RESEARCH

Echogenicity as a standalone nodule characteristic is not inferior to the TIRADS systems in the 10–20 mm nodule diameter range in patient selection for fine needle aspiration: a pilot study

Karoly Rucz[1], Laszlo Hegedűs[2], Steen Joop Bonnema[2], Andrea Frasoldati[3], Laszlo Jambor[4], Gabor Laszlo Kovacs[5], Enrico Papini[6], Gilles Russ [ID][7], Zsolt Karanyi[8], Endre V Nagy [ID][8] and Tamas Solymosi[8]

[1]Department of Endocrinology, Siofok Hospital, Siofok, Hungary

[2]Department of Endocrinology, Odense University Hospital, Odense, Denmark

[3]Endocrinology Unit of Arcispedale S. Maria Nuova, Reggio Emilia, Italy

[4]Institute of Medical Imaging, Faculty of Medicine, University of Debrecen, Debrecen, Hungary

[5]1st Department of Medicine, Flohr Ferenc Hospital, Kistarcsa, Hungary

[6]Regina Apostolorum Hospital in Albano, Rome, Italy

[7]Unité Thyroïde et Tumeurs Endocrines - Pr Leenhardt Hôpital La Pitie Salpetriere, Sorbonne Université, Paris, France

[8]Division of Endocrinology, Department of Medicine, Faculty of Medicine, University of Debrecen, Debrecen, Hungary

Correspondence should be addressed to E V Nagy: nagy.endre@med.unideb.hu

Abstract

Objective: The ultrasound evaluation of thyroid nodules (TNs) in patient selection for fine needle aspiration (FNA) requires both uniformly accepted definitions of each nodule characteristic and extensive experience from the examiner. We hypothesized that nodule echogenicity alone may provide comparable performance to more complex approaches in patient selection for FNA.

Patients and methods: Seven highly experienced investigators from four countries evaluated, online, the ultrasound (US) video recordings of 123 histologically verified TN by answering 17 nodule characteristics-related questions. The diagnostic performances of five TN image reporting and data systems (TIRADS) were compared to making decisions based solely on the echogenicity of the nodule for indicating FNA in 110 nodules ≥10 mm.

Results: In the 10–20 mm size range, the sensitivities and specificities of the five TIRADS systems in identifying malignant nodules were 80.5–91.0% and 31.4–50.9%, respectively. Had FNA been recommended for all hypoechoic nodules, disregarding other US characteristics, comparable sensitivity and specificity (87.5% and 43.4%, respectively) were obtained. Compared to nodules >20 mm, a higher proportion of cancers were hypoechoic in the 10–20 mm size range (87.2% vs 77.8%, *P* = 0.05). In the 10–20 mm size range, compared to hypoechoic nodules, a significantly lower proportion of isoechoic nodules demonstrated suspicious findings (70.7% vs 30.0%, *P* < 0.05).

Conclusion: In contrast to >20 mm diameter nodules, the recommendation of FNA may rely on a single US feature, echogenicity, in the 10–20 mm size range. If independently confirmed in larger cohorts, this may simplify nodule evaluation in this size range.

Keywords: echogenicity; ultrasound; thyroid nodule; TIRADS; fine needle aspiration

Introduction

The introduction of thyroid ultrasound (US) decades ago has altered the management of patients with thyroid nodules (TN). Since long it has been acknowledged that certain US characteristics, which include hypoechogenicity, microcalcifications, taller-than-wide shape, irregular margins, and extrathyroidal growth (1, 2, 3, 4, 5, 6, 7, 8), occur significantly more often in malignant compared with benign TN (1, 2, 9, 10, 11). On the other hand, none of these features alone, including the most sensitive one, hypoechogenicity, has proved to be sensitive enough to guide the clinical decision for or against performing fine needle aspiration (FNA) cytology in a given patient (3, 12, 13).

Since 2006, all guidelines have suggested the use of composite analyses of TN characteristics when indicating FNA. The basis of the thyroid nodule image reporting and data systems (TIRADS), introduced a decade later, is essentially the same (14). The main novelty of the various systems was the scoring of TNs according to their US characteristics, and the suggestion of different size limits for indicating FNA in the various TIRADS categories (9, 15, 16, 17, 18). Although only three of the five most commonly used classification systems use the TIRADS acronym (16, 17, 18), we will use this abbreviation for the other two (9, 15) as well.

All TIRADS share the view that FNA may be abandoned in <10 mm nodules, and all but one suggest diagnostic FNA for all nodules >20 mm, except for two rare forms of thyroid cysts (9, 15, 16, 17, 18). Thus, it is almost exclusively in the 10–20 mm size range where TIRADS impacts nodule management (13, 19, 20).

A limited number of publications compare the diagnostic performance of different TIRADS in surgically treated patients with histology proven diagnoses (21, 22, 23, 24). In this respect, the use of the FNA results as the reference standard for US performance calculation has limitations, as up to 30% of FNAs are either repeatedly non-diagnostic (Bethesda 1) or result in Bethesda 3 and 4 categories (25, 26). Of the individual US characteristics, even the most sensitive one, hypoechogenicity, has been found to have insufficient sensitivity when deciding on FNA (3, 12). However, we are not aware of any studies comparing the performance of echogenicity as a standalone nodule characteristic to that of the elaborate TIRADS in distinct nodule diameter ranges, i.e. <10 mm, 10–20 mm, and >20 mm. We hypothesized that, in the 10–20 mm diameter range, nodule echogenicity alone may provide comparable performance to the more complex approaches in patient selection for FNA.

Methods

The study design has been described in previous publications (27, 28). The same dataset (27, 28) has been re-analyzed for the current purpose. To summarize, US videos of 47 consecutively operated malignant and 76 consecutively operated benign TN cases (*n* = 123) were analyzed by seven investigators, with at least 15 years of experience in thyroid US, from European thyroid centers (SB, AF, LJ, GK, EP, KR, and GR). The number of cases (nodules) is 123. However, in the results section, the total is 110; this difference is attributed to nodules that were smaller than 10 mm and therefore excluded. The investigators were aware of clinical, laboratory, and palpation data but were unaware of the histopathological results. After viewing the US video of a patient, an online case report form (CRF) composed of 17 questions (Supplementary Table 1, see section on supplementary materials given at the end of this article) was used to report the findings. The CRF data enabled the electronic generation of five TIRADS scores: the American Thyroid Association TIRADS (ATA-T) (15), the American College of Radiology TIRADS (ACR-T) (16), the American Association of Clinical Endocrinologists/ American College of Endocrinology/Associazione Medici Endocrinologi TIRADS (AACE-T) (9), the European Thyroid Association TIRADS (EU-T) (17) and the Korean Society of Thyroid Radiology TIRADS (KSTR-T) (18) scores. The individual scores were automatically translated into recommendations 'for' or 'against' FNA according to the criteria of the respective TIRADS.

For comparison, nodule echogenicity (ECHOG) as a standalone determinant, based on the investigators' answers to a single question (question 13 of the CRF with the following choices: A. hyperechoic or isoechoic, B. hypoechoic, C. very hypoechoic, D. anechoic, E. cannot be determined), was used for the FNA recommendation. Thus, if the nodule was considered hypoechoic (responses B or C), FNA was 'recommended,' while for non-hypoechoic nodules, FNA was 'not recommended.' If the investigator was not able to assign a TN to any echogenicity category, it was considered non-hypoechoic for the purpose of the current analysis. In accordance with current guidelines, FNA was not indicated for pure cysts and spongiform cysts, irrespective of the echogenicity grade assigned by the examiner. The

investigators knew that comparison of the TIRADS was among the aims of the study, while they were not aware of the planned comparative analyses of echogenicity readings.

The diagnostic performance of the five different TIRADS and the ECHOG approach was compared and analyzed separately for the <10 mm, 10–20 mm, and >20 mm nodule diameter ranges. Sensitivities, specificities, and the 95% CI of the approaches tested were calculated by package 'epiR' (29) in the R project (30). For comparisons, Fisher's exact test was used.

Written consent was obtained from each patient after a full explanation of the purpose and nature of all procedures used. The study protocol was approved by the Regional and Institutional Ethics Committee of the University of Debrecen (DE RKEB/IKEB 5350-2019).

Results

Comparison of the ECHOG and TIRADS approaches

Lesions between 10 mm and 20 mm
There were 44 nodules with a maximum diameter in the 10–20 mm category (mean nodule size 14.8 ± 3.1 mm, range 10.1–19.6 mm).

In the 10–20 mm category (Fig. 1), the average sensitivity of the five TIRADS was 85.6% (ranging from 80.5% to 91.0%), while the average specificity was 41.9% (ranging from 31.4% to 50.9%). ECHOG performed comparably to the five TIRADS tested; its sensitivity and specificity were 87.2% and 43.4%, respectively, which was not statistically different from the mean of the five TIRADS. There was no TIRADS in which both the sensitivity and
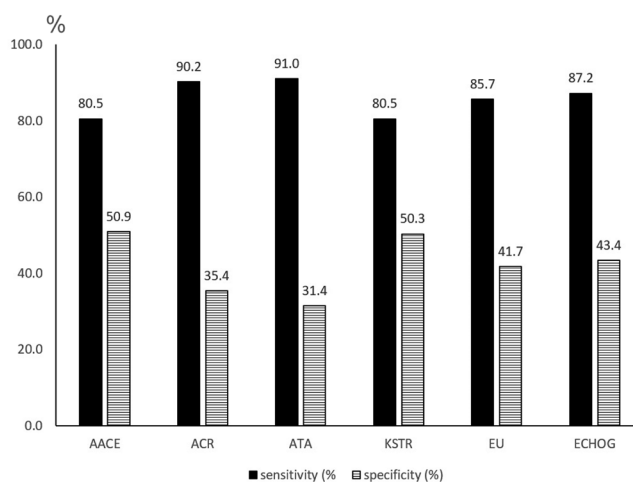
specificity were better or worse compared with each other. In 5.8% of cases (18/308 investigator responses; 44 nodules, 7 investigators), echogenicity could not be judged (response chosen: 'cannot be determined' to Question 13 in the CRF; Supplementary Table 1). If these responses are included among hypoechoic nodules, the sensitivity of echogenicity-based decision-making for FNA increases to 92.5%, while the specificity decreases to 37.1%.

If we exclude the three cases of follicular variant of papillary cancer (FV-PTC) and the single follicular thyroid cancer (FTC) case from the analysis of 10–20 mm nodules, the mean sensitivity of the five TIRADS would be 88.2%, and that of ECHOG would be 91.4%.

Lesions larger than 20 mm
There were 66 lesions >20 mm (mean nodule size: 37.4 ± 14.4 mm, range: 20.3–85.1 mm).

In this category, the mean sensitivity of the five TIRADS was 95.7% (ranging from 92.9% to 100%), while the average specificity was 9.4% (ranging from 0% to 18.4%) (Fig. 2). The sensitivity of ECHOG was 77.8%. Furthermore, 28.6% and 4.3% of cancers would not be recognized using the ECHOG and the mean of the five TIRADS, respectively ($P < 0.01$). On the other hand, ECHOG had higher specificity (59.8%) than the mean of the five TIRADS (9.3%) ($P < 0.01$).

In three cases of thyroid malignancy, where the US appearance was judged to be spongiform, one, three, and five (of the seven) examiners judged the nodule to be spongiform. This was the cause of the 7.1% false-negative rate in the EU-T and KSTR-T. The ACR-T does not recommend FNA for TIRADS 2 nodules (mixed, iso/hyperechoic nodules without suspicious signs) and TIRADS three nodules smaller than 25 mm, which is relevant for six and three cases, respectively.

If we did not include the two cases of FV-PTC and the single FTC case in the analysis of >20 mm nodules, then
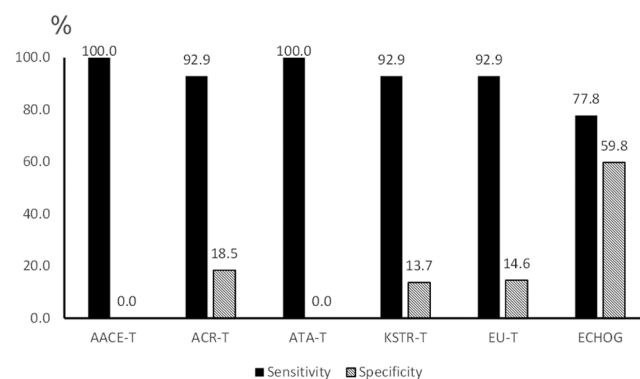

Figure 1

Diagnostic performance of five TIRADS and echogenicity alone in 10–20 mm nodules. Abbreviations refer to the respective TIRADS. ECHOG, echogenicity.

Figure 2

Diagnostic performance of five TIRADS and echogenicity alone in nodules larger than 20 mm. Abbreviations refer to the respective TIRADS. ECHOG, echogenicity.

the mean sensitivity of the five TIRADS would be 94.9% and that of ECHOG would be 79.0%.

All lesions larger than 10 mm combined

Analyzing the 10–20 mm and larger than 20 mm nodules together ($n = 110$), the mean nodule size was 28.4 ± 15.8 mm (range: 10.1–85.1 mm). The mean diagnostic sensitivity of the five TIRADS was 90.5%, while the mean specificity was 20.5% (Fig. 3). Compared to the mean of the five TIRADS studied, ECHOG had lower sensitivity (82.6%; $P < 0.01$) and higher specificity (54.2%; $P < 0.01$) (Fig. 3).

Nodule size and selected ultrasound characteristics which are cornerstones of TIRADS

The presence of cystic degeneration and the prevalence of any of the three generally used signs of suspicion (microcalcification, irregular border, and irregular shape) did not differ in malignant lesions between 10 and 20 mm and larger than 20 mm (Fig. 4). A significantly higher proportion of malignant nodules was hypoechoic in the 10–20 mm range (87.2%) compared with nodules >20 mm (77.8%; $P < 0.05$).

As for the level of ECHOG, a significantly higher proportion of very hypoechoic nodules presented with any of the three generally used signs of suspicion (microcalcification, irregular border, and irregular shape) than mildly to moderately hypoechoic nodules (79.6% vs 56.9, $P < 0.001$). A significantly higher proportion of mildly–moderately hypoechoic nodules than isoechoic nodules (56.9% vs 27.3%, $P < 0.001$; Table 1), presented with any of the three generally used signs of suspicion (microcalcification, irregular border, and irregular shape).

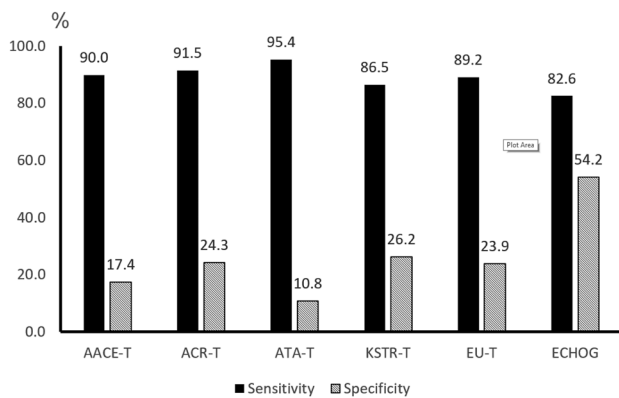Using any of the five TIRADS would have resulted in not recommending FNA in histologically verified thyroid

Figure 3

Diagnostic performance of the five TIRADS and echogenicity alone in nodules 10 mm and above. Abbreviations refer to the respective TIRADS. ECHOG, echogenicity.
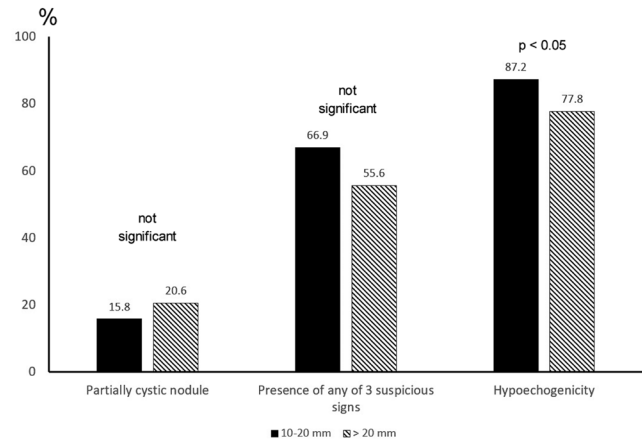

Figure 4

Malignant lesions at histology; prevalence of certain ultrasound characteristics in relation to nodule size. Suspicious signs include irregular borders, irregular shape, and microcalcifications. A significantly higher proportion of malignant nodules was hypoechoic in the 10–20 mm range (87.2%) compared with nodules >20 mm (77.8%; $P < 0.05$). There were no significant differences in the other comparisons of characteristics.

malignancies lacking all three generally used signs of suspicion (microcalcification, irregular border, and irregular shape) in 24 mildly–moderately hypoechoic and 34 hypoechoic (mildly–moderately and deeply hypoechogenic combined) nodules, respectively (Table 2). Using TIRADS would have resulted in the recognition of an additional three cancers among isoechoic nodules presenting with at least one of the three generally used signs of suspicion.

In the 10–20 mm nodule size range, the investigators found 64/133 and 34/177 microcalcifications in malignant and benign nodules, respectively. Punctate echogenic foci (microcalcifications and short comet tail artifacts combined as defined by ACR-T) were found in 41 benign and 78 malignant TNs.

Discussion

In order to test whether a simplification of the US evaluation of TN is possible without loss of sensitivity

Table 1 The prevalence of suspicious findings in thyroid cancers in relation to nodule echogenicity and size. Data are presented as *n* / total *n* (%).

	10–20 mm	>20 mm	Both sizes combined
Isoechoic	3/10 (30.0%)	6/23 (26.1%)	9/33 (27.3%)
Mildly–moderately hypoechoic	33/57 (57.9%)	33/59 (55.9%)	66/116 (56.9%)
Very hypoechoic	49/59 (83.1%)	29/39 (74.4%)	78/98 (79.6%)

Table 2 Cancers presenting as hypoechoic nodules in the 10–20 mm range, with no other signs of suspicion. Recommendation of FNA based on TIRADS. The size categories of different TIRADS are, by definition, different. Italics indicate that the given TIRADS defines a 15 mm or larger size limit for performing FNA.

	AACE-T	ACR-T	ATA-T	EU-T	KSTR-T
Mildly–moderately hypoechoic	>20 mm			*>15 mm*	*≥15 mm*
Solid, *n* = 20		*≥15 mm*	≥10 mm		
Partially cystic, *n* = 4		≥25 mm	≥20 mm		
Very hypoechoic	>10 mm	*≥15 mm*		>10 mm	*≥15 mm*
Solid, *n* = 9			≥10 mm		
Partially cystic, *n* = 1			≥20 mm		

and specificity in the diagnosis of thyroid malignancy, we have compared the diagnostic performance of five TIRADS systems with US echogenicity as a standalone feature in TN.

Although some US features are independent predictors of malignancy, it is widely accepted that any single US feature, including the most sensitive, hypoechogenicity, does not have both high enough sensitivity and specificity for the reliable detection of malignancies (18, 31, 32, 33, 34, 35). Our study confirmed this; if all nodules, irrespective of their size, are considered, the mean sensitivity of the five TIRADS systems in indicating FNA was 90.8%, while if we had performed FNA only in hypoechoic nodules, the sensitivity would be 82.6% (27). This difference is seemingly small. However, if nodules >20 mm are analyzed separately, ECHOG is inferior to TIRADS, since nearly 30% of cancers would escape US diagnosis, while this proportion is only 5% using TIRADS.

Our study focuses on nodules measuring 10–20 mm, the very size range where TIRADS does influence FNA decisions. Several novel findings merit attention. Most importantly, nodule hypoechogenicity has high sensitivity in this size range. This does not apply to the other TIRADS-determining descriptors. Secondly, the smaller the nodule, the higher the likelihood of malignant lesions being hypoechoic. This suggests that in the 10–20 mm size range, nodule echogenicity alone may guide the decision to offer FNA. Also, refraining from FNA in hypoechoic nodules in this size range carries a substantial risk of overlooking thyroid malignancy. Thirdly, the three suspicious characteristics (microcalcifications, irregular borders, and shape) are significantly less frequent in isoechoic than in mildly–moderately hypoechoic, and less frequent in mildly–moderately hypoechoic than in very hypoechoic malignant nodules; this is in accordance with the findings of other groups (32, 36, 37). In other words, where signs of suspicion are much needed, they are less often present, resulting in a high likelihood of overlooking malignancy in mildly–moderately hypoechoic nodules. In the 10–15 mm range, no other TIRADS than the ATA-T suggests FNA in mildly–moderately hypoechoic nodules lacking other characteristics of suspicion. Unfortunately, 30% of cancers in this category lack suspicious characteristics (27) and would remain unrecognized if any of the other

four TIRADS were used to indicate FNA. It is important to note that there was no single TIRADS that performed better in terms of both sensitivity and specificity simultaneously than any other. This means that the diagnostic performance of the mean of the five TIRADS was not influenced by the strength or weakness of any of the individual scoring systems.

Performing FNA in all hypoechoic nodules in the 10–20 mm range, but not in non-hypoechoic ones, resulted in comparable sensitivity and specificity compared to the more complex TIRADS-based decision for FNA. Adding merit to such an approach, interobserver agreement is good in discriminating between hypoechoic and isoechoic nodules, but is markedly weaker in differentiating between very hypoechoic and mildly–moderately hypoechoic nodules, as well as in the evaluation of the presence of microcalcifications and irregular nodule borders (27). One explanation might be that suspicious features, such as microcalcifications, irregular borders, and shape (taller-than-wide), are harder to discern as the nodule size decreases, despite the availability of US transducers operating at 18 MHz and above. If correct, one might expect ECHOG alone to be even more helpful with nodules <10 mm, but this was not assessed in our study, as FNA and surgery are rarely and almost never performed in hypoechoic and non-hypoechoic subcentimeter nodules, respectively. One may argue that the higher proportion of malignant hypoechoic nodules in the 10–20 mm range, compared to those >20 mm, may be due to the malignant outcomes of iso-echoic follicular lesions (FTC and FV-PTC) >20 mm. Although it had a minor influence on the sensitivity of TIRADS and ECHOG, the basic findings (i.e. ECHOG-based decision works in lesions between 10 and 20 mm but does not in nodules larger than 20 mm) remained the same after the exclusion of FTC and FV-PTC cases.

Strengths of our work include the use of video analysis instead of still images, the participation of ultrasound experts with different educational backgrounds, and the surgical–histopathological verification of all nodules. The relatively small number of cases (*n* = 110) studied is a strong limitation of the study, and therefore there is a need for confirmation of our findings in larger cohorts. Benign nodules evaluated only by US with or without FNA, and not resected surgically were not included,

which may result in falsely high sensitivity values for all methods in all comparisons.

We conclude that in the 10–20 mm TN range, in contrast to the >20 mm diameter range, the decision for FNA to detect thyroid malignancy may rely on a single US feature, echogenicity. All hypoechogenic nodules in this size range are candidates for FNA, and the consideration of suspicious signs in non-hypoechoic nodules is of less importance. If independently confirmed in larger cohorts, this may simplify the evaluation of 10–20 mm nodules. Currently, there is no general agreement on a 'tolerable' level of missed cancers or an accepted level of FNA rate in benign lesions. Therefore, we believe that the complex analysis of all nodule characteristics remains the basis of an optimal individual decision.

Author contribution statement

KR – investigator performing analysis of the ultrasound videos (equal), conceptualization (equal), methodology (equal), writing – review and editing (lead). LH – conceptualization (equal), review and editing (equal). SJB – investigator performing analysis of the study videos (equal). AF – investigator performing analysis of the study videos (equal). LJ, GLK, EP, GR – investigator performing analysis of the study videos (equal). ZK – methodology (equal), review and editing (equal); statistical planning and analysis (lead). EVN – conceptualization (equal), methodology (equal), review and editing (equal). TS – conceptualization (lead), methodology (equal), review and editing (equal).

References

1   Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Sherman SI & Tuttle RM. Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid* 2006 **16** 109–142. (https://doi.org/10.1089/thy.2006.16.109)

2   Gharib H, Papini E, Valcavi R, Baskin HJ, Crescenzi A, Dottorini ME, Duick DS, Guglielmi R, Hamilton CR Jr, Zeiger MA, *et al.* American Association of Clinical Endocrinologists and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules. *Endocrine Practice* 2006 **12** 63–102. (https://doi.org/10.4158/EP.12.1.63)

3   Papini E, Guglielmi R, Bianchini A, Crescenzi A, Taccogna S, Nardi F, Panunzi C, Rinaldi R, Toscano V & Pacella CM. Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound and color-Doppler features. *Journal of Clinical Endocrinology and Metabolism* 2002 **87** 1941–1946. (https://doi.org/10.1210/jcem.87.5.8504)

4   Tollin SR, Mery GM, Jelveh N, Fallon EF, Mikhail M, Blumenfeld W & Perlmutter S. The use of fine-needle aspiration biopsy under ultrasound guidance to assess the risk of malignancy in patients with a multinodular goiter. *Thyroid* 2000 **10** 235–241. (https://doi.org/10.1089/thy.2000.10.235)

5   Frates MC, Benson CB, Doubilet PM, Kunreuther E, Contreras M, Cibas ES, Orcutt J, Moore FD Jr, Larsen PR, Marqusee E, *et al.* Prevalence and distribution of carcinoma in patients with solitary and multiple thyroid nodules on sonography. *Journal of Clinical Endocrinology and Metabolism* 2006 **91** 3411–3417. (https://doi.org/10.1210/jc.2006-0690)

6   Orell SR & Philips J. The thyroid. Fine-needle biopsy and cytological diagnosis of thyroid lesions. *Monographs in Clinical Cytology* 1997 **14** I–IX, 1–205.

7   Moon WJ, Baek JH, Jung SL, Kim DW, Kim EK, Kim JY, Kwak JY, Lee JH, Lee JH, Lee YH, *et al.* Ultrasonography and the ultrasoundbased management of thyroid nodules: consensus statement and recommendations. *Korean Journal of Radiology* 2011 **12** 1–14. (https://doi.org/10.3348/kjr.2011.12.1.1)

8   Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Pacini F, Schlumberger M, *et al.* Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid* 2009 **19** 1167–1214. (https://doi.org/10.1089/thy.2009.0110)

9   Gharib H, Papini E, Garber JR, Duick DS, Harrell RM, Hegedüs L, Paschke R, Valcavi R & Vitti P. American Association of Clinical Endocrinologists, American College of Endocrinology, and Associazione Medici Endocrinologi Medical Guidelines for clinical practice for the diagnosis and management of thyroid nodules-2016 update. *Endocrine Practice* 2016 **22** 622–639. (https://doi.org/10.4158/EP161208.GL)

10 Hegedüs L. Clinical practice. The thyroid nodule. *New England Journal of Medicine* 2004 **351** 1764–1771. (https://doi.org/10.1056/NEJMcp031436)

11 Burman KD & Wartofsky L. Thyroid nodules. *New England Journal of Medicine* 2016 **374** 1294–1295. (https://doi.org/10.1056/NEJMc1600493)

12 Mandel SJ. Diagnostic use of ultrasonography in patients with nodular thyroid disease. *Endocrine Practice* 2004 **10** 246–252. (https://doi.org/10.4158/EP.10.3.246)

13 Durante C, Hegedüs L, Czarniecka A, Paschke R, Russ G, Schmitt F, Soares P, Solymosi T & Papini E. 2023 European Thyroid Association Clinical Practice Guidelines for thyroid nodule management. *European Thyroid Journal* 2023 **12**. (https://doi.org/10.1530/ETJ-23-0067)

14 Horvath E, Majlis S, Rossi R, Franco C, Niedmann JP, Castro A & Dominguez M. An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. *Journal of Clinical Endocrinology and Metabolism* 2009 **94** 1748–1751. (https://doi.org/10.1210/jc.2008-1724)

15 Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, *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** 1–133. (https://doi.org/10.1089/thy.2015.0020)

16 Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, Cronan JJ, Beland MD, Desser TS, Frates MC, *et al.* ACR thyroid

imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS committee. *Journal of the American College of Radiology* 2017 **14** 587–595. (https://doi.org/10.1016/j.jacr.2017.01.046)

17 Russ G, Bonnema SJ, Erdogan MF, Durante C, Ngu R & Leenhardt L. European Thyroid Association guidelines for ultrasound malignancy risk stratification of thyroid nodules in adults: the EU-TIRADS. *European Thyroid Journal* 2017 **6** 225–237. (https://doi.org/10.1159/000478927)

18 Shin JH, Baek JH, Chung J, Ha EJ, Kim JH, Lee YH, Lim HK, Moon WJ, Na DG, Park JS, *et al.* Ultrasonography diagnosis and imaging-based management of thyroid nodules: revised Korean society of thyroid radiology consensus statement and recommendations. *Korean Journal of Radiology* 2016 **17** 370–395. (https://doi.org/10.3348/kjr.2016.17.3.370)

19 Durante C, Hegedüs L, Na DG, Papini E, Sipos JA, Baek JH, Frasoldati A, Grani G, Grant E, Horvath E, *et al.* International expert consensus on US lexicon for thyroid nodules. *Radiology* 2023 **309** e231481. (https://doi.org/10.1148/radiol.231481)

20 Hoang JK, Asadollahi S, Durante C, Hegedüs L, Papini E & Tessler FN. An international survey on utilization of five thyroid nodule risk stratification systems: a needs assessment with future implications. *Thyroid* 2022 **32** 675–681. (https://doi.org/10.1089/thy.2021.0558)

21 Gao L, Xi X, Jiang Y, Yang X, Wang Y, Zhu S, Lai X, Zhang X, Zhao R & Zhang B. Comparison among TIRADS (ACR TI-RADS and KWAK-TI-RADS) and 2015 ATA Guidelines in the diagnostic efficiency of thyroid nodules. *Endocrine* 2019 **64** 90–96. (https://doi.org/10.1007/s12020-019-01843-x)

22 Skowrońska A, Milczarek-Banach J, Wiechno W, Chudziński W, Żach M, Mazurkiewicz M, Miśkiewicz P & Bednarczuk T. Accuracy of the European Thyroid Imaging Reporting and Data System (EU-TIRADS) in the valuation of thyroid nodule malignancy in reference to the post-surgery histological results. *Polish Journal of Radiology* 2018 **83** e579–e586. (https://doi.org/10.5114/pjr.2018.81556)

23 Chng CL, Tan HC, Too CW, Lim WY, Chiam PPS, Zhu L, Nadkarni NV & Lim AYY. Diagnostic performance of ATA, BTA and TIRADS sonographic patterns in the prediction of malignancy in histologically proven thyroid nodules. *Singapore Medical Journal* 2018 **59** 578–583. (https://doi.org/10.11622/smedj.2018062)

24 Trimboli P, Ngu R, Royer B, Giovanella L, Bigorgne C, Simo R, Carroll P & Russ G. A multicentre validation study for the EU-TIRADS using histological diagnosis as a gold standard. *Clinical Endocrinology* 2019 **91** 340–347. (https://doi.org/10.1111/cen.13997)

25 Chen Z, Wang JJ, Guo DM, Zhai YX, Dai ZZ & Su HH. Combined fine-needle aspiration with core needle biopsy for assessing thyroid nodules: a more valuable diagnostic method? *Ultrasonography* 2023 **42** 314–322. (https://doi.org/10.14366/usg.22112)

26 Cherella CE, Angell TE, Richman DM, Frates MC, Benson CB, Moore FD, Barletta JA, Hollowell M, Smith JR, Alexander EK, *et al.* Differences in thyroid nodule cytology and malignancy risk between children and adults. *Thyroid* 2019 **29** 1097–1104. (https://doi.org/10.1089/thy.2018.0728)

27 Solymosi T, Hegedüs L, Bonnema SJ, Frasoldati A, Jambor L, Kovacs GL, Papini E, Rucz K, Russ G, Karanyi Z, *et al.* Ultrasound-based indications for thyroid fine-needle aspiration: outcome of a TIRADS-based approach versus operators' expertise. *European Thyroid Journal* 2021 **10** 416–424. (https://doi.org/10.1159/000511183)

28 Solymosi T, Hegedűs L, Bonnema SJ, Frasoldati A, Jambor L, Karanyi Z, Kovacs GL, Papini E, Rucz K, Russ G, *et al.* Considerable interobserver variation calls for unambiguous definitions of thyroid nodule ultrasound characteristics. *European Thyroid Journal* 2023 **12**. (https://doi.org/10.1530/ETJ-22-0134)

29 Stevenson M, Nunes T, Heuer C, Marshall J, Sanchez J & Thornton R *epiR: Tools for the Analysis of Epidemiological Data*. R package, version 1.0-2, 2019. Available at: https://CRAN.R-project.org/package=epiR.

30 R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistifcal Computing, 2019. Available at: https://www.R-project.org/.

31 Kwak JY, Han KH, Yoon JH, Moon HJ, Son EJ, Park SH, Jung HK, Choi JS, Kim BM & Kim EK. Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk. *Radiology* 2011 **260** 892–899. (https://doi.org/10.1148/radiol.11110206)

32 Na DG, Baek JH, Sung JY, Kim JH, Kim JK, Choi YJ & Seo H. Thyroid imaging reporting and data system risk stratification of thyroid nodules: categorization based on solidity and echogenicity. *Thyroid* 2016 **26** 562–572. (https://doi.org/10.1089/thy.2015.0460)

33 Brito JP, Gionfriddo MR, Al Nofal A, Boehmer KR, Leppin AL, Reading C, Callstrom M, Elraiyah TA, Prokop LJ, Stan MN, *et al.* The accuracy of thyroid nodule ultrasound to predict thyroid cancer: systematic review and meta-analysis. *Journal of Clinical Endocrinology and Metabolism* 2014 **99** 1253–1263. (https://doi.org/10.1210/jc.2013-2928)

34 Remonti LR, Kramer CK, Leitão CB, Pinto LCF & Gross JL. Thyroid ultrasound features and risk of carcinoma: a systematic review and meta-analysis of observational studies. *Thyroid* 2015 **25** 538–550. (https://doi.org/10.1089/thy.2014.0353)

35 Campanella P, Ianni F, Rota CA, Corsello SM & Pontecorvi A. Quantification of cancer risk of each clinical and ultrasonographic suspicious feature of thyroid nodules: a systematic review and metaanalysis. *European Journal of Endocrinology* 2014 **170** R203–R211. (https://doi.org/10.1530/EJE-13-0995)

36 Lee JY, Lee CY, Hwang I, You SH, Park SW, Lee B, Yoon RG, Yim Y, Kim JH & Na DG. Malignancy risk stratification of thyroid nodules according to echotexture and degree of hypoechogenicity: a retrospective multicenter validation study. *Scientific Reports* 2022 **12** 16587. (https://doi.org/10.1038/s41598-022-21204-5)

37 Grani G, Gentili M, Siciliano F, Albano D, Zilioli V, Morelli S, Puxeddu E, Zatelli MC, Gagliardi I, Piovesan A, *et al.* A data-driven approach to refine predictions of differentiated thyroid cancer outcomes: a prospective multicenter study. *Journal of Clinical Endocrinology and Metabolism* 2023 **108** 1921–1928. (https://doi.org/10.1210/clinem/dgad075)