ORIGINAL ARTICLE

Diagnostic Performances of Ultrasound and Fine Needle Aspiration Cytology for Thyroid Cancer Using Biopsy Result as a Reference Standard, Central Ethiopia: A Paired Comparative Diagnostic Accuracy Study

Girmaye Tamrat Bogale[1], Endale Anberber[1], Agegnehu Berie Bayeh[2*]

OPEN ACCESS

Citation: Girmaye Tamrat Bogale, Endale Anberber, Agegnehu Berie Bayeh. Diagnostic Performances of Ultrasound and Fine Needle Aspiration Cytology for Thyroid Cancer Using Biopsy Result as a Reference Standard, Central Ethiopia: A Paired Comparative Diagnostic Accuracy Study. *Ethiop J Health Sci*. 2026;36(2):61. doi:http://dx.doi.org/10.4314/ejhs.v36i2.2

Received: December 30, 2024

Accepted: February 12, 2026

Published: March 1, 2026

Funding: Partial financial support from Addis Ababa University

Competing Interests: The authors declare that this manuscript was approved by all authors in its form and that no competing interest exists.

Affiliation and Correspondence:

[1]Department of Surgery, College of Health Sciences, Addis Ababa University, P.O Box 9086, Addis Ababa, Ethiopia

[2]Department of Surgery, College of Medicine and Health Sciences, Bahir Dar University, P.O Box 79, Bahir Dar, Ethiopia

*Email: agegnehuberie@gmail.com

ABSTRACT

BACKGROUND: *Thyroid nodules carry a substantial risk of malignancy. Ultrasound and fine needle aspiration cytology (FNAC) are the most important investigations used to risk-stratify such lesions. The performance of these tests is a scarcely researched area in Ethiopia. The main objective of this study was to assess the diagnostic performances of ultrasound and FNAC in identifying thyroid cancers, using final surgical pathology as the reference standard.*

METHODS AND MATERIALS: *A paired comparative diagnostic accuracy study using retrospective data was conducted at Tikur Anbessa Specialized Hospital and Lancet General Hospital. The collected data were analyzed using SPSS Windows version 21, Microsoft Excel 2019, and diagnostic accuracy–calculating spreadsheets.*

RESULTS: *Two hundred and sixty-six (266) cases were included in the analysis. Of these, 204 (76.7%) were females. The median age (in years) was 45.0 (IQR = 20.0). McNemar's test determined that the proportion of malignant lesions was statistically significantly different between the two tests ($P < 0.001$). Binary logistic regression models showed that the combination of ultrasound and FNAC had higher diagnostic performance metrics than the two tests used separately.*

CONCLUSION: *The study showed that FNAC outperforms ultrasound in diagnosing thyroid cancer. Thyroid ultrasound has higher sensitivity but significantly lower specificity than FNAC. This is an important and advantageous characteristic of a triaging test, but with a significant limitation as a rule-in test. Knowing the diagnostic performances of these tests greatly helps in triaging and clinical decision-making in the evaluation of individuals for thyroid cancer.*

KEYWORDS: *Thyroid nodules, Goiter, Thyroid cancer, Ultrasound, Cytology, Diagnostic accuracy*

INTRODUCTION

Thyroid nodules are common in the general population (1–3). The prevalence of these nodules varies depending on several factors, including the method of detection applied (4,5). Ultrasonographic assessment has reported a higher prevalence of thyroid nodules (6,7). The prevalence also shows wide variation across different regions and populations. Iodine deficiency, increasing age, female sex, and childhood radiation exposure to the head and neck region appear to increase the risk of developing thyroid nodules and cancer (8,9).

The most important concern with thyroid nodules is that 7–15% of them turn out to be malignant (10,11). Thyroid cancer has become a global public health problem, with rising incidence, mortality, and disability-adjusted life years (12–15). In the United States, its incidence in 2017 was found to be 2.69 times higher than in 1990. Mortality from thyroid cancer increased by 86.86% from 1990 to 2017. In women, it accounts for 5.1% of all cancers (16).

Ultrasound (US) and Fine Needle Aspiration Cytology (FNAC) are the two most important and widely used investigation modalities for assessing thyroid nodules (17,18). These tests accurately identify the presence or absence of malignancy in most nodules (19,20). Ultrasound is recommended as a first-line option in investigating suspected or known thyroid nodules (10,21). It is a non-invasive, rapid, and safe imaging modality used to stratify nodules based on their risk of malignancy. It is highly sensitive for detecting thyroid nodules and suspicious features, thereby guiding further management. Different methods of risk stratification are available in ultrasound reporting, including the American College of Radiology Thyroid Imaging Reporting and Data System (ACR TI-RADS) and the American Thyroid Association (ATA) guidelines (10,21–24). Fine Needle Aspiration Cytology is the single most important preoperative investigation for identifying malignant thyroid nodules (10). Its results are recommended to be reported according to The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC), which consists of six categories (25–27).

Diagnostic performance studies of medical investigations in general, and ultrasound and FNAC in particular, are not widely available in resource-limited settings such as our country. Although a few studies have examined patterns of thyroid FNAC (28,29) and agreement between FNAC and biopsy results (30) in Ethiopia, assessment of the diagnostic accuracy of ultrasound relative to FNAC—using final surgical biopsy as the reference standard—remains a scarcely researched area. The main objective of this study was to assess the diagnostic performances of ultrasound and FNAC in identifying thyroid cancers using final surgical pathology as the reference standard.

METHODS AND MATERIALS

Study design, period, and area: An institution-based retrospective paired diagnostic accuracy study was conducted from January to June 2024 at Tikur Anbessa Specialized Hospital (TASH) and Lancet General Hospital (LGH), located in Addis Ababa, the capital city of Ethiopia. The former is a government-owned hospital established in 1964 and is currently the largest teaching and clinical center in the country. LGH is a private hospital with endocrine surgery practices led by endocrine surgeons.

Source population, sample population, and sample size calculation:

The source population included all thyroid cases managed at TASH and LGH from January 2021 to January 2024. The sample population consisted of 266 post-thyroidectomy cases that met the inclusion criteria and were consecutively selected from the source population. The sample size (N = 266) was calculated to compare the diagnostic performances of two tests for paired groups, targeting a minimum probability of disagreement of 3%, $\alpha = 0.05$, and power of 80%. Yates' continuity correction was applied. This approach followed recommendations from a previous study, details of which are available elsewhere (31).

Sampling technique and inclusion and exclusion criteria: Clinical charts and records of all thyroid cases investigated and treated at TASH and LGH were retrieved and reviewed for eligibility. Two hundred sixty-six consecutive charts meeting the criteria were selected. Cases in which ultrasound and/or FNAC had not been performed or were not documented were excluded.

Data collection tools and procedures: Clinical charts and records of eligible thyroid cases were obtained, reviewed, and abstracted using a pre-developed data abstraction form. Variables included demographic characteristics (age and sex), clinical features (duration of goiter, recent rapid growth, voice change, consistency, family history of goiter and thyroid cancer, fixation), ultrasound characteristics, nodule size, FNAC Bethesda category, type of procedure performed, and final biopsy diagnosis.

Data processing and analysis: The abstracted data were entered, cleaned, and analyzed using SPSS version 21. Missing values were handled using the mode for categorical variables and the median for continuous variables. Diagnostic accuracy measures—such as sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), overall accuracy, and category-specific likelihood ratios—were calculated using Microsoft Excel 2019. ACR TI-RADS levels and FNAC Bethesda categories were compared using binomial logistic regression models (with histopathologic diagnosis as the dependent variable), area under the receiver operating characteristic curve (AUC) analysis (SPSS version 21, empirical non-parametric DeLong method), McNemar's test for paired proportions, and sensitivity and specificity differences calculated using published spreadsheets (32). Ultrasound findings were reported according to ACR TI-RADS (23), and FNAC results were reported using TBSRTC (25,26). Diagnostic performance metrics were estimated across multiple thresholds for both ordinal tests.

Data quality assurance and dissemination: Data collectors with at least an MD qualification received training on the study objectives and the use of the data abstraction form. The form was designed to be simple and clear. The entire research process was closely supervised by the principal investigator.

Ethical considerations: Ethical clearance was obtained from the Research & Ethics Committee of the Department of Surgery, School of Medicine, College of Health Sciences, Addis Ababa University (protocol number: DOS/REC/57/2024/2016). As this was a retrospective study, patient consent was not obtained; however, all identifying information was kept confidential. The study adhered to the Declaration of Helsinki.

This research was partially funded by Addis Ababa University. The funder had no role in the study design, data collection, analysis, reporting, or manuscript submission.

RESULTS

Sociodemographic data of participants: Two hundred and sixty-six (266) cases of thyroid pathologies with reports of thyroid ultrasound, fine needle aspiration cytology, and post-thyroidectomy histopathologic evaluation were included and analyzed in the study. Of these, 204 (76.7%) were females. The age (in years) of the participants ranged from 18 to 84, with a median of 45.0 (IQR = 20.0).

Clinical history and physical examination findings: Thirty-six (13.5%) of the participants reported recent rapid growth of their anterior neck swelling. A very small proportion of cases had a family history of goiter, 7 (2.6%), and thyroid cancer, 4 (1.5%). Twenty (9.4%) cases had voice change. The vast majority, 188 (70.7%), of the thyroid nodules were firm in consistency, and 20 (7.5%) were fixed. Firm-to-hard nodules constituted the least proportion, 17 (6.4%). The proportions of hard and soft nodules were 32 (12.0%) and 29 (10.9%), respectively.

Ultrasonographic characteristics of thyroid pathologies: One hundred and fifty-one (56.8%) of the cases had bilateral thyroid pathologies. Three or more nodules were found in the majority of cases. Twelve (4.5%) had diffuse enlargement of the thyroid gland. The proportions of patients with cervical lymph node and distant metastases were 46 (17.3%) and 8 (3.0%), respectively. The median anteroposterior and transverse dimensions of the 254 thyroid nodules that were subjected to fine needle aspiration cytology (excluding those with diffuse enlargement of the thyroid gland) were estimated to be 3.0 cm (IQR = 1.8 cm) and 2.4 cm (IQR = 1.6 cm), respectively. The different ultrasound features of the thyroid pathologies are summarized in Table 1.

Table 1: Ultrasonographic characteristics of thyroid lesions in TASH and TGH, Central Ethiopia, from Jan, 2021 to Jan, 2024. (N = 266).

Ultrasonographic characteristics	Frequency (%)
Laterality (n =266)	
Right	65 (24.4)
Left	50 (18.8)
Bilateral	151 (56.8)
Number of nodules (n = 266)	
One	67 (25.2)
Two	30 (11.3)
More than two	157 (59.0)
Diffuse enlargement	12 (4.5)
Composition (n = 254)	
Solid	138 (54.3)
Mixed	63 (27.7)
Spongiform	29 (11.4)
Cyst	24 (9.4)
Echogenicity (n = 254)	
Anechoic	22 (8.7)
Hyperechoic	71 (28.0)
Isoechoic	61 (24.0)
Hypoechoic	93 (36.6)
Very hypoechoic	0 (0)
Heterogeneous	7 (2.8)
Shape (n = 254)	
Taller than wide	34 (13.4)
Wider than tall	220 (86.6)
Margin (n = 254)	
Smooth	186 (73.2)
Ill-defined	47 (18.5)
Irregular	20 (7.9)
Extrathyroidal extension	1 (0.4)
Echogenic foci (n = 254)	
None	164 (64.6)
Popcorn, Comet tail	3 (1.2)
Macrocalcification	39 (15.4)
Rim calcification	6 (2.4)
Microcalcification	40 (15.7)
More than one type of calcification	2 (0.8)

Diagnostic performances of ACR TI-RADS, FNAC, and histopathology: One-fifth of the thyroid lesions were found to be benign on the basis of their ultrasound features. More than 70% of the cases were reported as ACR TI-RADS level 3 and above. Ultrasonographic evaluation based on the ACR TI-RADS threshold—considering levels 5, 4, and 3 as malignant and levels 2 and 1 as benign—showed that 191 (71.8%) had thyroid cancers. Fine needle aspiration cytology results showed that 141 (53.0%) were diagnosed as Bethesda categories IV and above. The most commonly diagnosed thyroid lesion was Bethesda category II, 125 (47.0%). The next most common category was Bethesda category IV, 58 (21.8%). The numbers of cases reported as Bethesda categories I and III were zero. Ninety-four (35.3%) of the thyroid lesions were diagnosed as malignant by histopathologic evaluation of the final surgical resection specimen. The higher the level and category of these tests, the greater the likelihood ratio. The ACR TI-RADS levels and FNAC Bethesda categories in relation to histopathologic diagnoses and category-specific likelihood ratios are summarized in Table 2.

Table 2: ACR TI-RADS Levels, FNAC Bethesda categories and histopathology reports of thyroid cases in TASH and LGH, Central Ethiopia, from Jan, 2021 to Jan, 2024.

Tests	Histopathologic Diagnosis, n (%)		Total, n (%)	LR	95% CI of LR
ACR TIRADS Level	Benign	Malignant			
1	22 (8.3)	0 (0.0)	22 (8.3)	0.00	0.00 – 0.68
2	47 (17.7)	6 (2.3)	53 (19.9)	0.16	0.06 – 0.42
3	56 (21.1)	14 (5.3)	70 (26.3)	0.46	0.27 – 0.78
4	37 (13.9)	43 (16.2)	80 (30.1)	2.23	1.56 – 3.18
5	10 (3.8)	31 (11.7)	41 (15.4)	5.67	2.91 – 11.05
Total	172 (64.7)	94 (35.3)	266 (100.0)	—	—
FNAC Bethesda category					
I	0 (0.0)	0 (0.0)	0 (0.0)	—	—
II	116 (43.6)	9 (3.4)	125 (47.0)	0.06	0.02 – 0.16
III	0 (0.0)	0 (0.0)	0 (0.0)	—	—
IV	41 (15.4)	17 (6.4)	58 (21.8)	1.01	0.64 – 1.59
V	9 (3.4)	22 (8.3)	31 (11.7)	4.47	2.15 – 9.32
VI	6 (2.3)	46 (17.3)	52 (19.5)	14.03	6.22 – 31.62
Total	172 (64.7)	94 (35.3)	266 (100.0)	—	—

ACR TI-RADS: American College of Radiology Thyroid Imaging Reporting and Data System; FNAC: Fine Needle Aspiration Cytology; CI: Confidence interval; LR: Likelihood ratio

Diagnostic performance metrics for the ACR TI-RADS levels and FNAC Bethesda categories were estimated using three different thresholds for each test. The overall accuracy of the thresholds showed wider variation as the thresholds changed. For both tests, lowering the thresholds increased sensitivity; however, this occurred at a significant cost of decreased specificity (Tables 3 and 4).

Table 3: Diagnostic performance of ACR TIRADS levels using different thresholds in identifying thyroid cancer in TASH and LGH, Central Ethiopia, from Jan, 2021 to Jan, 2024. (N=266).

Diagnostic performances	**Thresholds - Malignant VS Benign: Estimate (95% Confidence interval)**		
	ACR TIRADS 5 VS 4,3,2&1	**ACR TIRADS 5&4 VS 3,2&1**	**ACR TIRADS 5,4&3 VS 2&1**
Sensitivity	33.0 (23.5 - 42.5)	80.9(72.9 - 88.8)	95. 7(91.7 - 99.8)
Specificity	94.2 (90.7 - 97.7)	72.7(66.0 - 79.3)	40.1 (32.8 - 47.4)
PPV	75.6 (62.5 - 88.8)	61.8 (53.2 - 70.4)	46.6 (39.6 - 53.7)
NPV	72.0 (66.1 - 77.9)	87.4 (82.0 - 92.8)	94.5 (89.3 - 99.7)
LR (+)	5.7 (2.9 - 11.1)	3.0 (2.3 - 3.8)	1.6 (1.4 - 1.8)
LR (-)	0.7 (0.6 - 0.8)	0.3 (0.2 - 0.4)	0.1 (0.0 - 0.3)
Overall accuracy	72.6 (66.9 - 77.6)	75.6 (70.1 - 80.3)	59.8 (53.8 - 65.5)

ACR TI-RADS: American College of Radiology Thyroid Imaging Reporting and Data System; PPV: Positive predictive value; NPV: Negative predictive value; LR: Likelihood ratio

Table 4: Diagnostic performance of FNAC Bethesda categories using different thresholds in identifying thyroid cancer in TASH and LGH, Central Ethiopia, from Jan, 2021 to Jan, 2024. (N=266).

Diagnostic performances	**Thresholds - Malignant VS Benign: Estimate (95% Confidence interval)**		
	FNAC Bethesda 6 VS 5,4&2	**FNAC Bethesda 6&5 VS 4&2**	**FNAC Bethesda 6,5 &4 VS 2**
Sensitivity	48.9 (38.8 - 59.0)	72.3 (63.3 - 81.4)	95.7 (91.7 - 99.8)
Specificity	96.5 (93.8 - 99.3)	91.3 (87.1 - 95.5)	68.0 (61.1 - 75.0)
PPV	88.5 (79.8 - 97.1)	81.9 (73.7 - 90.2)	62.1 (54.2 - 70.0)
NPV	77.6 (72.0 - 83.2)	85.8 (80.7 - 90.9)	96.7 (93.5 - 99.9)
LR (+)	14.0 (6.2 - 31.6)	8.3 (5.0 - 13.7)	3.0 (2.4 - 3.7)
LR (-)	0.5 (0.4 - 0.6)	0.3 (0.2 - 0.4)	0.1 (0.0 - 0.2)
Overall accuracy	79.7 (74.5 – 84.1)	84.6 (79.8 – 88.4)	77.8 (72.5 – 82.4)

FNAC: Fine needle aspiration cytology; PPV: Positive predictive value; NPV: Negative predictive value; LR: Likelihood ratio

Binary logistic regression models showed that ultrasound and FNAC combined had higher diagnostic performance than the two tests used separately: ultrasound and FNAC [$\chi^2(7) = 180.6$; $P < 0.001$; overall accuracy = 88.0%], ultrasound [$\chi^2(4) = 89.0$; $P < 0.001$; overall accuracy = 75.6%], and FNAC [$\chi^2(3) = 155.2$; $P < 0.001$; overall accuracy = 84.6%]. Considering ACR TI-RADS levels 5, 4, and 3 and Bethesda categories 6, 5, and 4 as malignant yielded malignant concordance, benign concordance, and discordance rates of 30.1%, 22.2%, and 47.7%, respectively. McNemar's test for paired proportions determined that the proportion of malignant lesions was statistically significantly higher with FNAC than with ultrasound ($P < 0.001$). The two tests had similar sensitivity but significantly different specificity (Table 5).

Table 5: Thyroid cases, n (%) as diagnosed by ultrasound, FNAC and surgical biopsy in identifying thyroid cancer in TASH and LGH, Central Ethiopia, from Jan, 2021 to Jan, 2024. (N=266).

Ultrasound	Surgical Biopsy Result				Total
	Malignant		Benign		
	FNAC Malignant	FNAC Benign	FNAC Malignant	FNAC Benign	
Malignant	80 (30.1)	8 (3.0)	46 (17.3)	57 (21.4)	191 (71.8)
Benign	5 (1.9)	1 (0.4)	10 (3.8)	59 (22.2)	75 (28.2)
Total	85 (32.0)	9 (3.4)	56 (21.1)	116 (43.6)	266 (100.0)

FNAC: Fine Needle Aspiration Cytology

AUC was constructed for ACR TI-RADS levels and FNAC by treating the results as ordinal variables. It clearly showed that FNAC outperformed ACR TI-RADS in diagnosing thyroid cancer. ACR TI-RADS had an AUC of 0.82 (0.76–0.87) (P < 0.001). In contrast, FNAC had an AUC of 0.90 (0.86–0.94) (p < 0.001) (Figure 1).

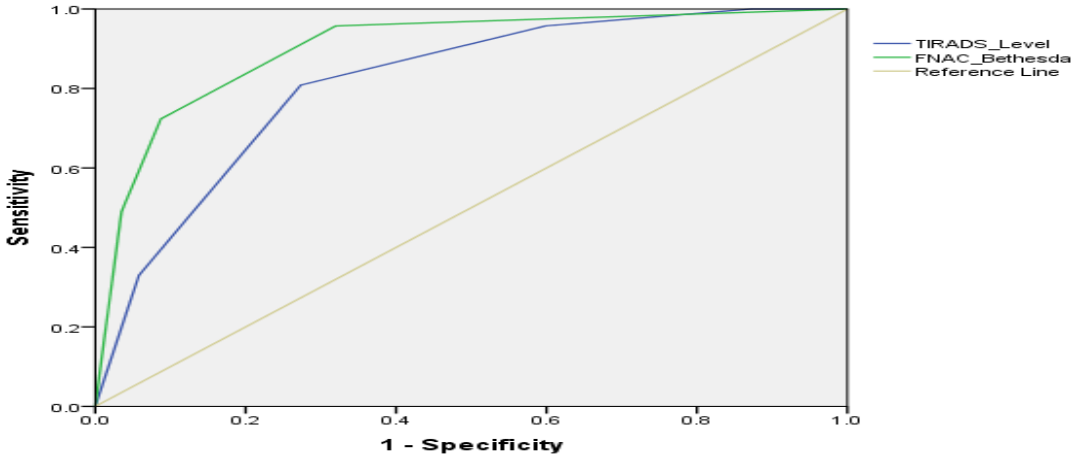

Figure 1*: Areas under receiver operating characteristics curves (AUC) for American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TIRADS) levels (AUC = 0.82; p < 0.001) and Fine Needle Aspiration Cytology (FNAC) Bethesda categories (AUC =0.90; p < 0.001).*

DISCUSSION

Thyroid nodules are common both in clinical practice and in the general population (1–3,33). These lesions are important because a significant proportion are eventually diagnosed and confirmed as cancers (10,11), and studies indicate that thyroid cancer has become a public health problem (1,2). The two most important investigations commonly applied to assess the malignancy risk of thyroid nodules are ultrasound and fine needle aspiration cytology (FNAC) (17,18). The American College of Radiologists has endorsed the ACR TI-RADS for stratifying thyroid nodules based on their risk of malignancy, with clear recommendations for further interventions (23). TBSRTC has been incorporated into the American Thyroid Association (ATA) guidelines (10,21).

A total of 266 cases of thyroid pathologies with reports of thyroid ultrasound, fine needle aspiration cytology, and post-thyroidectomy histopathologic evaluation were included and analyzed in this study. The vast majority, 204 (76.7%), were females. This finding is similar to reports from prior studies in Ethiopia (28–30,34,35) and India (36–38), with female preponderance ranging from 63.4% to 93.33%. The median age (in years) in the current study was 45.0 (IQR = 20.0). This is in line with the average age

and the most commonly affected age groups reported by previous studies (28,30,34,36–39).

Binary logistic regression models showed that the combination of ultrasound and FNAC had higher diagnostic performance than the two tests used separately. The diagnostic performance metrics of the two tests varied depending on the thresholds applied. The proportion of malignant lesions was statistically significantly higher with FNAC than with ultrasound ($P < 0.001$). The sensitivity of ACR TI-RADS in our study, considering levels 5, 4, and 3 as malignant, was 95.7% (91.7–99.8%). Previous studies have similarly reported high sensitivity of thyroid ultrasound in identifying malignant nodules (38–40). This high sensitivity is an important characteristic of an initial triaging test, helping to reduce the use of the more accurate but invasive FNAC, which is particularly relevant in low-resource settings where FNAC may not be readily available. However, this advantage comes at the cost of markedly low specificity, estimated at 40.1% (32.8–47.4%) in our study. This indicates that thyroid ultrasound is far from being a rule-in test. Debanu De et al. (37) also reported low specificity similar to our findings. However, comparison is difficult because ACR TI-RADS 1 and 2 nodules were not included in their study, and TI-RADS 3 nodules were considered benign.

In contrast, Periakaruppan G et al. (39) reported a significantly higher specificity for ultrasound (94.15%). However, unlike our study, which used surgical pathology as the gold standard, they used FNAC as the reference standard and also grouped ACR TI-RADS 3 nodules as benign lesions. Such classification is expected to improve the specificity of ultrasound in diagnosing thyroid cancer. As an initial triaging tool, ACR TI-RADS is intended to identify lesions that should undergo FNAC. In our study, TI-RADS 3 nodules were considered malignant (along with levels 5 and 4) because only post-thyroidectomy cases were included, and these nodules had both FNAC and histopathologic evaluations. This approach likely contributed to the low specificity observed. Although restricting the sample to post-thyroidectomy cases allowed the use of histopathology as the gold standard, it also introduced potential selection bias, as TI-RADS 3 nodules that were not aspirated or operated were excluded.

In the current study, thyroid ultrasound with ACR TI-RADS levels 5, 4, and 3 considered malignant showed a low PPV of 46.6% (39.6–53.7%) but a significantly higher NPV of 94.5% (89.3–99.7%). These findings are consistent with reports from studies in India (37,39). The study also demonstrated high diagnostic performance of thyroid FNAC, similar to previous reports (41–45), with an overall accuracy of 84.6% and an AUC of 0.90 (0.86–0.94), $p < 0.001$. However, specificity at the lower threshold was relatively low, at 68.0% (61.1–75.0%). The literature reports a wide range of specificity for thyroid FNAC, from 30.9% to 100% (42–45). The PPV and NPV of FNAC in this study also varied considerably depending on the threshold used. Previous studies have shown that factors such as FNAC technique, suboptimal specimens, pathologist experience, interpretive errors, and the nature of thyroid pathology contribute to lower specificity and higher false-positive rates (42). Due to the retrospective design of our study, these factors could not be adequately assessed. Although grouping atypia of undetermined significance as malignant can increase false positives, this was not a factor in our study, as no such cases were recorded.

Comparison of the diagnostic performances of thyroid ultrasound and FNAC showed that FNAC had higher overall accuracy than ultrasound. However, the addition of ultrasound substantially improved the diagnostic performance of FNAC. The concordance between the two tests was 69.9% (64.2–75.1%). Ultrasound demonstrated similar sensitivity but significantly lower specificity compared to FNAC. These findings are consistent with previous studies reporting high diagnostic accuracy of thyroid FNAC (41–45).

This study has important limitations. Due to its retrospective nature, detailed information on the techniques and procedures used for ultrasound, FNAC, and histopathologic evaluation was not available. Consequently, the potential effects of these factors on the findings could not be thoroughly explored. However, as the tests were conducted in academic and subspecialty settings,

major deviations from standard practices are unlikely. Another limitation arises from including only post-thyroidectomy cases, which enabled the use of histopathology as the gold standard but introduced selection bias, as not all thyroid cases undergo surgery.

In conclusion, the study demonstrated that thyroid ultrasound has high sensitivity and negative predictive value, making it an effective initial triaging tool. However, its similar sensitivity but lower specificity compared to FNAC indicates that a substantial proportion of ultrasound-diagnosed malignancies are ultimately benign on histopathology. This limitation should be clearly communicated during patient and physician counseling. Understanding the diagnostic performances of these tests is essential for effective triaging, clinical decision-making, evaluation of individuals for thyroid cancer, and improvement of local diagnostic practices.

ACKNOWLEDGMENT

We acknowledge Addis Ababa University for its partial financial support of this study.

REFERENCES

1. Mu C, Ming X, Tian Y, Liu Y, Yao M, Ni Y, et al. Mapping global epidemiology of thyroid nodules among general population: A systemic review and meta-analysis. *Front Oncol*. 2022;12:1029926. Doi: https://doi.org/10.3389/fonc.2022.1029926
2. Pellegriti G, Frasca F, Regalbuto C, Squatrito S, Vigneri R. Worldwide Increasing Incidence of Thyroid Cancer: Update on Epidemiology and Risk Factors. *Journal of Cancer Epidemiology*. 2013; vol.2013. Doi: https://doi.org/10.1155/2013/965212
3. Lucio Mango. Global Health Involvement in Thyroid Cancer Incidence Increase. *International Journal of Global Health*. 2020;1(1):14-17. Doi: https://doi.org/10.14302/issn.2693-1176.ijgh-20-3353
4. Sifontes-Dubón M, García-López JM, González-Ortega N, Pazos-Couselo M. Evaluation of a Clinical Pathway for Thyroid Nodular Disease: Timings and Delays in the Diagnosis and Treatment of Thyroid Cancer. *J Clin Med*. 2021;10(23):5681. Published 2021 Dec 1. Doi: https://doi.org/10.3390/jcm10235681
5. Sajisevi M, Caulley L, Eskander A, Du Y, Auh E, Karabachev A, et al. Evaluating the Rising Incidence of Thyroid Cancer and Thyroid Nodule Detection Modes A Multinational, Multi-institutional Analysis. *JAMA Otolaryngol Head Neck Surg*. 2022;148(9):811-818. Doi: https://doi.org/10.1001/jamaoto.2022.1743
6. Guth S, Aberle TJ, Galach A, Bamberger CM. Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. *Eur J Clin Invest* 2009; 39 (8): 699–706. Doi: https://doi.org/10.1111/j.1365-2362.2009.02162.x
7. Dean DS, Gharib H. Epidemiology of thyroid nodules. *Best Pract Res Clin Endocrinol Metab*. 2008;22(6):901-911. Doi: https://doi.org/10.1016/j.beem.2008.09.019
8. Knudsen N, Laurberg P, Perrild H, Bülow I, Ovesen L, Jørgensen T. Risk factors for goiter and thyroid nodules. *Thyroid*. 2002;12(10):879-888. Doi: https://doi.org/10.1089/105072502761016502
9. Ageeli RS, Mossery RA, Othathi RJ, Khawaji EA, Tarshi MM, Khormi GJ, et al. The Importance of the Thyroid Nodule Location in Determining the Risk of Malignancy: A Retrospective Study. *Cureus*. 2022 Sep; 14(9): e29421. Doi: https://doi.org/10.7759/cureus.29421
10. Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid*. 2016;26(1):1-133. Doi: https://doi.org/10.1089/thy.2015.0020
11. Gharib H, Papini E, Garber JR, et al. American Association of Clinical Endocrinologists, American College of Endocrinology, And Associazione Medici Endocrinologi Medical Guidelines for Clinical Practice for the Diagnosis and Management of Thyroid Nodules--2016 Update. *Endocr Pract*. 2016;22(5):622-639. Doi: https://doi.org/10.4158/EP161208.GL
12. Schuster-Bruce J, Jani C, Goodall R, Kim D, William Hughes W, Salciccioli JD, et al. A Comparison of the Burden of Thyroid Cancer Among the European Union 15+ Countries,

1990-2019 Estimates From the Global Burden of Disease Study. *JAMA Otolaryngol Head Neck Surg*. 2022;148(4):350-359. Doi: https://doi.org/10.1001/jamaoto.2021.4549
13. Wiltshire JJ, Drake TM, Uttley L, Balasubramanian SP. Systematic Review of Trends in the Incidence Rates of Thyroid Cancer. Thyroid. 2016;26(11) Doi: https://doi.org/10.1089/thy.2016.0100
14. Davies L, Welch HG. Current Thyroid Cancer Trends in the United States. *JAMA Otolaryngol Head Neck Surg*. 2014;140(4):317-322. Doi: https://doi.org/10.1001/jamaoto.2014.1
15. Deng Y, Li HT, Wang M, Li N, Tian T, Wu Y, et al. Global Burden of Thyroid Cancer From 1990 to 2017. *JAMA Network Open*. 2020;3(6): e208759. Doi: https://doi.org/10.1001jamanetworkopen.2020.8759
16. Zhai M, Zhang D, Long J, Gong Y, Ye F, Liu S, Li Y. The global burden of thyroid cancer and its attributable risk factor in 195 countries and territories: A systematic analysis for the Global Burden of Disease Study. *Cancer Med*. 2021 Jul;10(13):4542-4554. Doi: https://doi.org/10.1002/cam4.3970
17. Bomeli SR, LeBeau SO, Ferris RL. Evaluation of a thyroid nodule. *Otolaryngol Clin North Am*. 2010;43(2):229-vii. Doi: https://doi.org/10.1016/j.otc.2010.01.002
18. Bernet VJ, Ana-Maria Chindris A. Update on the Evaluation of Thyroid Nodules. *Journal of Nuclear Medicine* July 2021, 62 (Supplement 2) 13S-19S; DOI: https://doi.org/10.2967/jnumed.120.246025
19. Weller A, Sharif B, Qarib MH, St Leger D, De Silva HS, Lingam RK. British Thyroid Association 2014 classification ultrasound scoring of thyroid nodules in predicting malignancy: Diagnostic performance and inter-observer agreement. *Ultrasound*. 2020;28(1):4-13. Doi: https://doi.org/10.1177/1742271X19865001
20. Xie C, Cox P, Taylor N, LaPorte S. Ultrasonography of thyroid nodules: a pictorial review. *Insights Imaging*. 2016;7(1):77-86. Doi: https://doi.org/10.1007/s13244-015-0446-5
21. Haugen BR. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: What is new and what has changed? *Cancer*. 2017;123(3):372-381. Doi: https://doi.org/10.1002/cncr.30360
22. Russ G, Trimboli P, Buffet C. The New Era of TIRADSs to Stratify the Risk of Malignancy of Thyroid Nodules: Strengths, Weaknesses and Pitfalls. *Cancers (Basel)*. 2021;13(17):4316. Published 2021 Aug 26. Doi: https://doi.org/10.3390/cancers13174316
23. Tessler FN, Middleton WD, Grant EG, et al. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. *J Am Coll Radiol*. 2017;14(5):587-595. Doi: https://doi.org/10.1016/j.jacr.2017.01.046
24. Shimura H, Matsumoto Y, Murakami T, Fukunari N, Kitaoka M, Suzuki S. Diagnostic Strategies for Thyroid Nodules Based on Ultrasonographic Findings in Japan. *Cancers (Basel)*. 2021;13(18):4629. Published 2021 Sep 15. Doi: https://doi.org/10.3390/cancers13184629
25. Ali SZ, Baloch ZW, Cochand-Priollet B, Schmitt FC, Vielh P, VanderLaan PA. The 2023 Bethesda System for Reporting Thyroid Cytopathology. *Thyroid*. 2023;33(9):1039-1044. Doi: https://doi.org/10.1089/thy.2023.0141
26. Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. *Thyroid*. 2017;27(11):1341-1346. Doi: https://doi.org/10.1089/thy.2017.0500
27. Cochand-Priollet B, Vielh P, Royer B, et al. Cytopathologie thyroïdienne: le système de Bethesda 2010 [Thyroid cytopathology: Bethesda System 2010]. *Ann Pathol*. 2012;32(3):177-183. Doi: https://doi.org/10.1016/j.annpat.2012.02.015
28. Ebrahim H, Tilahun M, Fiseha T, Debash H, Bisetegn H, Alemayehu E, Fiseha M, Ebrahim E, Shibabaw A, Seid A, Getacher Feleke D, Mohammed O. Patterns of Fine Needle Aspiration Cytology Diagnosed Thyroid Nodules Among Clinically Suspected Patients in Northeast Ethiopia. *Pathology and Laboratory Medicine International*. 2023; 15:27-36 https://doi.org/10.2147/PLMI.S399682
29. Melak T, Abebe M, Tesfaye W, Melku M. Cytological patterns of thyroid lesions at Gondar University Hospital, Northwest Ethiopia: A retrospective study. *J Mol Pathophysiol*. 2015; 4(1):38-41. DOI: https://doi.org/10.5455/jmp.20150309080144

30. Ergete W, Abebe D. Discordance Rate between Thyroid Fine Needle Aspiration Cytology and Histopathologic Diagnosis. *Ethiop.J.Health Dev*. 2002;16(2):227-231 https://doi.org/10.4314/ejhd.v16i2.9815
31. Akoglu H. User's guide to sample size estimation in diagnostic accuracy studies. *Turk J Emerg Med*. 2022;22(4):177-185. Published 2022 Sep 30. https://doi.org/10.4103/2452-2473.357348
32. Newcombe RG. (2013). Confidence Intervals for Proportions and Related Measures of Effect Size. *Boca Raton: CRC Press/Taylor & Francis Group, LLC*.
33. Tadesse T, Hailu T, Getachew A, Abebe SM. Prevalence and Associated Factors of Thyroid Incidentaloma among Adult People Attending Gondar University Hospital, Northwest Ethiopia. *Ethiop J Health Sci*. 2020;30(1):23-30. https://doi.org/10.4314/ejhs.v30i1.4
34. Abdi D, Teshome ST, Melese B, et al. Prevalence of thyroid cancer among thyroid swelling in Jimma University Medical Center, South West Ethiopia: A five-year retrospective study. *International Journal of Medicine and Medical Sciences*. 2018. https://doi.org/10.5897/ijmms2018.1361
35. Melak T, Mathewos B, Enawgaw B, Damtie D. Prevalence and types of thyroid malignancies among thyroid enlarged patients in Gondar, Northwest Ethiopia: a three years institution based retrospective study. *BMC Cancer*. 2014; 14:899. Published 2014 Dec 2. https://doi.org/10.1186/1471-2407-14-899
36. De D, Dutta S, Tarafdar S, et al. Comparison between Sonographic Features and Fine Needle Aspiration Cytology with Histopathology in the Diagnosis of Solitary Thyroid Nodule. *Indian J Endocrinol Metab*. 2020;24(4):349-354. https://doi.org/10.4103/ijem.IJEM_349_20
37. Manoj PB, Innisai A, Hameed DS, Khader A, Gopanraj M, Ihare NH. Correlation of high-resolution ultrasonography findings of thyroid nodules with ultrasound-guided fine-needle aspiration cytology in detecting malignant nodules: A retrospective study in Malabar region of Kerala, South India. *J Family Med Prim Care*. 2019;8(5):1613-1616. https://doi.org/10.4103/jfmpc.jfmpc_163_19
38. Bhise SV, Shaikh A, Hippargekar PM, Kothule S. A Prospective Study of Ultrasonographic and FNAC Correlation of Thyroid Swellings with Histopathology. *Indian J Otolaryngol Head Neck Surg*. 2022;74(Suppl 2):1942-1948. https://doi.org/10.1007/s12070-020-01922-w
39. Periakaruppan G, Seshadri KG, Vignesh Krishna GM, Mandava R, Sai VPM, Rajendiran S. Correlation between Ultrasound-based TIRADS and Bethesda System for Reporting Thyroid-cytopathology: 2-year Experience at a Tertiary Care Center in India. *Indian J Endocrinol Metab*. 2018;22(5):651-655. https://doi.org/10.4103/ijem.IJEM_27_18
40. Nassar OH. Thyroid nodule diagnosis: A comparative analysis of ultrasonography and fine-needle aspiration cytology. *Indian J Med Sci*. 2024; 76:83-8. https://doi.org/10.25259/IJMS_212_2023
41. Sinna EA, Ezzat N. Diagnostic accuracy of fine needle aspiration cytology in thyroid lesions. *J Egypt Natl Canc Inst*. 2012;24(2):63-70. https://doi.org/10.1016/j.jnci.2012.01.001
42. Zhu Y, Song Y, Xu G, Fan Z, Ren W. Causes of misdiagnoses by thyroid fine-needle aspiration cytology (FNAC): our experience and a systematic review. *Diagn Pathol*. 2020;15(1):1. Published 2020 Jan 3. https://doi.org/10.1186/s13000-019-0924-z
43. Esmaili HA, Taghipour H. Fine-Needle Aspiration in the Diagnosis of Thyroid Diseases: An Appraisal in Our Institution. *ISRN Pathology*. 2012;2012(2012):1-4. https://doi.org/10.5402/2012/912728
44. Jamaiyar A, Yogesh K. How accurate is fine-needle aspiration cytology (FNAC) for thyroid lesion: A correlation of FNAC with histopathology. *J Family Med Prim Care*. 2023;12(1):15-20. https://doi.org/10.4103/jfmpc.jfmpc_1413_21
45. Gupta M, Acharya K, Jha A, et al. Diagnostic efficacy of ultrasonography-guided fine needle aspiration cytology on thyroid swellings. *Egypt J Otolaryngol*. 2022; 38:118. https://doi.org/10.1186/s43163-022-00298-8