The Research of Medical Science Review
ISSN: 3007-1208 & 3007-1216 Volume 3, Issue 6, 2025

DIAGNOSTIC ACCURACY OF ULTRASOUND BASED THYROID IMAGING REPORTING & DATA SYSTEM (TIRADS) IN DETECTING BENIGN & MALIGNANT THYROID NODULE TAKING BETHESDA SCORING AS GOLD STANDARD

Dr. Basit Arif[*1], Dr. Syeda Tehmina Ather[2], Dr. Khubeb Yousif[3], Dr. Imran Ali[4], Dr. Rajesh Kumar[5], Dr. Nitalya Raja[6]

*[*1,3,4,6]Dr. Ruth K. M. Pfau, Civil Hospital Karachi Postgraduate Resident MBBS*
[2,5]Dr. Ruth K. M. Pfau, Civil Hospital Karachi Assistant Professor MBBS, FCPS

[*1]basitarif_suriya@outlook.com, [2]tehminajunaid12@gmail.com, [3]khubebyousif49@gmail.com, [4]abroimran555@gmail.com, [5]vasandaniraj81@gmail.com, [6]dr.nia18@gmail.com

DOI: https://doi.org/10.5281/zenodo.15672997

Keywords
Diagnostic Imaging, Fine-Needle Biopsy, Fine-Needle Aspiration, TIRADS, Thyroid Nodule, Ultrasonography

Article History
Received on 06 May 2025
Accepted on 06 June 2025
Published on 16 June 2025

Corresponding Author: *
Dr. Basit Arif

Abstract

OBJECTIVE: *To evaluate the diagnostic accuracy of the TIRADS system in differentiating between benign and malignant thyroid nodules, using FNAC as the reference (gold) standard.*

METHODOLOGY: *This analytical, cross-sectional observational study was carried out in the ENT and Radiology Department of Dr. Ruth K.M. Pfau Civil Hospital, Karachi. A non-probability consecutive sampling technique was used to gather data, resulting in a total study sample of 164 individuals. Thyroid nodules were categorized based on TIRADS guidelines. Patients underwent USG-guided FNAC, following the Bethesda System for Reporting Thyroid Cytopathology. Only cases falling under Bethesda class II (benign), class V (suspicious for malignancy), and class VI (malignant) were included for analysis. Data collection and statistical evaluation were performed using SPSS version 26*

RESULTS: *The investigation encompassed a cohort of 164 individuals with a mean age of 42.31 ± 14.16 years. A significant proportion of the participants (58.5%) fell within the demographic of individuals exceeding the age of 40 years, while 70.1% of the subjects were identified as male. The parameters of sensitivity and specificity were recorded at 84.85% and 78.63%, respectively, accompanied by a positive predictive value (PPV) of 50.00%, a negative predictive value (NPV) of 95.37%, and an aggregate diagnostic accuracy of 79.88%.*

CONCLUSION: *The study demonstrates that the ultrasound-based TIRADS system exhibits good diagnostic accuracy in differentiating benign from malignant thyroid nodules when compared with the Bethesda scoring system. With high sensitivity, specificity, and particularly a strong negative predictive value, TIRADS proves to be an effective non-invasive tool for initial risk stratification, aiding in clinical decision-making and reducing unnecessary FNAC procedures in low-risk patients.*

INTRODUCTION

Thyroid nodule is one of the frequent clinical presentations of thyroid after examination and different diagnostic methods to find it out [1]. Studies have shown that the overall prevalence of thyroid nodules in the adult population by palpation is approximately 4–8%, by imaging techniques like ultrasonography 67%, while at autopsy the figure is around 50% [2]. Ultrasonography [3] is the most accurate method for the evaluation of thyroid nodules. One study reported the prevalence of thyroid nodules to be 61% [4]. A thyroid nodule is an abnormal growth of thyroid cells that form a lump within the thyroid gland that can be detected as a distinct area by imaging above background thyroid tissue [4]. Although thyroid nodules are common and usually benign, the main objective in clinical management is to distinguish malignant from benign lesions [5].

Although different stratification systems have been elaborated to help optimally manage those nodules, such systems have not been integrated in the radiological practice [3–5]. The American College of Radiology initiated stratified risk of TIRADS which had its first version in 2015 which were cemented by a call for standardized risk assessment method of a thyroid nodule based on ultrasound category [5]. In addition, ACR TIRADS has structured to show malignant nodules and to reduce the need for biopsy in benign nodules [6]. Thyroid nodule is a lump or swelling of solid or liquid nature, located inside the thyroid [6].

This distinction can be made by ultrasonographic evaluation which shows these nodules to separate from normal thyroid parenchyma [7]. Thyroid nodules management, especially with suspicion of malignancy, remain an issue of clinical controversy. It is estimated that 2–6% of thyroid nodules are palpable, 19–68% are diagnosed by ultrasound (USG) and 8–65% are found incidentally in autopsy [8]. When sensitivity and specificity of TIRADS in detecting benign and malignant thyroid nodules were combined in one study, sensitivity was 91.3% and specificity was 74.6% [9]. The EU-TIRADS classification system ranks the co-existence of four ultrasound (US) characteristics that play an equal role in determining the estimated rate of malignancy of thyroid nodules, including marked hypoechogenicity, irregular shape, microcalcifications, and irregular margins [9]. This classification relies on a different US pattern [10]. EU-TIRADS 1 (normal) — a thyroid gland without (or with clinically irrelevant) nodules Cystic or spongiform nodules without any suspicious ultrasound features are classified as EU-TIRADS 2 (benign). EU-TIRADS 3 (low risk) includes oval, hyper- or isoechoic nodules with smooth margins and the absence of suspicion.

Mildly hypoechoic nodules with oval shapes, smooth edges, and absence of suspicious features are categorized as EU-TIRADS 4 (intermediate risk). Finally, nodules exhibiting one or more suspicious ultrasound characteristics—such as marked hypoechogenicity, irregular shape, microcalcifications, or irregular margins—are classified as EU-TIRADS 5, indicating a high risk of malignancy [11–12].

The aim of this study is to initially scrutinize patients with thyroid nodules using ultrasound on the basis of TIRADS. TIRADS Score will help to calculate the risk of malignancy and will decide whether a patient needs FNAC or follow-up. This will help to reduce misleading FNAC performed for benign lesions. Furthermore, the result of this study will help the radiologist in decision-making regarding the detection of benign and malignant thyroid nodules in such suspected cases. In addition, by my study emphasis may be given on appropriate management plan to reduce morbidity, disease burden and hospital stay in this already compromised cohort of patients. Furthermore, this study will provide an efficient and pragmatic clinical tool for decision making for early suspicion in order to avoid invasive diagnostic delay.

METHODOLOGY

This descriptive, cross-sectional study was carried out over a span of six months in the Department of ENT and Radiology at Dr. Ruth K.M. Pfau Civil Hospital, Karachi. A total of 164 participants were enrolled using a non-probability consecutive sampling technique. Patients aged between 18 and 60 years of either gender were enrolled if they met the criteria of suspected thyroid nodule: having a clinically palpable nodule measuring ≥3 cm in the anterior or lateral aspect of the neck, persisting for at least six months, moving with deglutition.

Patients exhibiting a clinically or radiologically normal thyroid gland, who declined fine needle aspiration cytology (FNAC), presented with bleeding diathesis, or possessed a previously confirmed diagnosis of toxic or malignant thyroid nodules were systematically excluded from the investigation. All individuals deemed eligible underwent a comprehensive clinical evaluation followed by a neck ultrasonography utilizing a GE Logic S8 XD ultrasound apparatus, which was outfitted with a 5–15 MHz linear transducer. The thyroid gland was subjected to a thorough assessment for various parameters, encompassing size, composition, echogenicity, margins, and the number of nodules identified. Based on ultrasound findings, nodules were categorized following the Thyroid Imaging Reporting and Data System (TIRADS). Nodules classified under TIRADS categories 2, 3, and 4a were considered benign, while those falling under 4b and 5 were labeled as malignant. All patients subsequently underwent ultrasound-guided fine-needle aspiration cytology (FNAC), with sample analysis conducted by a qualified consultant cytopathologist in accordance with the Bethesda System for Reporting Thyroid Cytopathology. Only nodules classified as Bethesda class II (benign), class V (suspicious for malignancy), and class VI (malignant) were included in the analysis. The data were analyzed using SPSS software version 26.0. Continuous variables were presented as mean ± standard deviation (SD), and categorical data were summarized using frequencies and percentages. Diagnostic metrics, including sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall diagnostic accuracy of the TIRADS classification system in identifying malignant thyroid nodules, were calculated using FNAC results as the gold standard

RESULTS

The study included 164 participants, with an obvious majority of male subjects 70.1% (n=115) and female subjects belonging to 29.9% (n=49). The average age of the participants was 42.31 ± 14.16 years, the participants 18–40 years of age were accounted for as being 41.5% (n=68) and above 40 years old as 58.5% (n=96). The mean duration of the nodules was noted as 7.16 ± 3.57 months. Nodules had a duration of 2-7 months and beyond 7 months in 65.2% (n=107) and 34.8% (n=57) of the patients respectively (TABLE I).

Table II illustrates the comparison between TIRADS and FNAC findings in the evaluation of 164 thyroid nodules. Among the nodules classified as benign by TIRADS, 28 (84.8%) were confirmed as benign by FNAC, while 28 (21.4%) were actually malignant. Conversely, of the nodules labeled as malignant by TIRADS, 103 (78.6%) were verified as malignant on FNAC, and only 5 (15.2%) were found to be benign. Based on these results, the sensitivity of TIRADS in detecting malignant nodules was 78.6%, while the specificity was 84.8%. The positive predictive value (PPV) was high at 95.4%, indicating a strong likelihood that nodules identified as malignant by TIRADS were truly malignant. However, the negative predictive value (NPV) was relatively low at 50.0%, suggesting that a significant portion of nodules classified as benign might still be malignant. Overall, the TIRADS system shows good diagnostic performance, particularly in correctly identifying malignant thyroid nodules.

Table III elucidates the diagnostic efficacy of the TIRADS system in the identification of thyroid nodules across a cohort of 164 cases. The sensitivity was documented at 84.85% (95% CI: 72.62–97.08), signifying that TIRADS demonstrated a noteworthy capacity for the detection of malignant nodules. The specificity was reported at 78.63% (95% CI: 71.61–85.65), indicative of its commendable proficiency in accurately classifying benign nodules. The positive predictive value (PPV) was noted at 50.00% (95% CI: 36.90–63.10), implying that merely half of the nodules categorized as malignant by TIRADS were substantiated as such, Conversely, the negative predictive value (NPV) was extraordinarily high, at 95.37% (95% CI: 91.41–99.33), with the majority of nodules diagnosed as benign being truly non-malignant. The diagnostic accuracy was 79.88% (95% CI: 73.74–86.01). The positive likelihood ratio (LR) was 3.97 and the negative LR was 0.19 reflecting a moderate to strong effectiveness of TIRADS to differentiate between malignant and benign nodules.

Table I: Demographic Data of the Study Population (n=164)	
Variable	**Frequency (%)**
Gender	
Male	115 (70.1)
Female	49 (29.9)
Age (Mean ± SD) = 42.31 ± 14.16 years	
18 - 40 years	68 (41.5)
> 40 years	96 (58.5)
Duration of Nodules (Mean ± SD) = 7.16 ± 3.57 months	
2 - 7 months	107 (65.2)
> 7 months	57 (34.8)

Table II: Comparison of TIRADS & FNAC in Detecting Thyroid Nodules (n=164)		
TIRADS	**FNAC**	
	Benign	Malignant
Benign	28 (84.8)	28 (21.4)
Malignant	5 (15.2)	103 (78.6)

Table III: Diagnostic Accuracy of TIRADS in Detecting Thyroid Nodules (n=164)		
Diagnostic Variables	**TIRADS**	**95% Confidence Interval**
Sensitivity	84.85%	72.62 – 97.08
Specificity	78.63%	71.61 – 85.65
Positive Predictive Value	50.00%	36.90 – 63.10
Negative Predictive Value	95.37%	91.41 – 99.33
Diagnostic Accuracy	79.88%	73.74 – 86.01
Positive Likelihood Ratio	3.97	N/A
Negative Likelihood Ratio	0.19	N/A

DISCUSSION

Over the decades, the diagnosis of benign or malignant thyroid nodules has remained elusive. Thus, systematic imaging-based classification systems have received growing interest, particularly the Thyroid Imaging Reporting and Data System (TIRADS). TIRADS classifies thyroid nodules according to their specific ultrasound features and risk of malignancy. This classification framework assists clinicians in making appropriate choices for the additional diagnostic measures required, especially fine-needle aspiration cytology (FNAC), which is considered the gold standard in cytological evaluation.

When compared with FNAC, the sensitivity of TIRADS was obtained as 84.85% and specificity was found as 78.63% in the present study. These results are also in good agreement with those reported by Saqib et al. Sensitivity and specificity were 84.0% and 78.2%, respectively [13]. In addition, from the data obtained from our investigation, it has a positive predictive value (PPV) of 50.00%, a negative predictive value (NPV) of 95.37% and an overall diagnostic accuracy of 79.88%. The significant NPV, particularly crucial in preventing unnecessary biopsies, by identifying those patients as low-risk with TIRADS is reflected in this empirical data giving support to its utility in malignancy exclusion.

The relatively low PPV observed in our study (50.00%) represents a common weakness of TIRADS-based approaches, that is, the large number of benign nodules and the similarities in ultrasonographic characteristics of benign and malignant lesions. However, beyond this limitation, the strong NPV allows nodules considered as low risk on TIRADS to be very unlikely to have malignant features, thereby giving the clinician a valuable resource to keep an eye on in conservatively treated patients.

A comparison with the previous literature also supports our evidence. Kovatcheva et al. [9] performed a prospective cohort study and found that the sensitivity of the TIRADS classification system was marginally higher (91.3%) and the specificity was equivalent (74.6%). These differences may be influenced by size of the sample, the skill of the operator, the observer variability, and the demographic of the population. The differences found emphasize the importance of local factors and standardized training in TIRADS application in order to ensure equal diagnostic certainty in all clinical situations.

Furthermore, Yang et al. [14] and Shayganfar et al. [15] also evaluated the prognostic value of different frameworks of TIRADS (including ACR-TIRADS and K-TIRADS) for the prediction of malignancy, especially for nodules classified as the Bethesda category IV, a category known for its cytological undetermined. These findings support the versatility of TIRADS systems in increasing the diagnostic accuracy, particularly when combined with cytological examination. In the same vein, the work of Fernández et al. [16] validates EU-TIRADS for use in prepubertal age groups with excellent correlation to cyto-history, confirming the applicability of the system in various age ranges. Additionally, Khan et al. [17] compared TIRADS to Bethesda classification and found its usefulness in distinguishing benign from malignant lesions—thus, reassering the general opinion on the diagnostic value of TIRADS.

Finally, Malhi and Grant [18] also highlighted the structured nature of TIRADS and the potential to standardize US reporting communication between radiologists and clinicians. That structured reporting assists in decision-making tracks and leads to the optimized patient care. The present study present results which are valuable for the elucidation of the diagnostic accuracy of TIRADS compared with FNAC; however, it is not without limitation. One of the main limitations is the use of a non-probability sample model, thus generating selection bias and limiting the generalizability of the findings to the population at large. Furthermore, a cohort size of 164 patients is adequate for preliminary analysis, but may not encompass the entire spectrum of thyroid pathology or the variations seen in different demographics or regions.

A further limitation is the lack of indeterminate cytological categories, in particular Bethesda classes III and IV. These categories are often on the borderline of a diagnosis and of considerable clinical interest. By excluding such nodules, the study fails to reveal the actual measures of performance in the day-to-day use of TIRADS to support or reject the decision to sample nodules with indeterminate cytology. 24 -26 Furthermore, ICC of interobserver variabilities was not considered for ultrasound interpretation and for cytopathological evaluation, which may affect the reproducibility and diagnostic concordance.

Nevertheless, the study also has several strong points. It used a clear operational definition of suspected thyroid nodules and the use of unified reporting systems (the TIRADS for ultrasound and the Bethesda System for cytopathology) that made the results more reliable and easier to compare. The use of clinical and imaging criteria under the inclusion criteria, make the results more representative of the clinical reality. In addition, all FNACs were performed under ultrasonographic control by experienced radiologists, and interpreted by a senior consultant cytopathologist, potentially improving diagnostic accuracy.

For future studies, it would be advisable that larger, multicentric trials are performed with random sampling for improving external validity. Assessment of the diagnostic value of the TIRADS, if Bethesda categories III and IV cytopathologic results were included in the analysis, would have been more complete. Moreover, the inter- and intra-observer dependability of the radiologic and cytologic assessments should be analyzed to also define the standardization process and educational support required for consistent TIRADS application in clinical practice.

CONCLUSION

The study demonstrates that the ultrasound-based TIRADS system exhibits good diagnostic accuracy in differentiating benign from malignant thyroid nodules when compared with the Bethesda scoring system. With high sensitivity, specificity, and particularly a strong negative predictive value, TIRADS proves to be an effective non-invasive tool for initial risk stratification, aiding in clinical decision-making and reducing unnecessary FNAC procedures in low-risk patients.

REFERENCES

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):651-5.

Horvath E, Silva CF, Majlis S, Rodriguez I, Skoknic V, Castro A, et al. Prospective validation of the ultrasound based TIRADS (Thyroid Imaging Reporting and Data System) classification: results in surgically resected thyroid nodules. Eur Radiol. 2017;27:2619-28.

Hoang JK, Middleton WD, Farjat AE, Langer JE, Reading CC, Teefey SA, et al. Reduction in thyroid nodule biopsies and improved accuracy with American College of Radiology Thyroid Imaging Reporting and Data System. Radiology. 2018;287(1):185-93.

Kamran M, Hassan N, Ali M, Ahmad F, Shahzad S, Zehra N. Frequency of thyroid incidentalomas in Karachi population. Pak J Med Sci. 2014;30(4):793.

Weiss VL, Andreotti RF, Ely KA. Use of the thyroid imaging, reporting, and data system (TI-RADS) scoring system for the evaluation of subcentimeter thyroid nodules. Cancer Cytopathol. 2018;126(8):518-24.

Mohanty J, Mishra P. Role of ACR-TIRADS in risk stratification of thyroid nodules. Int J Res Med Sci. 2019;7(4):1039-43.

Alshaikh S, Harb Z, Aljusairi E, Almahari SA. Classification of thyroid fine-needle aspiration cytology into Bethesda categories: an institutional experience and review of the literature. Cyto J 2018;13:4.

Shin JH, Baek JH, Chung J, Ha EJ, Kim JH, Lee YH, et al. Ultrasonography diagnosis and imaging-based management of thyroid nodules: revised Korean Society of Thyroid Radiology consensus statement and recommendations. Korean J Radiol. 2016;17(3):370-95.

Kovatcheva RD, Shinkov AD, Dimitrova ID, Ivanova RB, Vidinov KN, Ivanova RS. Evaluation of the diagnostic performance of EU-TIRADS in discriminating benign from malignant thyroid nodules: a prospective study in one referral center. Eur Thyroid J. 2021;9(6):304-12.

Mahajan A, Vaidya T, Vaish R, Sable N. The journey of ultrasound-based thyroid nodule risk stratification scoring systems: Do all roads lead to Thyroid Imaging, Reporting and Data System (TIRADS)? J Head Neck Physicians Surg. 2017;5(2):57-65.

Russ G, Bonnema SJ, Erdogan MF, Durante C, Ngu R, Leenhardt L. European Thyroid Association guidelines for ultrasound malignancy risk stratification of thyroid nodules in adults: the EU-TIRADS. Eur Thyroid J. 2017;6(5):225–37.

Ha EJ, Baek JH, Na DG. Risk stratification of thyroid nodules on ultrasonography: current status and perspectives. Thyroid. 2017;27(12):1463–8.

Yang W, Fananapazir G, LaRoy J, Wilson M, Campbell MJ. Can the American Thyroid Association, K-TIRADS, and ACR-TIRADS ultrasound classification systems be used to predict malignancy in Bethesda category IV nodules?. Endocr Pract. 2020;26(9):945-52.

Shayganfar A, Hashemi P, Esfahani MM, Ghanei AM, Moghadam NA, Ebrahimian S. Prediction of thyroid nodule malignancy using thyroid imaging reporting and data system (TIRADS) and nodule size. Clin Imaging. 2020;60(2):222-7.

Fernández DY, Amenabar EV, Muñoz AC, Vallejo LA, León MC, Planes-Conangla M, Felip CI, et al. Ultrasound criteria (EU-TIRADS) to identify thyroid nodule malignancy risk in adolescents. Correlation with cyto-histological findings. Endocrinol Diabetes Nutr (Eng Ed.). 2021;68(10): 728-34.

Khan WS, Farooq SM, Nawaz S. Accuracy of ultrasound TIRADS classification for the differentiation between benign and malignant thyroid lesions taking Bethesda classification as gold standard. Imaging. 2024;16(2):94-100.

Malhi HS, Grant EG. Ultrasound of thyroid nodules and the thyroid imaging reporting and data system. Neuroimaging Clin. 2021;31(3):285-300.

Guo J, Du L, Bi W, Liu Y, Zhang C. Retrospective study from a single center to comparison of diagnostic value of three thyroid imaging reporting and data systems in Bethesda III/IV thyroid nodules. Front Oncol. 2025;15:1549646.