Article published online: 2023-10-13

Determining the Best Thyroid Imaging Reporting and Data System: A Prospective Study Comparing the Diagnostic Performance of ACR, EU, and K TIRADS in the Evaluation of Thyroid Nodules

Supraja Laguduva Mohan[1] Ramkumar Govindarajalou[1] Dukhabandhu Naik[2] Sunil Kumar Saxena[3] Pampa Ch. Toi[4] Gomathi Shankar V.[5]

[1] Department of Radiodiagnosis, Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry, India
[2] Department of Endocrinology, Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry, India
[3] Department of Otorhinolaryngology, Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry, India
[4] Department of Pathology, Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry, India
[5] Department of Surgery, Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry, India

Address for correspondence Ramkumar Govindarajalou, MD (Radiodiagnosis), Department of Radiodiagnosis, Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry 605006, India (e-mail: gramk80@gmail.com).

Indian J Radiol Imaging 2024;34:220–231.

Abstract

Background Many different risk stratification systems have been formulated for thyroid nodules, differing in their fine-needle aspiration cytology (FNAC) indication, suggesting a lack of consensus around the world.

Purpose This prospective study was conducted to find the best guideline for risk stratification, for a better malignancy yield, and with reduced rates of negative FNACs among three Thyroid Imaging, Reporting, and Data System (TIRADS) guidelines.

Materials and Methods A total of 625 thyroid nodules with conclusive FNAC or histopathological diagnosis were included in the study. Various sonographic parameters were recorded. They were classified into categories as per the three guidelines and compared with FNAC diagnosis. The guidelines were evaluated in terms of sensitivity, specificity, predictive values, and diagnostic accuracy. Sensitivity and specificity were compared by McNemar's test.

Results American College of Radiology (ACR) TIRADS had the highest diagnostic accuracy (56.8%), specificity (50.75%), positive predictive value (23.92%), lowest rates of negative FNACs (76.08%), and high negative predictive value (97.84 %). Korean (K) TIRADS had the maximum sensitivity (97.75%), highest negative predictive value (98.44%), and gross malignancy yield. European TIRADS was between the two other guidelines in most parameters with specificity like K TIRADS.

Keywords
- thyroid nodules
- Thyroid Imaging Reporting and Data System
- TIRADS
- thyroid ultrasound
- American College of Radiology
- Korean Society of Thyroid Radiology
- European Thyroid Association
- fine-needle aspiration cytology

article published online
October 13, 2023

DOI https://doi.org/10.1055/s-0043-1775862.
ISSN 0971-3026.

Thieme Medical and Scientific Publishers Pvt. Ltd., A-12, 2nd Floor, Sector 2, Noida-201301 UP, India

Conclusion All the three guidelines are very good screening tools, with comparable high sensitivity. ACR TIRADS is better in terms of specificity and reduced rates of negative FNACs. Including the presence of a suspicious cervical lymph node as a criterion and more frequent follow-up might further improve the diagnostic performance of the guideline.

Introduction

With the increasing use of high-resolution ultrasound (USG) of the neck, the detection of nonpalpable thyroid nodules has also increased.[1] Fine-needle aspiration cytology (FNAC) is the commonly used modality for differentiating benign and malignant thyroid nodules. However, more than half of such nodules are benign, and one-third are inconclusive by FNAC.[2] Though not as invasive as biopsy, FNAC can cause pain, intrathyroidal hemorrhage, and very rarely local hematoma, and can be a source of apprehension to patients. Ultrasonographic features can help differentiate benign and malignant lesions and can be used as a screening tool to decide which patients need to undergo FNAC, to get the maximum yield with reduced rates of negative FNAC.

Since 2009, various attempts have been made in developing a risk stratification system for thyroid nodules by different radiology, endocrine, and oncologic societies. Currently, there are up to around 14 different guidelines in stratifying thyroid nodules, including the first ever Thyroid Imaging, Reporting, and Data System (TIRADS) devised by Horvath et al,[3] followed by Kwak TIRADS,[4] British Thyroid Association guidelines,[5] American Thyroid Association guidelines,[6] American College of Radiology (ACR) guidelines,[7] European Thyroid Association (EU) guidelines[8], and Korean Society of Thyroid Radiology (KSThR) (K) guidelines.[9] A recent modification of K TIRADS was done in 2021.[10]

These guidelines vary in the sonographic features and patterns used to stratify nodules, and have different size cutoffs for FNAC, with varying diagnostic accuracies. Guidelines with higher sensitivity are accompanied by higher rates of negative FNACs. Guidelines with higher specificity for malignant nodules are associated with lower rates of negative FNACs, but can potentially miss malignant nodules. This study was undertaken to find out the better guideline among the three different recently developed TIRADS systems—the ACR, the EU, and the K (KSThR) guidelines.

Materials and Methods

This study was conducted in our tertiary care hospital from December 2018 to February 2021, after approval by the institute ethics committee. Written informed consent was obtained from all patients. Details such as age, gender, clinical features, and thyroid function (thyroid-stimulating hormone, T4 and T3) were noted down in the pro forma.

Patient population: A total of 1,052 thyroid nodules, from 986 patients, over 18 years of age were evaluated during the study period. B-mode USG and color Doppler were done in all these patients. FNAC was not recommended in patients if it was not indicated by at least any one of the three guidelines. Those without final cytological diagnosis or nondiagnostic (Bethesda I) or with FNAC inconclusive nodules (Bethesda III and IV) were excluded. Note that 625 nodules with definite diagnosis were finally included in the study. A detailed recruitment algorithm is depicted in ►**Fig. 1**.

Imaging technique: The conventional USG and color Doppler were performed in either Siemens Acuson S3000 USG system (Erlangen, Germany) or EsaoteMyLab 60 machine (Esaote, Italy). A 6 to 18 MHz linear transducer was used for most nodules and a 4- to 9-MHz transducer was used for larger lesions. Patients were examined in supine position with neck extended. The thyroid was examined in transverse and longitudinal planes. Screening of the entire neck was done to look for suspicious lymph nodes. Parameters like size, number, location, shape, margins, echogenicity, composition, vascularity, and presence of suspicious cervical lymphadenopathy (those with cystic change, hyperechogenicity, calcifications, and abnormal vascularity) were studied. In patients with multiple nodules, up to four nodules were evaluated, and a maximum of two nodules with the highest TIRADS score were indicated for FNAC. Each nodule was assigned TIRADS scores as per all the three guidelines.

Comparison with the composite reference standard: FNAC or histopathological examination (HPE) after excision was taken as gold standard, of which the former was routinely done for the diagnosis. Unguided FNAC was routinely done for large palpable thyroid nodules. USG-guided FNAC was done for smaller nodules that could not be palpated and in the nodule with the highest TIRADS category among multiple nodules. Nodules were categorized into six groups according to the Bethesda classification.[11] The nodules in which FNAC was not indicated by a particular guideline were considered as test negative for that guideline. Bethesda category II in FNAC or histopathologically benign lesions were considered true negative. Nodules in which FNAC was indicated were assumed to be test positive as per the guideline. Bethesda V and VI or histopathologically proven malignant lesions were considered true positive.

Statistical analysis: The data was entered in MS Excel 2019. The statistical analysis was carried out using the IBM SPSS (version 27) software. Continuous variables were summarized in terms of mean with standard deviation/median with interquartile range. Categorical variables were summarized in terms of frequency or percentages. Chi-square test was used to compare categorical variables. The comparison

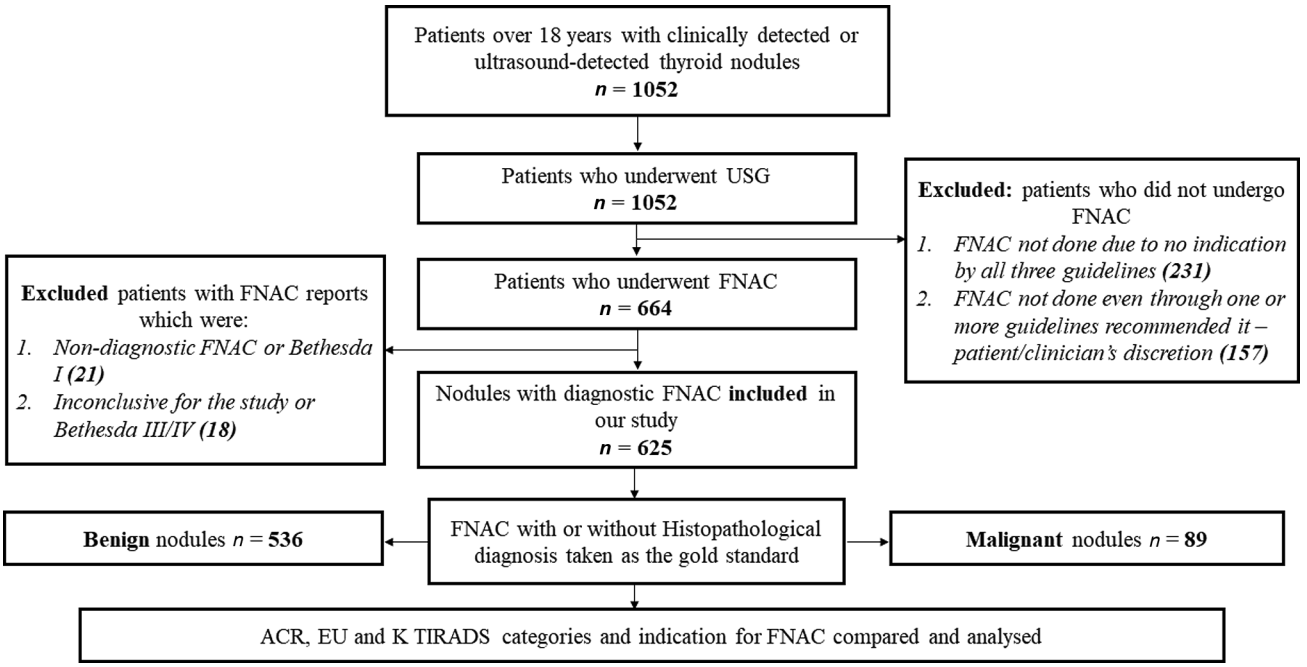

Fig. 1 Recruitment algorithm.

of continuous variables between benign and malignant lesions was carried out using the Mann–Whitney *U* test or Fisher's exact test. The diagnostic accuracy of each of the three guidelines was explored using sensitivity, specificity, predictive values, and likelihood ratios. The comparison of sensitivity and specificity of the three guidelines was done by McNemar's test. All statistical analyses were carried out at a 5% level of significance, and a p-value of < 0.05 was considered significant.

Results

Demographic characteristics: Of the 625 thyroid nodules (97 from men and 528 from women), there were 89 malignant and 536 benign nodules based on FNAC or HPE. The incidence of thyroid malignancy in our study population was 14.24%. The mean age of patients with benign nodules was 42.8 ± 12.6 years and that with malignant nodules was 45.4 ± 15.1 years. Most of the patients belonged to the 31 to 45 years age group. Among patients in whom the thyroid function status was known, there was equal distribution of benign and malignant nodules among euthyroid (13.1% malignant, 86.9% benign), hypothyroid (15.5% malignant, 84.5% benign), and hyperthyroid (14.5% malignant, 85.5% benign) states with no statistically significant difference ($p = 0.496$).

Sonographic features of benign and malignant thyroid nodules: Statistically significant difference was noted in the sonographic features between benign and malignant nodules. Features such as very hypoechoic echogenicity (84.6%), irregular shape (95%), taller than wide shape (76.9%), irregular margins (91.4%), microcalcifications (89.5%), suspicious lymph nodes (90.7%), and extrathyroid extension/capsular bulge were common in malignant nodules. Whereas features like anechoic cystic or spongiform composition, iso- or hyperechogenicity, oval shape (88.6%), wider than tall shape (89.1%), smooth margins (93.5%), no extrathyroid extension (86.5%), and no associated suspicious lymph nodes (91.4%) were commonly seen in benign nodules.

TIRADS category distribution: Most of the nodules included in the study fell under category 3 in all the three guidelines (►**Fig. 2**). There were fewer category 2 nodules (by all three guidelines) and category 1 nodules (by ACR TIRADS) included in the study, as most of them did not have an FNAC diagnosis, nor was it required by any guideline. ACR TIRADS had more proportion of nodules in category 4 ($n = 161$) compared to the other two guidelines. K TIRADS had maximum number of category 5 nodules ($n = 72$).

Diagnostic accuracy: The sensitivity (97.75%) and negative predictive value (98.44%) were maximum with K TIRADS, which had the best negative likelihood ratio (0.096). Specificity (50.75%) and positive predictive value (23.92%) were maximum with ACR TIRADS (►**Table 1**). The percentage of negative FNACs (proportion of benign nodules among those indicated for FNAC) was maximum with K TIRADS (82.49%) followed by EU TIRADS (82.21%) and least with ACR TIRADS (76.08%). Gross malignancy yield (number of malignant nodules that were indicated for FNAC) was maximum with K TIRADS ($n = 89$). Relative malignancy yield (proportion of malignant nodules among those indicated for FNAC) was least with K TIRADS (17.5%) followed by EU TIRADS (17.78%) due to higher number of negative FNACs. EU TIRADS scored between the other two guidelines in most parameters (►**Table 1**). There was a statistically significant difference in the specificity between the guidelines, with the ACR being the most specific guideline (McNemar's test, $p < 0.001$

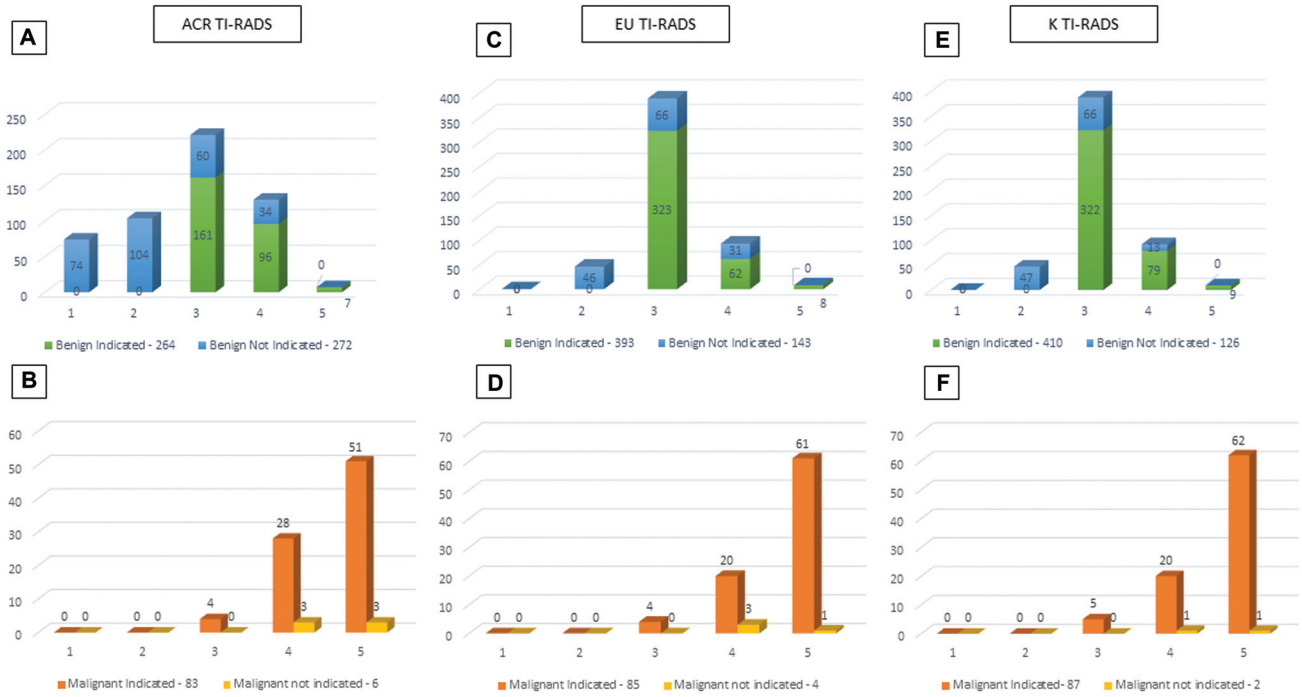

Fig. 2 Distribution of thyroid nodules as per the American College of Radiology (ACR) (**A** and **B**), European (EU) (**C** and **D**), and Korean (K) (**E** and **F**) Thyroid Imaging, Reporting, and Data System (TIRADS) in our study.

Table 1 Diagnostic performances of ACR, EU, and K TIRADS in differentiating benign and malignant thyroid nodules

Parameters	ACR TIRADS	EU TIRADS	K TIRADS
Sensitivity % (95% CI)	93.26 (85.9–97.49)	95.51 (88.89–98.76)	97.75 (92.12–99.73)
Specificity % (95% CI)	50.75 (46.43–55.06)	26.68 (22.98–30.64)	23.51 (19.98–27.33)
PPV % (95% CI)	23.92 (19.53–28.76)	17.78 (16.81–18.80)	17.51 (14.27–21.14)
NPV % (95% CI)	97.84 (95.36–99.20)	97.28 (93.14–98.95)	98.44 (94.47–98.81)
LR+ (95% CI)	1.893 (1.876–1.911)	1.303 (1.295–1.311)	1.278 (1.271–1.285)
LR– (95% CI)	0.133 (0.095–0.186)	0.168 (0.099–0.286	0.096 (0.034–0.268)
Diagnostic accuracy % (95% CI)	56.80 (52.81–60.72)	36.48 (32.70–40.39)	34.08 (30.37–37.94)
Gross malignancy yield (*n*)	83	85	87
Relative malignancy yield % (95% CI)	23.91 (83/347) (19.73–28.68)	17.78 (85/478) (14.62–21.46)	17.50 (87/497) (14.42–21.09)
Negative FNAC %	76.08	82.21	82.49

Abbreviations: ACR, American College of Radiology; CI, confidence interval; EU, European; FNAC, fine-needle aspiration cytology; K, Korean; LR+ , likelihood ratio of a positive test; LR–, likelihood ratio of a negative test; NPV, negative predictive value; PPV, positive predictive value; TIRADS, Thyroid Imaging, Reporting, and Data System.

between all the three guidelines). Comparison of sensitivity between them showed no statistically significant difference ($p = 0.1336$ for ACR vs. K, $p = 0.4795$ for EU vs. K TIRADS and ACR vs. EU TIRADS).

Discussion

Out of the 625 nodules in our study, 536 were proven to be benign and 89 were malignant. Malignant thyroid nodules

Table 2 Distribution of sonographic features between benign and malignant thyroid nodules

Sonographic criteria	Malignant 89	Benign 536	*p*-Value
Composition, *N* (%) - Solid - Almost entirely solid - Mixed solid cystic - Almost entirely cystic - Cystic - Spongiform	74 (22.4) 3 (3.5) 10 (8.1) 1 (2.3) 0 (0) 1(2.7)	256 (77.6) 83 (96.5) 113 (91.9) 43 (97.7) 5 (100) 36 (97.3)	< 0.0001
Echogenicity, *N* (%) - Isoechoic - Hyperechoic - Hypoechoic - Very hypoechoic - Anechoic	17 (4.5) 6 (8.8) 55 (35.7) 11 (84.6) 0 (0)	364 (95.5) 62 (91.2) 99 (64.3) 2 (15.4) 9 (100)	< 0.0001
Shape, *N* (%) - Oval - Round - Irregular	67 (11.4) 3 (20) 19 (95)	523 (88.6) 12(80) 1 (5)	< 0.0001
- Wider than tall - Taller than wide	79 (12.9) 10 (76.9)	533 (87.1) 3 (23.1)	< 0.0001
Margins, *N* (%) - Smooth - Ill defined - Lobulated - Irregular	32 (6.5) 17 (21.2) 8 (50) 32 (91.4)	462 (93.5) 63 (78.8) 8 (50) 3 (8.6)	< 0.0001
Extrathyroid extension, *N* (%) - Capsular bulge - Overt extra thyroid extension - No extra thyroid extension	3 (75) 4 (100) 83 (13.5)	1 (25) 0 (0) 534 (86.5)	< 0.0001
Calcifications, *N* (%) - None - Comet tail artifacts - Macrocalcification - Rim calcification - Microcalcification	19 (4.2) 0 (0) 18 (21.2) 1 (14.3) 51 (89.5)	436 (95.8) 21 (100) 67 (78.8) 6 (85.7) 6 (10.5)	< 0.0001
Suspicious lymph nodes, *N* (%) - Absent - Present	50 (8.6) 39 (90.7)	532 (91.4) 4 (9.3)	< 0.0001

were usually hypoechoic or very hypoechoic, solid, associated with irregular shape or margins, wider than taller shape, showed microcalcifications, and extrathyroidal extension, with a statistically significant difference (►**Table 2**). These correlated well with previous studies which also suggest that not a single sonographic feature, but a combination of features predict the risk of malignancy more accurately.[12–14] Smooth margins, anechoic/cystic or spongiform consistency, and comet tail artifacts are predictors of benign nodules.[13] Size of the nodule, however, could not reliably differentiate between benign and malignant pathologies. In our study, the mean size of malignant lesions (3.3 ± 2.14 cm) was not significantly higher than that for benign lesions (3.1 ± 1.59 cm). This was comparable to a previous study by Kamran et al, in which a threshold effect was found and the cancer risk was the same for all nodules more than 2 cm.[15]

ACR, EU, and K TIRADS differ in their size cutoff for FNAC as well as category distribution of nodules (►**Figs. 2** and **3**). K TIRADS has the lowest size cutoff among the three guidelines. In our study, ACR TIRADS had the maximum number of category 4 nodules (161 nodules) compared to EU and K TIRADS. This was due to macrocalcifications and rim calcifications (which are not considered as criteria by EU or K TIRADS) upgrading a solid isoechoic nodule from category 3 to category 4. K TIRADS and EU TIRADS had maximum numbers of category 3 nodules and have a separate category for no nodule (EU1 and K1), while purely spongiform and cystic nodules are described as category 2 (►**Fig. 4**). However, these category differences did not influence the number of FNACs performed or the final outcomes, as the number of negative FNACs done were still the lowest by ACR TIRADS due to its higher size cutoff (►**Figs. 4–12**).

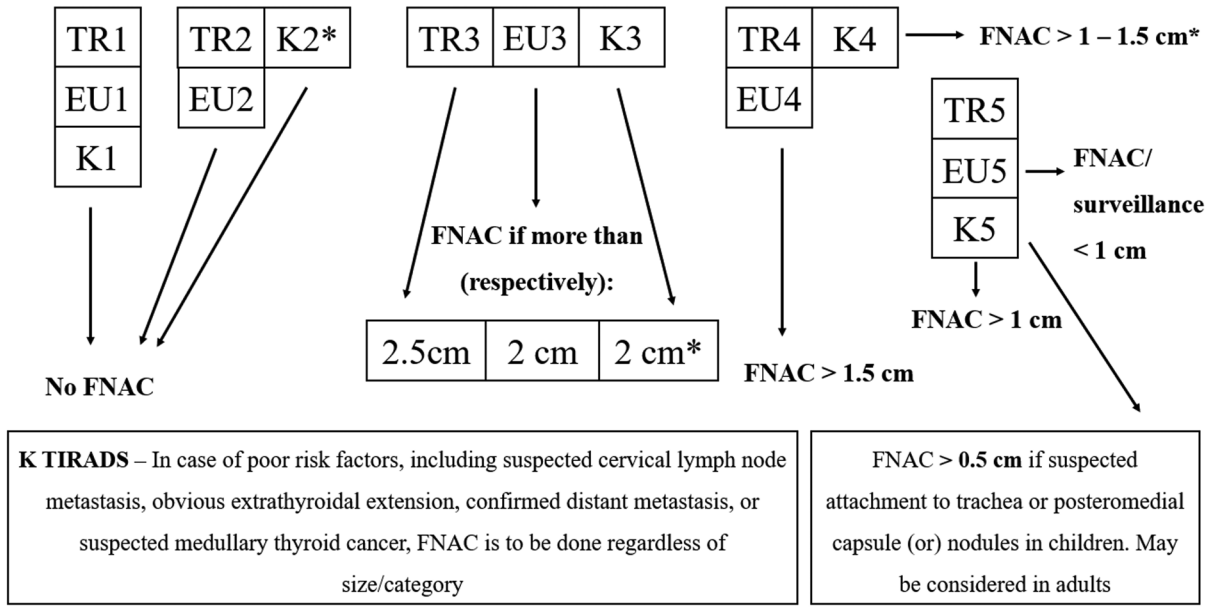

Fig. 3 Differences in size criteria for fine-needle aspiration cytology (FNAC) indication among the American College of Radiology (ACR), European (EU), and Korean (K) Thyroid Imaging, Reporting, and Data System (TIRADS). *Recent modifications in the 2021 version of K TIRADS.

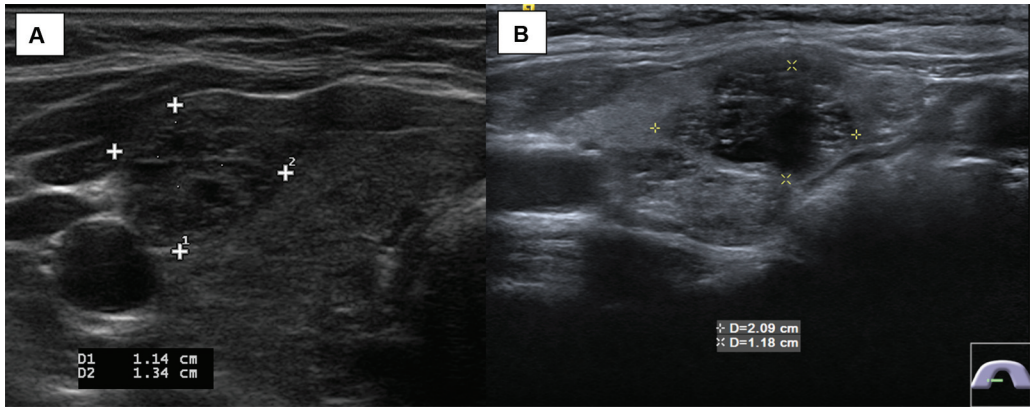

Fig. 4 Concordant fine-needle aspiration cytology (FNAC) indication. (**A** and **B**) Spongiform nodules (American College of Radiology [ACR] TR1, EU2, K2) with no indication for FNAC by any guideline (Bethesda II, colloid goiter).

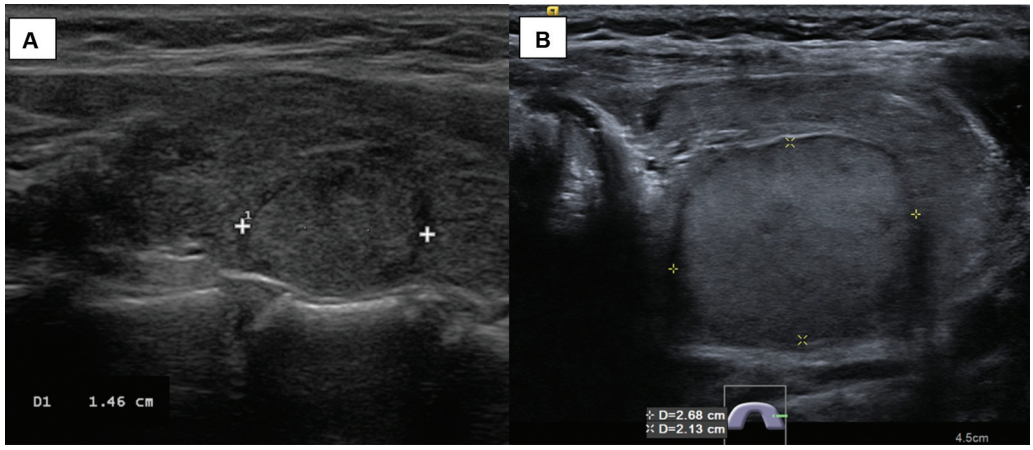

Fig. 5 Concordant fine-needle aspiration cytology (FNAC) indication. (**A**) Solid isoechoic nodule and (**B**) solid hyperechoic nodule (American College of Radiology [ACR] TR3, EU3, K3), the former nodule (**A**) not meeting the size cutoff for FNAC by any guideline, the latter (**B**) recommended for FNAC by all the three guidelines (Bethesda II, nodular goiter in both).

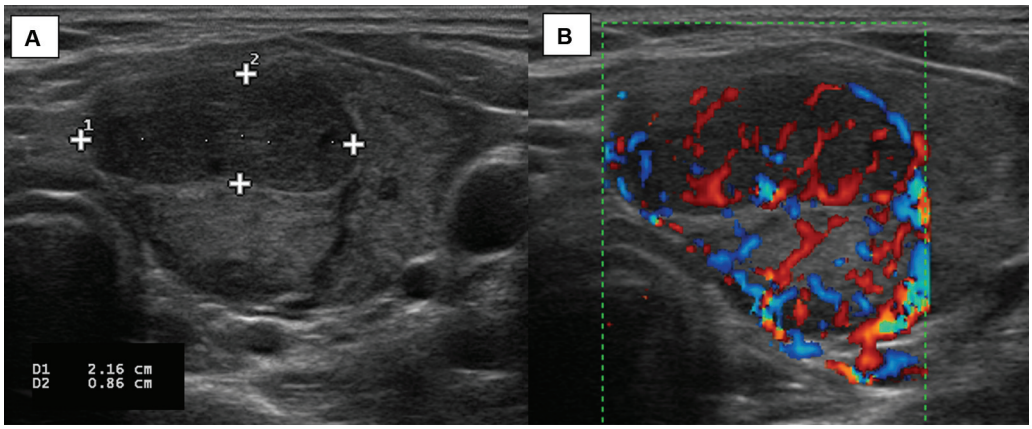

Fig. 6 Concordant fine-needle aspiration cytology (FNAC) indication. (**A** and **B**) Solid hypoechoic nodule > 1.5 cm with rich vascularity within the nodule and surrounding heterogeneous thyroid parenchyma (American College of Radiology [ACR] TR4, EU4, K4), FNAC indicated by all the three guidelines (Bethesda II, nodular goiter with lymphocytic thyroiditis).

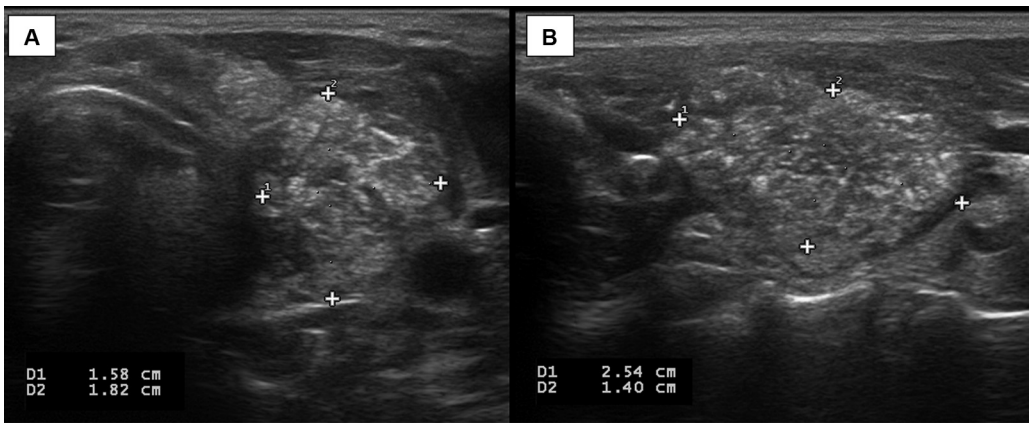

Fig. 7 Concordant fine-needle aspiration cytology (FNAC) indication. (**A** and **B**) Solid iso- and hyperechoic nodule with microcalcifications > 1 cm (American College of Radiology [ACR] TR5, EU5, K5), indicated for FNAC by all the three guidelines (Bethesda VI, papillary carcinoma).

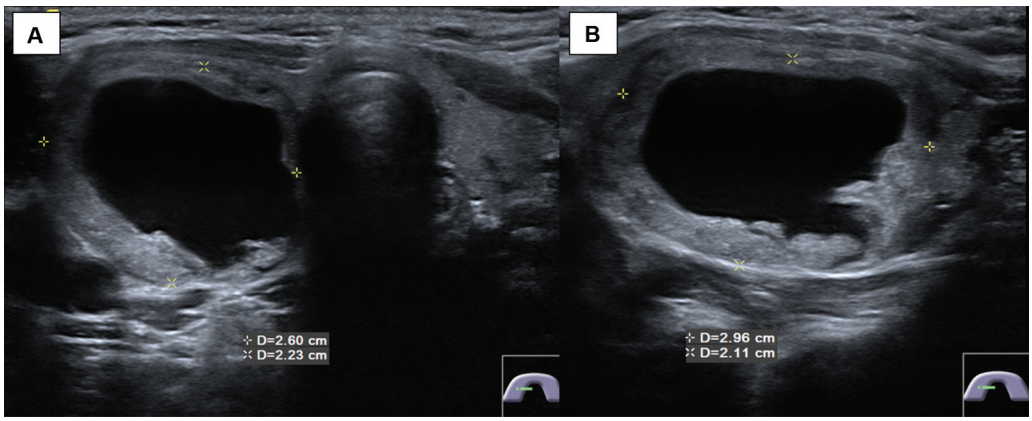

Fig. 8 Discordant fine-needle aspiration cytology (FNAC) indication. (**A** and **B**) Solid cystic nodule > 2 cm (American College of Radiology [ACR] TR2, EU3, K3) with FNAC indicated by European (EU) and Korean (K) Thyroid Imaging, Reporting, and Data System (TIRADS) (Bethesda II, nodular goiter).

ACR TIRADS had the highest overall diagnostic accuracy. The most sensitive guideline was K TIRADS and the most specific guideline was ACR TIRADS with high positive and negative predictive values. There was a statistically significant difference in specificity between all the three guidelines. Similar differences in sensitivity and specificity were seen by previous studies (►**Table 3**).[2,16–22] The high sensitivity of K TIRADS is due to the lower size cutoffs which leads to lesser number of missed malignancies. Among the malignant nodules that were missed in our study (false negatives), six were missed by ACR TIRADS, four by EU TIRADS, and two by K TIRADS (►**Table 4**). Three of the six nodules were

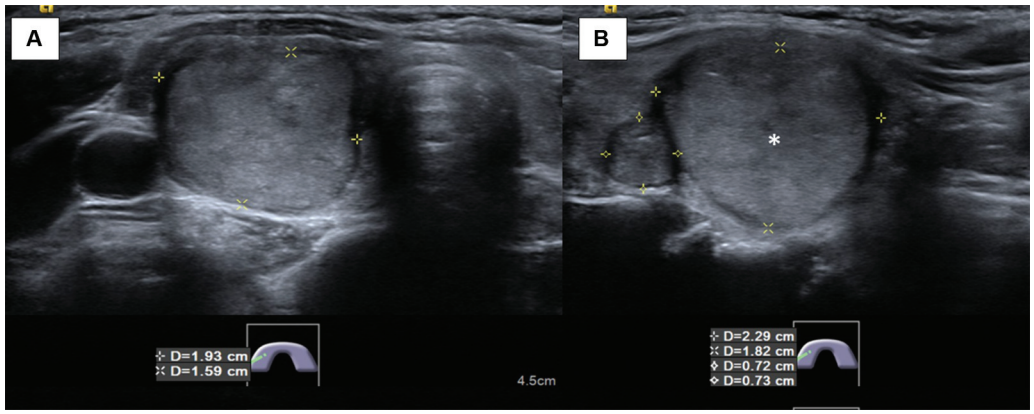

Fig. 9 Discordant fine-needle aspiration cytology (FNAC) indication. (**A** and **B**) Solid isoechoic nodule < 2.5 cm, asterisk in **B** (American College of Radiology [ACR] TR3, EU3, K3), FNAC indicated by Korean (K) and European (EU) Thyroid Imaging, Reporting, and Data System (TIRADS) (Bethesda II, nodular goiter).

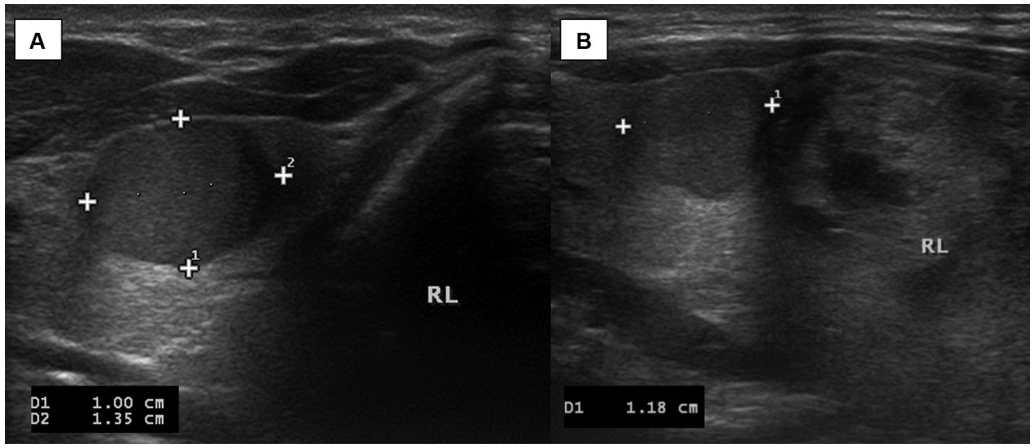

Fig. 10 Discordant fine-needle aspiration cytology (FNAC) indication. (**A** and **B**) Solid hypoechoic nodule < 1.5 cm (American College of Radiology [ACR] TR4, EU4, K4), with FNAC indicated only by Korean (K) Thyroid Imaging, Reporting, and Data System (TIRADS) (Bethesda II, colloid goiter).

TIRADS category 5 and three of them were category 4 by all the three guidelines (►**Figs. 11** and **12**). Six nodules were missed by ACR and four by EU TIRADS because they did not meet the size cutoff for FNAC. Even though two of the category 5 nodules and one category 4 nodule were subcentimetric, they were indicated for FNAC (the former by EU and K TIRADS, latter by K TIRADS) as these nodules were associated with suspicious lymph nodes (►**Fig. 12**). The K TIRADS missed only two nodules because they were very small and were also not associated with any suspicious lymph node or any other high-risk features.

Cervical lymph nodes are considered as a criterion for FNAC recommendation only by K and EU TIRADS. Suspicious features include cystic change, calcifications, hyperechogenicity, and abnormal peripheral or diffuse vascularity.[9] Among 43 nodules with associated suspicious lymph nodes, 39 were malignant nodules and 4 were benign. In three out of the four such benign nodules, the associated suspicious lymph nodes were proven to be metastatic carcinomas by FNAC of the lymph nodes, and the primary lesion was unknown by imaging in all three patients. The thyroid nodules were incidentally picked up on USG which was done for cervical lymphadenopathy in these patients. In the fourth benign thyroid nodule with associated suspicious lymph node, FNAC of the suspicious lymph node revealed metastatic papillary carcinoma, but repeated aspiration of the thyroid nodule only proved to be a benign cystic colloid goiter. There were no other focal lesions in the thyroid. About 50 out of the 89 malignant nodules were not associated with suspicious lymph nodes. But when associated, the presence of a suspicious lymph node increased the risk of malignancy in the thyroid nodule. Inclusion of suspicious cervical lymph nodes as a criterion for FNAC indication within ACR TIRADS increased the gross malignancy yield ($n = 86$) with slight increase in sensitivity, diagnostic accuracy, positive and negative predictive values, and improved the negative likelihood ratio (►**Table 5**).

The malignant nodules that were missed by the ACR and EU guidelines were, however, recommended to be followed up and might have eventually been diagnosed, had there been increase in size during follow-up. The follow-up recommendations by ACR TIRADS are every year up to 5 years for TR 5 nodules, 1, 2, 3, and 5 years for TR4 nodules, and 1, 3, and 5 years for TR3 nodules. If there is no increase in size, no

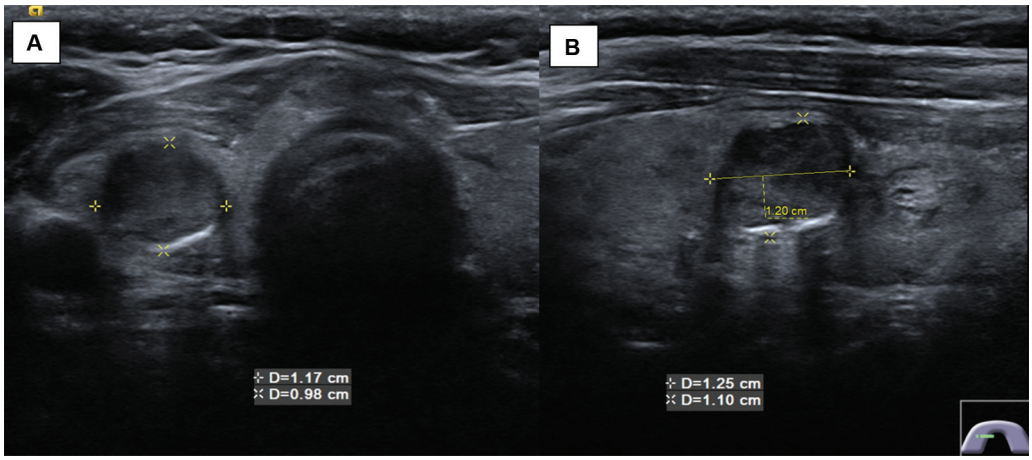

Fig. 11 Discordant fine-needle aspiration cytology (FNAC) indication. (**A** and **B**) Solid hypoechoic with macrocalcification < 1.5 cm (American College of Radiology [ACR] TR4, EU4, K4), FNAC indicated only by Korean (K) Thyroid Imaging, Reporting, and Data System (TIRADS) (Bethesda V, suspicious for papillary carcinoma).

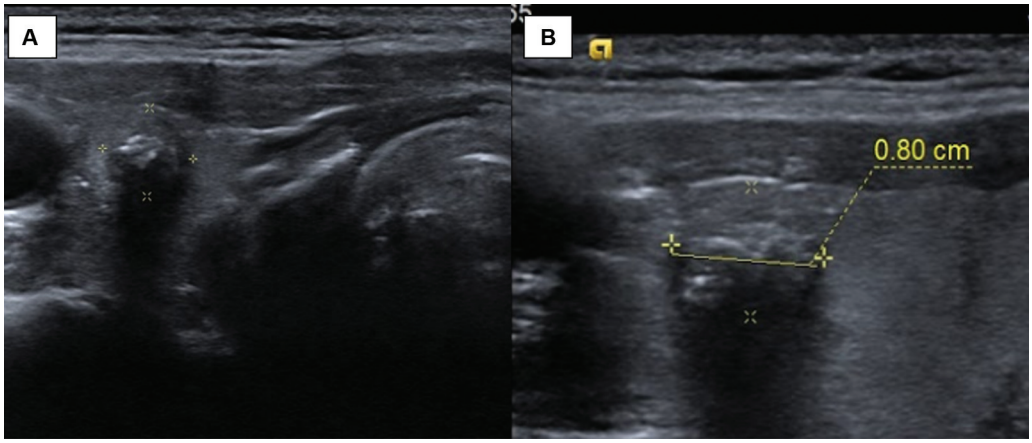

Fig. 12 Discordant fine-needle aspiration cytology (FNAC) indication. (**A** and **B**) Solid hypoechoic nodule with microcalcification < 1 cm (American College of Radiology [ACR] TR5, EU5, K5), FNAC indicated only by European (EU) and Korean (K) Thyroid Imaging, Reporting, and Data System (TIRADS) due to associated suspicious cervical lymph nodes (histopathological examination [HPE]; papillary microcarcinoma).

Table 3 Sensitivity and specificity of ACR, EU, and K TIRADS in differentiating thyroid nodules as benign and malignant by previous studies

Name of the study	ACR TIRADS		EU TIRADS		K TIRADS	
	Sensitivity %	Specificity %	Sensitivity %	Specificity %	Sensitivity %	Specificity %
Ha et al[17]	74.7	67.3	Not included		94.5	26.4
Grani et al[2]	83.3	56.2	86.1	32	91.7	17.8
Yoon et al[18]	77.3	67.7	87.4	38.9	95.7	23.6
Tan et al[19]	85.7	51.1	57.1	83.2	100	40.2
Huh et al[20]	80.4	62.2	95.2	28.1	Not included	
Gao et al[16]	81.6	79.7	Not included		Not included	
Na et al[21]	79.6	65.2	88.3	33.4	96.9	18.6
Kovatcheva et al[22]	Not included		69.9	56.3	Not included	
Ha et al[24]	76.1	61.8	84.6	39.3	91	39.7[a]

Abbreviations: ACR, American College of Radiology; EU, European; K, Korean; TIRADS, Thyroid Imaging, Reporting, and Data System.
[a]As per the 2021 modification of K TIRADS.

Table 4 Imaging characteristics of the six malignant thyroid nodules missed by the guidelines

Serial no.	Maximum size (cm)	Composition	Echoes	Shape	Margin	Calcification	Lymph nodes	TIRADS category		
								ACR	EU	K
1	0.80	Solid	Hypoechoic	Oval W > T	Ill defined	Micro	Suspicious	5	5[a]	5[a]
2	1.25	Solid	Hypoechoic	Oval W > T	Ill defined	Macro	Not suspicious	4	4	4[a]
3	0.65	Solid	Hypoechoic	Oval W > T	Smooth	Micro	Suspicious	5	5[a]	5[a]
4	0.91	Solid	Hypoechoic	Oval W > T	Smooth	Rim	Suspicious	4	4	4[a]
5	0.51	Solid	Hypoechoic	Oval W > T	Irregular	Micro	Not suspicious	5	5	5
6	0.98	Solid	Hypoechoic	Oval W > T	Ill defined	Nil	Not suspicious	4	4	4

Abbreviations: ACR, American College of Radiology; EU, European; FNAC, fine-needle aspiration cytology; K, Korean; TIRADS, Thyroid Imaging, Reporting, and Data System.
Note: W > T = wider than tall, T > W = taller than wide.
[a]FNAC was recommended only by these guidelines.

Table 5 Diagnostic performances of ACR TIRADS and ACR TIRADS including suspicious lymph nodes as criterion

Parameters	ACR TIRADS	ACR TIRADS + lymph node criterion[a]
Sensitivity % (95% CI)	93.26 (85.9–97.49)	96.63 (90.46–99.3)
Specificity % (95% CI)	50.75 (46.43–55.06)	50.75 (46.43–55.06)
PPV % (95% CI)	23.92 (19.53–28.76)	24.57 (22.87–26.36)
NPV % (95% CI)	97.84 (95.36–99.20)	98.91 (96.74–99.64)
LR+ (95% CI)	1.89 (1.87–1.91)	1.96 (1.79–2.16)
LR– (95% CI)	0.13 (0.095–0.186)	0.07 (0.02–0.20)
Diagnostic accuracy % (95% CI)	56.80 (52.81–60.72)	57.28 (53.3–61.20)
Negative FNAC %	76.08	75.42
Gross malignancy yield (*n*)	83	86
Relative malignancy yield %	23.91	24.57

Abbreviations: ACR, American College of Radiology; CI, confidence interval; FNAC, fine-needle aspiration cytology; LR +, likelihood ratio of a positive test; LR–, likelihood ratio of a negative test; NPV, negative predictive value; PPV, positive predictive value; TIRADS, Thyroid Imaging, Reporting, and Data System.
[a]Including FNAC for any nodule with suspicious lymph node regardless of size or category.

further follow-up is needed after 5 years.[23] As per EU TIRADS, active surveillance is recommended for category 5 lesions that are less than 10 mm in size and if there is proven growth or suspicious lymph nodes are picked up during follow-up, FNAC is recommended.[8] With regards to K TIRADS, 6-monthly follow-ups are recommended in K TIRADS 5 nodules, for 1 to 2 years followed by once a year; at 1, 3, and 5 years for K TIRADS 3 or 4, followed by once in 3 to 5 years and once in 5 years for category 3 and 4, respectively; and at 2 to 5 years for K TIRADS 2 nodules. If there is no growth at 5 years, no further follow-up is needed for K TIRADS 2 or 3 nodules.[10]

Among benign nodules, 264, 393, and 410 nodules were wrongly presumed as test positive by ACR, EU, and K TIRADS, respectively, the differences attributed mainly to the size cutoffs. Previously, purely spongiform nodules > 2 cm were indicated for FNAC by the K TIRADS.[9] This was removed in the 2021 modification along with other changes like increase in

size cutoffs for K TIRADS 3 and 4 nodules. This has led to increase in the specificity of K TIRADS.[24] One spongiform nodule was found to be malignant in our study and the same nodule also had microcalcifications and was category 4 by ACR and category 5 by EU and K TIRADS, and was hence indicated for FNAC by all three guidelines. This is also in agreement with previous studies,[25] in which almost all the spongiform nodules were benign. The modified K TIRADS suggests 1 to 1.5 cm size cutoff for category 4 nodules which is to be decided depending on the age, nodule location, clinical risk factor, preference, etc. All such patients in our study underwent FNAC by clinician's discretion and patient preference. We also observed that a specific subset of solid homogeneously hypoechoic nodules without any other suspicious feature, with features of thyroiditis in the background parenchyma (22 nodules), were all benign (►**Fig. 6**). But due to the solid and hypoechoic nature, they were given a higher category of 4. A subcategorization with higher size criteria for FNAC may be useful in such nodules.

One of the limitations of our study was nonavailability of histopathological diagnosis in all patients (available in 77 nodules), as FNAC is the diagnostic modality of choice and only a few symptomatic patients with benign thyroid nodules undergo surgery. Also, there was no indication for FNAC by any guideline in 231 patients, who subsequently did not undergo FNAC and had to be excluded from the study. Almost all such nodules would have been benign (true negatives), which suggests that the actual specificity of the guidelines would have been much higher. Possible selection bias due to the study being conducted in a tertiary referral institute could be another limitation. Also, the objectivity of the guidelines was not evaluated as we did not study interobserver variation.

Comparison of the three guidelines in our study showed all of them to be very good screening tools to identify malignant thyroid nodules, with very high sensitivity and negative predictive value and low negative likelihood ratio. However, ACR TIRADS seemed to perform better in terms of specificity and positive predictive value and thereby reducing the number of negative FNACs, which could also result in a socioeconomic advantage on a larger scale. The statistically insignificant slightly lower sensitivity of ACR TIRADS would have eventually been overcome by its follow-up recommendations. Nevertheless, including the presence of suspicious cervical lymph nodes as a risk factor to determine the cutoff criteria for FNAC, as in EU and K TIRADS, and more frequent follow-up might improve the diagnostic performance of the guideline.

Authors' Contributions

S.L.M.: Data acquisition and analysis, literature search, and manuscript preparation. R.G.: Design, data analysis, manuscript editing, and manuscript review. D.N.: Manuscript editing and review. S.K.S.: Manuscript editing and review. P.C.T.: Manuscript editing and review. G.S.V.: Manuscript editing and review.

Note

This work (before completion of the study) has been presented at the 72nd TN & PY state IRIA conference as an e-poster and at the SonoCon 2020 as a paper.

Funding

None.

Conflict of Interest

None declared.

Acknowledgment

The authors wish to extend their gratitude to Dr. Palanivel C. for his contribution in statistical analysis.

References

1 Cappelli C, Castellano M, Pirola I, et al. The predictive value of ultrasound findings in the management of thyroid nodules. QJM 2007;100(01):29–35
2 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 (01):95–102
3 Horvath E, Majlis S, Rossi R, et al. An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. J Clin Endocrinol Metab 2009;94(05):1748–1751
4 Kwak JY, Han KH, Yoon JH, et al. Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk. Radiology 2011;260(03):892–899
5 Perros P, Boelaert K, Colley S, et al; British Thyroid Association. Guidelines for the management of thyroid cancer. Clin Endocrinol (Oxf) 2014;81(Suppl 1):1–122
6 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(01):1–133
7 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(05):587–595
8 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. Eur Thyroid J 2017;6(05):225–237
9 Shin JH, Baek JH, Chung J, et al; Korean Society of Thyroid Radiology (KSThR) and Korean Society of Radiology. Ultrasonography diagnosis and imaging-based management of thyroid nodules: revised Korean Society of Thyroid Radiology Consensus Statement and Recommendations. Korean J Radiol 2016;17(03): 370–395
10 Ha EJ, Chung SR, Na DG, et al. 2021 Korean Thyroid Imaging Reporting and Data System and Imaging-Based Management of Thyroid Nodules: Korean Society of Thyroid Radiology Consensus Statement and Recommendations. Korean J Radiol 2021;22(12): 2094–2123
11 Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. Thyroid 2017;27(11):1341–1346
12 Anil G, Hegde A, Chong FHV. Thyroid nodules: risk stratification for malignancy with ultrasound and guided biopsy. Cancer Imaging 2011;11(01):209–223
13 Moon WJ, Jung SL, Lee JH, et al; Thyroid Study Group, Korean Society of Neuro- and Head and Neck Radiology. Benign and

malignant thyroid nodules: US differentiation–multicenter retrospective study. Radiology 2008;247(03):762–770
14 Hoang JK, Lee WK, Lee M, Johnson D, Farrell S. US features of thyroid malignancy: pearls and pitfalls. Radiographics 2007;27 (03):847–860, discussion 861–865
15 Kamran SC, Marqusee E, Kim MI, et al. Thyroid nodule size and prediction of cancer. J Clin Endocrinol Metab 2013;98(02): 564–570
16 Gao L, Xi X, Jiang Y, et al. Comparison among TIRADS (ACR TI-RADS and KWAK- TI-RADS) and 2015 ATA guidelines in the diagnostic efficiency of thyroid nodules. Endocrine 2019;64(01):90–96
17 Ha EJ, Na DG, Baek JH, Sung JY, Kim JH, Kang SY. US fine-needle aspiration biopsy for thyroid malignancy: diagnostic performance of seven society guidelines applied to 2000 thyroid nodules. Radiology 2018;287(03):893–900
18 Yoon SJ, Na DG, Gwon HY, et al. Similarities and differences between Thyroid Imaging Reporting and Data Systems. AJR Am J Roentgenol 2019;213(02):W76–W84
19 Tan L, Tan YS, Tan S. Diagnostic accuracy and ability to reduce unnecessary FNAC: a comparison between four Thyroid Imaging Reporting Data System (TI-RADS) versions. Clin Imaging 2020; 65:133–137
20 Huh S, Lee HS, Yoon J, et al. Diagnostic performances and unnecessary US-FNA rates of various TIRADS after application of equal size thresholds. Sci Rep 2020;10(01):10632
21 Na DG, Paik W, Cha J, Gwon HY, Kim SY, Yoo RE. Diagnostic performance of the modified Korean Thyroid Imaging Reporting and Data System for thyroid malignancy according to nodule size: a comparison with five society guidelines. Ultrasonography 2021; 40(04):474–485
22 Kovatcheva RD, Shinkov AD, Dimitrova ID, Ivanova RB, Vidinov KN, Ivanova RS. Evaluation of the diagnostic performance of EU-TIRADS in discriminating benign from malignant thyroid nodules: a prospective study in one referral center. Eur Thyroid J 2021;9(06):304–312
23 Tappouni RR, Itri JN, McQueen TS, Lalwani N, Ou JJ. ACR TI-RADS: pitfalls, solutions, and future directions. Radiographics 2019;39 (07):2040–2052
24 Ha EJ, Shin JH, Na DG, et al. Comparison of the diagnostic performance of the modified Korean Thyroid Imaging Reporting and Data System for thyroid malignancy with three international guidelines. Ultrasonography 2021;40(04):594–601
25 Aydoğan Bİ, Ceyhan K, Şahin M, Çorapçıoğlu D. Are thyroid nodules with spongiform morphology always benign? Cytopathology 2019;30(01):46–50

