International Journal of Radiology Research
www.radiologyjournal.in
E-ISSN: 2663-4562, P-ISSN: 2663-4554
Received: 08-11-2021 Accepted: 24-11-2021; Published: 09-12-2021
Volume 3, Issue 2, 2021, Page No. 18-26

Comparisons of ACR TI-RADS, Kwak TI-RADS and EU-TIRADS in malignancy risk stratification of thyroid nodules in Filipino patients

Er Jan Niko Malimban, Reza Koa Sales
Department of Radiology, Makati Medical Center, Amorsolo St., Legaspi Village, Makati City, Philippines

Abstract
Purpose: To give reliable and applicable clinical data and evidence in constructing a suitable TIRADS scoring system for the population of the Philippines.
Methodology: Retrospective study of patient records who underwent thyroidectomy at our institution from 2018- 2019 were reviewed and included patients of all ages with records of total or nearly total thyroidectomy or lobectomy performed; complete preoperative ultrasound of thyroid nodules; and has surgical pathology. All patient medical record and ultrasound images were correlated to pathological results. The nodules were characterized by size, number, composition, echogenicity, orientation, margins and calcification. Nodules were then scored and TI-RADS was then calculated.
Results: Among the three scoring systems, ACR TIRADS is the most specific while EU and Kwak TIRADS are more sensitive than ACR. Results show that the suspicious ultrasound features accurately predicted malignancy, with malignancy odds significantly higher in nodules with a high level of suspicion by ultrasound (TR5 highly suspicious), however, no significant difference exists between benign and malignant in terms of size.
Conclusion: TI-RADS may correlate that in Filipinos malignancy is better predicted by ultrasound features than size. There are no difference in diagnostic efficiency among the three scoring systems. In Filipinos, it is recommended that sonographically suspicious nodules undergo FNAB regardless of size because ultrasound features better predict malignancy.
Significance: Filipinos have a higher incidence, recurrence rate, and worse outcomes compared to other populations when diagnosed with thyroid cancer. It is imperative to adopt a scoring system that is reliable in malignancy risk stratification of thyroid nodules for the Philippine population.

Keywords: ACR TI-RADS, Kwak TI-RADS and EU-TIRADS, thyroid nodules

Introduction

Thyroid nodules are a common finding in the general population, currently 10% with still increasing malignancy risk [1]. In the general population, the prognosis of well-differentiated thyroid cancer is promising, with an overall survival rate of more than 90% in 10-years. However, studies show that Filipinos have a higher incidence, recurrence rate, and worse outcomes when diagnosed with thyroid cancer. Compared with non-Filipino Asians and non-Asians, Filipinos have an increased risk of death from thyroid cancer, regardless of age. Furthermore, Filipino women have a higher proportionate risk of death from thyroid cancer [2]. To date, no clinical guidelines have included Filipino ethnicity as a potential prognostic factor.

Ultrasonography (US) is the primary modality for thyroid examination because it can distinguish benign and malignant nodules [3] using a thyroid imaging reporting and data system (TI-RADS). This new development has led to identify benign and malignant nodules better and optimize clinical management [4, 11]. Among these scoring systems is the American College of Radiology (ACR) guidelines (Fig. 1), guidelines proposed by Kwak *et al*.(K-TI-RADS) (Fig. 2) and the European Thyroid Association (ETA) (Fig 3).

There are similarities and differences among these guidelines in structure, risk stratification, size thresholds, and diagnostic performance. These scoring systems and their recommendations have been essential for the diagnosis of thyroid nodules [13]. However, these scoring systems have been formulated in the setting wherein an ultrasound-guided biopsy is utilized. In the Philippines, the majority of thyroid FNAs are performed solely by palpation, resulting in low sensitivity due to inadequate tissue sampling, which leads to unnecessary re-biopsy and surgery [14].

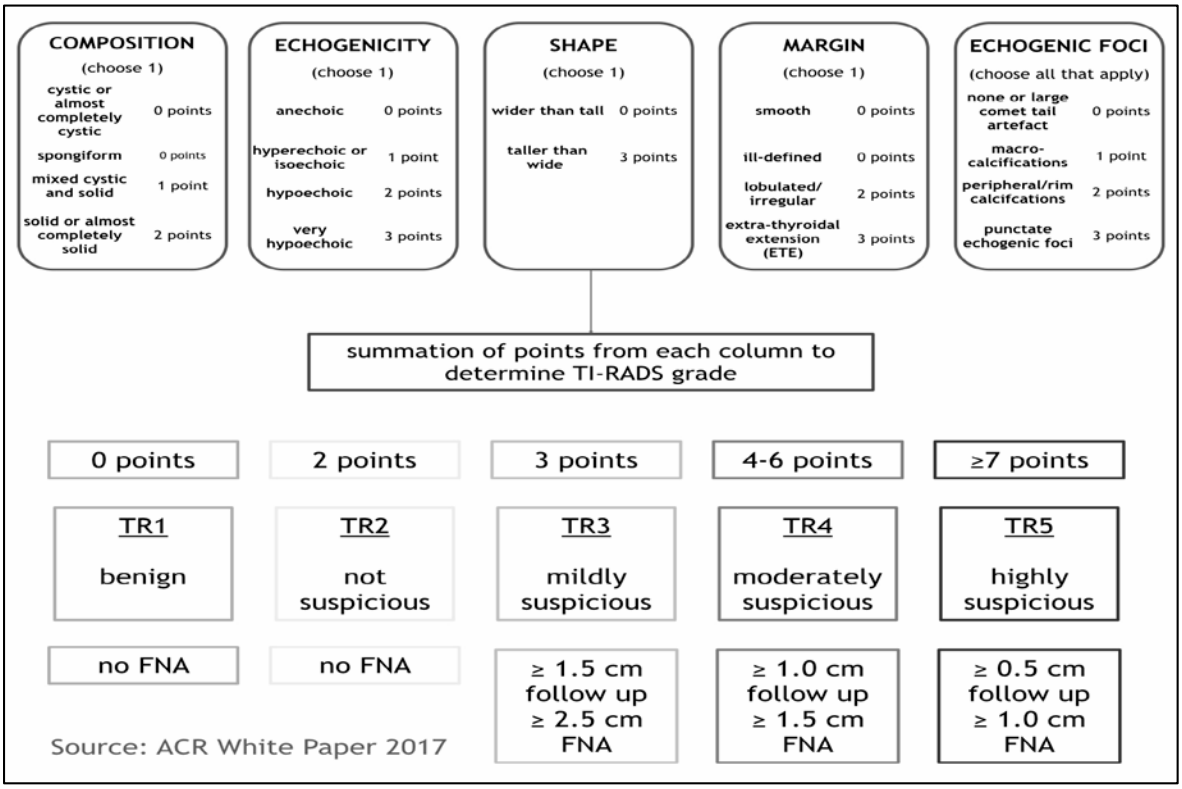

Fig 1: American College of Radiology TI-RADS guideline

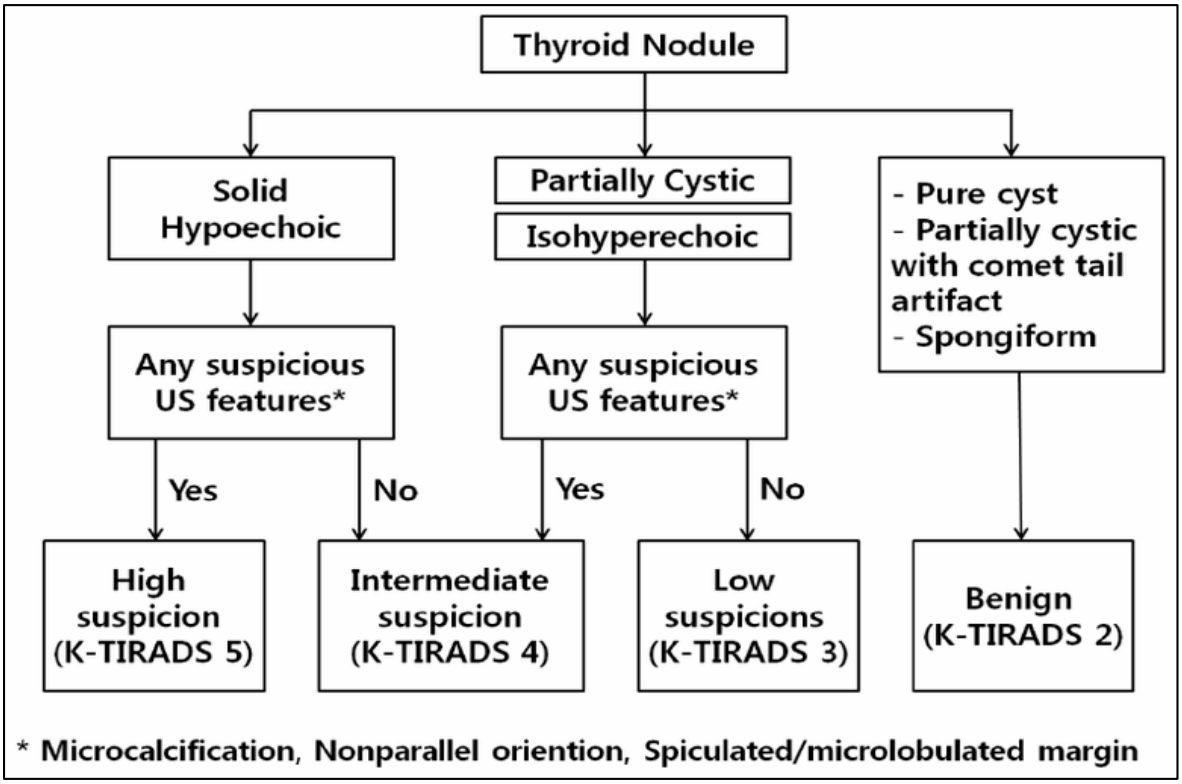

Fig 2: TI-RADS proposed by Kwak (Kwak TI-RADS)

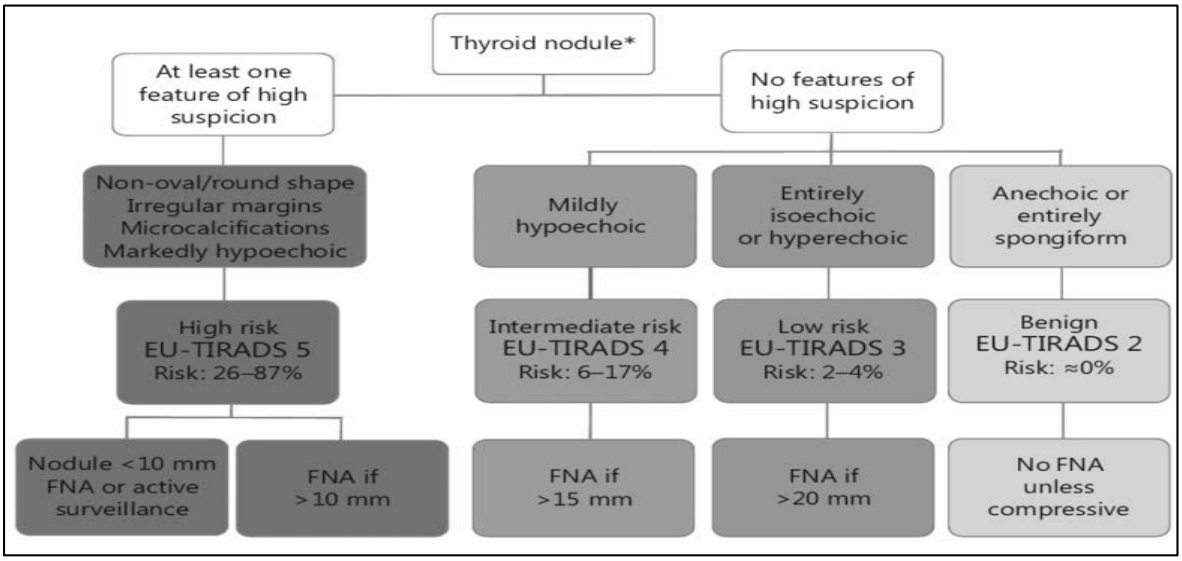

Fig 3: European Thyroid Association TI-RADS guideline

Comparative studies have been done with regards to their sensitivity and specificity to different populations. These studies have yielded variable results that may lead to different specificities and suggestions of management. Furthermore, none of these have been studied when it comes to the Filipino population [4, 11, 12].

Given that some journals show worse prognosis and higher incidence of thyroid nodules to Filipinos, a reliable and accurate diagnosis of thyroid nodules is critical for prognosticating, assessing treatment strategies. It is imperative to adopt a scoring system that is reliable in malignancy risk stratification of thyroid nodules for the Philippine population and setting [2].

Objectives

The primary objective of this study is to retrospectively analyze and compare the ultrasound diagnostic efficacy of the K-TI-RADS, ACR-TI-RADS, and EU-TI-RADS in detecting malignancy in the Philippine population to serve as a guide for accurate and efficient diagnosis and prognostic assessment of thyroid nodules in vulnerable populations such as Filipinos.

The specific objectives of the study include determining the impact of nodule size and recommendations on the performance of the three classification systems further. To give reliable and applicable clinical data and evidence in constructing a suitable TI-RADS scoring system for the population of the Philippines.

Methodology

Research Design/ Data Collection A retrospective study of the medical records of patients who underwent thyroidectomy at Makati Medical Center between January 1, 2018- December 2019 Patients of all ages will be included in this study. Only patients who m*et al*l the following criteria were included in this study: (1) total or nearly total thyroidectomy or lobectomy performed; (2) complete preoperative US of thyroid nodules; and (3) surgical pathology. Non-mass-forming lesions and nodules that failed to meet the criteria for any pattern were excluded. Moreover, patients who underwent other diagnostic procedures aside from the ones mentioned above (e.g., fine needle biopsy) were not included in the study. Thyroid nodules were divided into two groups according to the maximal diameter.

All patient medical record data (i.e., age and sex), ultrasound images will be extracted using the Makati Medical Center system (RIS/PACS) and chart records creating a census and the requirement for random sampling procedure. Patient data extracted will be written in a standard data collection form (Appendix A). The said data gathered will then be tabulated as seen in the dummy tables in Appendix B.

All US examinations were performed with GE Logiq s8, GE Logiq e9, and GE Logiq p6, devices equipped with either a 5–12MHz or an 8–15 MHz linear-array transducer. US images will be retrospectively reviewed by two radiologists who are experienced in thyroid US and blinded to the patients' clinical data and pathological results.

The characteristics of nodules that will be documented size, number (single, multinodular), composition (cystic/spongiform, mixed, solid), echogenicity (anechoic, hyperechoic/isoechoic, hypoechoic, markedly hypoechoic), Orientation (wider-than-tall, taller-than-wide), margins (well defined/ill-defined, lobulated/irregular,extra-thyroidal extension), calcification (none, micro-calcifications, macro-calcifications, rim-calcification)

Two experienced radiologists classified the degree of suspicion of thyroid nodule according to TI-RADS (proposed by ACR, Kwak, and EU) independently. If there were differences, they discussed to get an agreement.

Sample Size

The sample size was computed using OpenEpi online software, Version 3. Based on the study of Liu, S. Y., *et al*., out of the 1,612 nodules included in the study, 773 (48.0%) were malignant. The margin of error used is 8%. At a 95% confidence level, the result showed that the sample size is 150.

Accounting for the missing data, there should be added 10%. The computation is as follows:

N_final=150+.10N_final

0.90N_final=150

N_final=166.67

The final sample size is 167 (always round up).

Statistical Analysis

Plan for data processing and analysis Quantitative data are presented as the mean ± standard deviation (SD). Qualitative data are presented as frequencies. The $\chi 2$ test with Yates' correction and Fisher's exact test will compare categorical variables while t-test when comparing quantitative data. The sensitivity, specificity, and area under the curve (AUC) were computed to evacuate the diagnostic accuracy of the three TI-RADS. The best cut-off score was determined using the suggestion of youden index. A pairwise comparison of AUC was made using Hanley and McNeil method. A value of $P < 0.05$ will be considered statistically significant. Medcalc Statistical software version 20.008 was used to carry out statistical calculations.

Ethical Considerations

This study will be conducted following the ethical principles based on the Declaration of Helsinki and the National Guidelines for Biomedical Research of the National Ethics Committee (NEC) of the Philippines. The research proposal will undergo approval from the Institutional Review Board (IRB).

Data recorded will only be in writing and not recorded via video nor audio. Collected data will be kept strictly confidential by the primary investigator with password protection. Each research subject and the subjects' names will not appear in any of the data collection tools. All identifiable information and data will be given a code number. Only the research team will have access to data safety, privacy, and confidentiality. Data of vulnerable subjects included in the study with be gathered. A master list with a corresponding code number linking to the subject identity will be used and kept separately from the research data. Only the primary investigator will have access to the patient's names and other pertinent information to ensure patient confidentiality at all times. Used data collection forms will be shredded after all data have been encoded into an Excel spreadsheet. Specific names will not appear on the documents or spreadsheets for data analysis and in the final paper. Individually identifiable research data will not be shared with others outside of the research and analysis team. The primary investigator and all research team members will have updated and completed the Good Clinical Practice (GCP) training. The investigator agrees to hold in trust and confidence all information obtained, which will only be used for the contemplated purpose and none other. The investigators will neither discuss nor disclose the obtained information to anyone outside of the research team. The data will be stored in the primary investigator's database, password-protected, and the projected duration of storage is at least ten years.

Results and Data Analysis

A total of 190 patients were included in the study where 27.9% are malignant. The average age of the patients is 50.9 (±13.7) and no significant difference on the age of malignant patients (50 years old) when compared to benign (51.3 years old). Table 1 also shows that only 16.6% are single nodular, among benign 17.2% are singular while 15.1% among malignant, and no significant difference exist (p=.7323). On size with suspicious features, mostly are from 2.5 to 5cm (45.7%) while 18.1% are from 1.5 to 3%. No significant difference exist between benign and malignant on the size (p=.8590).

Table 1: Profile of Respondents

	Overall (n=190)	Benign (n=137)	Malignant (n=53)	p value
Age (years)	50.9 ± 13.7	51.3 ± 13.0	49.8 ± 15.4	0.5228ns
Gender				
Male	34 (17.9)	23 (16.8)	11 (20.8)	0.3833 ns
Female	156 (82.1)	114 (83.2)	42 (79.2)	
Single/Multinodular				
Single	31 (16.6)	23 (17.2)	8 (15.1)	0.7323 ns
Modular	156 (83.4)	111 (82.8)	45 (84.9)	
Size with suspicous features				
>0.5- 1	23 (12.2)	17 (12.6)	6 (11.3)	0.8590 ns
>1.0- 2	31 (16.5)	22 (16.3)	9 (17.0)	
>1.5- 3	34 (18.1)	22 (16.3)	12 (22.6)	
>2.0- 4	14 (7.4)	11 (8.1)	3 (5.7)	
>2.5- 5	86 (45.7)	63 (46.7)	23 (43.4)	

*significant, ns not significant

Table 2 shows that majority are solid in terms of composition (85.2%) while mostly are either smooth or ill-defined on margins (78.9%). Moreover, on echogenicity, 54.1% are classified as hyperchoic or isoechoic while 28.1% are hypoechoic. On echogenic foci, 40% are none, but 43.8% are punctuate. On shape, results show that only 15.2% are described to be non-oval or taller than wide. This sonographic findings except for margins are significantly associated with malignancy classification. Specifically, results show that malignant is significantly associated with solid composition, while mixed and cystic or spongiform are associated with benign. Moreover, both marked hypoechoic and hypoechoic are significantly associated with malignant results while hyperechoic or isoechoic or anechoic are significantly associated with benign. Results also show that on echogenic foci, puncture is significantly associated with malignant while the rest such as none, macro and peripheral are significantly associated with benign.

On shape, non-oval or taller than wide is significantly associated with malignant while being oval is significantly associated with benign results.

Table 2: Sonographic Findings

	Overall (n=191)	Benign (n=137)	Malignant (n=54)	p value
Composition				
Cystic/spongiform	4 (2.1)	3 (2.2)	1 (1.9)	0.0197*
Mixed	24 (12.7))	23 (16.9)	1 (1.9)	
Solid	161 (85.2)	110 (80.9)	51 (96.2)	
Margins				
Smooth/Ill defined	146 (78.9)	110 (82.7)	36 (69.2)	0.0596ns
Lobulated/Irregular	38 (20.5)	22 (16.5)	16 (30.8)	
Extrathyroidal	1 (0.5)	1 (0.8)	0 (0.0)	
Echogenicity				
Anechoic	4 (2.2)	4 (3.0)	0 (0.0)	0.0002*
Hyperechoic or isoechoic	100 (54.1)	83 (62.4)	17 (32.7)	
Hypoechoic	52 (28.1)	31 (23.3)	21 (40.4)	
Marked hypoechoic	29 (15.7)	15 (11.3)	14 (26.9)	
Echogenic foci				
None	74 (40.0)	59 (44.4)	15 (28.8)	0.0019*
Macro	13 (7.0)	12 (9.0)	1 (1.9)	
Peripheral	17 (9.2)	15 (11.3)	2 (3.8)	
Punctuate	81 (43.8)	47 (35.3)	34 (65.4)	
Shape				
Oval/ wider than tall	156 (84.8)	119 (90.2)	37 (71.2)	0.0013*
Non oval/ Taller than wide	28 (15.2)	13 (9.8)	15 (28.8)	

*significant, ns not significant

Table 3 shows that on ACR TI-RADS Classification, 38.9% of them are highly suspicious while only 3.2% are benign. Results show that only TR5 or highly suspicious is significantly associated with malignancy while TR1 to TR4 are associated with benign. On EU TI-RADS classification, 55.3% are high risk while only 3.7% are benign. Results show that only TR5 or high risk is significantly associated with malignant while the rest are towards benign. On Kwak classification, 47.9% are highly suspicious, while 41.6% are intermediate. Results show that only high suspicious are significantly associated with malignant.

Table 3: Malignant rates in the categories of ACR, EU, and Kwak TI-RADS

	Overall (n=191)	Benign (n=137)	Malignant (n=54)	p value
ACR				
TR1 - Benign	6 (3.2)	5 (3.6)	1 (1.9)	0.0001*
TR2 - Not Suspicious	14 (7.4)	13 (9.5)	1 (1.9)	
TR3 - Mildly Suspicious	33 (17.4)	26 (19.0)	7 (13.2)	
TR4 - Moderately Suspicious	63 (33.2)	54 (39.4)	9 (17.0)	
TR5 - Highly Suspicious	74 (38.9)	39 (28.5)	35 (66.0)	
EU				
TR2 Benign	7 (3.7)	6 (4.4)	1 (1.9)	0.0182*
TR3 - Low risk	62 (32.6)	52 (38.0)	10 (18.9)	
TR4 - Intermediate risk	16 (8.4)	13 (9.5)	3 (5.7)	
TR5 - High risk	105 (55.3)	66 (48.2)	39 (73.6)	
Kwak				
TR2 - Benign	5 (2.6)	4 (2.9)	1 (1.9)	0.0002*
TR3 - Low suspicions	15 (7.9)	14 (10.2)	1 (1.9)	
TR4 - Intermediate suspicion	79 (41.6)	67 (48.9)	12 (22.6)	
TR5 - High suspicion	91 (47.9)	52 (38.0)	39 (73.6)	

*significant, ns not significant

Table 4 shows that based on ROC, TI-RADS ACR best cut-off score is >4 or referring to TR5 highly suspicious, that can predict malignant outcome. AUC of 0.69 (95% CI 0.62 to 0.75), p value of 0.0001 denotes significant discrimination of ACR in predicting malignancy. Sensitivity is at 66% (95% CI 51.7% to 78.5%) while specificity is 71.5% (95% CI 63.2% to 78.9%). On TI-RADS EU, resulting AUC of 0.63 (95% CI 0.55 to 0.70) also denotes significant discrimination. It has slightly higher sensitivity of 73.6% (95% CI 59.7% to 84.7%) but very low specificity of only 51.8% (95% CI 43.1% to 60.4%). On TI-RADS Kwak Classification, AUC of 0.68 (95% CI 0.61 to 0.74) is also significant in discriminating between benign n and malignant.

Its sensitivity is 73.6% (95% CI 59.7% to 84.7%) while specificity of 62% (95% CI 53.4% to 70.2%). Figure 1 shows the three different AUC and no much difference as curves are overlapping at some points. Specifically, ACR has slightly higher AUC than EU but difference (0.061) is not significant (p=.3264), while same AUC when compared to Kwak (p=.8703) with only a difference of 0.010. Likewise, Kwak has slightly higher AUC (0.051) than EU, but its difference is also not significant (p=.4066).

Table 4: Diagnostic performance of ACR, EU, and Kwak TI-RAD

TIRADS	Best cut-off	Sensitivity (95% CI)	Specificity (95% CI)	NPV	PPV	AUC (95% CI)	p value
ACR	>4	66.0 (51.7 – 78.5)	71.5 (63.2 - 78.9)	84.5 (78.7 - 88.9)	47.3 (39.3 - 55.5)	0.69 (0.62 - 0.75)	0.0001*
EU	>4	73.6 (59.7 – 84.7)	51.8 (43.1 - 60.4)	83.5 (75.9 - 89.1)	37.1 (31.8 - 42.8)	0.63 (0.55 - 0.70)	0.0004*
KWAK	>4	73.6 (59.7 – 84.7)	62 (53.4 - 70.2)	85.9 (79.2 - 90.7)	42.9 (36.5 - 49.5)	0.68 (0.61 - 0.74)	0.0001*

*significant, ns not significant

Table 5: Comparison of diagnostic performance AUC of ACR, EU and Kwak TI-RADS

	Difference	p value
ACR vs EU	0.061	0.3264
ACR vs Kwak	0.010	0.8703
EU vs Kwak	-0.051	0.4066

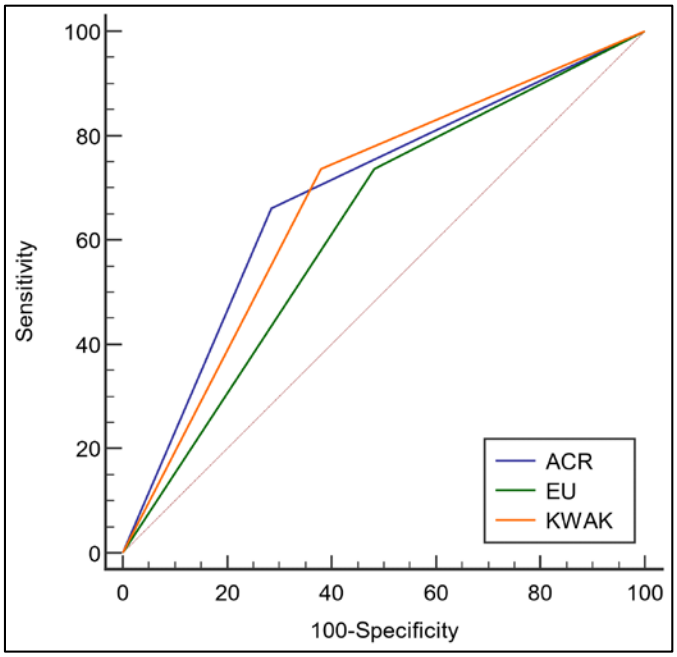

Fig 1: Comparison of ROC of ACR, EU and Kwak TI-RADS

At most 1.5 cm size, the three ACR, EU, Kwak turns out to have the same AUC curve. This implies that their diagnostic capability is also the same. Although it can be said, that EU has slightly lower AUC of only 0.54 as compared to 0.62 for ACR and Kwak. Additionally, EU has slightly lower specificity of only 41%, as compared to 56.4% for ACR and Kwak.

Each ACR (p=.1755), EU (p=.6602) and Kwak (p=.1755) turs out to be not significant discriminatory in predicting malignant among less than 1.5 size

Table 6: Diagnostic performance to size <1.5 cm of ACR, EU, and Kwak TI-RADS

	n	SN	SP	AUC	p value
ACR	54	66.7 (38.4 to 88.2)	56.4 (39.6 to 72.2)	0.62 (0.5 to 0.7)	0.1755ns
EU	54	66.7 (38.4 to 88.2)	41 (25.6 to 57.9)	0.54 (0.4 to 0.7)	0.6602 ns
KWAK	54	66.7 (38.4 to 88.2)	56.4 (39.6 to 72.2)	0.62 (0.5 to 0.7)	0.1755 ns

Table 7: Comparison of Diagnostic performance to size >1.5 cm AUC of ACR, EU and KWAK TI-RADS

	Difference	p value
ACR vs EU	0.08	0.5124ns
ACR vs Kwak	0.00	1.0000 ns
EU vs Kwak	-0.08	0.5124 ns

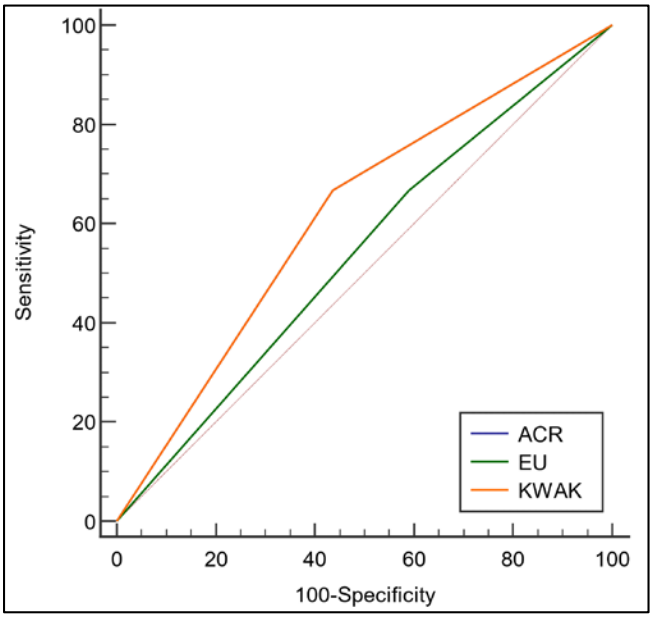

Fig 2: Comparison of size <1.5 cm ROC of ACR, EU and Kwak TI-RADS

At least 1.5 cm size, the three ACR, EU, Kwak turns out to have the same AUC curve. This implies that their diagnostic capability is also the same. Although it can be said, that EU has slightly lower AUC of only 0.66 as compared to 0.71 for ACR and 0.70 for Kwak. Moreover, EU and Kwak are slightly more sensitive (76.3% vs 65.8%) than ACR, while ACR is slightly higher specificity of 77.1%. Each ACR (p=.0001), EU (p=.0002) and Kwak (p=.0001) turns out to be significant discriminatory in predicting malignant among above 1.5 size.

Table 8: Diagnostic performance to size >1.5 cm of ACR, EU, and Kwak TI-RADS

	n	SN	SP	AUC	p value
ACR	134	65.8 (48.6 to 80.4)	77.1 (67.4 to 85)	0.71 (0.6 to 0.8)	0.0001*
EU	134	76.3 (59.8 to 88.6)	55.2 (44.7 to 65.4)	0.66 (0.6 to 0.7)	0.0002*
KWAK	134	76.3 (59.8 to 88.6)	63.5 (53.1 to 73.1)	0.70 (0.6 to 0.8)	0.0001*

Table 9: Comparison of Diagnostic performance to size >1.5 cm AUC of ACR, EU and Kwak TI-RADS

	Difference	p value
ACR vs EU	0.05	0.4899ns
ACR vs KWAK	0.01	0.8887 ns
EU vs KWAK	-0.04	0.5739 ns

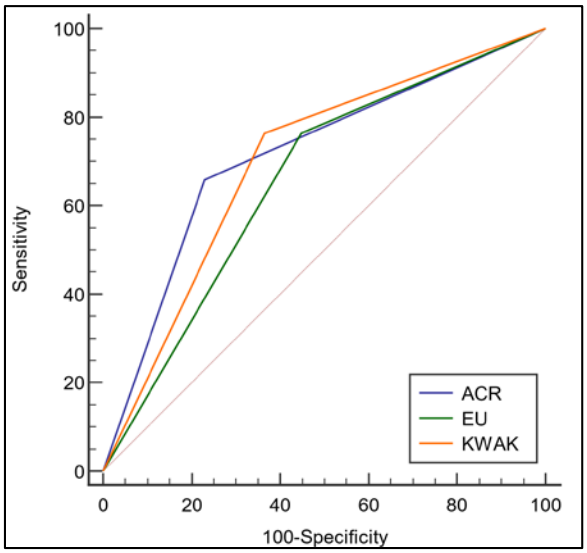

Fig 3: Comparison of size <1.5 cm ROC of ACR, EU and Kwak TI-RADS

Discussion

Risk stratification systems such as ACR, EU, and Kwak TI-RADS have all been compared before. However, their efficacy and accuracy to high-risk populations, such as Filipinos, are not well evaluated. The main reason we sought in the current study is to find a starting point and appraise the TI-RADS developed by the ACR, EU, and Kwak and the advantages they may have over one another concerning the Filipino population. We found that ACR, EU, and Kwak risk stratifications systems exhibited similar diagnostic performance. The study results reveal that ACR, EU, and Kwak TI-RADS show no significant difference in discriminating between benign and malignant nodules in a Filipino population. This is confirmed by the AUC estimate of 0.69, 0.63, and 0.68, respectively, with overlapping points at some point. In the study, Among the three scoring systems, ACR TI-RADS is the most specific with a specificity of 71.5% (95% CI 63.2% to 78.9%).

EU and Kwak TI-RADS are more sensitive than ACR, with a sensitivity of 73.6% (95% CI 59.7% to 84.7%) versus sensitivity of 66% (95% CI 51.7% to 78.5%).

Results show that the suspicious ultrasound features for ACR TI-RADS, EU TI-RADS, and Kwak TI-RADS accurately predicted malignancy, with malignancy odds significantly higher in nodules with a high level of suspicion by ultrasound (TR5 highly suspicious) as seen in ROC curves in Figure 1.

These results coincide with studies that have validated the ACR, EU, and Kwak TI-RADS in non-Filipino adult populations. Higher TI-RADS categories remained associated with an increased likelihood of malignancy on multivariable analysis. However, no significant difference exists between benign and malignant in terms of size, as seen in tables1 and 2. The three scoring systems do not consider size in the assignment of TI-RADS score, but ACR TI-RADS does use size to suggest the management of nodules. TI-RADS may correlate that in Filipinos, it is recommended that sonographically suspicious nodules undergo FNAB regardless of size because malignancy is better predicted by ultrasound features, clinical factors, and possibly genetics and race rather than size alone. We also found that ACR, EU, and Kwak TI-RADS guidelines showed better diagnostic efficiency in differentiating nodules >1.5 cm. But no difference in diagnostic efficiency among the three guidelines for nodules with a size ≤1 cm.

A low threshold of suspicion should be applied in Filipino patients, and a high specificity scoring system should be adapted. This study shows that the ACR TI-RADS, EU-TI-RADS, and Kwak TI-RADS discriminate well between malignant and benign nodules, useful in malignancy risk stratification, particularly when at TI-RADS category 5, and is applicable in the Filipino population. However, practitioners should be aware of the increased rate of false positives at higher TIRADS categories, as well as the possibility of false negatives, especially in subcentimete.

Conclusion

The three TI-RADS scoring systems discriminate between malignant and benign nodules equally in the Filipino population, particularly at TI-RADS category 5.

In Filipinos, it is recommended that sonographically suspicious nodules undergo FNAB regardless of size because ultrasound features better predict malignancy.

Recommendations

We, therefore, suggest that any TIRADS can be helpful as a decision-making tool but leans towards ACR over EU and Kwak TI-RADS in the management of Filipino thyroid nodules for the sole reason of having a higher specificity in the high-risk population.

Because of the different biologic behavior of thyroid cancer in the Filipinos versus non-Filipino population, protocol adjustment should be applied and looked upon. Limitations of our study include its retrospective design that resulted in different ultrasound machines and potential variation in image quality. Analyses were based on the recorded images and thus may have led to misdiagnosis by TI-RADS. Second, all of the patients underwent thyroidectomy, which may have led to selection bias, resulting in underestimating or overestimating variables or ACR TI-RADS, EU TI-RADS, and KWAK-TIRADS. The small study population is another restriction and an even smaller number of malignant nodules (n = 53), possibly limiting the reflection of the actual population effect. Differences between studies may be partly due to the reference standards, inter-observer variability, and the study population.

References

1. Kim J, Gosnell J, Roman S. Geographic influences in the global rise of thyroid cancer. Nature Reviews Endocrinology,2019:16(1):17-29.
2. Nguyen M, Hu J, Hastings K, Daza E, Cullen M, Orloff L, Palaniappan L. Thyroid cancer mortality is higher in Filipinos in the United States: An analysis using national mortality records from 2003 through 2012. Cancer,2017:123(24):4860-4867.
3. Huang BL, Ebner SA, Makkar JS, Bentley-Hibbert S, McConnell RJ, Lee JA *et al*. A Multidisciplinary Head-to-Head Comparison of American College of Radiology Thyroid Imaging and Reporting Data System and American Thyroid Association Ultrasound Risk Stratification Systems. Oncologist,2020P:25(5):398-403. doi: 10.1634/theoncologist.2019-0362. Epub 2019 Nov 19. PMID: 31740569; PMCID: PMC7216459.
4. Yang R, Zou X, Zeng H, Zhao Y, Ma X. Comparison of Diagnostic Performance of Five Different Ultrasound TI-RADS Classification Guidelines for Thyroid Nodules. Frontiers in Oncology, 2020, 10.
5. Horvath E, Majlis S, Rossi R *et al*. An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. J ClinEndocrinolMetab,2009:94:1748-51.
6. 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,2011:260:892-9.
7. Park JY, Lee HJ, Jang HW *et al*. A proposal for a thyroid imaging reporting and data system for ultrasound features of thyroid carcinoma. Thyroid,2009:19:1257-64.
8. 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 CollRadiol,2017:14:587-95.
9. Russ G, Bonnema SJ, Erdogan MF *et al*. European Thyroid Association Guidelines for Ultrasound Malignancy Risk Stratification of Thyroid Nodules in Adults: The EU-TIRADS. Eur Thyroid J,2017:6:225-37.
10. Modi L, Sun W, Shafizadeh N, Negron R, Yee-Chang M, Zhou F *et al*. Does a higher American College of Radiology Thyroid Imaging Reporting and Data System (ACR TI-RADS) score forecast an increased risk of malignancy? A correlation study of ACR TI-RADS with FNA cytology in the evaluation of thyroid nodules. Cancer Cytopathology,2020:128(7):470-481.

11. Shi Y, Chen L, Liu Y, Zhan J, Diao X, Fang L *et al*. Differences among the Thyroid Imaging Reporting and Data System proposed by Korean, the American College of Radiology, and the European Thyroid Association in the diagnostic performance of thyroid nodules. Translational Cancer Research,2020:9(8):4958-4967.
12. Zhang W, Xu H, Zhang Y, Guo L, Xu S, Zhao C *et al*. Comparisons of ACR TI-RADS, ATA guidelines, Kwak TI-RADS, and KTA/KSThR guidelines in malignancy risk stratification of thyroid nodules. Clinical Hemorheology and Microcirculation, 2020, 1-14.
13. Vargas-Uricoechea H, Meza-Cabrera I, Herrera-Chaparro J. Concordance between the TIRADS ultrasound criteria and the BETHESDA cytology criteria on the nontoxic thyroid nodule. Thyroid Research, 2017, 10(1).
14. Lo T, Uy A, Maningat P. Well-differentiated thyroid cancer: the Philippine General Hospital experience. Endocrine Abstracts, 2015.
15. Baloch ZW, LiVolsi VA. Fine-needle aspiration of thyroid nodules: past, present, and future. EndocrPract,2004:10:234-41.
16. American College of Radiology Committee on TI-RADS™. AACR TI-RADS Assessment categories. Available at: https://www.acr.org/-/media/ACR/Files/RADS/TI-RADS/TI-RADS-chart.pdf?la=en [Accessed 6 April 2021].