THYROID
Volume 33, Number 5, 2023

DOI: 10.1089/thy.2022.0672

Malignancy Risk of Thyroid Nodules That Are Not Classifiable by the American Thyroid Association Ultrasound Risk Stratification System: A Systematic Review and Meta-Analysis

Daniel Kwon,[1] Marta Kulich,[1] Wendy J. Mack,[2] Rodrigo Martinez Monedero,[1] Eri Joyo,[3] and Trevor E. Angell[3]

Background: Sonographic evaluation is fundamental to thyroid nodule assessment. The American Thyroid Association (ATA) ultrasound risk stratification system (USRSS) is widely used, but the appearance of some nodules has been considered nonclassifiable (NC-ATA). The risk of malignancy (RoM) of NC-ATA nodules varies widely between studies, leading to uncertainty in clinical management. The aim of this study was to comprehensively evaluate the prevalence and malignancy risk of NC-ATA nodules.
Methods: A systematic review was performed searching PubMed/MEDLINE and EMBASE to identify original studies of thyroid nodules classified using the ATA USRSS from 2016 to 2022 and reporting the outcome of NC-ATA nodules. Meta-analysis was conducted to obtain pooled RoM estimates and meta-regression sensitivity analyses were used to explore sources of between-study heterogeneity.
Results: Of 6377 screened studies, 135 underwent full-text review, and 16 studies reporting 21,271 nodules were included. Within these, the pooled prevalence of NC-ATA nodules was 7.8% (1872 nodules; [confidence interval; CI 5.1–11.1]). The pooled RoM estimate for NC-ATA nodules was 20.3% [CI 13.0–28.7] and there was significant heterogeneity between studies ($I^2 = 92.8\%$, $p < 0.001$). NC-ATA nodule RoM estimates were significantly different by study type: single-center versus multicenter studies (24.8% vs. 12.3%, respectively, $p = 0.031$) and study design: retrospective versus prospective studies (25.1% vs. 8.5%, respectively, $p = 0.003$). No significant difference was observed in RoM based on inclusion of <1 cm nodules or geographic region. Meta-regression analysis showed study design and use of surgical histology for diagnostic criteria contributed significantly to differences in the reported RoM estimates.
Conclusion: In this first meta-analysis comprehensively assessing the RoM of NC-ATA nodules, the malignancy risk was found to be comparable with the current ATA USRSS intermediate suspicion category. Significant heterogeneity was observed between studies and limits the interpretation of these results. In future iterations of the ATA USRSS that seek into incorporate categorization of NC-ATA nodules, these meta-analysis data may help to inform proper malignancy risk stratification. The study protocol was registered on PROSPERO, the international prospective register of systematic reviews (CRD42020182498), on July 14, 2020.

Keywords: thyroid neoplasms, papillary thyroid cancer, ultrasound, thyroid nodule, risk stratification

[1]Department of Otolaryngology—Head and Neck Surgery; [2]Population and Public Health Sciences; [3]Division of Endocrinology and Diabetes; Keck Medicine of the University of Southern California, Los Angeles, California, USA.

This study was presented as an abstract at the 90th Annual Meeting of the American Thyroid Association; September 30, 2021–October 3, 2021; held virtually.

Introduction

THYROID NODULES ARE highly prevalent and there has been a significant increase in their detection and evaluation for the past several decades.[1] Thyroid ultrasound evaluation has emerged as an integral part of thyroid nodule assessment to stratify the risk of malignancy based on sonographic features. To improve implementation, standardized ultrasound risk stratification systems (USRSS) have been devised, including in the 2015 American Thyroid Association (ATA) Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer.[2] This system utilizes sonographic patterns to categorize nodules (with associated malignancy risk) as benign (<1%), very low suspicion (<3%), low suspicion (5–10%), intermediate suspicion (10–20%), or high suspicion (>70–90%). Based on their respective risk of malignancy (RoM), each category has an associated size cutoff at which fine needle aspiration (FNA) is recommended.

While the ATA USRSS has performed well in validation studies, a subset of nodules does not fit any of the described patterns. These ''nonclassifiable'' (NC-ATA) nodules generally consist of a combination of lower suspicion and higher suspicion features, such as iso- or hyperechoic or mixed cystic/solid nodules that have one or more high-risk features (e.g., punctate echogenic foci, taller than wide shape, and irregular margins). Some prior studies have suggested that NC-ATA nodules may be encountered relatively frequently[3–7] and have high RoM,[5,8–11] while other data indicate that NC-ATA nodules are uncommon[8,12–15] and RoM is low.[3,4,6,16,17]

This wide variability in the reported findings among individual studies, which is perhaps due to differences in study populations or methods, results in uncertainty as to the true prevalence and RoM of NC-ATA nodules. Therefore, this systematic review and meta-analysis aims to comprehensively determine the frequency and RoM for NC-ATA nodules described in the literature and assess variables influencing malignancy risk estimates.

Methods

Search strategy

We performed a literature search according to the Preferred Reporting Items for a Systematic Review and Meta-analysis (PRISMA) guidelines. A systematic search was done from January 2015 until August 2022 of PubMed/MEDLINE and EMBASE databases to identify English language publications that recognized and reported on NC-ATA nodules (Supplementary Data). Identified references with abstracts were exported to Covidence (Covidence Systematic Review Software, Veritas Health Innovation, Melbourne, Australia; www.covidence.org) for initial screening. This study was considered exempt by the governing institutional review board.

Study selection

After removal of duplicates, at least two authors (T.E.A., D.K., M.K., E.J., and/or R.R.M.) independently performed a title and abstract screen followed by a full-text review by two independent reviewers (T.E.A., D.K., M.K., and/or R.R.M.) using previously agreed upon inclusion and exclusion criteria. The reference lists for all studies that underwent full-text evaluation were screened for additional relevant articles for review. Any disagreements were resolved by discussion and review by two authors (T.E.A. and D.K.). Studies met initial inclusion criteria if they were original publications of thyroid nodules in adult subjects; provided categorization using the ATA USRSS; and reported pathologic outcomes of NC-ATA nodules, either by histopathology and/or definitive biopsy result (i.e., Benign [Bethesda II] or Malignancy [Bethesda VI] cytology).

Studies were excluded if only a specific cytopathologic subgroup (e.g., indeterminate cytology) or histopathology subgroup (e.g., follicular thyroid carcinoma) was evaluated. In addition, studies were excluded if they did not report the existence of NC-ATA nodules or explicitly recategorized them into another ATA USRSS category. Studies that recognized the presence of NC-ATA nodules but did not report their pathologic outcomes were excluded. In studies that included both <1 and >1 cm nodules, only data for >1 cm were included if these data were separately reported. Studies of <1 cm nodules were generally excluded, but to allow for the greatest number of relevant data after initial review, studies that had <15% of nodules <1 cm were considered to have an acceptably low proportion and were included.

When inclusion or exclusion criteria were unanswered in the article, personal correspondence to the authors was attempted. If additional data were collected from study authors, the study was assessed for inclusion/exclusion as aforementioned. Studies were not included if additional data could not be collected or if criteria could not be confirmed. The following data were extracted from the selected articles: study design and location, diagnostic criteria (histopathology vs. histopathology and definitive FNA), inclusion of subcentimeter nodules, number of patients and nodules investigated, proportion of nodules with indeterminate cytology and no surgical pathology that were excluded, proportion of NC-ATA nodules, proportional of nodules with histopathologic diagnosis, and RoM of NC-ATA nodules.

Risk of bias assessment

Risk of bias and assessment of applicability was conducted by two authors (D.K., R.R.M., and/or M.K.) with disagreements resolved by discussion between authors (T.E.A. and D.K.) using the revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool.[18]

Statistical analyses

Meta-analysis to obtain pooled estimates for the proportion of NC-ATA nodules that were malignant used the metaprop procedure in Stata (V17.0, College Station, TX, USA).[19] Each study summarized the number of NC-ATA nodules and the number of malignancies. The proportion of NC-ATA malignancies was computed for each study, with a 95% score-based confidence interval [CI] on binomial proportions. Random effects meta-analysis of proportions used the DerSimonian and Laird approach[20] on Freeman-Tukey double arcsine normalizing transformations of the study proportions; the study weights used the inverse variance of the transformed values. Between-study heterogeneity in malignancy proportions was assessed by I^2 statistics.

Prevalence of malignancy was also summarized by subgroups: study type (single- vs. multicentered); study design (prospective vs. retrospective); inclusion of <1 cm nodules; and geographic location (Asia, Europe, and America). To investigate sources of between-study heterogeneity in estimates of malignancy proportions, mixed effects logistic regression with a random study effect was conducted, testing for possible heterogeneity by single- versus multicentered studies, prospective versus retrospective study designs, diagnostic criteria used, total number of patients evaluated, total number of NC-ATA nodules evaluated, percentage of nodules histologically verified, and percentage of nodules that were excluded for because of indeterminate cytology without verified diagnosis. Continuous data were evaluated as dichotomous variables with cutoff at the median value.

Results

Study selection

As depicted in Figure 1, a total of 6377 titles/abstracts were screened after duplicate removal. Of those, 126 articles underwent full-text review. Reasons for study exclusion are shown in Supplementary Table S1. Ultimately, 16 articles fulfilled the inclusion and exclusion criteria and were analyzed for this study.[3–17,21] The authors of two articles were contacted for additional information, of which one author provided additional data that led to inclusion of the study.

Study characteristics

Study characteristics are summarized in Table 1. Of the included studies, 12 were retrospective analyses, 3 were prospective,[4,14,17] and 1 study reported on both retrospective and prospective cohorts.[6] Single-center analyses represented 11/16 (68.8%) of included studies[3,5–8,10–13,15,16,21] and most of the studies (11/16, 68.8%) reported data on populations from centers in Asia.[3,5,6,8,10–13,15,16,21] Six studies (37.5%) included a subset of <15% subcentimeter nodules,[4,7,14–16] and four studies (25%) reported only thyroid nodules that underwent surgical resection.[8,10,11,15]

Sample size varied from 167 to 3685 thyroid nodules, with a total pooled sample of 21,271 nodules. A total of 1872 nodules were classified as NC-ATA, representing 1.75–19.6%

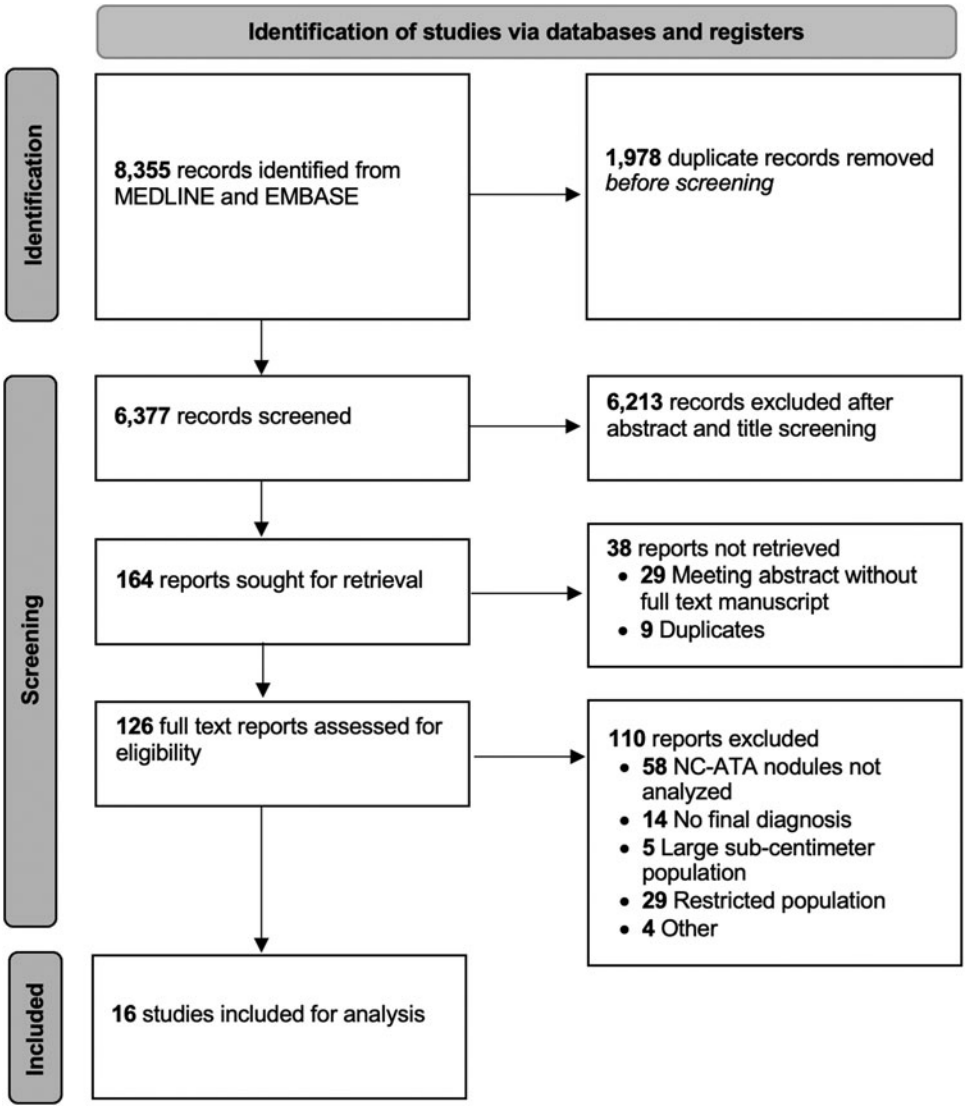

FIG. 1. PRISMA diagram. PRISMA, Preferred Reporting Items for a Systematic Review and Meta-Analysis.

Table 1. Characteristics of Included Studies

Study	*Country*	*Design*	*Diagnostic criteria*	*Includes nodules <1 cm*	*Total nodules*	*NC-ATA nodules,* n *(%)*	*NC-ATA RoM,* n *(%)*
Yoon et al.[12]	South Korea	Multicenter, retrospective	Surgery + FNA	No	1293	44 (3.4)	8 (18.2)
Middleton et al.[3]	United States	Multicenter, retrospective	Surgery + FNA	Yes	3422	477 (13.9)	45 (9.4)
Chnget al.[8]	Singapore	Single-center, retrospective	Surgery only	No	167	3 (1.8)	1 (33.3)
Ha et al.[16]	South Korea	Multicenter, retrospective	Surgery + FNA	Yes	586	53 (9.0)	4 (7.5)
Ha et al.[21]	South Korea	Multicenter, retrospective	Surgery + FNA	No	2000	100 (5.0)	19 (19.0)
Persichetti et al.[14]	Italy	Single-center, prospective	Surgery + FNA	Yes	987	31 (3.1)	6 (19.4)
Gao et al.[15]	China	Single-center, retrospective	Surgery only	Yes	2614	70 (2.7)	13 (18.6)
Grani et al.[4]	Italy	Single-center, prospective	Surgery + FNA	No	502	90 (17.9)	9 (10.0)
Ha et al.[13]	South Korea	Single-center, retrospective	Surgery + FNA	No	1938	34 (1.8)	7 (20.6)
Peng et al.[5]	China	Single-center, retrospective	Surgery + FNA	Yes	230	45 (19.6)	31 (68.9)
Slowinska-Klencka et al.[9]	Poland	Single-center, retrospective	Surgery + FNA	No	1000	51 (5.1)	16 (31.4)
Merhav et al.[17]	Israel	Single-center, prospective	Surgery + FNA	No	255	15 (5.9)	0 (0.0)
Paker et al.[10]	Israel	Single-center, retrospective	Surgery only	No	238	16 (6.7)	9 (56.3)
Kuru et al.[11]	Turkey	Single-center, retrospective	Surgery only	No	1143	87 (7.6)	55 (63.2)
Seifert et al.[7]	Germany	Multicenter, retrospective	Surgery + FNA	Yes	1211	135 (11.1)	16 (11.9)
Yang and Na[6]	South Korea	Single-center, retro- and prospective	Surgery + FNA	No	3685	621 (16.8)	62 (10.0)

ATA, American Thyroid Association; FNA, fine needle aspiration; NC-ATA, nonclassifiable; RoM, risk of malignancy.

of all nodules among studies. The random effects pooled prevalence was 7.8% [CI 5.1–11.1]. The criteria used to categorize nodules as NC-ATA were described in all publications. NC-ATA nodules were described as isoechoic with at least one high-risk feature, hyperechoic with at least one high-risk feature, mixed solid/cystic with at least one high-risk feature, or when composition, echogenicity or margins could not be determined. The RoM for NC-ATA nodules ranged from 0% to 68.9%.

Meta-analysis of malignancy rate

As shown in Figure 2, the overall pooled RoM estimate for NC-ATA nodules in the 16 included studies was 20.3% [CI 13.0–28.7]. There was significant heterogeneity among study estimates ($I^2 = 92.8\%$, $p < 0.001$). In subgroup analyses, single-center studies showed higher pooled RoM estimate of 24.8% [CI 13.0–38.7] compared with multicenter studies (12.3% [CI 8.3–16.8], $p = 0.031$; Fig. 3A). There was significant interstudy heterogeneity found across single-center studies ($I^2 = 94.4\%$, $p < 0.001$), but not multicenter studies ($I^2 = 57.1\%$, $p = 0.053$). Furthermore, the pooled RoM was higher in retrospective versus prospective studies (25.1% [CI 15.2–36.3] vs. 8.5% [CI 4.2–14.1], $p = 0.003$; Fig. 3B). Significant interstudy heterogeneity was seen among retrospective studies ($I^2 = 93.8\%$, $p < 0.001$), but not prospective studies ($I^2 = 47.6\%$, $p = 0.13$).

The pooled RoM for studies that only included nodules with surgical pathology was 42.9% [CI 13.6–74.9] compared with 16.1% [CI 10.6–22.3] for studies including both surgical and other nonsurgical outcomes (e.g., definitive biopsy result; trend at $p = 0.057$, Fig. 3C). Of note, studies that only included nodules >1 cm had similar pooled RoM (20.7% [CI 10.7–32.7]) to studies that included a subset of nodules <1 cm (20.0% [8.6–34.4]) (Fig. 3D). Finally, with respect to different geographic regions (Fig. 3E), we found no difference in pooled RoM by geographic region between studies from Asia (23.3% [CI 12.1–36.5]) versus Europe/America (14.4% [CI 8.5–21.1]; $p = 0.142$).

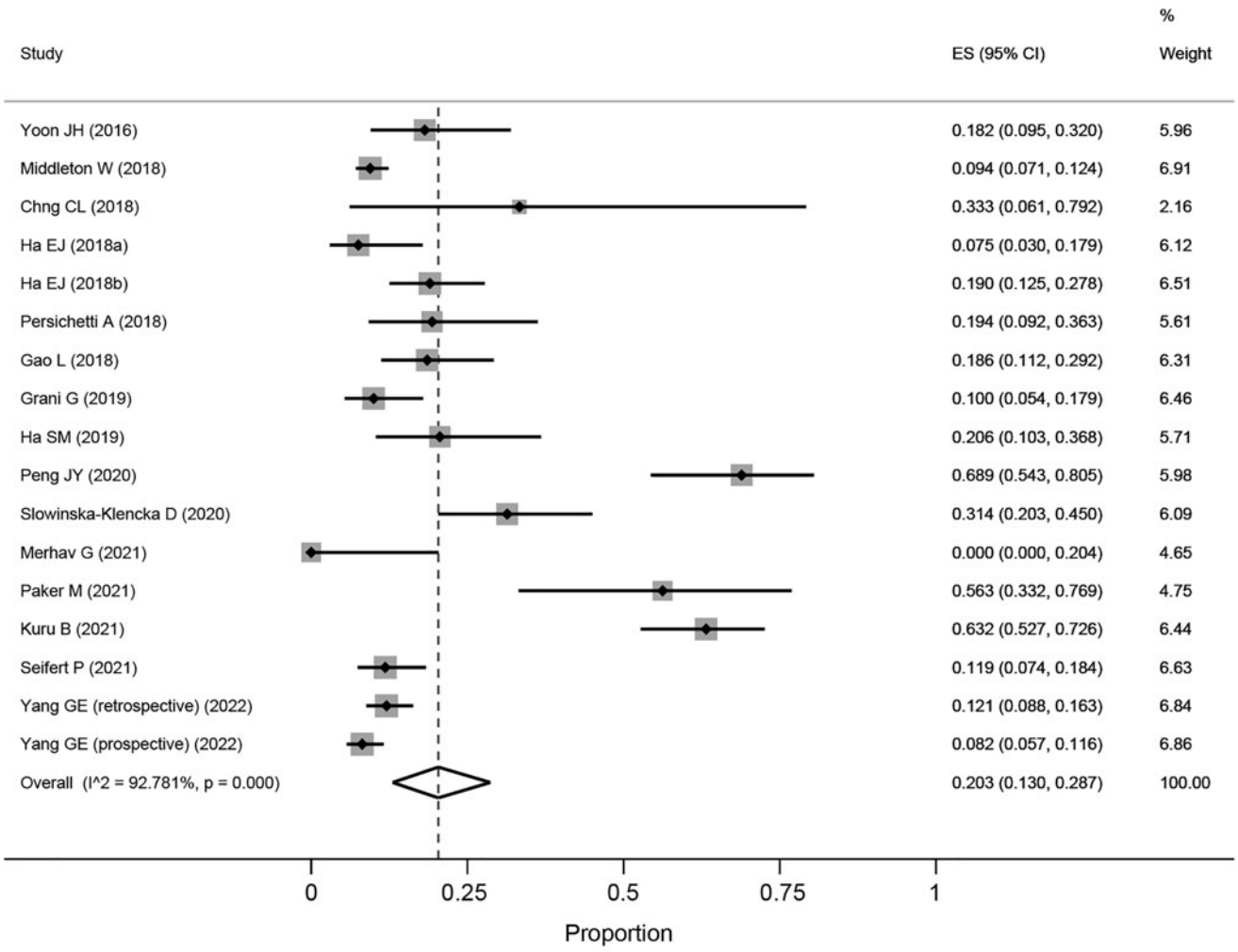

FIG. 2. Meta-analysis of NC-ATA nodule RoM. The RoM for thyroid nodules evaluated using the ATA USRSS with sonographic patterns deemed nonclassifiable (NC-ATA) from 16 included studies. The pooled RoM for NC-ATA nodules was 20.3% [CI 13.0–28.7], with significant heterogeneity identified among study estimates ($p < 0.001$). Ha 2018a[16] and Ha 2018b.[21] ATA, American Thyroid Association; CI, confidence interval; RoM, risk of malignancy; USRSS, ultrasound risk stratification system.

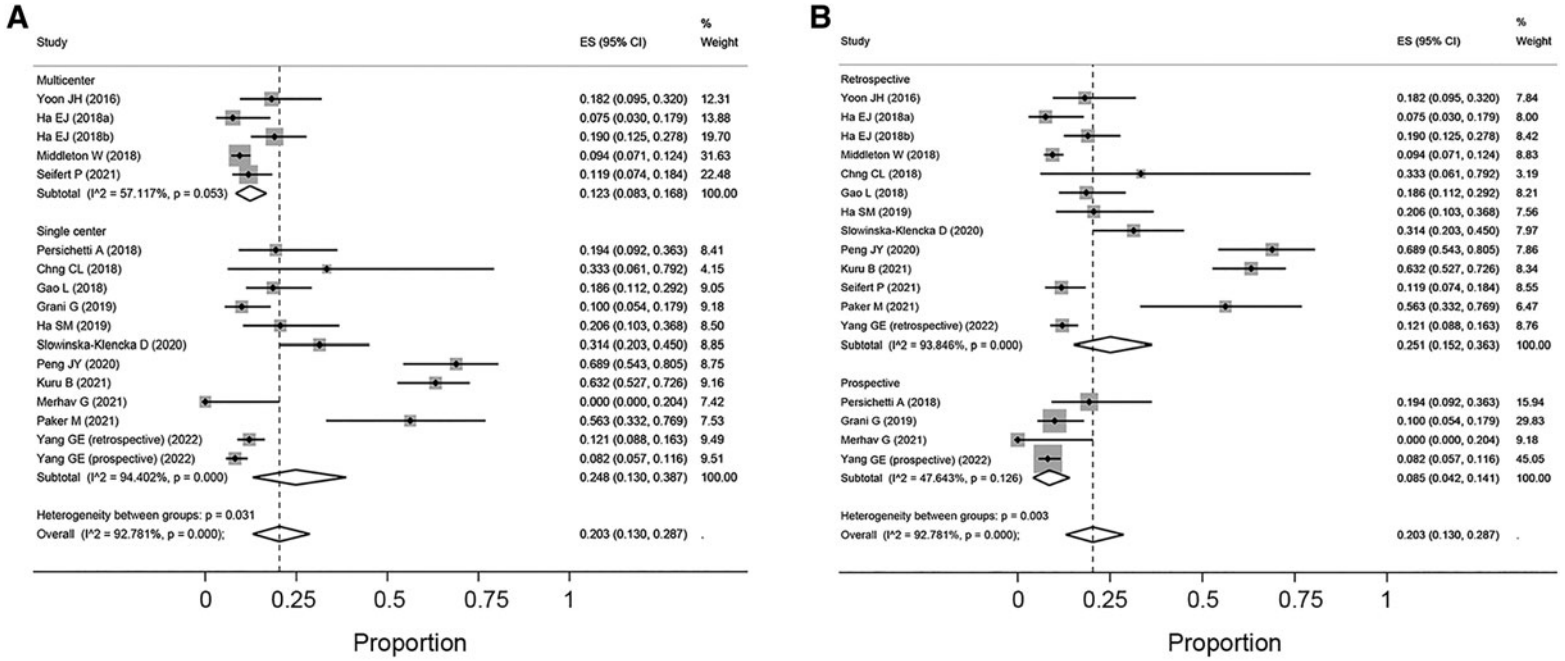

FIG. 3. Meta-analysis of RoM for NC-ATA nodules. Subanalyses of RoM for ATA USRSS nonclassifiable (NC-ATA) nodules for 16 included studies, stratified by (**A**) study centers: multicenter versus single-center studies; (**B**) study design: retrospective versus prospective; (**C**) diagnostic criteria: Histopathology + Cytology versus Histopathology alone; (**D**) nodule size inclusion: only ≥1 cm nodules versus inclusion of <1 cm nodules, and; (**E**) geographic location: Asian versus Europe/American sites.

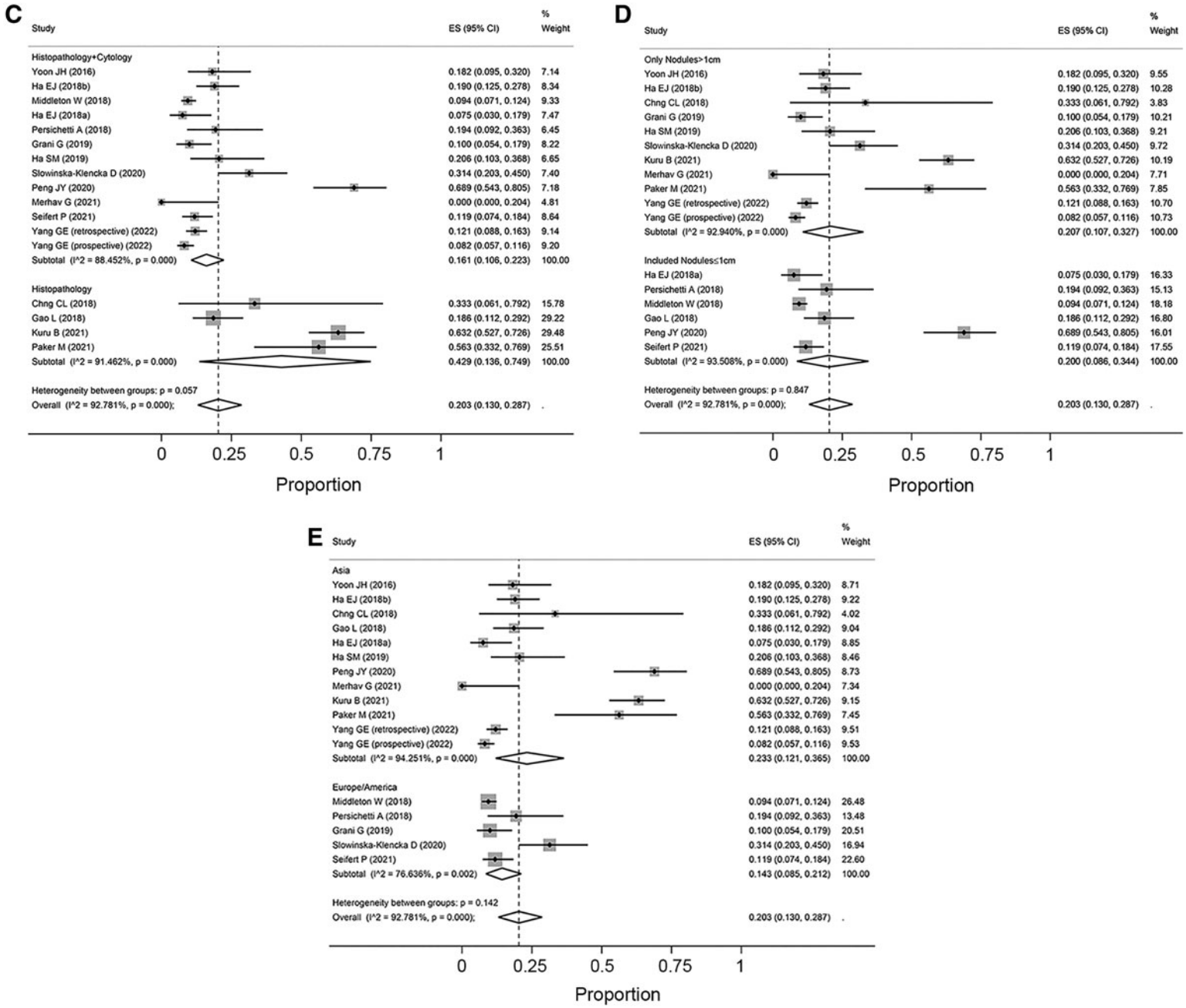

FIG. 3. (*Continued*).

Because of significant heterogeneity observed in the primary analysis, sensitivity analyses were performed to identify sources of between-study heterogeneity that contributed to differences in the RoM estimates. After removal of one study representing a statistical outlier,[5] the overall pooled RoM estimate for NC-ATA nodules in these studies was 19% [CI 13.0–25.0], with heterogeneity remaining significant ($I^2 = 83.0\%$, $p < 0.001$). Mixed effects logistic regression performed to evaluate study variables included single- versus multicentered studies, prospective versus retrospective design, study location (Europe/America vs. Asia), use of histology for diagnostic outcome, total and NC-ATA nodule sample size, nodule size threshold, and rate of exclusion due to unconfirmed nodule diagnosis.

This analysis showed that RoM was lower in multicenter versus single-center studies ($p = 0.049$) and in prospective versus retrospective studies ($p = 0.015$), while higher RoM was observed in studies with only surgically resected nodules ($p = 0.013$) and studies with >35% of nodules with histopathologic diagnosis ($p = 0.01$). Other variables did not show significant contribution to differences in RoM.

Quality assessment

The methodological quality and risk of bias in the included studies were assessed using the revised QUADAS-2 and are reported in Table 2. All studies selected patients with thyroid nodules who underwent evaluation and treatment and, therefore, inherently did not evaluate patients with nodules who did not undergo additional testing. Within the parameters of patients receiving treatment for thyroid nodules, most studies were considered low risk of impactful bias in patient selection, testing, and final diagnosis. We did note higher risk of bias for patient selection and patient selection applicability in Slowinska-Klencka et al.,[9] due to unclear recruitment of their patient cohort, as well as an unclear risk of bias in the index test for two studies in which blinding was not clearly stated.[8,11] The four studies that reported only nodules with surgical pathology were also graded to have higher risk of bias in patient selection.[8,10,11,15]

Discussion

Ultrasonography is a primary modality for assessment of thyroid nodule risk and guiding clinical management. Despite our clinical reliance upon ultrasound evaluation, there remain limitations to sonographic systems for nodule evaluation. Here we address the subset of nonclassifiable nodules within the widely used ATA USRSS system to attempt to better delineate their prevalence and malignancy risk through a comprehensive systematic review and meta-analysis comprising 16 studies and >20,000 nodules.

Importantly, we found that NC-ATA nodules are common with a pooled prevalence of 7.8% [CI 5.1–11.1] and portend a malignancy risk of 20.3% [CI 13.0–28.7]. This malignancy risk approximates the upper range assigned to ATA USRSS intermediate suspicion nodules (10–20%) and may overlap with the risk implied for the American College of Radiology Thyroid Imaging Reporting and Data System (ACR-TIRADS)[22] category 4, as well as some in category 5 (those with 7 points).[7,16,23] In subgroup analyses, risk estimates varied significantly by aspects of study design.

We found significantly lower RoM reported in multicentered (12.3%) and prospective studies (8.5%) compared with single-center and retrospective studies, respectively. No identified studies were both prospective and multicentered. Even if the true malignancy risk of NC-ATA nodules is closer to the lower rate seen in these subgroups, it is comparable with that assigned to intermediate suspicion nodules and represents a non-negligible risk to patients. These findings broadly indicate that FNA should be considered in patients with NC-ATA nodules in a manner comparable with those with an intermediate suspicion ATA sonographic pattern.

Significant heterogeneity observed among the studies should be considered when interpreting the pooled RoM estimates in this meta-analysis. The analysis of study characteristics to assess sources of heterogeneity revealed that lower NC-ATA nodule RoM was observed in multicenter studies and prospective studies, whereas higher RoM was found for studies that only used histopathology diagnosis and in those

Table 2. Quality Assessment of Included Studies (QUADAS-2)

	Risk of bias				*Applicability concerns*		
	Patient selection	*Index test*	*Reference standard*	*Flow and timing*	*Patient selection*	*Index test*	*Reference standard*
Yoon et al.[12]	LR	LR	LR	LR	LR	LR	LR
Middleton et al.[3]	LR	LR	LR	LR	LR	LR	LR
Chng et al.[8]	LR	UC	LR	LR	HR	LR	LR
Ha et al.[16]	LR	LR	LR	LR	LR	LR	LR
Ha et al.[21]	LR	LR	LR	LR	LR	LR	LR
Persichetti et al.[14]	LR	LR	LR	LR	LR	LR	LR
Gao et al.[15]	LR	LR	LR	LR	HR	LR	LR
Grani et al.[4]	LR	LR	LR	LR	LR	LR	LR
Ha et al.[13]	LR	LR	LR	LR	LR	LR	LR
Peng et al.[5]	LR	LR	LR	LR	LR	LR	LR
Slowinska-Klencka et al.[9]	HR	LR	LR	LR	HR	LR	LR
Merhav et al.[17]	LR	LR	LR	LR	LR	LR	LR
Paker et al.[10]	LR	LR	LR	LR	HR	LR	LR
Kuru et al.[11]	LR	UC	LR	LR	HR	LR	LR
Seifert et al.[7]	LR	LR	LR	LR	LR	LR	LR
Yang and Na[6]	LR	LR	LR	LR	LR	LR	LR

HR, high risk (dark gray boxes); LR, low risk (gray boxes); UC, unclear (light gray boxes).

that had a percentage of diagnosis from histopathology that was above the median (>35%). In addition, there may be other aspects of thyroid nodule assessment that were not reported in the studies that affected the individual RoMs.

Differences in thyroid ultrasound expertise could lead to variability in the criteria for classifying nodules as NC-ATA. If nodules with ATA low or ATA high suspicion sonographic phenotypes were included as NC-ATA nodules, this would lead to lower or higher RoM, respectively. Similarly, differences in cytological reporting and use of indeterminate (Bethesda III–V) categories, as well as how patients with indeterminate cytology were referred for surgical resection could lead to heterogeneity in the observed RoMs. These were not data available in the published studies. Although whether and how these variables could alter the pooled RoM estimates is not known, the majority of studies indicate standard of care criteria and clinical decision-making that likely would be applicable to general populations.

The strengths of this study include the large population of nodules derived from systematic review, well-defined criteria to include a broad and high-quality sample, and subanalyses to evaluate variables that could impact RoM in NC-ATA nodules. We acknowledge some limitations as well. The included studies only reported on nodules with a confirmed diagnosis, either by surgical histopathology or definitive biopsy (Benign [Bethesda II or Malignancy [Bethesda VI] cytology), which excluded indeterminate cytology nodules without histopathologic standard and, therefore, our results may include some selection bias.

Cytologically indeterminate nodules that were determined to be more suspicious may have been resected while less concerning indeterminate nodules were managed conservatively and not reported, which could result in a higher observed RoM than if all indeterminate nodules were included. No studies reported use of molecular testing. Thyroid malignancies other than PTC (e.g., follicular thyroid carcinoma and medullary thyroid carcinoma) may be under-reported in studies evaluating USRSSs that rely on confirmed diagnosis, perhaps due to less specific sonographic features and higher chances of indeterminate cytology in these less common thyroid cancers.[24]

In addition, in older studies, tumors now defined as noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP), which is an indolent nonmalignant tumor, may have been reported as follicular variant of PTC and been counted as malignant. Since NIFTP may present as a thyroid nodule with NC-ATA appearance,[25] the reported malignancy risk for NC-ATA nodules may be overestimated. Because the diagnosis of NIFTP can be made only upon histopathologic assessment, surgical resection is still accepted management for nodules that are ultimately determined to be NIFTP.

Moving forward, these data further support the development of more robust criteria for thyroid nodule evaluation. Previous studies have attributed much of the interobserver variability seen during use of the ATA USRSS system to the presence of nonclassifiable isoechoic or hyperechoic nodules with high-risk features.[26] USRSS that evaluate nodule features individually, such as ACR-TIRADS, are able to assess risk for any combination of nodule characteristics and, therefore, can categorize all nodules. In contrast, more holistic nodule stratification systems such as the ATA USRSS can help to avoid pitfalls of feature misclassification and subsequent inaccurate risk assessment. In future iterations of the ATA USRSS that seek into incorporate categorization of NC-ATA nodules, these meta-analysis data can inform proper malignancy risk.

Conclusion

Data from this meta-analysis provide a pooled estimate for the prevalence and RoM of NC-ATA nodules. Although significant heterogeneity observed between studies indicates the need for caution when interpreting these results, they indicate that the RoM is comparable with that of the current ATA USRSS intermediate suspicion category. This suggests that for NC-ATA nodules, use of similar criteria for performing FNA is likely warranted.

Authors' Contributions

Conceptualization (equal), methodology (equal), investigation (equal), data curation (lead), supervision, writing—original draft (lead), and writing—review and editing (equal) by D.K. Investigation (equal) and writing—original draft (supporting) by M.K. Formal analysis (lead) and writing—review and editing (supporting) by W.J.M. Investigation (equal) and writing—review and editing (supporting) by R.M.M. and E.J. Conceptualization (equal), methodology (equal), investigation (equal), supervision (lead), visualization (lead), writing—original draft (supporting), and writing—review and editing (equal) by T.E.A.

Disclaimer

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Funding Information

This study was supported by grants UL1TR001855 from the National Center for Advancing Translational Science (NCATS) of the U.S. National Institutes of Health.

Supplementary Material

Supplementary Data
Supplementary Table S1

References

1. Davies L, Welch HG. Current thyroid cancer trends in the United States. JAMA Otolaryngol Head Neck Surg 2014; 140(4):317–322; doi: 10.1001/jamaoto.2014.1
2. 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
3. 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(5):1148–1154; doi: 10.2214/AJR.17.18822

4. 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
5. Peng JY, Pan FS, Wang W, et al. Malignancy risk stratification and FNA recommendations for thyroid nodules: A comparison of ACR TI-RADS, AACE/ACE/AME and ATA guidelines. Am J Otolaryngol 2020;41(6):102625; doi: 10.1016/j.amjoto.2020.102625
6. Yang GE, Na DG. Impact of the ultrasonography assessment method on the malignancy risk and diagnostic performance of five risk stratification systems in thyroid nodules. Endocrine 2022;75(1):137–148; doi: 10.1007/s12020-021-02795-x
7. Seifert P, Schenke S, Zimny M, et al. Diagnostic Performance of Kwak, EU, ACR, and Korean TIRADS as Well as ATA Guidelines for the ultrasound risk stratification of non-autonomously functioning thyroid nodules in a region with long history of iodine deficiency: A German Multicenter Trial. Cancers (Basel) 2021;13(17):4467; doi: 10.3390/cancers13174467
8. Chng CL, Tan HC, Too CW, et al. Diagnostic performance of ATA, BTA and TIRADS sonographic patterns in the prediction of malignancy in histologically proven thyroid nodules. Singapore Med J 2018;59(11):578–583; doi: 10.11622/smedj.2018062
9. Słowińska-Klencka D, Wysocka-Konieczna K, Klencki M, et al. Diagnostic Value of Six Thyroid Imaging Reporting and Data Systems (TIRADS) in cytologically equivocal thyroid nodules. J Clin Med 2020;9(7):2281; doi: 10.3390/jcm9072281
10. Paker M, Goldman T, Masalha M, et al. A comparison of two widely used risk stratification systems for thyroid nodule sonographic evaluation. Isr Med Assoc J 2021;23(11):714–719.
11. Kuru B, Kefeli M, Danaci M. Comparison of 5 Thyroid Ultrasound Stratification Systems for differentiation of benign and malignant nodules and to avoid biopsy using histology as reference standard. Endocr Pract 2021;27(11):1093–1099; doi: 10.1016/j.eprac.2021.04.411
12. 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(3):917–924; doi: 10.1148/radiol.2015150056
13. 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
14. Persichetti A, Di Stasio E, Guglielmi R, et al. Predictive value of malignancy of thyroid nodule ultrasound classification systems: A prospective study. J Clin Endocrinol Metab 2018;103(4):1359–1368; doi: 10.1210/jc.2017-01708
15. Gao L, Xi X, Wang J, et al. Ultrasound risk evaluation of thyroid nodules that are ''unspecified'' in the 2015 American Thyroid Association management guidelines: A retrospective study. Medicine (Baltimore) 2018;97(52):e13914; doi: 10.1097/MD.0000000000013914
16. Ha EJ, Na DG, Moon WJ, et al. 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; doi: 10.1089/thy.2018.0094
17. Merhav G, Zolotov S, Mahagneh A, et al. Validation of TIRADS ACR Risk Assessment of Thyroid Nodules in Comparison to the ATA Guidelines. J Clin Imaging Sci 2021;11:37; doi: 10.25259/JCIS_99_2021
18. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155(8):529–536; doi: 10.7326/0003-4819-155-8-201110180-00009
19. Nyaga VN, Arbyn M, Aerts M. Metaprop: A Stata command to perform meta-analysis of binomial data. Arch Public Health 2014;72(1):39; doi: 10.1186/2049-3258-72-39
20. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177–188; doi: 10.1016/0197-2456(86)90046-2
21. 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
22. 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; doi: 10.1016/j.jacr.2017.01.046
23. Middleton WD, Teefey SA, Reading CC, et al. Multiinstitutional analysis of thyroid nodule risk stratification using the American College of Radiology Thyroid Imaging Reporting and Data System. AJR Am J Roentgenol 2017;208(6):1331–1341; doi: 10.2214/AJR.16.17613
24. Trimboli P, Castellana M, Piccardo A, et al. The ultrasound risk stratification systems for thyroid nodule have been evaluated against papillary carcinoma. A meta-analysis. Rev Endocr Metab Disord 2021;22(2):453–460; doi: 10.1007/s11154-020-09592-3
25. Liu R, Gao L, Xia Y, et al. Three ultrasound phenotypes of non-invasive follicular thyroid neoplasm with papillary-like nuclear features proposed for imaging-pathology analysis: single center experience. Gland Surg 2021;10(1):307–318; doi: 10.21037/gs-20-612
26. Pang Z, Margolis M, Menezes RJ, et al. Diagnostic performance of 2015 American Thyroid Association guidelines and inter-observer variability in assigning risk category. Eur J Radiol Open 2019;6:122–127; doi: 10.1016/j.ejro.2019.03.002

Address correspondence to:
Trevor E. Angell, MD
Division of Endocrinology and Diabetes
Keck Medicine of the University of Southern California
2020 Zonal Avenue, IRD620
Los Angeles, CA 90033
USA

E-mail: trevor.angell@med.usc.edu