DOI: https://doi.org/10.54393/pjhs.v6i3.2685

PAKISTAN JOURNAL OF HEALTH SCIENCES
(LAHORE)
https://thejas.com.pk/index.php/pjhs
ISSN (E): 2790-9352, (P): 2790-9344
Volume 6, Issue 03 (March 2025)

Original Article

OPEN ACCESS

Determination of Diagnostic Accuracy of ACR-TI-RADS in Detecting Malignancy in Thyroid Nodules On Ultrasonography, Keeping Bethesda Cytological Score at FNAC as Gold Standard

Touseef Bilal Khan[1], Tarique Ali Shaikh[2], Sajjad Akram[3], Ramzan Ali[4], Werdah Zia[5], Tarique Khan[1] and Sania Maqbool[6*]

[1]Department of ENT, Mayo Hospital, Lahore, Pakistan
[2]Department of ENT, Sharif Medical and Dental College, Sharif Medical City Hospital, Lahore, Pakistan
[3]Department of ENT, Jinnah Hospital, Lahore, Pakistan
[4]Department of Pediatric ENT, University of Child Health Sciences, Lahore, Pakistan
[5]Department of ENT, Arif Memorial Teaching Hospital, Rashid Latif Medical College, Lahore, Pakistan
[6]Department of Physical Therapy, Lahore College for Women University, Lahore, Pakistan

ARTICLE INFO

Keywords:
Thyroid Nodules, ACR TI-RADS, Bethesda System, Fine Needle Aspiration Cytology

How to Cite:
Khan, T. B., Shaikh, T. A., Akram, S., Ali, R., Zia, W., Khan, T., & Maqbool, S. (2025). Determination of Diagnostic Accuracy of ACR-TI-RADS in Detecting Malignancy in Thyroid Nodules On Ultrasonography, Keeping Bethesda Cytological Score at FNAC as Gold Standard: ACR-TI-RADS in Detecting Malignancy in Thyroid Nodules, Bethesda Cytological Score. Pakistan Journal of Health Sciences, 6(3), 297-302. https://doi.org/10.54393/pjhs.v6i3.2685

***Corresponding Author:**
Sania Maqbool
Department of Physical Therapy, Lahore College for Women University, Lahore, Pakistan
saniamaqbool28@gmail.com

Received date: 23rd December, 2024
Acceptance date: 20th March, 2025
Published date: 31st March, 2025

ABSTRACT

Thyroid nodules are a common clinical concern, requiring precise evaluation to differentiate benign from malignant cases. **Objective:** To compare the diagnostic accuracy of the ACR TI-RADS with the Bethesda System, the gold standard for thyroid cytopathology. **Methods:** This cross-sectional study was carried out within the Diagnostic Radiology Division of Mayo Hospital Lahore between August and November 2024, and included 224 patients who had thyroid nodules on ultrasonography and were planned for fine needle aspiration cytology (FNAC). A radiologist assessed the ACR-TIRADS scores after data were accumulated employing a standardized case report form.The FNACs were categorized utilizing the Bethesda framework. SPSS was used to calculate the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy of the ACR TI-RADS system. **Results:** Their study had a total of 224 patients, with a prevalence of female (186, 83.0%). The sensitivity of ACR TI-RADS in identifying malignant nodules was 87.5% (SE=0.049, 95% CI (77.9%, 97.1%)), while specificity was lower at 31.6% (SE=0.039, 95% CI (23.9%, 39.3%)). The PPV and NPV were 33.6% and 86.5%, respectively. Overall, the accuracy of ACR TI-RADS was 47.4%. **Conclusions:** It was concluded that the ACR TI-RADS system showed excellent sensitivity but low specificity in identifying malignant thyroid nodules. This implies that ACR TI-RADS is more useful for excluding cancer than identifying malignant instances. To improve the ACR TI-RADS system's predictability across a range of populations, more research is required.

INTRODUCTION

Thyroid nodules, also known as incidentalomas, are frequently encountered in routine clinical practice, with their prevalence varying based on the screening method and population studied. Research indicates a prevalence of 2–7% through palpation, 19–68% via ultrasound, and 8–65% in autopsy findings [1-3].While over 90% of detected nodules are benign and clinically insignificant, they hold clinical significance due to the potential risk of malignancy, with thyroid cancer observed in approximately 4.0% to 10% of cases [2, 4]. Thus, accurately distinguishing between benign and potentially malignant nodules is crucial for clinicians, as early identification of malignancy can significantly improve patient prognosis [5, 6]. A robust system for stratifying nodules based on malignancy risk is therefore essential. B-mode ultrasound has proven effective in detecting thyroid nodules [1].The American

DOI: https://doi.org/10.54393/pjhs.v6i3.2685

College of Radiology (ACR) created the ACR Thyroid Imaging Reporting and Data System (TI-RADS) to aid in precise risk assessment. Based on certain ultrasonography characteristics, counting as composition, echogenicity, shape, borders, and echogenic foci, ACR TI-RADS classifies nodules into five grades (I through V), with higher numbers indicating a greater chance of malignancy [7, 8]. This efficient approach makes a difference radiologist quickly identify and report knobs which will warrant advance examination. For thyroid nodules, Fine Needle Aspiration Cytology (FNAC) is regularly utilized as a preparatory symptomatic strategy. The Bethesda System for Reporting Thyroid Cytopathology (TBS) remains the most widely recognized standard for evaluating FNAC data [9, 10]. This framework categorizes nodules into six levels (I through VI), from benign (category II, with no danger chance) to highly suspicious (category VI, with up to a 99% threat chance) [9, 10]. TBS, moreover, helps in clinical decision-making, from suggesting follow-up for benign cases to proposing thyroidectomy for high-risk cases. Given its unwavering quality, the Bethesda framework has been broadly embraced by pathologists [9-11]. The ACR-TIRADS scoring system is extensively utilized in local practice in Pakistan to categorize thyroid nodules and guide decisions on the need for further cytological assessment. Research gives limited data on the effectiveness of the Bethesda System and TI-RADS in detecting malignant thyroid nodules within the Pakistani population.

Despite widespread use of ACR TI-RADS and Bethesda systems for thyroid nodule evaluation, limited local evidence exists in Pakistan directly comparing ultrasound-based ACR TI-RADS with cytology-based Bethesda classification for diagnostic accuracy. Previous studies show inconsistent sensitivity, specificity, and accuracy across different populations, highlighting variability in performance and lack of population-specific validation. Moreover, most available literature. This study aims to examine and assess the sensitivity of the TI-RADS risk stratification tool by directly contrasting its results with Bethesda, the "Gold Standard,". By assessing its alignment with biopsy outcomes, we can accurately gauge the sensitivity of the imaging system, thereby strengthening clinicians' confidence in its use. With limited local studies on this topic, our findings will contribute valuable insights to both national and international literature.e.

METHODS

This cross-sectional study was conducted in the Diagnostic Radiology Department of Mayo Hospital Lahore from August to November 2024 after approval from the Institutional Review Board of King Edward Medical University, CPSP/REU/RAD-20Z 7-066-3343. 224 cases were chosen using a non-probability sampling technique for this cross-sectional survey-based investigation; the sample size was determined to be 224 with a 95% confidence level and a 5% margin of error, using an online sample size calculator (calculator.net). This estimation was based on a 16.58% prevalence of goiter in the Pakistani population, with an expected sensitivity of 80 ± 13% and specificity of 92.7 ± 5% [12]. Data were collected using a structured case report form. Informed consent was taken from all patients. Patients of any gender who presented with thyroid nodules found on B-mode ultrasonography and were scheduled for fine-needle aspiration cytology (FNAC) at the Mayo Hospital Radiology Department were the focus of the inclusion criteria. Participants with biopsy-confirmed thyroid cancer (Bethesda 6) or those with a normal thyroid ultrasound scan were not included in this study. A radiologist with at least a fellowship certification performed an ultrasound scan on each patient for ACR-TIRADS assessment. Scoring was done by assessing the parameters outlined by the TIRADS scoring system, with nodules assessed for their composition, echogenicity, shape, margins and echogenic foci. Each parameter received a minimum score of zero and a maximum score of three. These individual parameter scores were added to achieve the final TIRADS grade. All patients who underwent ultrasound examination for thyroid nodules had FNAC. Ultrasound-guided FNACs were performed by a fellowship-qualified radiologist. The samples were air dried, then sent to a pathologist for staining and further evaluation. The pathologist determined the cytological features of each sample under a microscope and assigned the relevant Bethesda category. ACR-TIRADS subcategories were removed to streamline the analysis, and scores were arranged as follows: TR3, TR4, and TR5 were categorized as "Malignant," whilst TR1 and TR2 were categorized as "Benign". Similarly, Bethesda categories I and II were classified as "Benign", while categories III to VI were categorized as "Malignant". The results from TIRADS and Bethesda were summed up into these two broad categories, and then compared in a 2 x 2 table. True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) were assessed from this table. Further tests were applied to calculate the Positive Predictive Value (PPV), Negative Predictive Value (NPV) and the Diagnostic Accuracy. For data analysis, SPSS version 29 was used to evaluate the accuracy, sensitivity, and specificity of the ACR-TIRADS score relative to the Bethesda cytology scoring. Frequencies and percentages were reported for categorical variables, such as gender. Data were stratified by age and gender, and associations with ACR-TIRADS and Bethesda scores were assessed.

RESULTS

A total of 224 patients were enrolled in the study, with 186 female and 38 male. The age of the participants ranged

DOI: https://doi.org/10.54393/pjhs.v6i3.2685

from 17 to 77 years, with a mean age of 40.12 ± 11.135 years in females and 47.32 ± 17.96 years in males.The age distribution across genders is detailed in table 1.

Table 1: Frequency of Gender in Each Age Group

Age Groups	Female	Male	n (%)
16-25 Years	15	6	21 (9.3%)
26-35 Years	55	4	59 (26.3%)
36-45 Years	66	8	79 (33.6%)
46-55 Years	26	9	35 (15.6%)
56-65 Years	19	0	19 (8.4%)
66-75 Years	1	8	9 (4.0%)
76-85 Years	0	2	2 (0.8%)
Total	186	38	224

The results for frequency of benign versus malignant lesions as per the Bethesda classification are shown in Figure 1

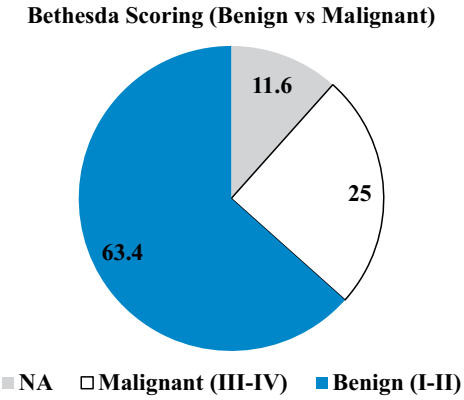

Figure 1: Frequency of Benign versus Malignant Lesions as per the Bethesda Classification

The results for the frequency of benign versus malignant lesions as per ACR-TIRADS classification are shown in Figure 2.

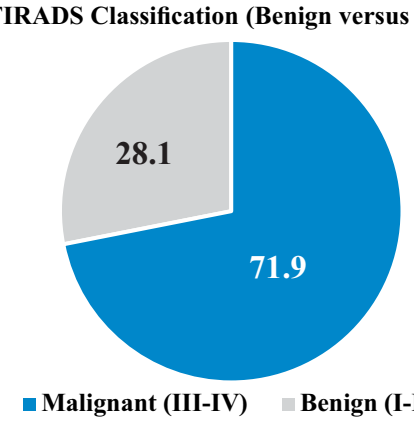

Figure 2: Frequency of Benign versus Malignant Lesions as per ACR-TIRADS Classification

Results present a 2×2 contingency analysis comparing the ACR TI-RADS classification with the Bethesda System for reporting thyroid cytopathology, categorizing nodules into benign and malignant groups, as shown in Table 2.

Table 2: 2 x 2 Contingency Analysis for ACR-TIRADS Classification versus Bethesda

Bethesda Class	ACR-TIRADS Class		Total
	III-V	I-II	
III-VI	49 (TP)	7 (FN)	56
I-II	97 (FP)	45 (TN)	142
Total	146	52	198

TP = True positive, FN = False negative, FP = False positive, TN = True Negative.

The diagnostic performance of the ACR TI-RADS classification in identifying malignant cases compared to the Bethesda classification was evaluated. The sensitivity of the ACR TI-RADS was found to be 87.5% (SE=0.049, 95% CI (77.9%, 97.1%)), indicating a high probability of correctly identifying malignant lesions. However, the specificity was lower at 31.6% (SE=0.039, 95% CI (23.9%, 39.3%)), reflecting a limited ability to accurately identify benign lesions. The positive predictive value (PPV) was calculated at 33.6% (SE=0.038, 95% CI (26.1%, 41.1%)), suggesting that a notable proportion of positive results may not indicate true malignancy. In contrast, the negative predictive value (NPV) was observed to be 86.5% (SE=0.047, 95% CI (77.3%, 95.7%)), indicating that a substantial majority of negative results accurately reflected the absence of malignancy. Overall, the accuracy of the ACR TI-RADS classification in this sample was 47.4% (SE=0.035, 95% CI (40.5%, 54.3%)), having minimal physical performance.

DISCUSSION

The precise preoperative diagnosis of thyroid nodules requires the use of both ultrasonography and fine needle aspiration cytology (FNAC). Since most of the nodules are painless, it is crucial to choose nodules carefully for FNAC to prevent needless invasive operations. Thyroid ultrasonography, the main imaging modality, is essential for the first evaluation of these nodules, and the ACR TI-RADS grading system offers an evidence-based method for classifying cancer risk and offering therapeutic recommendations [13]. Ultimately, ACR TI-RADS aims to improve patient care by reducing unwarranted FNACs and minimizing over-surveillance of low-risk nodules [14, 15]. A study evaluated US TI-RADS while comparing results with Bethesda scoring, reporting values of 70.6% and 90.4%, respectively [16]. Another study done in 2022 observed a similar sensitivity of 72.3% but a lower specificity of 66.7% [17]. In our study, ACR TI-RADS showed a higher sensitivity of 87.5%, indicating a robust ability to detect true positives. However, the specificity in our study was lower at 31.6%, suggesting a higher rate of false positives compared to these studies. Variability in specificity across studies may reflect differences in patient populations, ultrasound interpretation criteria, or diagnostic thresholds, underscoring the importance of context-specific

DOI: https://doi.org/10.54393/pjhs.v6i3.2685

evaluations when implementing ACR TI-RADS in clinical practice. In terms of diagnostic accuracy, a study reported a 60% accuracy for US TI-RADS in predicting malignancy in thyroid nodules [16]. Another study found a higher accuracy of 85.7% [17]. By comparison, our study showed a diagnostic accuracy of 47.4% for ACR TI-RADS, indicating lower precision in malignancy prediction. This discrepancy may result from variations in patient populations, nodule characteristics, or ultrasound interpretation methods. These findings suggest that TI-RADS accuracy may vary based on clinical setting and demographic factors, and further studies are required to refine TI-RADS criteria and improve predictive reliability across diverse populations. While studies directly comparing TI-RADS scores with biopsy results are limited, most research has relied on comparing scores with FNAC rather than definitive histopathology from resected specimens. Our study uniquely compared the ACR TI-RADS classification system with the Bethesda scoring system, providing a comprehensive evaluation of both ultrasound-based risk stratification and cytological assessment in thyroid nodule management. For instance, a study done in 2019 from Pakistan compared ultrasonography with FNAC in managing thyroid nodules and found high concordance between the two. On FNAC, 1.6% of the 124 patients whose nodules were determined to be benign by ultrasonography were discovered to be malignant, while 98.38% of them were confirmed to be benign. On the other hand, FNAC revealed benign results in 44.4% of nine individuals whose lesions were classified as malignant by ultrasonography and confirmed malignancy in 55.6% of them. They concluded that ultrasonography is a highly effective non-invasive diagnostic technique that has significant diagnostic accuracy in differentiating between benign and malignant nodules [18]. Our study consists of a comparison of ACR TI-RADS and Bethesda systems, further aiming to clarify each system's strengths, enhancing clinical decision-making accuracy. In a comparable study conducted in Pakistan, 201 patients' TI-RADS scores and FNAC results were compared. They reported a 76.1% overall diagnostic accuracy for predicting malignancy, with a TI-RADS sensitivity of 77.8%, specificity of 75.5%, positive predictive value (PPV) of 53.8%, and negative predictive value (NPV) of 90.2% [19]. In comparison, our study observed a lower PPV of 33.6% but a similarly high NPV of 86.5%. This difference in PPV suggests that, while ACR TI-RADS effectively rules out malignancy in benign nodules, its ability to predict malignant cases accurately may be limited. These findings highlight the need for further research to refine the ACR TI-RADS criteria and enhance its predictive capabilities in diverse clinical settings. A 2023 hospital-based cross-sectional study included 132 patients with thyroid nodules to evaluate the diagnostic accuracy of ACR TI-RADS. The system showed a strong correlation with cytology findings and high sensitivity and specificity. It proved to be a reliable, noninvasive tool for thyroid nodule assessment, reducing unnecessary fine-needle aspirations and supporting its use as a standardized screening method, particularly in resource-limited settings [3]. In another study, ACR TI-RADS had the highest sensitivity, at 94.5%, when the relative accuracy of five distinct TI-RADS systems was assessed on 939 thyroid nodules. The study concluded that ACR TI-RADS performs well in recognizing threat and is the foremost brief strategy for thyroid nodules of all sizes. When compared to histopathology results, it was seen that ACR TI-RADS appeared higher in efficiency than the Bethesda system, suggesting its value as a preliminary diagnostic approach [20]. Our study found a similarly high sensitivity for ACR TI-RADS at 87.5%, aligning with the findings of a study done in Goztepe City Hospital of Turkey [20], and reinforcing its reliability in thyroid nodule assessment. These results support the integration of ACR TI-RADS as an effective component of diagnostic workflows, particularly when emphasizing sensitivity in the early detection of malignancy.

This study has several limitations, including its single-center design and use of non-probability sampling, which may limit generalizability of findings. The study also relied on FNAC (Bethesda system) rather than final histopathology, which is the true gold standard for malignancy confirmation. Additionally, inter-observer variability in ultrasound interpretation may have influenced TI-RADS scoring. Future research should include multi-center studies with larger, more diverse populations and should compare ACR TI-RADS directly with histopathological outcomes. Standardization of ultrasound interpretation and integration of advanced imaging or AI-based tools may further improve diagnostic accuracy and reduce false-positive rates.

CONCLUSIONS

It was concluded that compared to the Bethesda classification, the ACR TI-RADS scoring system demonstrated high sensitivity but low specificity in detecting malignant thyroid nodules.This indicates that while ACR TI-RADS is useful for ruling out cancer, its positive predictive value suggests it may not accurately diagnose malignant cases. Further research is needed to enhance the predictive accuracy of the ACR TI-RADS system across diverse populations.

DOI: https://doi.org/10.54393/pjhs.v6i3.2685

Authors' Contribution

Conceptualization: TBK
Methodology: TBK, RA, WZ, TK
Formal analysis: SA
Writing and Drafting: TAS, SM
Review and Editing: TAS, SM

All authors approved the final manuscript and take responsibility for the integrity of the work

Conflicts of Interest

All the authors declare no conflict of interest.

Source of Funding

The author received no financial support for the research, authorship and/or publication of this article.

REFERENCES

[1] Dean DS and Gharib H. Epidemiology of Thyroid Nodules. Best Practice and Research Clinical Endocrinology and Metabolism.2008Dec;22(6):901-11.doi:10.1016/j.beem.2008.09.019.

[2] Alexander EK, Doherty GM, Barletta JA. Management of Thyroid Nodules.The Lancet Diabetes and Endocrinology.2022Jul;10(7):540-8.doi:10.1016/S2213-8587(22)00139-5.

[3] 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 Research.2023Sep;1 6(1):26.doi:10.1186/s13044-023-00169-1.

[4] Joseph-Auguste J, Lin L, Demar M, Duffas O, Molinie V, Sulpicy C *et al.* Epidemiologic, Clinical, Ultrasonographic, and Cytological Features of Thyroid Nodules In Predicting Malignancy Risk:A Retrospective Study of 442 French Afro-Caribbean Patients.International Journal of Endocrinology. 2020;2020(1): 4039290.doi: 10.1155/2020/4039290.

[5] Lin B, Ma H, Ma M, Zhang Z, Sun Z, Hsieh IY *et al.* The Incidence and Survival Analysis for Anaplastic Thyroid Cancer: A SEER Database Analysis. American Journal of Translational Research.2019 Sep;11(9): 5888.

[6] Bailey S and Wallwork B. Differentiating Between Benign and Malignant Thyroid Nodules:'An Evidence-Based Approach in General Practice'.Australian Journal of General Practice.2018Nov;47(11):770-4. doi:10.31128/AJGP-03-18-4518.

[7] 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.Journal of the American College of Radiology.2017May;14(5):587-95.doi:10.1016/j.jacr.2017.01.046.

[8] Azab EA, Abdelrahman AS, Ibrahim ME. A Practical Trial to Use Thyroid Imaging Reporting and Data System (TI-RADS) in Differentiation Between Benign and Malignant Thyroid Nodules.Egyptian Journal of Radiology and Nuclear Medicine.2019Dec;50:1-8. doi: 10.1186/s43055-019-0020-0.

[9] Alshaikh S, Harb Z, Aljufairi E, Almahari SA. Classification of Thyroid Fine-Needle Aspiration Cytology into Bethesda Categories: An Institutional Experience and Review of the Literature. Cytojournal.2018 Feb;15:4.doi:10.4103/cytojournal.cytojournal_32_17.

[10] Liu X, Medici M, Kwong N, Angell TE, Marqusee E, Kim MI *et al.* Bethesda Categorization of Thyroid Nodule Cytology and Prediction of Thyroid Cancer Type and Prognosis.Thyroid.2016Feb;26(2):256-61.doi:10.1089/thy.2015.0376.

[11] Reuters KB, Mamone MC, Ikejiri ES, Camacho CP, Nakabashi CC, Janovsky CC *et al.* Bethesda Classification and Cytohistological Correlation of Thyroid Nodules in A Brazilian Thyroid Disease Center.European Thyroid Journal.2018Jun;7(3):133-8.doi:10.1159/000488104.

[12] Ullah I, Asif M, Alam N, Ali S, Haq IU, Khan I. Study on the Prevalence of Goiter and Associated Factors Among Hospitalized Patients of District Timergara, Dir Lower Pakistan. Biomedical Journal of Scientific and Technical Research.2022;41(2):32458-63.doi: 10.26717/BJSTR.2022.41.006563.

[13] Barbosa TL, Junior CO, Graf H, Cavalvanti T, Trippia MA, da Silveira Ugino RT *et al.* ACR TI-RADS and ATA US Scores Are Helpful for the Management of Thyroid Nodules with Indeterminate Cytology.BioMed Central Endocrine Disorders.2019Dec;19:1-1.doi:10.1186/s12902-019-0429-5.

[14] Giles WH, Maclellan RA, Gawande AA, Ruan DT, Alexander EK, Moore FD *et al.* False Negative Cytology in Large Thyroid Nodules. Annals of Surgical Oncology.2015Jan;22:152-7.doi:10.1245/s10434-014-3952-7.

[15] AlSaedi AH, Almalki DS, ElKady RM, ElKady R. Approach to Thyroid Nodules:Diagnosis and Treatment.Cureus.2024Jan;16(1).doi:10.7759/cureus.52232.

[16] Singaporewalla RM, Hwee J, Lang TU, Desai V. Clinico-Pathological Correlation of Thyroid Nodule Ultrasound and Cytology Using the TIRADS and Bethesda Classifications. World Journal of Surgery. 2017 Jul; 41: 1807-11. doi: 10.1007/s00268-017-3919-5.

[17] George NA, Suresh S, Jiji V, Renu S, Thomas S, Janardhan D *et al.* Correlation of TIRADS and Bethesda Scoring Systems with Final Histopathology of Thyroid Nodules–An Institutional Experience.

DOI: https://doi.org/10.54393/pjhs.v6i3.2685

Indian Journal of Otolaryngology and Head and Neck Surgery.2021Jan:1-6.doi:10.1007/s12070-021-02380-8.

[18] Alshoabi SA and Binnuhaid AA. Diagnostic Accuracy of Ultrasonography Versus Fine-Needle Aspiration Cytology for Predicting Benign Thyroid Lesions. Pakistan Journal of Medical Sciences.2019May; 35(3):630.doi:10.12669/pjms.35.3.292.

[19] Nighat S, Zahra M, Javed AM, Ahmad S, Anwar S, Kamal UH. Diagnostic Accuracy of TI-RADS Classification in Differentiating Benign and Malignant Thyroid Nodules Study from Southern Punjab, Pakistan.Biomedica.2021Sep;37(3).doi:10.51441/BioMedica/5-504.

[20] Orhan Soylemez UP, Gunduz N. Diagnostic Accuracy of Five Different Classification Systems for Thyroid Nodules: A Prospective, Comparative Study.Journal of Ultrasound in Medicine.2022May;41(5):1125-36. doi:10.1002/jum.15802.