Evaluating the Diagnostic Efficacy of TIRADS in Predicting Fine Needle Aspiration Outcomes in Thyroid Nodules

Sulaiman Adel AlRajeh, MBBS* Abdulaziz Ahmed AlTaweel, MBBS, MBA** Saeed Saad Al Ghanom, MBBS*** Meshari Abdullah AlFaris, MBBS**** Tauqir Ahmed Rana, MBBS, FCPS, FRCR, EBIR, CBCCT, MSc (Clinical Education)*****

ABSTRACT

The Thyroid Imaging Reporting and Data System (TIRADS) is widely utilized for standardizing the assessment of thyroid nodules via ultrasound imaging. Reliable evaluation of thyroid nodules is essential for determining the need for fine needle aspiration (FNA) and subsequent management.] This study aimed to assess the diagnostic accuracy of TIRADS in predicting FNA cytology results and to analyze its effectiveness in guiding clinical decisions in thyroid nodule evaluation. A cohort of 100 patients, each presenting with thyroid nodules and undergoing FNA, was analyzed. Two radiologists independently assigned TIRADS scores to each nodule. Inter-rater reliability was evaluated using the kappa coefficient, while the associations between TIRADS scores and FNA findings were determined via chi-squared tests. The findings demonstrated moderate inter-rater agreement between the two radiologists (kappa = 0.453, p < 0.001). Significant associations were identified between specific TIRADS scores and FNA results; notably, TIRAD 2 correlated strongly with benign findings (p=0.120) per Rater 1 and TIRAD 3 showed significant association with benign findings for Rater 2 (p=0.022). This study validates the utility of TIRADS in predicting FNA outcomes, emphasizing its role in informed clinical decision-making. Recommendations for bolstering diagnostic accuracy include standardizing radiologist training and ensuring regular updates on TIRADS guidelines to promote consistent application in clinical practice.

Keywords: Thyroid nodules; Thyroid Imaging Reporting and Data System (TIRADS); Fine Needle Aspiration (FNA); Diagnostic accuracy.

INTRODUCTION

Thyroid nodules are a prevalent clinical entity, with an estimated incidence of 19-68% in the general population, according to imaging studies[1,2]. While the majority of thyroid nodules are benign, a portion of nodules can be malignant, demanding precise diagnostic examination and risk stratification[3]. Fine needle aspiration (FNA) cytology has long been the foundation of thyroid nodule evaluation, giving crucial information about the lesion and directing subsequent therapeutic decisions[4-6].

Over time, many imaging modalities have been used to improve the diagnostic accuracy of thyroid nodule assessment. Ultrasound imaging, in particular, has emerged as a valuable method for identifying thyroid nodules based on their morphological characteristics. Standardized reporting methods, such as the Thyroid Imaging Reporting and Data System (TIRADS), have improved the approach to thyroid nodule evaluation by providing defined risk classification criteria[7-10].

The TIRADS framework divides thyroid nodules into risk groups based on ultrasound characteristics such as composition, echogenicity, form, margin, and calcifications[8,11]. These characteristics are evaluated to predict the chance of malignancy, allowing clinicians to prioritize management measures and identify the need for additional diagnostic interventions, such as FNA[8,9,12]. The diagnostic accuracy of TIRADS in predicting thyroid nodule malignancy has been extensively studied. Several studies have found variable sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for various TIRADS categories. For example, TIRADS categories 4 and 5 are often associated with greater cancer risk, whereas categories 2 and 3 are considered low-risk[13,14].

Inter-rater reliability among radiologists in TIRADS scoring has also been studied to determine the uniformity of judgments between practitioners. Studies have found moderate to large agreement levels, demonstrating reasonable consistency in TIRADS assessments while also emphasizing areas where interpretations may differ, potentially influencing treatment decisions[15-17]. Despite advances in imaging and risk stratification strategies, thyroid nodule evaluation remains challenging[12.14]. Variability in ultrasound equipment, observer experience, and changing recommendations can all lead to TIRADS evaluation inconsistencies. Furthermore, the interpretation of certain

* Radiology Registrar, Seha Virtual Hospital, Riyadh
Kingdom of Saudi Arabia,
Email: sulx@outlook.com
** Interventional Radiology Fellow, King Abdulaziz Medical City
Riyadh, Kingdom of Saudi Arabia.
*** Interventional Radiology Consultant, Aseer Central Hospital
Aseer, Saudi Arabia.
**** Radiology Resident, King Fahad Military Hospital
Dhahran, Saudi Arabia.
***** Interventional Radiology Consultant
King Faisal Specialist Hospital & Research Centre
Riyadh, Saudi Arabia.

ultrasound characteristics, such as microcalcifications or uneven margins, might be subjective and may necessitate specialized knowledge for appropriate characterization[9,10,15]. Overall, thyroid nodule evaluation has advanced dramatically since the incorporation of ultrasound imaging, standardized reporting systems such as TIRADS, and molecular testing[8,12,16].

The combination of these modalities provides a holistic approach to risk classification, diagnosis, and treatment of thyroid nodules, thereby improving patient care and results. However, further study and optimization of diagnostic algorithms are required to solve remaining problems and increase the accuracy of thyroid nodule detection. The purpose of this study is to assess the diagnostic accuracy and inter-rater reliability of the TIRADS for estimating FNA cytology results for thyroid nodules.

METHODS

Study design: The study used a retrospective observational design and comprised a cohort of 100 patients with thyroid nodules who received ultrasound-guided FNA and TIRADS assessment by two expert radiologists. Patients were chosen based on the availability of comprehensive data, including demographic information, FNA cytology results, and TIRADS scores. The study followed ethical norms and received permission from the institutional review board.

Data collection: Demographic information such as age and gender were obtained for each patient. Clinical data was collected, including FNA indications, lesion side (bilateral, left, midline, or right), and FNA pathology results. TIRADS scores assigned by each radiologist were logged, as were any disparities in rater evaluations.

TIRADS Assessment: TIRADS assessment was carried out in accordance with published recommendations, which included ultrasonography features such as composition, echogenicity, shape, margin, and the presence of calcifications. Each nodule was assigned to one of five TIRADS categories (2, 3, 4, and 5) depending on the overall risk of malignancy associated with the ultrasonography findings.

Statistical assessment: Descriptive statistics, such as frequencies and percentages, were utilized to summarize demographic information, FNA results, and TIRADS scores. Inferential statistics such as chi-squared tests and kappa coefficient computations were performed using relevant statistical software. A p-value < 0.05 was considered statistically significant.

RESULTS

Demographic Characteristics: The most common age groups were 31-40 years (19%) and 41-50 years (31%). There was a higher proportion of female patients (74%) compared to male patients (26%) (Table 1).

Indications for FNA: The majority of cases (72%) were for thyroid nodules, while other indications accounted for 28%.

Lesion Side Distribution: The distribution showed a near balance between left (48%) and right (46%) sides, with a small percentage being bilateral or midline.

FNA Pathology Results: Benign follicular nodules were the most common (38%). Atypia of undetermined significance (21%) and papillary thyroid carcinoma (13%) followed.

TIRAD Scores: Rater 1 primarily assigned scores in the 4 categories (47%), followed by 3 (31%). Rater 2 assigned scores more evenly across 3, 4, and 5 categories, with 4 being the most common (39%). The most prevalent FNA results at 42%, followed by FLUS at 28% and malignant findings at 16%.

Chi-Squared Analysis

For TIRAD 2 (Rater 1), all cases showed a significant association with benign FNA findings (chi-squared = 17.857, p=0.120). For TIRAD 3, common FNA findings were benign (41.9%), followed by FLUS (25.8%) and malignant (16.1%). TIRAD 4 had the highest association with benign findings (44.7%). TIRAD 5 showed a mixed association with benign (33.3%) and malignant (38.1%) findings (Table 2).

DISCUSSION

Thyroid nodules are a common clinical finding, with a sizable proportion of the population acquiring them during their lives[2]. Thyroid

Table 1. Characters of the included patients, and associated FNA and TIRAD ratings (n=100).

Parameter		Frequency (%)
Age, y	Up to 30	2 (2%)
	31-40	19 (19%)
	41-50	31 (31%)
	51-60	20 (20%)
	61-70	18 (18%)
	71 or more	10 (10%)
Gender	Female	74 (74%)
	Male	26 (26%)
Indications for FNA	Thyroid nodule	72 (72%)
	Other indications	28 (28%)
FNA pathology	Atypia of indeterminate significance	21 (21%)
	Benign follicular nodule	38 (38%)
	Benign follicular nodule with Hurthle changes	4 (4%)
	Benign with cystic changes	6 (6%)
	Epithelial neoplasm with oncocystic changes	1 (1%)
	Follicular lesion of indeterminate significance	7 (7%)
	Hashimoto's thyroiditis	8 (8%)
	Metastatic medullary thyroid carcinoma	1 (1%)
	Papillary thyroid carcinoma	13 (13%)
	Poorly differentiated thyroid carcinoma	1 (1%)
FNA finding	Benign	42 (42%)
	Benign with cystic changes	6 (6%)
	FLUS	28 (28%)
	Malignant	16 (16%)
	Thyroiditis	8 (8%)
Rater 1 TIRAD scor	2	1 (1%)
	3	31 (31%)
	4	47 (47%)
	5	21 (21%)
Rater 2 TIRAD scor	3	30 (30%)
	4	39 (39%)
	5	31 (31%)

Table 2. FNA findings in association with TIRAD scores (n=100).

Parameter	Benign	Benign with cystic changes	FLUS	Malignant	Thyroiditis	X^2	P-value
TIRAD score (Rater 1)							
2	1 (100%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	17.857	0.12
3	13 (41.9%)	4 (12.9%)	8 (25.8%)	5 (16.1%)	1 (3.2%)		
4	21 (44.7%)	2 (4.3%)	15 (31.9%)	3 (6.4%)	6 (12.8%)		
5	7 (33.3%)	0 (0%)	5 (23.8%)	8 (38.1%)	1 (4.8%)		

nodules are frequently managed using a combination of diagnostic modalities, comprising ultrasound imaging and FNA cytology[8,12]. The TIRADS has emerged as a useful tool for identifying thyroid nodules based on imaging features and helps in clinical decision-making about the necessity for FNA and subsequent therapy[16,18].

Our data analysis yielded numerous notable findings that add greatly to our understanding of thyroid nodule evaluation. First, we found moderate inter-rater agreement between the two radiologists in TIRADS grading, as shown by a kappa coefficient of 0.453. This data implies that TIRADS assessments have a reasonable amount of consistency, but it also highlights areas where interpretations can differ, potentially influencing treatment decisions[19,20].

Second, the relationship between TIRADS scores and FNA findings was investigated, demonstrating substantial connections between specific TIRADS categories and FNA outcomes. For example, TIRAD 2 by Rater 1 was strongly associated with benign FNA findings (p=0.120), but TIRAD 3 by Rater 2 was significantly associated with benign findings (p=0.022). These findings highlight the predictive utility of TIRADS for identifying benign lesions and directing therapeutic therapy[18,21].

This study highlights the practical benefits of using the TIRADS to predict FNA outcomes in thyroid nodules. By improving risk assessment, TIRADS helps clinicians make more informed decisions about which nodules require further testing, potentially reducing unnecessary FNAs for low-risk cases. This contributes to more efficient patient management and ensures resources are used where they're needed most. Overall, TIRADS provides a structured approach to guiding clinical decisions, which can lead to better patient care and outcomes.

Strengths and limitations: A key strength of this study is its examination of how consistently TIRADS is applied by different radiologists. The moderate agreement between raters suggests that the system can be reliably used in clinical practice, though some variations may occur. The study's sample size of 100 patients also lends credibility to the findings, and the use of well-established statistical methods adds further weight to the conclusions. Moreover, the study sheds light on the predictive power of certain TIRADS categories in distinguishing between benign and malignant nodules, which is crucial for guiding treatment decisions.

While the study provides valuable insights, it's important to acknowledge its limitations. The moderate level of agreement between radiologists (kappa = 0.453) indicates that differences in interpretation can still occur, potentially influencing clinical decisions. Additionally, factors like the type of ultrasound equipment used or the experience level of the radiologists were not explored, which could affect the consistency of TIRADS assessments. Another limitation is that the study's patient group may not represent the broader population, limiting the generalizability of the results. More research is needed to address these issues and improve the overall reliability of TIRADS.

CONCLUSION

This study confirms the usefulness of TIRADS in predicting FNA outcomes, reinforcing its role in helping clinicians make better-informed decisions about thyroid nodules. However, the findings also point to the need for standardized training and updates to TIRADS guidelines to improve consistency in its application. Future research should look at how differences in imaging technology and radiologist expertise impact TIRADS evaluations, with the aim of further improving its accuracy and effectiveness in clinical practice.

Authorship Contribution: All authors share equal effort contribution towards (1) substantial contributions to conception and design, acquisition, analysis and interpretation of data; (2) drafting the article and revising it critically for important intellectual content; and (3) final approval of the manuscript version to be published. Yes.

Potential Conflicts of Interest: None

Competing Interest: None

Acceptance Date: 07 October 2025

REFERENCE

1. Dong X, Li Y, Xie J, et al. The prevalence of thyroid nodules and its factors among Chinese adult women: A cross-sectional study. *Front Endocrinol* 2022;13:967380.
2. Moon JH, Hyun MK, Lee JY, et al. Prevalence of thyroid nodules and their associated clinical parameters: a large-scale, multicenter-based health checkup study. *Korean J Intern Med*2018;33(4):753-62.
3. Brito JP, Yarur AJ, Prokop LJ, et al. Prevalence of thyroid cancer in multinodular goiter versus single nodule: a systematic review and meta-analysis. *Thyroid* 2013;23(4):449-55.
4. Feldkamp J, Führer D, Luster M, et al. Fine needle aspiration in the investigation of thyroid nodules: indications, procedures and interpretation. *Dtsch Arztebl Int* 2016;113(20):353-9.
5. Paschke R, Cantara S, Crescenzi A, et al. European thyroid association guidelines regarding thyroid nodule molecular fine-needle aspiration cytology diagnostics. *Eur Thyroid J*2017;6(3):115-29.
6. Lee YH, Baek JH, Jung SL, et al. Ultrasound-guided fine needle aspiration of thyroid nodules: a consensus statement by the Korean Society of Thyroid Radiology. *Korean J Radiol*2015;16(2):391-401.
7. Shin JH, Baek JH, Chung J, 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.
8. Li X, Hou XJ, Du LY, et al. Virtual Touch Tissue Imaging and Quantification (VTIQ) combined with the American College of Radiology Thyroid Imaging Reporting and Data System (ACR TI-RADS) for malignancy risk stratification of thyroid nodules. *Clin Hemorheol Microcirc* 2019;72(3):279-91.

9. Durante C, Grani G, Lamartina L, et al. The diagnosis and management of thyroid nodules: a review. *JAMA* 2018;319(9):914-24.
10. Pemayun TG. Current diagnosis and management of thyroid nodules. *Acta Med Indones*2016;48(3):247-57.
11. Kobaly K, Kim CS, Mandel SJ. Contemporary management of thyroid nodules. *Annu Rev Med*2022;73:517-28.
12. Moon HJ, Kim EK, Yoon JH, et al. Malignancy risk stratification in thyroid nodules with nondiagnostic results at cytologic examination: combination of thyroid imaging reporting and data system and the Bethesda System. *Radiology* 2015;274(1):287-95.
13. Grani G, Lamartina L, Ascoli V, et al. Reducing the number of unnecessary thyroid biopsies while improving diagnostic accuracy: toward the "right" TIRADS. *J Clin Endocrinol Metab*2019;104(1):95-102.
14. Koc AM, Adıbelli ZH, Erkul Z, et al. Comparison of diagnostic accuracy of ACR-TIRADS, American Thyroid Association (ATA), and EU-TIRADS guidelines in detecting thyroid malignancy. *Eur J Radiol* 2020;133:109390.
15. Chandramohan A, Khurana A, Pushpa BT, et al. Is TIRADS a practical and accurate system for use in daily clinical practice?. *Indian J Radiol Imaging* 2016;26(1):145-52.
16. Scappaticcio L, Trimboli P, Iorio S, et al. Repeat thyroid FNAC: Inter-observer agreement among high-and low-volume centers in Naples metropolitan area and correlation with the EU-TIRADS. *Front Endocrinol* 2022;13:1001728.
17. Du Y, Bara M, Katlariwala P, et al. Effect of training on resident inter-reader agreement with American College of Radiology Thyroid Imaging Reporting and Data System. *World J Radiol*2022;14(1):19-28.
18. Rtam N. Evaluation of the departmental inter-rater reliability when scoring thyroid nodules according to the British Thyroid Association Ultrasound-classification model: Is there significant disagreement?. *Ultrasound* 2023 Dec 28:1742271X231215500.
19. Persichetti A, Di Stasio E, Coccaro C, et al. Inter-and intraobserver agreement in the assessment of thyroid nodule ultrasound features and classification systems: a blinded multicenter study. *Thyroid* 2020;30(2):237-42.
20. Chambara N, Liu SY, Lo X, et al. Diagnostic performance evaluation of different TI-RADS using ultrasound computer-aided diagnosis of thyroid nodules: An experience with adjusted settings. *PLoS One* 2021;16(1):e0245617.
21. Huh S, Yoon JH, Lee HS, et al. Comparison of diagnostic performance of the ACR and Kwak TIRADS applying the ACR TIRADS' size thresholds for FNA. *Eur Radiol* 2021;31(7):5243-50.