AMERICAN JOURNAL OF DIAGNOSTIC IMAGING
2026; 12(3): 98-112
10.5455/ajdi.20260114121548

WisdomGale

ORIGINAL ARTICLE

Open Access

Diagnostic accuracy of thyroid imaging and data system ultrasound classification of thyroid nodules compared with cytological findings among patients in Tanzania

Chausiku Samwel Kisunte[1,2], Clement Kalambo[1,2], Patrick Amsi[3], Adnan Sadiq[1,2], Felister Uisso[1,2]

[1]Department of Radiology and Medical Imaging, School of Medicine, KCM University, Moshi, Tanzania

[2]Department of Radiology and Medical Imaging, Kilimanjaro Christian Medical Centre (KCMC), Moshi, Tanzania

[3]Department of Pathology, Kilimanjaro Christian Medical Centre, Moshi, Tanzania

ABSTRACT

Background: Thyroid nodules (TNs) are common worldwide, particularly among women and in iodine-deficient regions. Although most nodules are benign, a significant proportion may be malignant. Ultrasound (US) is the primary imaging tool for evaluation, and the American College of Radiology Thyroid Imaging Reporting and Data System (ACR TI-RADS) has shown strong performance in risk stratification. This study assessed the diagnostic accuracy of ACR TI-RADS compared with cytological findings at Kilimanjaro Christian Medical Centre.

Methods: A cross-sectional study was conducted among 255 patients who underwent thyroid US and fine-needle aspiration cytology (FNAC) between January 2022 and April 2025. For diagnostic accuracy analysis, Bethesda II–III were grouped as benign, Bethesda IV–VI as malignant, and Bethesda I cases were excluded. ACR TI-RADS categories TR1–TR3 were considered low risk (test-negative), while TR4–TR5 were considered high risk (test-positive). Sensitivity, specificity, predictive values, and overall accuracy were calculated, and receiver operating characteristic curve analysis was performed.

Results: The median age was 47 years, and 86.3% of patients were female. Most nodules were mixed cystic/solid (58.8%), hyperechoic (60.4%), wider-than-tall (89.0%), and had smooth margins (67.5%). Based on ACR TI-RADS, 82% of nodules were low risk and 18% high risk. FNAC classified 83.9% as benign, 11% malignant, and 6.3% suspicious. Malignancy rates rose with higher TI-RADS categories, reaching 88% in TR5. ACR TI-RADS demonstrated high diagnostic accuracy with 90.6% sensitivity, 92.6% specificity, and an area under the curve of 0. 951.

Conclusion: ACR TI-RADS is a highly sensitive and specific, non-invasive tool for TN evaluation. While FNAC remains the reference standard, this single-center study suggests TI-RADS has the potential to reduce unnecessary FNACs, particularly in benign nodules. Limitations include reliance on FNAC, single-center design, and operator dependence, which may affect reproducibility.

ARTICLE HISTORY

Received 14 January 2026
Revised 10 June 2026
Accepted 10 June 2026
Published 26 June 2026

KEYWORDS

Thyroid nodules; ultrasound; ACR TI-RADS; cytology; diagnostic accuracy; Tanzania

Introduction

Thyroid nodules (TNs) are common endocrine abnormalities defined as discrete lesions within the thyroid gland, often detected incidentally during clinical or imaging evaluations [1,2]. The global prevalence of TNs varies depending on detection method and population characteristics. Palpation-based studies report rates of 4%–7%, whereas high-resolution ultrasound (US) detects nodules in up to 50%–68% of adults, particularly among women and iodine-deficient populations [3,4]. Recent systematic reviews and meta-analyses further indicate that imaging-based screening identifies TNs in a significant proportion of the adult population, with pooled estimates suggesting that

Contact Chausiku Samwel Kisunte chausikusamwel64@gmail.com Department of Radiology and Medical Imaging, School of Medicine, KCM University, Moshi, Tanzania.

© The Author(s) 2026. Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY). http://creativecommons.org/licenses/by/4.0/.

nearly half of adults may harbor at least one lesion on US examination [5,6].

While most nodules are benign and asymptomatic, a small proportion, less than 15%, may be malignant, predominantly papillary carcinoma [7–9]. In Africa, 11% of TNs are malignant, a figure influenced by widespread iodine deficiency [10]. In Tanzania, thyroid cancer accounts for 1.2% of all cancers treated at the Ocean Road Cancer Institute [11].

US remains the primary imaging modality for evaluating TNs due to its accessibility, safety, and high diagnostic performance [12–14]. To enhance consistency in US reporting and risk stratification, standardized systems like the thyroid imaging reporting and data system (TI-RADS) have been developed. Among these, the American College of Radiology (ACR) TI-RADS has gained prominence for its high sensitivity (75%–97%) and utility in minimizing unnecessary fine needle aspiration of benign nodules [7,15,16].

The ACR TI-RADS categorizes nodules based on US features such as composition, echogenicity, shape, margins, and calcifications, assigning a malignancy risk score (TR1–TR5). Higher TR categories correlate with increasing malignancy risk [17–20]. Despite its growing global use, no prior study in Tanzania has assessed the diagnostic accuracy of ACR-TIRADS against cytological findings. Given the operator-dependent nature of US and variability in equipment resolution across diagnostic centers, local validation is essential to ensure reproducibility and generalizability.

Fine needle aspiration cytology (FNAC), guided by systems such as the Bethesda classification, remains the gold standard for initial evaluation of malignancy in TNs [21,22]. However, FNAC has limitations, including sampling errors, indeterminate results, and reliance on cytological interpretation. Accurate US-based risk stratification can therefore reduce overdiagnosis and overtreatment of benign nodules, particularly in resource-limited settings where histopathology services may be constrained.

Considering the financial burden on healthcare systems and the anxiety for patients, it is neither feasible nor appropriate to perform FNAC on every TN. This study aims to evaluate the diagnostic accuracy of ACR TI-RADS US classification of TNs compared to cytological findings at Kilimanjaro Christian Medical Centre (KCMC). By clarifying its predictive performance in this setting, the results may inform better diagnostic pathways, improve patient outcomes, and support efficient use of limited resources in low- and middle-income countries [8,13,23].

Material and Methods

Study design, site, and population

A hospital-based retrospective cross-sectional study was conducted at KCMC, a tertiary referral and teaching hospital in Moshi, Tanzania, over a 3-year period from January 2022 to April 2025. The study included patients aged 18 years and above who underwent both thyroid US and FNAC during the study period. FNAC was performed based on clinician discretion guided by ACR TI-RADS recommendations and local protocols.

Sample size and sampling procedures

The minimum sample size of 255 was calculated using Buderer's formula for diagnostic accuracy studies, assuming a sensitivity of 82% and disease prevalence of 89%. The assumed prevalence of 89% used in Buderer's formula reflected the proportion of TNs undergoing FNAC in our tertiary referral setting rather than malignancy prevalence. However, the observed cytological malignancy rate in the study population was substantially lower (11%). As the study included all eligible patients during the study period, the final sample represents a retrospective convenience cohort.

Variables assessed included US features (composition, echogenicity, shape, margins, calcifications, and vascularity), demographic data (age, sex, and residence), and cytological findings.

Patient electronic records (Electronic Hospital Management System), US image archives, and pathology reports were systematically reviewed. Only cases with complete and matching US and FNAC reports were included to ensure data integrity and diagnostic comparability. A flow diagram (Fig. 2) summarizes patient inclusion and exclusion.

Data collection

A structured data collection tool was used to retrospectively capture patient demographics, clinical presentation, US characteristics based on ACR TI-RADS, and corresponding FNAC results reported using the Bethesda System for Reporting Thyroid Cytopathology (TBSRTC). Data entry was

cross-checked by two independent reviewers to minimize transcription errors.

US assessment

Thyroid US examinations were performed using two Siemens ACUSON NX3 Elite machines equipped with high-frequency linear transducers (7–12 MHz). Routine quality control checks were conducted to ensure consistent image resolution across machines. Patients were scanned in the supine position with the neck extended, and images were obtained in both transverse and longitudinal planes.

Sonographic features were recorded and scored according to ACR TI-RADS criteria. The sum of points determined the TI-RADS category: TR1 (0 points), TR2 (2 points), TR3 (3 points), TR4 (4–6 points), and TR5 (≥7 points) (Fig. 1).

In addition to the five standard ACR TI-RADS lexicon features, Doppler vascularity patterns and presence of cervical lymphadenopathy were recorded and analyzed separately as complementary US findings. These variables are not components of the ACR TI-RADS scoring system and were evaluated independently for exploratory association with malignancy.

Images were interpreted independently by the principal investigator and subsequently reviewed by a senior radiologist during weekly consensus meetings. Discrepancies were resolved by joint discussion to reach agreement. Interobserver agreement was not formally quantified using kappa statistics, which is acknowledged as a limitation.

Cytology assessment

FNAC results were interpreted according to the TBSRTC. For diagnostic accuracy analysis, Bethesda I (non-diagnostic) cases were excluded. Bethesda II–III were grouped as benign, while Bethesda IV–VI were grouped as malignant to allow construction of 2 × 2 contingency tables for sensitivity and specificity calculations. This methodological choice explains the difference between descriptive malignant case counts (Bethesda V–VI only) and the higher number used in diagnostic accuracy analysis (Bethesda IV–VI) (Table 1).

Data analysis

Data were analyzed using SPSS version 25.

- **Descriptive statistics:** Continuous variables were summarized using measures of central tendency and dispersion; categorical variables were presented as frequencies and proportions.
- **Comparative analysis:** Cross-tabulation of ACR TI-RADS categories versus Bethesda cytology results was performed. Associations between categorical variables were tested using chi-square, as appropriate.
- **Diagnostic accuracy analysis:** ACR TI-RADS categories TR1–TR3 were considered TI-RADS-negative (low risk), while TR4–TR5 were considered TI-RADS-positive (high risk). Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall accuracy were calculated using 2 × 2 contingency tables.
- **Receiver operating characteristic (ROC) analysis:** ROC curve analysis was performed to evaluate diagnostic performance and determine the optimal TI-RADS cutoff value for predicting malignancy. The area under the curve (AUC) was used to assess the discriminatory ability of US in differentiating benign from malignant nodules. For ROC analysis, ACR TI-RADS categories were treated as an ordinal variable (TR1–TR5). The optimal threshold for predicting malignancy was determined using the Youden index (J = sensitivity + specificity − 1). The cutoff corresponding to TR4 provided the best balance between sensitivity and specificity. No adjustment for covariates was performed.

Results

Patient inclusion

A total of 605 patients were assessed for eligibility. Of these, 350 were excluded: 64 had neck pathologies unrelated to TNs, 86 lacked cytological results, and 100 had missing US images or reports. The remaining 255 patients were included in the study. Among these, 7 cases were non-diagnostic (Bethesda I), 214 were benign (Bethesda II), 6 were indeterminate (Bethesda III–IV), and 28 were malignant (Bethesda V–VI). For diagnostic accuracy analysis, non-diagnostic cases were excluded, leaving 248 patients. In this analysis, Bethesda II–III were grouped as benign, while Bethesda IV–VI were grouped as malignant, resulting in 32 malignant nodules compared to 28 malignant nodules (Bethesda V–VI only) in descriptive reporting. This

methodological adjustment ensured consistency in constructing 2 × 2 contingency tables for sensitivity and specificity calculations. The patient selection process is summarized in Figure 2.

Demographic and clinical characteristics of study participants

The median age of participants was 47 years (IQR: 35–58), with most aged between 39 and 58 years (117, 45.9%). The majority were female (220, 86.3%) and managed as outpatients (232, 91.0%). Symptom duration varied, with 112 (43.9%) reporting symptoms for 1–5 years. Nearly half resided in Moshi Urban (123, 48.2%), followed by Arusha (35, 13.7%) and Rombo (19, 7.5%) (Table 2)

US characteristics by ACR TI-RADS

Most nodules were mixed cystic-solid (150, 58.8%) and hyperechoic (154, 60.4%). The majority were wider-than-tall (227, 89.0%), had smooth margins

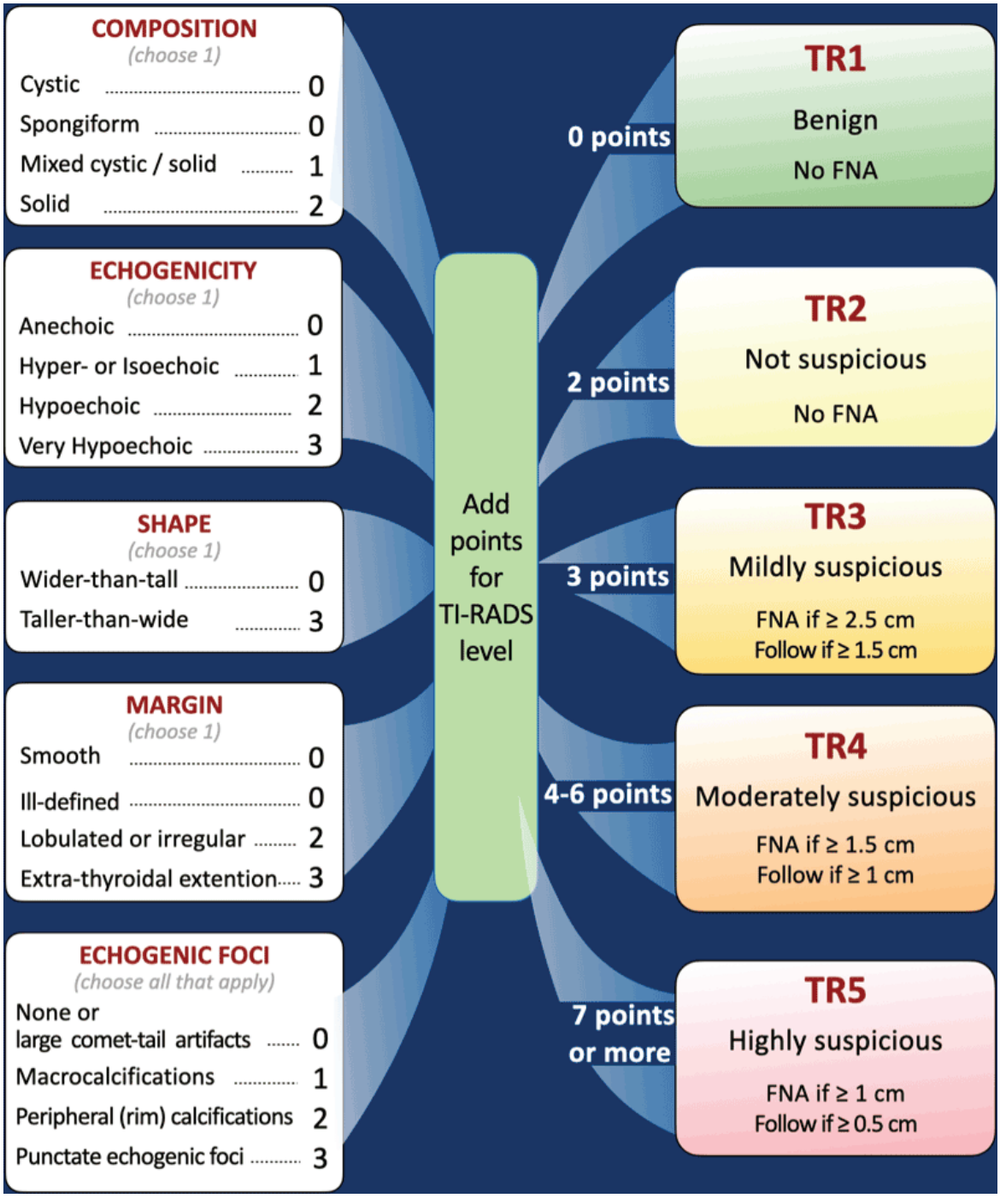


Figure 1. Showing five categories on the basis of the ACR TI-RADS lexicon [12,15].

Table 1. TBSRTC categories and grouping used for diagnostic accuracy analysis.

Diagnostic category	Description	ROM (%)	Management
I	Non-diagnostic/unsatisfactory	1–4	Repeat FNAC by US guidance
II	Benign	0–3	Clinical follow-up and US monitoring
III	Atypia or follicular lesion of undetermined significance	5–15	Repeat FNAC, molecular testing, or surgery
IV	Follicular neoplasm or suspicious for follicular neoplasm	15–30	Surgical lobectomy for definitive diagnosis
V	Suspicious for malignancy	60–75	Lobectomy or total thyroidectomy
VI	Malignant	97–99	Total thyroidectomy + adjuvant therapy

Source: The 2017 TBSRTC [22].

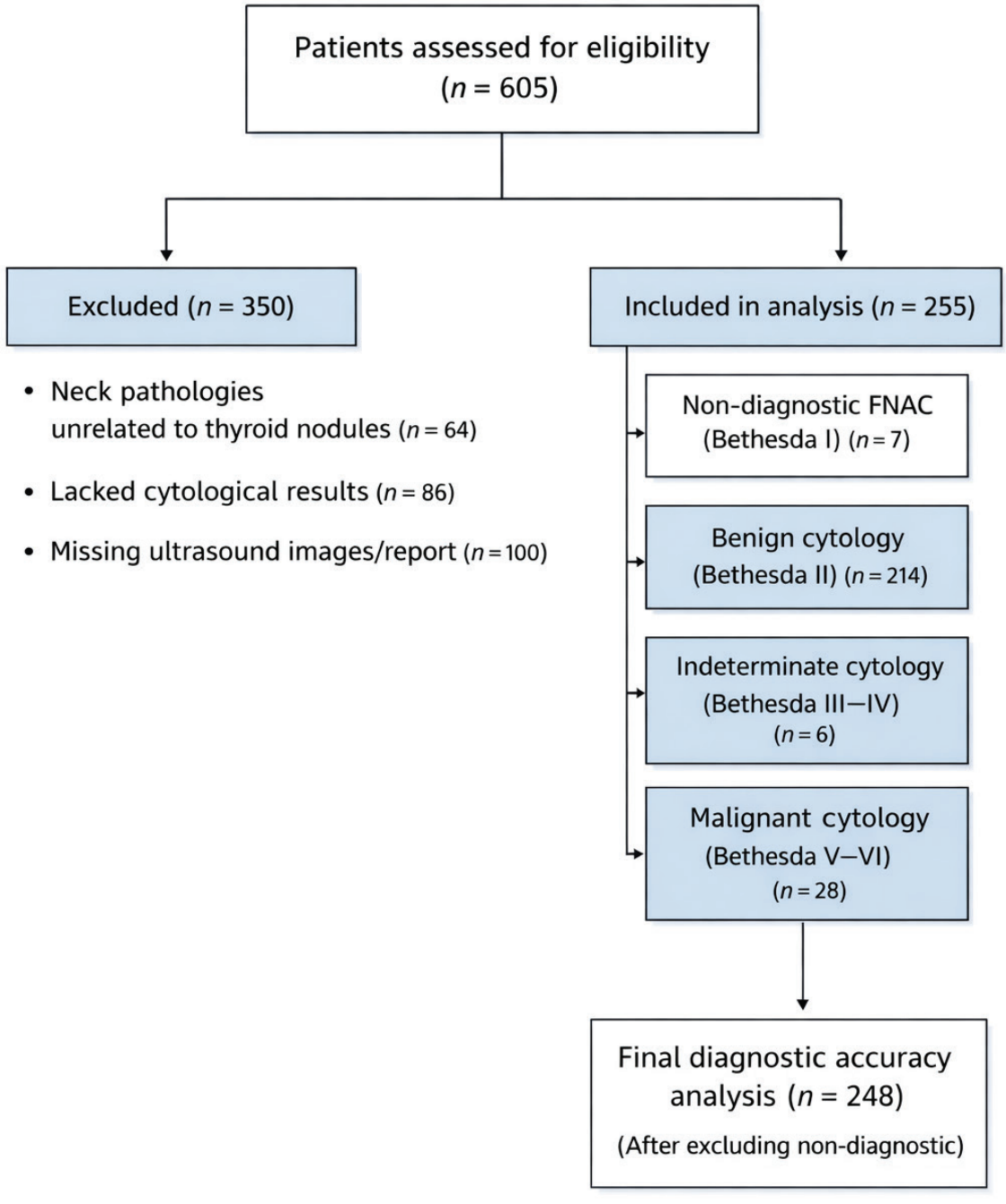


Figure 2. Participant recruitment flow chart diagram.

Table 2. Demographic and clinical characteristics of study participants with TNs (*N* = 255).

Demographic characteristics	*n*	%
Age (Years), Median (IQR)	47 (35–58)	-
Age category (years)		
19–38	76	29.8
39–58	**117**	**45.9**
59–68	40	15.7
≥69	22	8.6
Sex		
Female	**220**	**86.3**
Male	35	13.7
In/Outpatient		
In-patient	23	9.0
Outpatient	**232**	**91.0**
Duration of symptoms (years)		
< 1year	63	24.7
1–5	**112**	**43.9**
6–10	16	6.3
11–15	31	12.2
>15	33	12.9
Region/Residence		
Arusha	35	13.7
Moshi Urban	**123**	**48.2**
Moshi Rural	5	2.0
Same	5	2.0
Mwanga	8	3.1
Hai	16	6.3
Siha	10	3.9
Rombo	19	7.5
Others	34	13.3

Bold values denote the predominant category within each variable. IQR = Interquartile Range.

(172, 67.5%), and lacked echogenic foci or exhibited comet-tail artefacts (204, 80.0%). Multiple nodules were present in 168 patients (65.9%), and 207 nodules (81.2%) measured more than 2 cm. A regular halo sign was observed in 24 cases (9.4%). On Doppler, peripheral vascularity was most common (97, 38.0%), followed by absent (75, 29.4%), mixed (46, 18.0%), and internal vascularity (37, 14.5%). Cervical lymphadenopathy was noted in 25 patients (9.8%). According to ACR TI-RADS, most nodules were categorized as TR3 (40%), followed by TR2 (31%) and TR1 (11%). TR4 accounted for 5% of nodules, while TR5 comprised 13%. Overall, 82% of nodules were low risk (TR1–TR3), and 18% were high risk (TR4–TR5) (Tables 3 and 5).

Cytological findings (Bethesda system)

FNAC revealed that most nodules were benign (Bethesda II, 214 cases, 83.9%). Malignant diagnoses (Bethesda V–VI) accounted for 28 cases (11.0%), including papillary thyroid carcinoma (12, 4.7%), follicular thyroid carcinoma (14, 5.5%), and neuroendocrine carcinoma (2, 0.8%). Indeterminate findings (Bethesda III–IV) were observed in six cases (2.4%), while non-diagnostic results (Bethesda I) occurred in seven cases (2.7%). The most common benign diagnosis was follicular nodule (90, 35.3%), followed by simple nodular goiter (37, 14.5%) and benign follicular neoplasm (30, 11.8%) (Table 4 and 5).

Table 3. US characteristics of TNs by ACR TI-RADS (*n* = 255).

US feature	*n*	%
Composition		
Cystic	9	3.5
Spongiform	24	9.4
Mixed cystic/solid	150	58.8
Solid	72	28.2
Echogenicity		
Anechoic	18	7.1
Hyperechoic	154	60.4
Hypoechoic	70	27.5
Very hypoechoic	13	5.1
Shape		
Wider-than-tall	227	89.0
Taller-than-wide	28	11.0
Margins		
Smooth	172	67.5
Ill-defined	52	20.4
Lobulated/Irregular	21	8.2
Extra-thyroidal extension	10	3.9
Echogenic foci		
None/comet-tail artefact	204	80.0
Macrocalcification	28	11.0
Peripheral calcification	2	0.8
Microcalcifications	21	8.2
Nodule size		
> 2 cm	207	81.2
	48	18.8
Number of nodules		
Single	87	34.1
Multiple	168	65.9
Halo sign		
Present (regular halo)	24	9.4
Absent	231	90.6
Vascularity (Doppler)		
Internal	37	14.5
Peripheral	97	38.0
Mixed	46	18.0
None	75	29.4
Cervical lymphadenopathy		
Present	25	9.8
Absent	230	90.2
ACR TI-RADS category		
TR1	28	11.0
TR2	79	31.0
TR3	102	40.0
TR4	12	5.0
TR5	34	13.0

Note: Percentages are based on total nodules (*n* = 255).

Table 4. Distribution of cytological findings according to TBSRTC (*N* = 255).

Cytological diagnosis	Malignant	Benign	Indeterminate	Non-diagnostic	Total *n* (%)
PTC	12	-	-	-	12 (4.7)
Follicular thyroid carcinoma	14	-	-	-	14 (5.5)
Neuroendocrine carcinoma grade III	2	-	-	-	2 (0.8)
Benign follicular nodule	-	90	-	-	90 (35.3)
Simple cystic goitre	-	22	-	-	22 (8.6)
Cystic colloid goitre	-	20	-	-	20 (7.8)
Benign follicular neoplasm	-	30	-	-	30 (11.8)
Simple nodular goitre	-	37	-	-	37 (14.5)
Multinodular goitre	-	6	-	-	6 (2.4)
Benign follicular hyperplasia	-	8	-	-	8 (3.1)
Lipomatous tumor	-	1	-	-	1 (0.4)
Atypia or follicular lesion of indeterminate significance	-	-	6	-	6 (2.4)
Non-diagnostic	-	-	-	7	7 (2.7)
Total *n* (%)	28 (10.9)	214 (83.9)	6 (2.4)	7 (2.7)	255

Table 5. Direct comparative table aligning ACR TI-RADS categories with Bethesda categories (*n* = 255).

ACR TI-RADS	Number of nodules (*n*, %)	Bethesda I (*n*, %)	Bethesda II (*n*, %)	Bethesda III (*n*, %)	Bethesda IV (*n*, %)	Bethesda V (*n*, %)	Bethesda VI (*n*, %)
TR1	28 (11.0%)	2 (7.1%)	25 (89.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
TR2	79 (31.0%)	5 (6.3%)	70 (88.6%)	2 (2.5%)	0 (0.0%)	0 (1.3%)	0 (0.0%)
TR3	102 (40.0%)	0 (0.0%)	90 (88.2%)	0 (0.0%)	0 (0.0%)	4 (3.9%)	0 (0.0%)
TR4	12 (5.0%)	0 (0.0%)	4 (33.3%)	0 (0.0%)	4 (33.3%)	5 (41.7%)	3 (25.0%)
TR5	34 (13.0%)	0 (0.0%)	25 (73.5%)	0 (0.0%)	0 (0.0%)	7 (20.6%)	8 (23.5%)
Total	255 (100%)	7 (2.7%)	214 (83.9%)	2 (0.8%)	4 (1.6%)	16 (6.3%)	12 (4.7%)

Association of US features with cytology

A linear-by-linear association test confirmed a statistically significant increasing trend in malignancy rates across ascending ACR TI-RADS categories ($p < 0.001$), with malignancy rates rising from 0% in TR1–TR2 to 88% in TR5. Solid composition, very hypoechoic echogenicity, taller-than-wide shape, irregular or lobulated margins, and microcalcifications were strongly associated with malignancy (all $p < 0.001$). All cystic and spongiform nodules were benign, while 23 of 72 solid nodules (31.9%) were malignant. Very hypoechoic nodules ($n = 13$) had a 100% malignancy rate, though this subgroup was small and should be interpreted cautiously. Malignancy was more common in single nodules (25.6%) than multiple nodules (6.6%) ($p < 0.001$). Internal vascularity and cervical lymphadenopathy were strongly associated with malignancy, with subgroup malignancy rates of 75.7% and 92.0%, respectively ($p < 0.001$). These findings reflect prevalence within subgroups rather than formally calculated predictive values. Nodule size and halo sign were not significantly associated with malignancy ($p > 0.05$) (Table 6).

Diagnostic accuracy of ACR TI-RADS

After excluding non-diagnostic cases for diagnostic accuracy analysis, TI-RADS categories TR4–TR5 were classified as TI-RADS-positive and TR1–TR3 as TI-RADS-negative. ACR TI-RADS demonstrated high diagnostic performance among 248 patients, with sensitivity of 90.6%, specificity of 92.6%, PPV of 64.4%, NPV of 98.5%, and overall accuracy of 92.3%. Nodules classified as TI-RADS 1–3 were significantly less likely to be malignant. The odds of a benign diagnosis were 120.8 times higher in the TI-RADS-negative group, with a relative risk of 2.77 for benign and 0.023 for malignant outcomes ($p < 0.001$) (Table 7).

ROC curve analysis

The predictive performance of ACR TI-RADS was further evaluated using ROC analysis. The

Table 6. Association of US features with cytological findings (n = 255).

US feature	Benign *n* (%)	Malignant *n* (%)	χ^2 (*p*-value)
Composition			23.076 (<0.001)*
Cystic	8 (100.0)	0 (0.0)	
Spongiform	22 (100.0)	0 (0.0)	
Mixed cystic/solid	139 (93.3)	10 (6.7)	
Solid	47 (68.1)	22 (31.9)	
Echogenicity			71.692 (<0.001)*
Anechoic	17 (100.0)	0 (0.0)	
Hyperechoic	146 (97.3)	4 (2.7)	
Hypoechoic	53 (77.9)	15 (22.1)	
Very hypoechoic	0 (0.0)	13 (100.0)	
Shape			190.482 (<0.001)*
Wider-than-tall	214 (97.3)	6 (2.7)	
Taller-than-wide	2 (7.1)	26 (92.9)	
Margins			125.413 (<0.001)*
Smooth	164 (98.8)	2 (1.2)	
Ill-defined	47 (92.2)	4 (7.8)	
Lobulated/Irregular	4 (19.0)	17 (81.0)	
Extra-thyroidal extension	1 (10.0)	9 (90.0)	
Echogenic foci			82.965 (<0.001)*
None/comet-tail artefact	188 (94.9)	10 (5.1)	
Macrocalcification	24 (88.9)	3 (11.1)	
Peripheral calcification	1 (50.0)	1 (50.0)	
Microcalcifications	3 (14.3)	18 (85.7)	
Number of nodules			13.543 (<0.001)*
Single	61 (74.4)	21 (25.6)	
Multiple	155 (93.4)	11 (6.6)	
Size			1.419 (0.234)
> 2 cm	173 (85.6)	29 (14.4)	
	43 (93.5)	3 (6.5)	
Halo sign			2.206 (0.217)
Present	23 (95.8)	1 (4.2)	
Absent	193 (86.2)	31 (13.8)	
Vascularity (Doppler)			123.943 (<0.001)*
Internal	9 (24.3)	28 (75.7)	
Peripheral	91 (96.8)	3 (3.2)	
Mixed	45 (100.0)	0 (0.0)	
None	71 (98.6)	1 (1.4)	
Cervical lymphadenopathy			137.928 (<0.001)*
Present	2 (8.0)	23 (92.0)	
Absent	214 (96.0)	9 (4.0)	

Note: χ^2 = Chi-square test. *$p < 0.001$ indicates statistical significance.

Youden index identified TR4 as the optimal cut-off point for malignancy prediction, corresponding to the binary classification of TR1–TR3 as test-negative and TR4–TR5 as test-positive used in the diagnostic accuracy analysis. The ROC curve (AUC) was 0.951 (95% CI: 0.901–1.000, $p < 0.001$), indicating excellent discriminatory ability (Fig. 3).

Table 7. Diagnostic accuracy of ACR TI-RADS US classification compared to cytological findings (*N* = 248).

Category	Cytological findings		Total	%	95% confidence interval	Formula
	Malignant *n* (%)	Benign *n* (%)				
ACR-TIRAD score						
Malignant (positive)	29 (64.4)	16 (35.6)	45 (100.0)	64.4 (PPV)		
Benign (Negative)	3 (1.5)	200 (98.5)	203 (100.0)	98.5 (NPV)		
Total	32 (12.9)	216 (87.1)	248 (100.0)			
Diagnostic metrics						
Sensitivity	29	3	32	90.6		TP/(TP + FN)
Specificity	200	16	216	92.6		TN/(TN + FP)
PPV	29	16	45	64.4		TP/(TP + FP)
NPV	200	3	203	98.5		TN/(TN + FN)
Accuracy	229	19	248	92.3		(TP + TN)/Total
Risk estimates						
Odds ratio (ACR TI-RADS_Neg / ACR TI-RADS_Pos)				120.833	33.157–440.349	
Cytological findings = Benign				2.771	1.869–4.108	
cytological findings = malignant				0.023	0.007–0.072	
Chi-square				129.950	$p < 0.001$*	

Abbreviations: TP = True Positive; TN = True Negative; FP = False Positive; FN = False Negative; PPV = Positive Predictive Value; NPV = Negative Predictive Value; CI = Confidence Interval.
*Note: $p < 0.05$ considered statistically significant; * indicates $p < 0.001$.

Discussion

This study assessed the diagnostic accuracy of the ACR TI-RADS US classification system in stratifying TNs at KCMC by comparing sonographic features with cytological findings. The findings reaffirmed the high prevalence of TNs, predominantly benign, with malignant nodules being relatively rare but clinically significant. Effective triaging is therefore essential to avoid overdiagnosis and unnecessary interventions while ensuring timely identification of malignancy [15].

The study population was predominantly female, with a mean age of 47 years (range 39–58 years). This pattern may be attributed to hormonal influences, reproductive factors, and differences in healthcare-seeking behavior. Most participants were from Moshi Urban, likely reflecting KCMC's proximity, lower travel costs, and greater health awareness and access to care among urban residents. These findings are consistent with similar studies from Uganda, India, and Mexico, which reported comparable sex distribution, age profile, and urban predominance [10,24,25].

Most TNs demonstrated benign US features corresponding to low-risk ACR TI-RADS categories (TR1–TR3), with only 18% classified as TR4–TR5, requiring FNAC. The predominance of benign nodules likely reflects referral patterns and the generally non-malignant nature of thyroid disease in this setting, consistent with findings from similar international studies [24,26]. The predominance of benign features may help reduce unnecessary procedures, while the reliable identification of suspicious nodules supports early intervention and appropriate management, especially in settings with limited histopathological resources [27].

Most TNs were cytologically benign (83.9%), with benign follicular nodules predominating, while indeterminate and non-diagnostic results were uncommon. These findings mirror those in South Africa and Uganda [10,18]. Interestingly, the study found a relatively high proportion of follicular thyroid carcinoma (FTC, 5.5%), which differs from most literature where papillary thyroid carcinoma (PTC) is more common [28–30]. This may reflect the limitation of cytology in differentiating FTC from follicular adenomas, as definitive diagnosis requires histopathological assessment [21,26,31].

A strong correlation between ACR TI-RADS score and malignancy was observed ($p < 0.001$), with malignancy rates escalating from TR3 (2%) to TR5

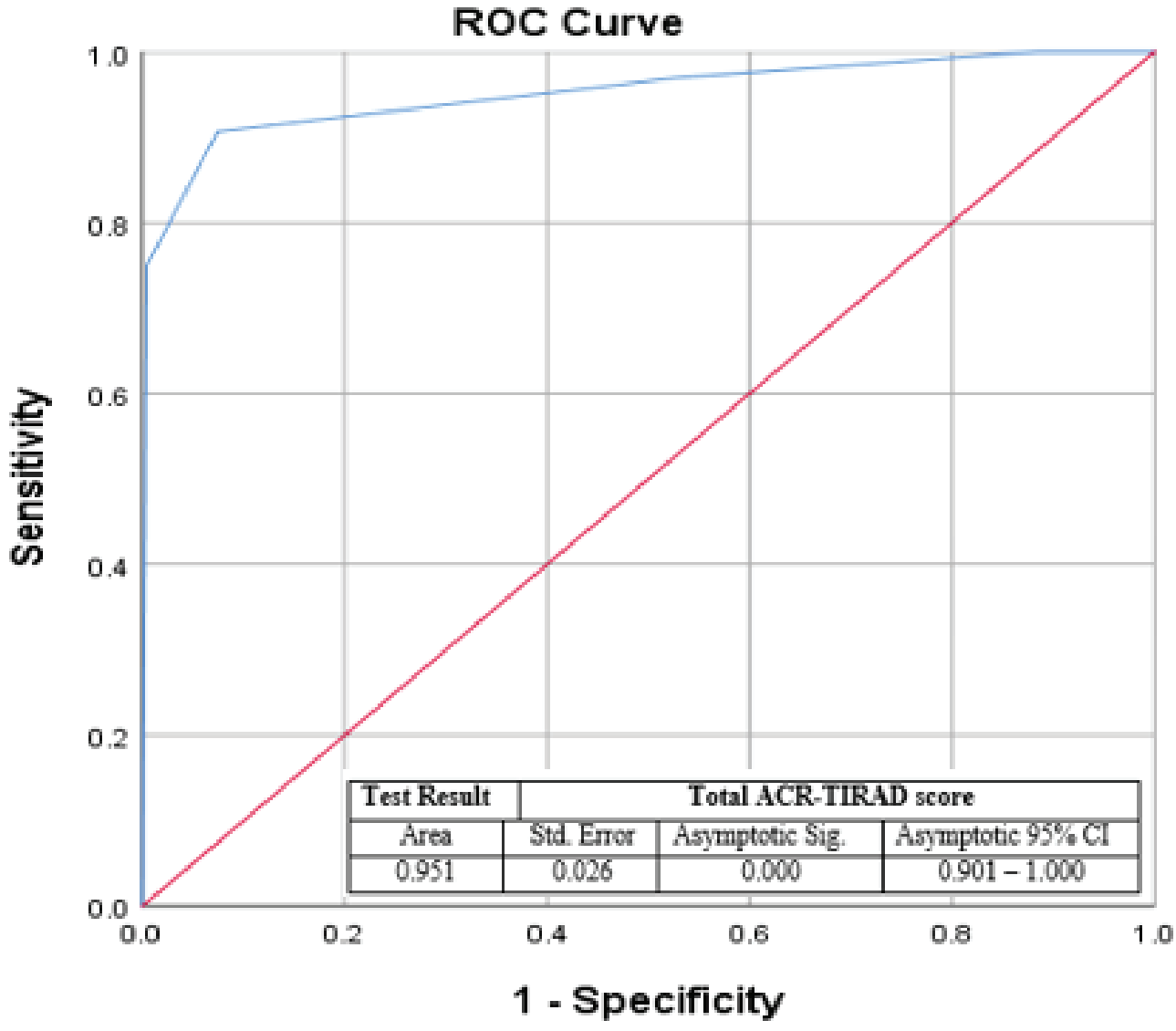


Figure 3. ROC curve for ACR TI-RADS predictive performance.

(88%), validating ACR TI-RADS as a reliable risk stratification tool [10,15,20].

ACR TI-RADS showed excellent diagnostic performance with a sensitivity of 90.6%, specificity of 92.6%, NPV of 98.5%, and accuracy of 92.3%. This supports previous studies in Uganda, the Philippines, and Canada [2,10,32]. Although the PPV (64.4%) was moderate, the high NPV confirms ACR TI-RADS's utility in safely ruling out malignancy in low-risk nodules. The AUC of 0.951 further confirms excellent performance, consistent with findings from China [33,34].

The apparent discrepancy between 28 (11%) malignant nodules in descriptive reporting and 32 (12.9%) malignant nodules in diagnostic accuracy analysis arises from methodological grouping. While Bethesda V–VI represent unequivocally malignant cytology, Bethesda IV lesions (follicular neoplasms) were grouped with malignant categories for statistical purposes, consistent with prior TI-RADS validation studies [25]. This approach facilitated diagnostic accuracy calculations but may slightly overestimate true malignancy, as definitive diagnosis of follicular neoplasms requires histopathological confirmation. By excluding Bethesda I cases and grouping Bethesda IV–VI as malignant, we ensured methodological consistency and maximized the robustness of sensitivity and specificity estimates.

US features significantly associated with malignancy included solid composition, hypoechogenicity,

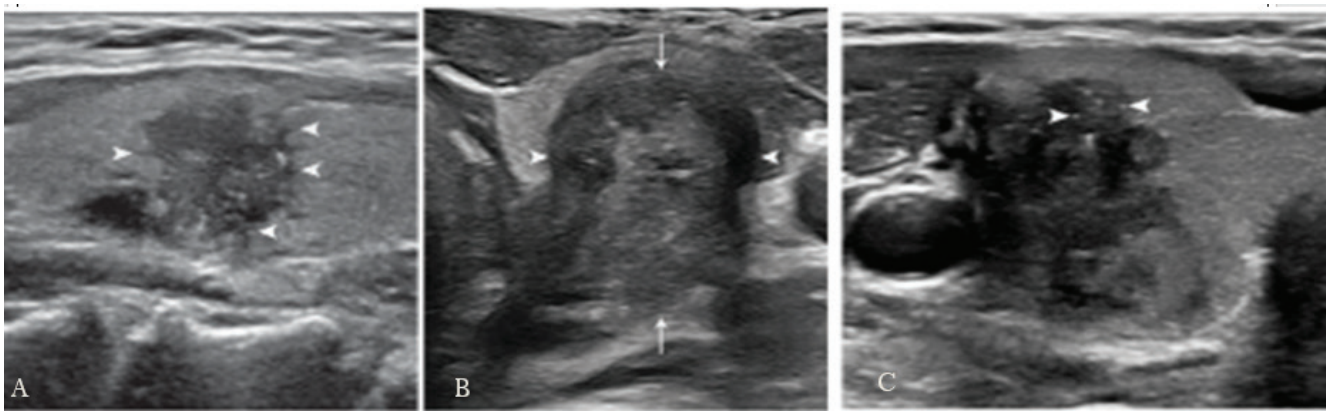


Figure 4. US features suggestive of malignant TNs. (A) Solid markedly hypoechoic TN with irregular and ill-defined margins with spiculations (arrowheads). (B) Nodule demonstrating a taller-than-wide configuration on the transverse view with marked hypoechogenicity (arrows), a feature highly suspicious for malignancy. (C) Heterogeneous hypoechoic nodule with microcalcifications and irregular margins (arrowheads), findings commonly associated with malignant thyroid lesions.

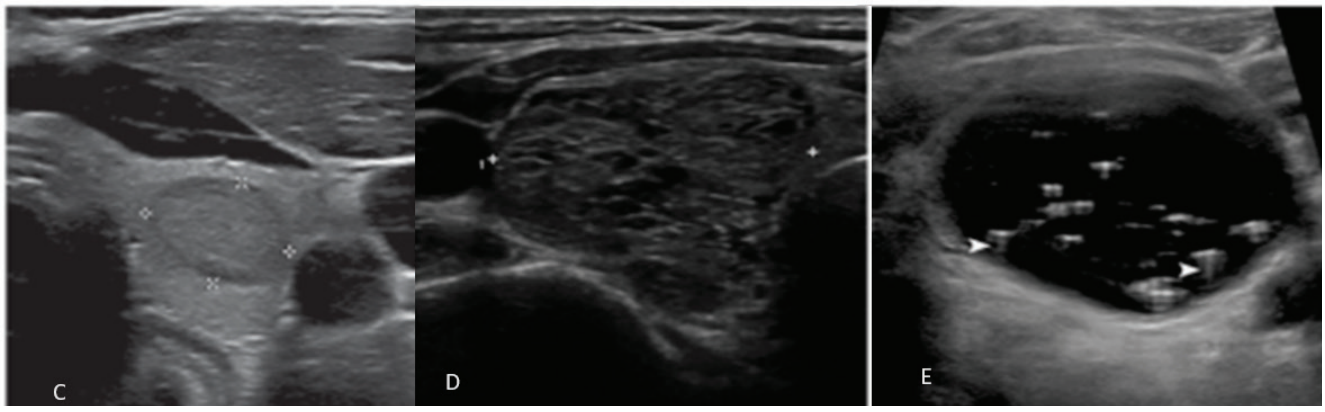


Figure 5. US features suggestive of benign TNs. (A) Oval isoechoic TN with smooth, well-defined margins and a wider-than-tall orientation (calipers), features commonly associated with benign lesions. (B) Spongiform TN composed of multiple tiny cystic spaces giving a honeycomb (calipers) appearance, a characteristic benign pattern. (C) Predominantly cystic nodule with comet-tail (ring-down) artifact (arrowhead) from colloid crystals, a typical benign finding.

irregular margins, taller-than-wide shape, and microcalcifications [13,20] Figures 4 and 5. Additional predictors such as internal vascularity, solitary nodules, and cervical lymphadenopathy also showed strong associations [4,35]. Conversely, nodule size and halo sign were not statistically significant, in line with studies from India and Brazil [36].

However, some subgroup findings (e.g., 100% malignancy in very hypoechoic nodules) were based on small absolute numbers and should be interpreted cautiously.

Methodological considerations and limitations

Several limitations must be acknowledged. First, FNAC was used as the reference standard, but it is imperfect, particularly in differentiating follicular adenomas from carcinomas. Histopathology remains the definitive diagnostic standard. Second, only patients who underwent both US and FNAC were included, introducing potential selection bias, as FNAC is typically performed selectively based on clinical judgment. This may have inflated diagnostic performance metrics through verification bias. Third, non-diagnostic cases (Bethesda I) were the only group excluded from diagnostic accuracy calculations, while Bethesda II (benign) and Bethesda III (atypia/follicular lesion of undetermined significance) were grouped as benign, and Bethesda IV–VI were grouped as malignant for statistical purposes. Because FNAC cannot reliably distinguish follicular adenoma from follicular carcinoma without

capsular or vascular invasion assessment, grouping Bethesda IV–VI as malignant may overestimate true histologically confirmed malignancy. This limitation may affect PPV estimates and introduce potential misclassification bias, and should be interpreted with caution, as it may slightly inflate diagnostic performance compared with studies that use histopathological confirmation as the reference standard. This approach facilitated construction of the 2 × 2 contingency table but may affect generalizability. A flow diagram has been added to clarify patient inclusion and exclusion.

US is inherently operator-dependent. Although images were reviewed by a principal investigator and confirmed by a senior radiologist, inter-observer variability was not formally assessed. This limits reproducibility and should be addressed in future studies. In addition, while high-resolution Siemens ACUSON NX3 Elite machines were used, equipment quality may differ across centers in sub-Saharan Africa, affecting generalizability.

Implications for practice

Despite these limitations, the high NPV supports ACR TI-RADS as a safe rule-out tool for malignancy in low-risk nodules. In resource-limited settings, its use has the potential to reduce unnecessary FNACs and streamline patient management. However, this implication was not directly tested in the study design and should be interpreted as a potential benefit rather than a demonstrated outcome.

Future directions

Future research should include multicenter, prospective studies with standardized imaging protocols, inter-observer variability assessment, and integration of clinical features such as cervical lymphadenopathy. Such studies would strengthen evidence for broader adoption of ACR TI-RADS in African and other resource-limited contexts.

Conclusion

This study demonstrates that the ACR TI-RADS classification system is a reliable, non-invasive tool for stratifying malignancy risk in TNs in Tanzania. Among 605 patients screened, 248 were analyzed for diagnostic accuracy, with the system showing excellent performance (sensitivity 90.6%, specificity 92.6%, AUC 0.951).

Key US features such as hypoechogenicity, irregular margins, microcalcifications, and a taller-than-wide shape were strongly associated with malignancy. The high NPV (98.5%) highlights its utility in safely ruling out cancer in low-risk nodules (TR1–TR3), thereby reducing unnecessary interventions. These findings provide valuable regional data and support standardized application of ACR TI-RADS in resource-limited settings. Future multicenter studies incorporating histopathological confirmation and inter-observer variability assessment are recommended to validate and extend these results.

Acknowledgments

Special thanks to Dr. Clement Kalambo, Patrick Amsi, Dr. Adnan Sadiq, and Dr. Felister Uisso for their supervision, the Radiology and Pathology departments at KCMC, and all study participants.

Conflict of interest

The author declares no competing interests.

Funding

The principal author provided financial support for this research.

Ethical considerations

Ethical approval was obtained from the KCMUCo Ethical Review Committee, with ethical clearance No.PG.106/2024. Also, permission to conduct this study was sought from the director of KCMC hospital and heads of radiology and histopathology departments. Patient confidentiality was maintained by anonymizing all records.

Availability of data and materials

All study data and related materials are available from the corresponding author upon request.

References

1. Nabahati M, Moazezi Z, Fartookzadeh S, Mehraeen R, Ghaemian N, Sharbatdaran M. The comparison of accuracy of ultrasonographic features versus ultrasound-guided fine-needle aspiration cytology in diagnosis of malignant thyroid nodules. J Ultrasound 2019; 22(3):315–21.
2. Floridi C, Cellina M, Buccimazza G, Arrichiello A, Sacrini A, Arrigoni F, et al. Ultrasound imaging classifications of thyroid nodules for malignancy risk stratification and clinical management: state of the art. Gland Surg 2019; 8:S233–44.

3. Woliński K, Szkudlarek M, Szczepanek-Parulska E, Ruchała M. Usefulness of different ultrasound features of malignancy in predicting the type of thyroid lesions: a meta-analysis of prospective studies. Pol Arch Med Wewn 2014; 124(3):97–104.

4. Sultan SR. B-mode ultrasound characteristics of thyroid nodules with high-benign probability and nodules with risk of malignancy. Cureus 2023; 15(5):e39281; doi:10.7759/cureus.39281

5. Mu C, Ming X, Tian Y, Liu Y, Yao M, Ni Y, et al. Mapping global epidemiology of thyroid nodules among general population: a systematic review and meta-analysis. Front Oncol 2022; 12:1029926; doi:10.3389/fonc.2022.1029926

6. Liao LJ, Hung SF, Cheng PC, Hsu WL. Ultrasonography of incidental thyroid nodules: a systematic review and meta-analysis of prevalence. J Med Ultrasound 2025; 33(4):365–70; doi:10.4103/jmu.jmu_62_24.

7. Wettasinghe MC, Rosairo S, Ratnatunga N, Wickramasinghe ND. Diagnostic accuracy of ultrasound characteristics in the identification of malignant thyroid nodules. BMC Res Notes 2019; 12:193; doi:10.1186/s13104-019-4235-y

8. Al-Ghanimi I, Al-Sharydah A, Al-Mulhim S, Faisal S, Al-Abdulwahab A, Al-Aftan M, et al. Diagnostic accuracy of ultrasonography in classifying thyroid nodules compared with fine-needle aspiration. Saudi J Med Med Sci 2020; 8(1):25–30; doi:10.4103/sjmms.sjmms_123_19

9. Cozzolino A, Filardi T, Simonelli I, Grani G, Virili C, Stramazzo I, et al. Diagnostic accuracy of ultrasonographic features in detecting thyroid cancer in the transition age: a meta-analysis. Eur Thyroid J 2022; 11(3):e220039; doi:10.1530/ETJ-22-0039

10. Isse HM, Lukande R, Sereke SG, Odubu FJ, Nassanga R, Bugeza S. Correlation of the ultrasound thyroid imaging reporting and data system with cytology findings among patients in Uganda. Thyroid Res 2023; 16(1):26; doi:10.1186/s13044-023-00169-1

11. Sakafu L, Mselle T, Mwaiselage J, Msengwa A, Maunda K, Eddin BS. Clinician's awareness and knowledge on the management of differentiated thyroid cancer and the use of radioactive iodine in Tanzania. Tanzan J Health Res 2016; 18(2):1–8.

12. Sencha AN, Patrunov YN, Pavlovich SV, Timofeyeva LA, Tukhbatullin MG, Smetnik AA. Thyroid ultrasound. Cham, Switzerland: Springer, p 1, 2019.

13. Durante C, Hegedüs L, Na DG, Papini E, Sipos JA, Baek JH, et al. International expert consensus on US lexicon for thyroid nodules. Radiology 2023; 309(1):e231481; doi:10.1148/radiol.231481

14. Naidu K, Saksenberg V, Goolam Mahyoodeen N. Clinical and ultrasound characteristics distinguishing benign and malignant thyroid nodules in Johannesburg, South Africa. J Endocrinol Metab Diabetes South Afr 2023; 28(2):62–8; doi:10.1080/16089677.2023.2201234

15. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, et al. ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS Committee. J Am Coll Radiol 2017; 14(5):587–95.

16. Giovanella L, Campennì A, Tuncel M, Petranović Ovčariček P. Integrated diagnostics of thyroid nodules. Cancers (Basel) 2024; 16(2):1–26.

17. Hoang J. Thyroid nodules and evaluation of thyroid cancer risk. Australas J Ultrasound Med 2010; 13(4):33–6.

18. Nicolaou MA, Jacobs K, Bhana S, Naidu K, Nicolaou V. A retrospective study correlating sonographic features of thyroid nodules with fine-needle aspiration cytology in a South African setting. S Afr J Radiol 2019; 23(1):1749; doi:10.4102/sajr.v23i1.1749

19. Li F, Pan D, Wu Y, Peng J, Li Q, Gui X, et al. Ultrasound characteristics of thyroid nodules facilitate interpretation of thyroid fine-needle aspiration cytology. Cancer Manag Res 2020; 12:11883–92; doi:10.2147/CMAR.S278781

20. Hussein MA, Elesawy YF, Ghoweba DEAAR, Mousa S. Correlation of ultrasound features in the TIRADS scoring system with cytological findings in the FNAC of thyroid nodules and their association with the metabolic status. Egypt J Intern Med 2024; 36(1):12; doi:10.1186/s43162-024-00290-z

21. Cibas ES, Ali SZ. The 2017 Bethesda system for reporting thyroid cytopathology. J Am Soc Cytopathol 2017; 6(6):217–22.

22. Delfim RLC, Veiga LCGD, Vidal APA, Lopes FPPL, Vaisman M, Teixeira PDFDS. Likelihood of malignancy in thyroid nodules according to a proposed thyroid imaging reporting and data system (TI-RADS) classification merging suspicious and benign ultrasound features. Arch Endocrinol Metab 2017; 61(3):211–21.

23. Fresilli D, David E, Pacini P, Del Gaudio G, Dolcetti V, Lucarelli GT, et al. Thyroid nodule characterization: how to assess the malignancy risk. Update of the literature. Diagnostics 2021; 11(10):2251.

24. Mandal D, Mallik R, Dey S, Lama KY. Sonographic evaluation of thyroid nodule based on thyroid imaging reporting and data system classification and its correlation with fine needle aspiration cytology findings. Int J Res Med Sci 2024; 12(6):1957–63.

25. Castilla Villanueva MA, Solis Cano DG, Amador Martínez A, Téliz Meneses MA, Baquera-Heredia J, Vallin Orozco CE, et al. Individual ultrasonographic characteristics of thyroid nodules and their cytopathological correlation to determine malignancy risk. Cureus 2024; 16(7):e12345; doi:10.7759/cureus.12345

26. Gharib H, Papini E, Garber JR, Duick DS, Harrell RM, Hegedüs L, et al. AACE/ACE/AME guidelines for clinical practice for the diagnosis and management of thyroid nodules – 2016 update. Endocr Pract 2016; 22(5):622–39; doi:10.4158/EP161208.GL

27. Tuladhar AS, Pudasaini S, Simkhada S, Shrestha A, Pradhan S. Correlation of American College of Radiology (ACR)-Thyroid Imaging Reporting and Data System (TIRADS) findings in Ultrasonogram (USG) of thyroid nodules with FNAC or Biopsy findings. Nepal Med Coll J 2022; 24(1):23–9.

28. Dawish MAA, Robert AA, Muna A, Eyad A, Ghamdi AA, Hajeri KA, et al. Bethesda system for reporting thyroid cytopathology: a three-year study at a tertiary care referral center in Saudi Arabia. World J Clin Oncol 2017; 8(2):151–7.

29. Zahid A, Shafiq W, Nasir KS, Loya A, Abbas Raza S, Sohail S, et al. Malignancy rates in thyroid nodules classified as Bethesda categories III and IV; a subcontinent perspective. J Clin Translational Endocrinol 2021; 23:100250.

30. Brazeal A, Wobenjo A, Macharia C, Otido M, Kamau R, MacLeod J. Diagnostic accuracy of Bethesda classification of thyroid nodules at a Kenyan Hospital: a retrospective study. Ann Afr Surg 2024; 21(4):120–6.

31. Bongiovanni M, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda system for reporting thyroid cytopathology: a meta-analysis. Acta Cytol 2012; 56(4):333–9.

32. Aribon PA, Teope E, Corneja-Egwolf AL, Maningat MPD. Diagnostic accuracy of American College of radiology thyroid imaging reporting data system: a single-center cross-sectional study. J ASEAN Fed Endocr Soc 2024; 39(1):61–8.

33. Shi M, Nong D, Xin M, Lin L. Accuracy of ultrasound diagnosis of benign and malignant thyroid nodules: a systematic review and meta-analysis. Int J Clin Pract 2022; 2022:5056082.

34. Nie W, Zhu L, Yan P, Sun J. Thyroid nodule ultrasound accuracy in predicting thyroid malignancy based on TIRADS system. Adv Clin Exp Med 2022; 31(6):597–606.

35. Mossus Y, Biouélé RCM, Atanga LC, Nyeki ARN, Eko DM, Pouka ONNM, et al. Predictive values of ultrasound characters associated with malignant thyroid nodules in Yaoundé: a cross-sectional study. Pan Afr Med J 2024; 47:38.

36. Sakajiri RK, Rahal A, Francisco MJ, Queiroz MRGD, Garcia RG, Martins LAL, et al. Ultrasound classification of thyroid nodules: does size matter?. Einstein (São Paulo) 2022; 20:6747; doi:10.31744/einstein_journal/2022AO6747