADS category for the nodule, or 3. nodule was a lymphoma. All nodules could be classified under ATA. Eight nodules were missing elements of ultrasound features to determine an ACR-TIRADS score. Of the 417 thyroid nodules included for analysis, most were benign (Figure 1).

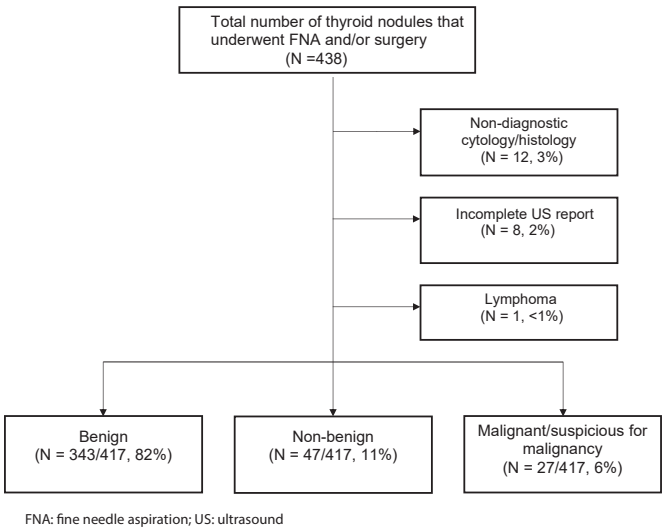


Figure 1. *Inclusion, exclusion criteria, and number of benign, non-benign, and malignant thyroid nodules*

Table 1. *Diagnostic performance characteristics of the ATA and ACR-TIRADS scoring systems for thyroid nodules*

	ATA	ACR-TIRADS
Sensitivity	97%	70%
Specificity	11%	29%
Positive predictive value	9%	18%
Negative predictive value	98%	81%

A total of 334 (80%) nodules occurred in females and 83 (20%) occurred in males. Males were not significantly more likely to have a malignant or non-benign nodule as compared to females in our cohort (odds ratio [OR] 1.41, 95% confidence interval [CI] 0.78-2.57). The overall mean age of patients was 58.7 ± 14 years. Those with a malignant or non-benign nodule were younger (mean age 55.6 ± 13.1 years) as compared to those with a benign nodule (mean age 59.4 ± 14.1 years, p = 0.03). Mean size of the largest dimension of thyroid nodule was 2.2 ± 1.1 cm in benign nodules and 2.0 ± 1.2 cm in malignant or non-benign nodules (p = 0.16).

The sensitivity, specificity, PPV, and NPV of the ATA and ACR-TIRADS scoring systems are listed in Table 1. While ATA had superior sensitivity and NPV, ACR-TIRADS had better specificity and PPV.

Agreement between the two radiologists regarding the need for FNA based on ATA criteria was moderate (kappa 0.66).

ACR-TIRADS reduced unnecessary investigations as compared to ATA however TIRADS missed more cancerous and non-benign nodules

Of the 343 benign FNAs, 98 (28%) could have been avoided if TIRADS criteria were applied compared to 37 (11%) based on ATA criteria. Of the 306 FNAs recommended based on ATA criteria, 64 (21%) would have been omitted if applying TIRADS criteria instead (Figure 2). Three cytologically/histologically benign nodules (<1%) would have been biopsied if using TIRADS that otherwise would not have been recommended for biopsy if using ATA.

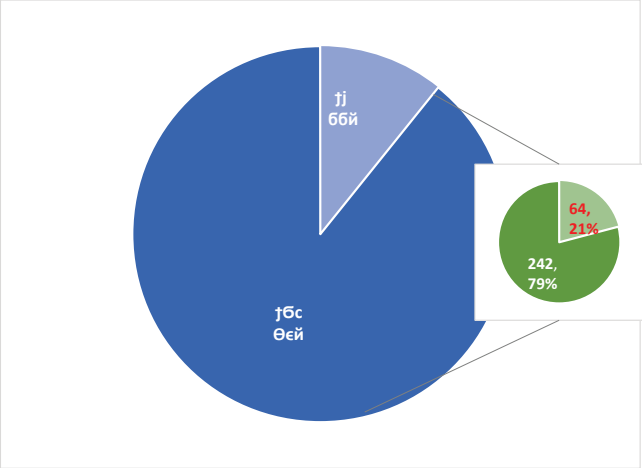
A total of 8/47 (17%) non-benign and 0/27 (0%) cancerous thyroid nodules would have been missed if using ATA whereas 15/47 (32%) non-benign and 7/27 (26%) of malignant thyroid nodules would have been missed initially if applying ACR-TIRADS. One case of FLUS would have been captured by TIRADS but would not have been captured if using ATA.

Of the 22 initially undetected malignant (n = 7) and non-benign cases (n = 15) by ACR-TIRADS, 3 were TIRADS risk category II, 6 were TIRADS III, and 13 were TIRADS IV. About half of these nodules (12/22, 55%) qualified for ultrasound surveillance according to TIRADS. Of the 10 nodules that did not need ultrasound follow up based on TIRADS criteria, only 2 were malignant, the rest were FLUS. Comparing these recommendations with our actual follow up data, we found that the 2 cases of papillary thyroid carcinoma (PTC) were clinically and radiologically monitored and eventually resected. Radioactive iodine (RAI) was not required in either case, indicating that these were low risk thyroid cancer associated with a favourable prognosis. No thyroid cancer specific death was observed among the 10 nodules not requiring follow up as per ACR-TIRADS.

Adding vascularity to TIRADS (TIRADS-Vasc) improved the sensitivity of TIRADS in predicting malignancy risk

If vascularity was added to TIRADS (TIRADS-Vasc), the number of malignant cases missed could have been reduced by 43% (from 7 to 4 cases). In the cases where TIRADS-Vasc made a difference in clinical management, the nodules were all confirmed to be vascular in the second round of reporting. However, ATA still demonstrated superior sensitivity in cancer detection compared to both TIRADS and TIRADS-Vasc (Figure 3).

The inter-rater reliability in reporting vascularity comparing CB to KH was fair (kappa 0.30). The intra-rater reliability for author CB overtime was also fair (kappa 0.32).



The blue pie chart represents the total number of benign FNAs in the study. The light blue pie represents the number of benign FNAs that could be avoided if using ATA only. The dark blue pie shows the number of benign FNAs deemed necessary by ATA. The light green pie shows out of the FNAs deemed necessary by ATA, the number of aspirations that could be omitted if TIRADS criteria were applied instead. The dark green pie shows the number of cytologically/histologically benign nodules that would have been recommended for FNA by both TIRADS and ATA.

Figure 2. *A total of 21% of FNAs recommended by ATA may be avoided if using TIRADS criteria instead*

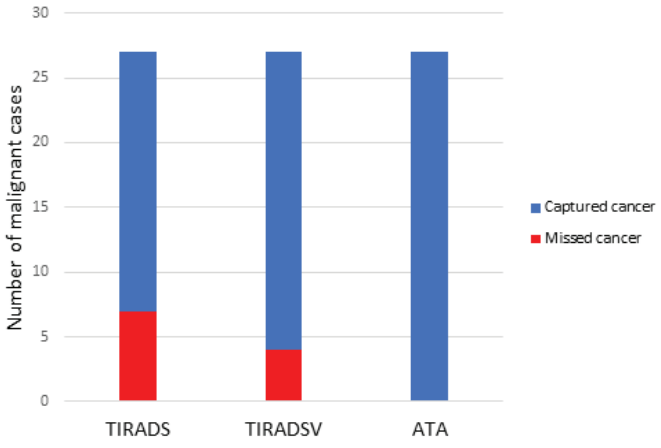


Figure 3. *TIRADS-Vasc reduced the number of malignant cases missed by TIRADS alone but ATA still showed greater sensitivity in cancer detection*

Discussion

In this study, we compared the diagnostic accuracy of the ATA and ACR-TIRADS risk stratification systems for a cohort of primarily benign thyroid nodules. We found that ATA had superior sensitivity and negative predictive value but had a lower specificity and positive predictive value as compared to the ACR-TIRADS system. A total of 28% (98 of 343) of the benign FNAs could have been avoided if applying ACR-TIRADS criteria compared to 11% (11%) based on ATA criteria. Of the nodules recommended for biopsy based on ATA, 21% (64/306) FNAs could have been avoided if TIRADS instead of ATA criteria were used to guide management. However, a greater number of non-benign and malignant nodules would have been undetected initially by ACR-TIRADS, although about half of these would have qualified for ultrasound surveillance. A total of 2/27 malignant cases were missed entirely by ACR-TIRADS (no follow up indicated if applying ACR-TIRADS criteria) whereas none of the malignant cases were missed by ATA. If vascularity was added to TIRADS as an additional prognostic marker, the number of initially undetected malignant nodules could be reduced from 7 to 4 cases.

The sensitivity of ACR-TIRADS in our series is similar to what has been published in the literature. A meta-analysis of 12 studies that examined the performance characteristics of ACR-TIRADS showed a pooled sensitivity of 79% [21]. This was similar to an updated meta-analysis [22], which showed a pooled sensitivity of 74% (61-83%) for ACR-TIRADS and 87% (75%-94%) for ATA. In contrast to previous studies, the specificity of ACR-TIRADS in our study was lower at approximately 30%, compared to 56-80% reported in the literature [10,14,19,20,22] Variable PPV (12-88%) [10,14,19] and NPV (68-98%) [10,14,19] for TIRADS have been found compared to a PPV of 18% and an NPV of 81% in our study. These findings reflect differences in study populations, baseline malignancy rates, and variability in ultrasound operator techniques and image interpretation.

ACR-TIRADS has been shown by multiple groups to have the greatest impact on reducing unnecessary investigations when compared to other commonly used classification systems [10,14,15,23]. We confirmed these previous findings and also found that ACR-TIRADS missed more cases of malignant and non-benign nodules. However, over 55% (12/22) of the cases qualified for ultrasound surveillance and if the nodules grew sufficiently, FNAs would have been performed during follow up. Most papillary thyroid carcinomas are small (<2 cm in size) without evidence of distant metastasis at the time of diagnosis [24]. Given that these usually grow slowly and most will have an indolent course [25], a slight delay in diagnosis is probably unlikely to result in significant morbidity and mortality. In fact, several prospective studies have shown that observation of these small tumours may be safe, with low rates of tumour growth, and no increased risk of distant metastases or death over a period of several years [26-28].

Nonetheless, it is important to note that in our study, a small number of nodules (2/10) that would not have qualified for surveillance based on TIRADS were malignant (2/27 of the malignant tumours were entirely missed by ACR-TIRADS). ATA did not miss any malignant tumours. Both cases had ongoing follow up and were eventually surgically removed without requiring RAI, indicating that these were low risk tumours associated with a favourable prognosis. This underscores the importance of using clinical judgment when deciding whether a nodule requires biopsy and/or further follow up as well as the frequency of radiologic monitoring as relying on TIRADS alone may result in missed malignancy of unclear significance. Our results indicate that TIRADS and ATA each have their own advantages and disadvantages, and that TIRADS does not entirely supplant the ATA.

There are four key areas of discrepancy between ATA and TIRADS (Table 2). For clinicians, this means that for cases that fall under one of these categories where there is disagreement between TIRADS and ATA, clinical risk factors and patient preferences should be considered in making a management decision. It would be worthwhile to determine whether the sensitivity of ACR-TIRADS could be improved if combined with clinical risk factors in future studies.

We showed in our study that if vascularity was added to TIRADS, the sensitivity of TIRADS could be improved. However, the number of cases where TIRADS-Vasc changed management was small. Furthermore, vascularity has not been consistently found to be a reliable marker of malignancy in previous studies [15-17], which may be attributable to the fact that the presence of intra-nodular vascularity is highly operator and observer-dependent [29]. This is supported by our own data showing only modest inter and intra-rater agreement for vascularity. Nonetheless, this intriguing finding raises the possibility that vascularity could be used as a tie-breaker in situations where TIRADS and ATA recommendations disagree. Clearly, more prospective, long term studies with larger sample sizes are needed before vascularity can be recommended for use as a key prognostic feature.

A major strength of our study is that this is the first study to characterize the diagnostic performance of the two commonly used risk stratification systems for thyroid nodules using real-world data in a Canadian context. Furthermore, we describe the follow up course and outcomes of nodules not deemed necessary for follow up as per ACR-TIRADS. There are several limitations. First, this was a retrospective, observational study in which we included a cohort of nodules that were already felt to have required FNA and/or surgery by a physician. PPV and NPV should be thus be interpreted with caution in this regard, given selection bias. However, the malignancy rate in our series is comparable to those reported by other selected and unselected cohorts [14,30]. Second, this was a single centre study conducted at a large, academic teaching hospital so findings may not be generalizable to other settings. Third, the ACR-TIRADS scores were converted from the original descriptions of the lesions with respect to their ATA classification, rather than re-scoring the nodules with TIRADS through independent observation. However, many aspects of ACR-TIRADS were captured in the first round of reporting and only 8 nodules were excluded as they could not be classified using ACR-TIRADS. Lastly, most of the nodules were assessed by a single observer due to limited resources. For the small number of nodules scored by two observers, the inter-rater agreement was fair.

In conclusion, we compared the diagnostic efficacies of ATA and ACR-TIRADS risk stratification systems for thyroid nodules, and found that ATA was more sensitive in cancer detection but TIRADS had better specificity. A greater number of unnecessary FNAs could be reduced if applying TIRADS

Table 2. *ACR TI-RADS vs ATA: Comparison of the Recommendations for Biopsy*

ATA Pattern	ATA US features Brackets = TI-RADS Score for USS findings	ATA size cut-off to recommend FNA	TI-RAD Score & Recommendation for Biopsy
High suspicion	Solid (2) hypoechoic (2-3) nodule with one of the following high-risk features: = microcalcification (3) = taller than wide (3) = irregular margins (2) = ETE (3) = rim calcification with small extrusive soft tissue component (2) Note = TIRADS gives macrocalcification 1 point and ATA states coarse calcification is a low risk feature	≥1cm ** Some nodules less than 1 cm with ++ high risk features or have a significant increase in size are biopsied	Range of score = 6 to ≥7 TI-RAD Score = TR4 (bx at ≥ 1.5cm) TR5 (bx at ≥1cm) Agreement between ATA & TI-RADS: - TR4 and ATA are concordant for nodules ≥1.5cm - TR4 and ATA are discordant for nodules 1-1.4cm - TR5 and ATA concordant
Intermediate suspicion	Solid (2) Hypoechoic nodules (2-3) without any high-risk features	≥1 cm	Range of score = 4-5 TIRAD Score = 4 (bx at ≥1.5cm) Agreement between ATA & TI-RADS: - TR4 and ATA are concordant for nodules ≥1.5cm - TR4 and ATA are discordant for nodules 1-1.4cm
Low suspicion	Solid (2) Isoechoic (1) or hyperechoic (1) nodule without any high-risk features OR Partially cystic nodule (1) with eccentric solid component (1 for hyper/isoechoic of solid component) Note = TI-RADS does not score eccentric solid component higher than if non-eccentric	≥1.5cm	Range of score = 2-3 TI-RAD Score = TR2 (no bx) TR3 (bx at ≥2.5cm) Agreement between ATA & TI-RADS: - TR2 and ATA are concordant for nodules < 1.5cm - TR2 and ATA are discordant for nodules >=1.5cm - TR3 and ATA are concordant for nodules ≥2.5 cm - TR3 and ATA discordant for nodules 1.5-2.4cm
Very low suspicion	Spongiform (0) Or Partially cystic (1) without any high-risk features or eccentric solid area (1 for hyper/hypo of solid component)	Can observe Vs FNA ≥2cm	Range of score 0-2 TI-RAD Score = TR1/2 (no bx) Agreement between ATA & TI-RADS: - Concordant
Benign	Pure cyst (0)	No biopsy	Range of score = 0 TI-RAD Score = TR1 (no bx) Agreement between ATA & TI-RADS: - Concordant

Bx: biopsy; US: ultrasound; ETE: extra-thyroidal extension

instead of ATA criteria, but TIRADS also missed more cases of malignant and non-benign nodules. However, the malignant cases were low risk tumours associated with a favourable prognosis. Nonetheless, clinical judgement is important in deciding which nodules require further investigation, especially in cases where management recommendations differ between different radiologic classification systems.

Statements

Statement of ethics

This work was conducted ethically in accordance with the World Medical Association Declaration of Helsinki. Ethics approval was obtained from our institution.

Conflict of interest statement

The authors have no conflicts of interest to declare.

Author contributions

Authors JQ and IH contributed to the conception of this study, data analysis, and drafted the manuscript. CB, KH, MM, ZG, KB contributed to data collection. JG contributed to manuscript content. KMH and all authors contributed to critical review of the manuscript.

Financial disclosure

Dr. Qiang is supported by Canadian Institutes of Health Research Postdoctoral Fellowship and the Peter Gilgan Centre for Women's Cancers.

References

1. Ezzat S, Sarti DA, Cain DR, et al. Thyroid incidentalomas. Prevalence by palpation and ultrasonography. Arch Intern Med. 1994;154(16):1838.
2. Gharib H. Changing concepts in the diagnosis and management of thyroid nodules. Endocrinol Metab Clin North Am. 1997;26(4):777-800.
3. Leech JV, Smith LW, Clute HM. Aberrant Thyroid Glands. Am J Pathol. 1928;4(5):481-492.
4. Rojeski MT, Gharib H. Nodular thyroid disease. Evaluation and management. N Engl J Med. 1985;313(7):428-436.
5. Burguera B, Gharib H. Thyroid incidentalomas. Prevalence, diagnosis, significance, and management. Endocrinol Metab Clin North Am. 2000;29(1):187-203.
6. Dean DS, Gharib H. Epidemiology of thyroid nodules. Best Pract Res Clin Endocrinol Metab. 2008;22(6):901-11.
7. 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.
8. 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.
9. Tessler FN, Middleton WD, Grant EG, et al. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. J Am Coll Radiol. 2017;14(5):587-595.
10. Ha EJ, Na DG, Moon WJ, Lee YH, Choi N. Diagnostic Performance of Ultrasound-Based Risk-Stratification Systems for Thyroid Nodules: Comparison of the 2015 American Thyroid Association Guidelines with the 2016 Korean Thyroid Association/Korean Society of Thyroid Radiology and 2017 American College of Radiology Guidelines. Thyroid. 2018;28(11):1532-1537.
11. 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.
12. Xu T, Gu JY, Ye XH, 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.
13. Wu XL, Du JR, Wang H, et al. Comparison and preliminary discussion of the reasons for the differences in diagnostic performance and unnecessary FNA biopsies between the ACR TIRADS and 2015 ATA guidelines. Endocrine. 2019;65(1):121-131.
14. 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;104:95-102.
15. Pandya A, Caoili E, Jawad-Makki F, et al. Retrospective cohort study of 1947 thyroid nodules: a comparison of the 2017 American College of Radiology TI-TRADS and the 2015 American Thyroid Association Classifications. AJR Am J Roentgenol. 2020;214(4):900-906
16. 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):1941-1946.
17. 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.
18. Moon HJ, Kwak JY, Kim MJ, Son EJ, Kim EK. 2010 Can vascularity at power Doppler US help predict thyroid malignancy? Radiology. 2010;255:260–269.
19. 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;64(1):90-96.
20. Cibas ES, Ali SZ. The Bethesda system for reporting thyroid cytopathology. Thyroid. 2009;19(11):1159-65.
21. 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(2):309-15.
22. Castellana M, Castellana C, Treglia G, et al. Performance of Five Ultrasound Risk Stratification Systems in Selecting Thyroid Nodules for FNA. J Clin Endocrinol Metab. 2020;105(5):dgz170.
23. Middleton WD, Teefey SA, Reading CC, et al. Comparison of Performance Characteristics of American College of Radiology TI-RADS, Korean Society of Thyroid Radiology TIRADS, and American Thyroid Association Guidelines. AJR Am J Roentgenol. 2018;210:1148-54.

24. Kent WD, Hall SF, Isotalo PA, et al. Increased incidence of differentiated thyroid carcinoma and detection of subclinical disease. CMAJ. 2007; 177(11):1357-61.
25. Mazzaferri EL, Jhiang SM. Long-term impact of initial surgical and medical therapy on papillary and follicular thyroid cancer. Am J Med. 1994;97(5):418.
26. Sugitani I, Toda K, Yamada K, et al. Three distinctly different kinds of papillary thyroid microcarcinoma should be recognized: our treatment strategies and outcomes. World J Surg. 2010;34:1222–31.
27. Ito Y, Miyauchi A, Kihara M, et al. Patient age is significantly related to the progression of papillary microcarcinoma of the thyroid under observation. Thyroid. 2014;24:27-34.
28. Tuttle RM, Fagin JA, Minkowitz G, et al. Natural history and tumor volume kinetics of papillary thyroid cancers during active surveillance. JAMA Otolaryngol Head Neck Surg. 2017;143:1015-1020.
29. Floridi C, Cellina M, Buccimazza G, et al. Ultrasound imaging classifications of thyroid nodules for malignancy risk stratification and clinical management: state of the art. Gland Surg. 2019;8(Suppl 3):S233-S244.
30. Durante C, Grani G, Lamartina L, et al. The diagnosis and management of thyroid nodules: a review. JAMA. 2018;319(9): 914-924.

SUPPLEMENTARY DATA

Composition		Echogenicity		Shape		Margin		Echogenic Foci	
Cystic or almost completely cystic	0 point	Anachoic	0 point	Wider than tall	0 point	Smooth	0 point	None or artifact	0 point
Spongiform	0 point	Hyperechoic or isoechoic	1 point	Taller than wide	3 points	Ill-defined	0 points	Macrocalcifications	1 point
Mixed cystic and solid	1 point	Hypoechoic	2 points		Lobulated or irregular	2 points	Peripheral calcifications	2 points	
Solid or almost completely solid	2 points	Very hypoechoic	3 points		Extrathyroidal extension	3 points	Punctate foci	3 points	

Total points:

- 0 = TIRADS 1, benign, no FNA
- 2 = TIRADS 2, not suspicious, no FNA
- 3 = TIRADS 3, mildly suspicious, FNA if ≥ 2.5 cm, follow if ≥ 1.5 cm
- 4-6 = TIRADS 4, moderately suspicious, FNA if ≥ 1.5 cm, follow if ≥ 1.0 cm
- 7 or more = TIRADS 5, highly suspicious, FNA ≥ 1.0 cm, follow if ≥ 0.5 cm