Published in final edited form as:

J Surg Res. 2020 November ; 255: 469–474. doi:10.1016/j.jss.2020.05.074.

Discordance Between ATA and ACR Guideline Systems for Thyroid Nodule Biopsy

Eric Kim, BS^a, Roshini Pudhucode, BS^b, Herbert Chen, MD FACS^a, Brenessa Lindeman, MD MEHP FACS^a

^aUniversity of Alabama at Birmingham, Division of Surgical Oncology, Department of Surgery, Birmingham, Alabama, USA

^bAlabama College of Osteopathic Medicine, Dothan, Alabama, USA

Abstract

Background: Previously, fine-needle aspiration (FNA) biopsy was recommended for any thyroid nodule >1.0cm in size. In 2015, the American Thyroid Association (ATA) introduced a pattern-based approach for biopsy recommendations based on size and ultrasound (US) characteristics. In 2016, the American College of Radiology (ACR) published the Thyroid Imaging Reporting and Data System (TIRADS), using a point-based system that assesses risk of ultrasound characteristics.

Methods: This study aims to compare recommendations for thyroid nodule biopsy between the ATA and ACR systems and identify outcomes of nodules with discordant recommendations (DR). US characteristics, FNA biopsy, and surgical pathology results were evaluated for all patients with >1.0cm thyroid nodules treated at a single tertiary-care institution from 2010–2018.

Results: Inclusion criteria were met by 1100 nodules from 687 patients; 42.8% (n=471) had DR between the ATA and ACR guidelines. All (100%) DR nodules were not recommended for biopsy by ACR, though 53% were recommended to have follow-up. A majority (79%) of DR nodules were recommended for biopsy by ATA, with the remaining 21% recommended for follow-up. Among surgically excised DR nodules (n=292), 10.3% (n=30) nodules were found to be malignant, with the vast majority (90.3%) being well-differentiated carcinoma. Among malignant nodules, the ACR would not have recommended biopsy or follow-up for 26.7% (n=8).

Corresponding Author: Brenessa Lindeman, MD, MEHP, 1808 7th Ave S, BDB 603, Birmingham, AL 35294, Phone: 205-975-5000, Fax: 205-975-5791, blindeman@uabmc.edu.

All authors provided contributions in regards to conceptualization and study design, data collection, data analysis, editing, and providing final approval of the submission. All authors agree to be accountable for the work and its accuracy and integrity.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declarations of Interest: None

Disclosure:

The authors report no proprietary or commercial interest in any product mentioned or concept discussed in this article.

Conclusion: The ACR classification system is more restrictive compared to the ATA system for recommending thyroid nodule biopsy. This discrepancy could result in confusion for clinicians and delay in diagnosis or therapy for patients with thyroid cancer.

Keywords

thyroid cancer; ATA; ACR TI-RADS; thyroid ultrasound; thyroid guidelines; FNA biopsy

Introduction:

Thyroid nodules are routinely found in clinical practice. It is generally accepted that the prevalence of a clinically palpable thyroid nodule is between 4 to 7 percent of the adult population.^{1, 2} The true prevalence is much greater when including microscopic nodules, with numerous studies across different populations citing increases in prevalence to 20–68% with ultrasonography^{3–7} and up to 60% with autopsy.^{4, 8, 9} However, most thyroid nodules are benign, progress slowly, and do not warrant surgical intervention. This has resulted in well-justified concern regarding overdiagnosis and overtreatment of benign thyroid nodules that would have otherwise led an indolent course without intervention. Therefore, guidelines for managing these lesions have evolved from biopsy of all nodules larger than 1cm¹⁰ to a more selective approach based on cytological assessment of both size and ultrasonography characteristics using ultrasound-guided fine-needle aspiration (FNA) biopsy. This strategy has been shown to double the yield of detecting carcinoma with a corresponding decline in unnecessary surgical procedures for benign nodules.¹¹

Thyroid ultrasonography has gained widespread use as the initial evaluation of a thyroid nodule. Although imaging alone cannot distinguish between benign and malignant lesions, it provides useful information about a nodule's size and morphology which can be used to guide clinical decision-making.¹² In order to capture clinically suspicious thyroid nodules, guidelines have been established by international organizations that detail when a nodule warrants biopsy based on ultrasound characteristics. In the United States, guidelines published by the American College of Radiology (ACR)¹³ and American Thyroid Association (ATA)¹⁴ are commonly used to identify those nodules that merit FNA biopsy. The ACR-Thyroid Imaging Reporting and Data System (TI-RADS) utilizes a point system while the ATA system uses a pattern-based system for risk-stratification of nodules. Although the two systems share many overlapping features, their differing approaches often lead to discordant recommendations, especially for smaller nodules with indeterminate risk features. Few studies have examined and characterized those nodules given discordant recommendations by the two systems.

In this study, we aim to categorize all thyroid nodules greater than 1cm according to both ATA and ACR TI-RADS guidelines, with further focus on nodules with discordant biopsy and follow-up recommendations between the two systems, hypothesizing that the more sensitive ATA system will recommend biopsy for a greater number of nodules and consequently detect more malignancies than the more specific ACR TI-RADS system which will recommend fewer biopsies for benign nodules but detect fewer malignancies.

Methods:

We conducted a retrospective cohort review of patients evaluated for a thyroid nodule at the University of Alabama at Birmingham (UAB) from 2010 to 2018. Patient characteristics including age, sex, race, and visit date and nodule characteristics including size, composition, margin regularity, echogenicity, shape, and echogenic foci were collected for further analysis. The protocol of this study was approved as exempt by the UAB Institutional Review Board.

Participants

Patients were identified using the UAB Endocrine Surgery Registry, a comprehensive cohort of patients with different endocrine pathologies treated at our institution. The registry was queried for all patients with diagnosis codes corresponding to thyroid nodule, nontoxic unimodular goiter, and nontoxic multinodular goiter or procedural code of thyroid surgery. Patients with a diagnosis of Graves' disease, Hashimoto's thyroiditis, thyrotoxicosis, or incomplete ultrasonography reports were excluded (Figure 1).

Ultrasonography

The majority of nodules were accompanied with a report by a radiologist at our institution that scored each ultrasound characteristic according to ACR TI-RADS guidelines. For nodules with incomplete or absent reports, each ultrasound characteristic was determined by the authors in accordance with the guidelines set forth by the ACR TI-RADS White Paper.¹³ For nodules with low-resolution images or ambiguous characteristics, two raters (EK, BL) assigned characteristics independently and compared results. There was high inter-rater reliability (95%) using this method. Although conflicting readings were rare, raters discussed suspicious ultrasounds until consensus was reached. For patients with multiple ultrasonography reports and FNA biopsies, the most recent report preceding surgical intervention was used. If the patient did not undergo surgery, the most recent available US report was used. If multiple nodules were found within the thyroid of one patient, each nodule was evaluated independently provided it met all inclusion criteria.

Statistical Analysis

Each nodule was evaluated independently, and its size in the longitudinal, anterior-posterior, and transverse axes as well as ultrasound characteristics were compiled using Microsoft Excel (Redmond, WA). Using nested functions, nodules were scored and stratified by risk in accordance with both ATA and ACR TI-RADS classification systems, as shown in Table 1. Nodules that are neither solid nor hypoechoic but have one or more suspicious features lack definitive recommendations by ATA guidelines. These nodules were categorized as high suspicion in our study due to the correlation between suspicious US features and malignancy. We analyzed the data using SPSS, version 26.0 (Armonk, NY: IBM Corp). Independent samples T-tests were used to compare differences between average size of nodules. A p -value <0.05 was considered to be statistically significant.

Results:

The initial query from the UAB Endocrine Surgery Registry identified 1657 potential participants. Of these, 687 patients with 1100 thyroid nodules greater than 1cm met all inclusion criteria (Figure 1). The median age of this cohort was 59 years, and 83.0% (n=570) were female. Participant race was 62.7% (n=431) non-Hispanic White, 35.4% (n=243) black, with the remaining 1.9% (n=13) Hispanic, Asian, and those who declined to disclose their race. The mean nodule size was 2.6cm $\pm$ 1.5cm.

Of the 1100 nodules included in the study, 471 (42.8%) were found to have discordant recommendations (DR) for FNA biopsy between the ATA and ACR TI-RADS systems. These recommendations fell into three categories: ATA follow-up and ACR no follow-up, ATA Biopsy and ACR no follow-up, and ATA biopsy with ACR follow-up (Figure 2). The majority of these DR nodules were on the smaller side with 53.1% (n=250) falling between 1.0–1.49cm in size, 20.2% (n=95) between 1.5–1.99cm, 15.7% (n=74) between 2.0–2.49, and the remaining 11.0% (n=52) greater than 2.5cm in size (Figure 2). Among the 471 discordant nodules, ATA recommended biopsy for 79.0% (n=372). The remainder (n=99, 21.0%), all of which were nodules characterized as low suspicion and were concomitantly 1.0–1.5cm in size, were recommended for one-year follow-up by the ATA. The ACR TI-RADS system recommended against biopsy for 100% of the 471 discordant nodules, but advised follow-up for 53% (n=250) of them, resulting in neither FNA biopsy nor follow-up recommendations for the remaining 47% (n=221).

Total or partial thyroidectomy was performed on 62.0% (n=292) of nodules with discordant recommendations. Nodules removed surgically had a mean size of 2.0cm $\pm$ 1.1cm compared to 1.4cm $\pm$ 0.4 for those that were not (p<0.01). The majority of surgically-resected nodules were benign with 27.7% (n=81) classified as nodular goiter and 54.4% (n=159) classified as nodular hyperplasia (Figure 3). Surgical pathology confirmed malignancy for 10.3% (n=30) of nodules. Of these, 90.3% (n=27) were papillary thyroid carcinomas, 6.7% (n=2) were medullary thyroid carcinomas, and one (3.3%) was Hurthle cell carcinoma.

Among those with discordant biopsy recommendations, benign nodules had a significantly larger mean size compared to malignant nodules (2.08 $\pm$ 1.18 vs 1.52 $\pm$ 0.52, p=0.01). In regards to ultrasound characteristics, most discordant malignant nodules had only 1 or 2 suspicious features (n=25, 83%). Most were solid (n=24, 80%) with the remainder being mixed cystic and solid (n=6, 20%). Of the 30 discordant malignant nodules, 26.7% (n=8) had irregular margins and 3.3% (n=1) had lobulated borders. The majority of nodules were isoechoic (n=16, 53.3%) or hypoechoic (n=12, 40%), and only 33% (n=10) showed echogenic foci. None were taller than wide. If surgical decision-making relied exclusively on the recommendations of either system, these 30 malignant nodules would have widely conflicting treatment strategies. The ATA system would recommend biopsy for 90.0% (n=27) of these malignant nodules and follow-up for the remaining 10.0% (n=3). Conversely, the ACR TI-RADS system would recommend follow-up for 73.3% (n=22) and neither biopsy nor follow-up for the remaining 26.7% (n=8). Of the 8 malignant nodules that would have been missed by ACR TI-RADS, 87.5% (n=7) were papillary thyroid carcinoma with one (12.5%) Hurthle cell carcinoma (Figure 3).

Discussion:

The role of ultrasonography for detection and risk prediction of thyroid nodules should not be understated.¹⁵ The specific ultrasound characteristics of nodules help guide clinicians and can impact management strategies.¹⁶ Therefore, standardized guidelines set forth by organizations like the American Thyroid Association and American College of Radiology are important for optimizing treatment and reducing variability across healthcare providers. However, this study and others have shown that the rate of discordance in FNA biopsy recommendations between these systems is significant and may lead to confusion when treating patients.^{17–19} Recommendations from these systems are particularly important when considering the historical concern of over-diagnosis and overtreatment of thyroid cancers that are likely subclinical diseases. Numerous studies have described an increasing rate of early-stage thyroid cancer incidence without a corresponding rise in mortality.^{20–22} In their study on long-term outcomes of untreated benign nodules, Kuma et al. found that 92% of nodules remained benign with no advance in cytological classification after 9 to 11 years.²³

The current paradigm to avoid over-diagnosis and over-treatment of thyroid malignancy in the United States has led to fewer biopsies and the potential to miss thyroid malignancies. The concern of over-treatment is warranted given the slow-growing and indolent nature of thyroid malignancies. However, in a more recent study, Lim et al. found a statistically significant 3.6% annual percentage increase in thyroid carcinoma incidence and a related 2.9% annual percentage increase in mortality for advanced-stage PTCs in the United States from 1974 to 2013. Despite not reaching significance, annual percentage increases were increased for PTCs of all known stages. Taken together, these results challenge the notion that thyroid malignancy rates are solely the result of over-diagnosis.²⁴ This suggests that more conservative approaches may lead to under-detection of thyroid malignancies that should be actively surveilled – whether with biopsy or follow-up. The results of our study support this as the ACR TI-RADS failed to recommend preoperative biopsy for discordant nodules that were found to be malignant.

In our study, the majority of discordant recommendations between the ATA and ACR TI-RADS systems arose from the varying size thresholds for both biopsy and follow-up. Yet, numerous studies have shown that nodule size is a poor predictor of malignancy when compared to the presence of suspicious ultrasound characteristics.^{25–27} In fact, some guidelines have found success with recommendations based solely on ultrasound characteristics.²⁸ This favors ATA's pattern-based approach to risk stratification which places more emphasis on suspicious ultrasound features. Our results also indicate that classification systems may benefit from assigning greater weight to these features and loosening size requirements especially with respect to follow-up. Although most patients receive follow-up care irrespective of these recommendations, this may or may not be true for all patient populations and clinicians should exercise caution in discontinuing surveillance of less suspicious nodules, which was recommended for 47% of DR nodules by the ACR-TIRADS system in our study. This is especially important for nodules on the smaller end of the size range which were more likely to have discordant recommendations and significantly associated with malignancy in our study population of discordant nodules.

Limitations of this study include those inherent to retrospective studies, such as the possibility of incomplete data capture and omission of potentially important variables which we attempted to circumvent by including only those patients with complete records. This study included patients from a single tertiary-care hospital, which may limit the generalizability of its findings to other settings and contexts. Furthermore, rates of malignancy could only be calculated from the discordant nodules that were surgically removed, resulting in a potential selection bias that overestimated malignancy. Another limitation is inherent to the subjective analysis of ultrasounds by a team of radiologists who may have received different training. This was addressed by the use of multiple raters for any discrepant ultrasound findings. By assigning high suspicion to those nodules that could not be categorized by ATA, our model may overestimate the number of nodules recommended for biopsy and follow-up by the ATA guidelines. Due to the observational nature of the study, long-term outcomes associated with conflicting guideline recommendations and earlier management of suspicious nodules were not assessed. This is particularly important for identification of indolent malignancies which may provide the basis for future study.

Conclusion:

Nearly half (42.8%) of all thyroid nodules identified at a single center over a 9-year period were given discordant recommendations for further management by the ATA and ACR TI-RADS guidelines, with the ATA guidelines recommending biopsy at a greater rate than ACR TI-RADS. Among the discordant nodules that were surgically removed, malignancy was identified in 10.3%. More than one-quarter of malignant nodules – or 3% of all nodules – would not have been recommended for biopsy or ongoing surveillance if the ACR TI-RADS system alone had been used. A large proportion of discordance occurred within nodules at the smaller end of the size range (1.0–1.5cm), so a more proactive approach to ongoing surveillance and biopsy among these nodules may prevent missing a small but significant subset of patients with thyroid malignancy.

Funding:

This work was supported in part by the T35 National Research Service Award (NRSA) Short-Term Institutional Research Training Grant (PA-18-404), funded by the National Institutes of Health.

References:

1. Singer PA, et al., Treatment guidelines for patients with thyroid nodules and well-differentiated thyroid cancer. American Thyroid Association. Arch Intern Med, 1996 156(19): p. 2165–72. [PubMed: 8885814]
2. Mazzaferri EL, Management of a solitary thyroid nodule. N Engl J Med, 1993 328(8): p. 553–9 DOI: 10.1056/NEJM199302253280807. [PubMed: 8426623]
3. Ezzat S, et al., Thyroid incidentalomas. Prevalence by palpation and ultrasonography. Arch Intern Med, 1994 154(16): p. 1838–40 DOI: 10.1001/archinte.154.16.1838. [PubMed: 8053752]
4. Tan GH and Gharib H, Thyroid incidentalomas: management approaches to nonpalpable nodules discovered incidentally on thyroid imaging. Ann Intern Med, 1997 126(3): p. 226–31 DOI: 10.7326/0003-4819-126-3-199702010-00009. [PubMed: 9027275]

5. Reiners C, et al., Prevalence of thyroid disorders in the working population of Germany: ultrasonography screening in 96,278 unselected employees. *Thyroid*, 2004 14(11): p. 926–32 DOI: 10.1089/thy.2004.14.926. [PubMed: 15671771]
6. Brander A, et al., Thyroid gland: US screening in a random adult population. *Radiology*, 1991 181(3): p. 683–7 DOI: 10.1148/radiology.181.3.1947082. [PubMed: 1947082]
7. Guth S, et al., Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. *Eur J Clin Invest*, 2009 39(8): p. 699–706 DOI: 10.1111/j.1365-2362.2009.02162.x. [PubMed: 19601965]
8. Wang C and Crapo LM, The epidemiology of thyroid disease and implications for screening. *Endocrinol Metab Clin North Am*, 1997 26(1): p. 189–218 DOI: 10.1016/s0889-8529(05)70240-1. [PubMed: 9074859]
9. Dean DS and Gharib H, Epidemiology of thyroid nodules. *Best Pract Res Clin Endocrinol Metab*, 2008 22(6): p. 901–11 DOI: 10.1016/j.beem.2008.09.019. [PubMed: 19041821]
10. American Thyroid Association Guidelines Taskforce on Thyroid, N., et al., Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*, 2009 19(11): p. 1167–214 DOI: 10.1089/thy.2009.0110. [PubMed: 19860577]
11. Gharib H, Fine-needle aspiration biopsy of thyroid nodules: advantages, limitations, and effect. *Mayo Clin Proc*, 1994 69(1): p. 44–9 DOI: 10.1016/s0025-6196(12)61611-5. [PubMed: 8271850]
12. Hegedus L, Thyroid ultrasound. *Endocrinol Metab Clin North Am*, 2001 30(2): p. 339–60, viii-ix DOI: 10.1016/s0889-8529(05)70190-0. [PubMed: 11444166]
13. Tessler FN, 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): p. 587–595 DOI: 10.1016/j.jacr.2017.01.046. [PubMed: 28372962]
14. Haugen BR, 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: What is new and what has changed? *Cancer*, 2017 123(3): p. 372–381 DOI: 10.1002/cncr.30360. [PubMed: 27741354]
15. Hegedus L, Clinical practice. The thyroid nodule. *N Engl J Med*, 2004 351(17): p. 1764–71 DOI: 10.1056/NEJMcp031436. [PubMed: 15496625]
16. Marqusee E, et al., Usefulness of ultrasonography in the management of nodular thyroid disease. *Ann Intern Med*, 2000 133(9): p. 696–700 DOI: 10.7326/0003-4819-133-9-200011070-00011. [PubMed: 11074902]
17. Ha SM, et al., Diagnostic Performance of Practice Guidelines for Thyroid Nodules: Thyroid Nodule Size versus Biopsy Rates. *Radiology*, 2019 291(1): p. 92–99 DOI: 10.1148/radiol.2019181723. [PubMed: 30777805]
18. Lauria Pantano A, et al., Differences between ATA, AACE/ACE/AME and ACR TI-RADS ultrasound classifications performance in identifying cytological high-risk thyroid nodules. *Eur J Endocrinol*, 2018 178(6): p. 595–603 DOI: 10.1530/EJE-18-0083. [PubMed: 29626008]
19. Gao L, 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): p. 90–96 DOI: 10.1007/s12020-019-01843-x. [PubMed: 30659427]
20. Davies L and Welch HG, Current thyroid cancer trends in the United States. *JAMA Otolaryngol Head Neck Surg*, 2014 140(4): p. 317–22 DOI: 10.1001/jamaoto.2014.1. [PubMed: 24557566]
21. Hall SF, et al., Access, excess, and overdiagnosis: the case for thyroid cancer. *Cancer Med*, 2014 3(1): p. 154–61 DOI: 10.1002/cam4.184. [PubMed: 24408145]
22. Powers AE, et al., Changes in Trends in Thyroid Cancer Incidence in the United States, 1992 to 2016. *JAMA*, 2019 322(24): p. 2440–2441 DOI: 10.1001/jama.2019.18528. [PubMed: 31860035]
23. Kuma K, et al., Fate of untreated benign thyroid nodules: results of long-term follow-up. *World J Surg*, 1994 18(4): p. 495–8; discussion 499 DOI: 10.1007/bf00353745. [PubMed: 7725734]
24. Lim H, et al., Trends in Thyroid Cancer Incidence and Mortality in the United States, 1974–2013. *JAMA*, 2017 317(13): p. 1338–1348 DOI: 10.1001/jama.2017.2719. [PubMed: 28362912]
25. Cavallo A, et al., Thyroid Nodule Size at Ultrasound as a Predictor of Malignancy and Final Pathologic Size. *Thyroid*, 2017 27(5): p. 641–650 DOI: 10.1089/thy.2016.0336. [PubMed: 28052718]

26. Leenhardt L, et al., Indications and limits of ultrasound-guided cytology in the management of nonpalpable thyroid nodules. *J Clin Endocrinol Metab*, 1999 84(1): p. 24–8 DOI: 10.1210/jcem.84.1.5418. [PubMed: 9920057]
27. Papini E, et al., Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound and color-Doppler features. *J Clin Endocrinol Metab*, 2002 87(5): p. 1941–6 DOI: 10.1210/jcem.87.5.8504. [PubMed: 11994321]
28. Kim EK, et al., New sonographic criteria for recommending fine-needle aspiration biopsy of nonpalpable solid nodules of the thyroid. *AJR Am J Roentgenol*, 2002 178(3): p. 687–91 DOI: 10.2214/ajr.178.3.1780687. [PubMed: 11856699]

Figure 1.
Flowchart of patient selection and nodule inclusion.

Figure 2.
ATA and ACR TI-RADS biopsy and follow-up recommendations for discordant nodules stratified by size

Figure 3. Surgical pathology results of discordant nodules with corresponding ATA and ACR TI-RADS recommendations.

NIFTP: Non-invasive follicular thyroid neoplasm with papillary-like nuclear features

Table 1.

ATA and ACR TI-RADS recommendations for FNA biopsy and follow-up for nodules

ACR Level	FNAB Size Criteria	Follow-Up Recommendation	ATA Level	FNAB Size Criteria	Follow-Up Recommendation
TR1	No FNAB	None	Benign	No FNAB	None
TR2	No FNAB	None	Very Low Suspicion	2.0 cm	None
TR3	2.5 cm	Follow if 1.5 cm	Low Suspicion	1.5 cm	Repeat US 12–24 mo.
TR4	1.5 cm	Follow if 1 cm	Intermediate Suspicion	1.0 cm	Repeat US 12–24 mo.
TR5	1.0 cm	Follow if 0.5 cm	High Suspicion	1.0cm	Repeat US 6–12 mo.