The Research of Medical Science Review
ISSN: 3007-1208 & 3007-1216 Volume 3, Issue 5, 2025

BRITISH THYROID ASSOCIATION 2014 CLASSIFICATION ULTRASOUND SCORING OF THYROID NODULES IN PREDICTING MALIGNANCY

Dr. Mahnoor Ayub[*1], Dr. Khubaib Shahid[2]

*[*1]Post Graduate Trainee (FCPS Diagnostic Radiology), Fatima Memorial Hospital, Lahore*
[2]Professor / Head of Radiology Department (Study Supervisor), Fatima Memorial Hospital, Lahore.

[*1]mahnoorrao57@gmail.com

DOI: https://doi.org/10.5281/zenodo.15488282

Keywords
British Thyroid Association 2014 Classification, Diagnostic Accuracy, Malignancy, NPV, PPV, Sensitivity, Specificity, Thyroid Nodule

Article History
Received on 14 April 2025
Accepted on 14 May 2025
Published on 22 May 2025

Corresponding Author: *
Dr. Mahnoor Ayub

Abstract

Background: *Thyroid nodules are commonly detected, but histopathology, the gold standard for malignancy diagnosis, is invasive and costly. The 2014 British Thyroid Association ultrasound scoring system offers a non-invasive alternative, though it's diagnostic accuracy remains controversial, prompting this study to assess its reliability.*

Objective: *To assess diagnostic accuracy of British thyroid association 2014 classification ultrasound scoring in predicting malignant thyroid nodules taking histopathology as gold standard.*

Duration: *Six months w.e.f 21-12-2023 to 20-06-2024.*

Methodology: *After ethical approval, 145 patients who met the inclusion criteria were enrolled from the OPD and referred to the radiology department. Informed consent was obtained. Demographics such as age, BMI, neck circumference, smoking history, diabetes, hypertension, residence, and occupation were recorded. Patients underwent ultrasonography and were classified based on the British Thyroid Association 2014 guidelines. Biopsy samples were sent for cytology. Findings were documented, and data was analyzed using SPSS version 22.0.*

Results: *The study included 145 participants, with a mean age of 45.5 years, 60% aged 20-50 years, and 66.2% female, with 63.4% classified as overweight/obese. Health conditions like hypertension (56.6%) and diabetes (43.4%) were common. Ultrasound scoring identified 40.7% as malignant, while histopathology confirmed 42.1%. The diagnostic performance of ultrasound scoring showed a sensitivity of 90.16%, a positive predictive value (PPV) of 93.23%, a negative predictive value (NPV) of 93.01%, and an accuracy of 93.10%. Stratification by subgroups showed no significant differences in malignancy prevalence.*

Conclusion: *Ultrasound scoring proved to be a reliable method for predicting malignant thyroid nodules, offering high diagnostic performance. The study showed no significant differences in malignancy prevalence across subgroups. These results suggest that ultrasound scoring is a valuable non-invasive tool for thyroid nodule evaluation.*

INTRODUCTION

Thyroid gland malignancies are the most prevalent among endocrine system cancers, with thyroid nodules being approximately four times more common in women than in men.[1] The long-term survival rates, even in cases of recurrent disease, have prompted questions regarding the extent of thyroidectomy required for patients with well-differentiated thyroid carcinomas.[1,2] Over the last 40 years, the incidence of thyroid cancer has increased significantly in the United States, with similar trends observed globally. Thyroid cancer is currently the 13th most common cancer diagnosis overall and the 6th most common among women.[3] A study found that 30.5% of the thyroid nodules studied were malignant.[4]

Thyroid carcinoma is classified into three main histologic types: differentiated (including papillary, follicular, and Hurthle cell), medullary, and anaplastic, the latter being an aggressive, undifferentiated tumor. Mortality rates for thyroid cancer remain low, with differentiated thyroid tumors generally offering an excellent prognosis, as evidenced by a 10-year survival rate exceeding 90-95%.[5] In response to numerous queries regarding the management of thyroid disease amidst the global pandemic, members of the British Thyroid Association (BTA) and the Society for Endocrinology have provided guidelines to assist endocrine and primary care practitioners during this extraordinary period.[6,7] Regarding diagnostic tools, one study highlighted that the BTA-U score demonstrated a sensitivity and negative predictive value of 100% in identifying and predicting malignancy, although its specificity and positive predictive value were 34% and 32%, respectively.[8] Another study showed the BTA-U score's sensitivity at 74% and specificity at 92% in detecting malignancy.[9] The varying diagnostic accuracy of the BTA-U score across different studies highlights need for further research to establish a more reliable and consistent understanding of its sensitivity, specificity, and diagnostic accuracy. Before its widespread implementation in clinical settings, clear-cut evidence is necessary to validate its performance, ensuring that it can be trusted as an effective tool for diagnosing thyroid malignancies.

METHODOLOGY

This cross-sectional study was conducted at Fatima Memorial Hospital Lahore over six months after ethical approval. It aimed to assess the diagnostic performance of ultrasound scoring (BTA 2014) in identifying malignant thyroid nodules. Using the WHO calculator, a sample size of 145 patients was determined with a 95% confidence level, 30.5% prevalence of thyroid malignancy, a sensitivity of 74% for BTA-U, and 92% specificity, with a margin of error of 13% for sensitivity and 6% for specificity, based on previous studies.[9] Participants aged 20 to 70 years, both genders, presenting with thyroid lesions for biopsy, were enrolled. Exclusion criteria included recurrent lesions, chemotherapy or radiotherapy, inappropriate biopsy samples, metastatic disease, and pregnancy. After obtaining informed consent, participants' demographics, including age, gender, BMI, neck circumference, smoking, diabetes, hypertension, and occupation, were recorded. Patients underwent ultrasonography, classifying nodules as malignant or benign per BTA 2014. Biopsy samples were sent for cytological analysis. The operational definitions used in this study were as follows: A Thyroid Nodule was classified as positive if the BTA-U score was between 3-5 (malignant) and as benign if the score was between 1-2. On cytology, a nodule was considered positive if more than 10% of tumor cells absorbed the H&E stain. A True Positive was defined as a BTA-U score of 3-5 with malignant histopathology. A True Negative was a BTA-U score of 1-2 with benign histopathology. A False Positive was when the BTA-U score was 3-5 but the histopathology was benign, while a False Negative occurred when the BTA-U score was 1-2, but the histopathology showed malignancy. Data was analyzed using SPSS version 22.0, checking normality with the Shapiro-Wilk test. Descriptive statistics for continuous variables (age, BMI, etc.) and frequencies for categorical variables (gender, smoking, etc.) were calculated. A 2x2 contingency table was generated to evaluate the sensitivity, specificity, PPV, NPV, and diagnostic accuracy of BTA-U against cytology as the gold standard. Stratification by age, gender, BMI, smoking, and other factors was performed, and diagnostic performance was assessed within each subgroup. This methodology provided a comprehensive evaluation of the ultrasound scoring

system's effectiveness in diagnosing thyroid malignancy.

RESULTS

The study included 145 participants with a mean age of 45.54 ± 14.98 years. The majority were aged 20-50 years (60%) and female (66.2%). The mean BMI was 26.78 ± 3.91, with 63.4% categorized as overweight/obese. Neck circumference averaged 37.75 ± 2.84 cm, with 73.8% at increased risk. Regarding health factors, 26% were smokers, 43.4% had diabetes, and 56.6% had hypertension. Most participants resided in rural areas (52.4%) and were unemployed (54.5%). The average symptom duration was 14.48 ± 6.88 months, with 55.2% experiencing symptoms for over a year. Lesion sizes ranged from 1-2 cm (35.9%) to 3-4 cm (64.1%). Data is given in Table 1.0. The data was found normally distributed. Out of the 145 participants, 40.7% were classified as malignant based on ultrasound scoring, while 42.1% were diagnosed as malignant through histopathology. The remaining participants were classified as non-malignant in both modalities, as given in Table 2.0. Moreover, stratification of frequency of malignant nodules as per histopathology on the basis of sub groups produced insignificant difference for all the sub groups (p-value>0.05). Table 3.0 presents the 2x2 contingency table used to evaluate the diagnostic performance of ultrasound scoring in diagnosing malignant thyroid nodules, with histopathology findings as the gold standard. In this analysis, 55 true positives (malignant nodules correctly identified), 4 false positives (benign nodules incorrectly identified as malignant), 6 false negatives (malignant nodules incorrectly identified as benign), and 80 true negatives (benign nodules correctly identified) were recorded. The results show a sensitivity of 90.16%, specificity of 95.24%, and accuracy of 93.10%. The disease prevalence was 42.10%, with a positive predictive value of 93.23% and a negative predictive value of 93.01%. Output from 2x2 contingency table of diagnostic accuracy showed almost similar output with little variable as given in Table 3.0 that however makes way for future research also.

Table 1.0: Demographic Characteristics of Patients with Suspicion of Thyroid Malignancy

Characteristics	Participants (n=145)
Age (years)	45.54±14.98
• **20-50** years	88 (60.0%)
• **51-70** years	56 (38.6%)
Gender	
• Male	49 (33.8%)
• Female	96 (66.2%)
BMI (kg/m^2)	26.78±3.91
• **Normal Weight**	53 (36.6%)
• **Overweight / Obese**	92 (63.4%)
Neck Circumference (cm)	37.75±2.84
• **Normal**	38 (26.0%)
• **Increased Risk**	107 (73.8%)
Smoking	
• **Yes**	38 (26.0%)
• **No**	107 (73.8%)
Diabetes	
• **Yes**	63 (43.4%)
• **No**	82 (56.6%)
Hypertension	
• **Yes**	82 (56.6%)
• **No**	63 (43.4%)

Residence	
• **Urban**	69 (47.6%)
• **Rural**	76 (52.4%)
Occupation	
• **Employed**	66 (45.5%)
• **Unemployed**	79 (54.5%)
Duration of Symptoms	14.48±6.88
• **≤12 months**	65 (44.8%)
• **>12 months**	80 (55.2%)
Lesion Size	1.64±0.48
• **1-2 cm**	52 (35.9%)
• **3-4 cm**	93 (64.1%)

Table 2.0: Frequency of Malignant Thyroid Nodule upon Ultrasound Scoring and Histopathology

Modality	Malignant Thyroid Nodule	Frequency (n)	Percent (%)
Ultrasound Scoring	Yes	59	40.7 %
	No	86	59.3 %
	Total	145	100.0 %
Histopathology	Yes	61	42.1 %
	No	84	57.9 %
	Total	145	100.0 %

Table 3.0: 2x2 Contingency Table to Determine Diagnostic Performance of Ultrasound Scoring in Diagnosing Malignant Thyroid Nodule Taking Histopathology Findings as Gold Standard

Ultrasound Scoring	Histopathology		Total
	Malignant	Benign	
Malignant	**55**[a]	**4**[c]	**59**
Benign	**6**[b]	**80**[d]	**86**
Total	**61**	**84**	**145**

[a]True Positive = 55, [c]False Positive = 4, [b]False Negative = 6, [d]True Negative = 80

Statistic	Formula	Value
Sensitivity	$\frac{a}{a+b}$	90.16%
Specificity	$\frac{d}{c+d}$	95.24%
Accuracy	$\frac{a+d}{a+b+c+d}$	93.10%
Disease prevalence	$\frac{a+b}{a+b+c+d}$	42.10%

Positive Predictive Value	$\frac{a}{a+c}$	93.23%
Negative Predictive Value	$\frac{d}{b+d}$	93.01%

Table 4.0: Output from 2x2 Contingency Table to Determine Diagnostic Performance of Ultrasound Scoring in Diagnosing Malignant Thyroid Nodule Taking Histopathology Findings as Gold Standard Across Sub Groups

Group	Sub Groups	Sensitivity	Specificity	PPV	NPV	Diagnostic Accuracy	Prevalence
Age	20-50 years	91.30%	94.55%	90.94%	94.77%	93.33%	37.5%
	51-70 years	85.71%	96.43%	96.00%	87.10%	91.07%	50.0%
Gender	Male	85.00%	100.0%	100.0%	90.63%	93.88%	40.8%
	Female	92.68%	92.73%	90.47%	94.45%	92.71%	42.7%
BMI	Normal Weight	96.15%	100.0%	100.0%	96.42%	98.11%	49.1%
	Overweight / Obese	85.71%	92.98%	88.22%	91.39%	90.22%	38.0%
Neck Circum.	Normal	86.67%	91.30%	86.68%	91.30%	89.47%	39.5%
	Increased Risk	91.30%	96.72%	95.46%	93.65%	94.39%	43.0%
Smoking	Yes	85.71%	100.0%	100.0%	92.32%	94.74%	36.8%
	No	91.49%	93.33%	91.48%	93.34%	92.52%	43.9%
Diabetes	Yes	100.0%	100.0%	100.0%	100.0%	100.0%	39.7%
	No	83.33%	91.30%	88.23%	87.50%	87.81%	43.9%
Hypertension	Yes	86.84%	95.45%	94.28%	89.38%	91.47%	46.3%
	No	95.65%	95.00%	91.33%	97.54%	95.23%	35.5%
Residence	Yes	87.10%	92.11%	89.99%	89.75%	89.86%	44.9%
	No	93.33%	97.83%	96.56%	95.74%	96.05%	39.5%
Occupation	Employed	88.46%	97.50%	95.83%	92.86%	93.94%	39.4%
	Unemployed	91.43%	93.18%	91.43%	93.18%	92.41%	44.3%
Duration of Symptoms	≤12 months	82.14%	91.89%	88.47%	87.17%	87.69%	43.1%
	>12 months	96.97 %	97.87%	96.98%	97.87%	97.50%	41.3%
Lesion Size	1-2 cm	90.91%	90.00%	86.95%	93.11%	90.38%	42.3%
	3-4 cm	89.74 %	98.15%	97.22%	92.99%	94.63%	41.9%

DISCUSSION

Thyroid nodules are commonly detected, but accurately predicting malignancy remains a challenge.[10] Histopathology has been the gold standard for diagnosing malignancy; however, it is invasive, costly, and time-consuming. The British Thyroid Association 2014 ultrasound scoring system was introduced as a non-invasive method to predict malignancy,[11,12] but variations in diagnostic accuracy have been observed across studies.[8,9] This inconsistency in results prompted the planning of this study to assess the reliability and diagnostic accuracy of the 2014 British scoring system.

In this study, the mean age of participants was 45.54±14.98 years, with the majority being between 20-50 years (60%) and 66.2% female. A similar mean age of 46.06±10.34 years was reported by Amin et al.,[13] while other studies reported 51±15 years by Weller et al.[8] and 42.24±9.47 years by Ain et al.[14] Wahid et al.[15] found 26.8% of participants were in the 20-30 age group, with 73.2% in the 31-50 age group. The 66.2% female proportion in this study aligns with findings from Weller et al. (42.4%), Amin et al.[13] (27.19%), Wahid et al.[15] (37.8%), and Ain et al.[14] (32.0%). The mean BMI in this study was 26.78±3.91 kg/m², while Ain et al.[14] reported a mean BMI of

24.34±4.65 kg/m². Regarding malignancy, 40.7% of participants were classified as malignant based on ultrasound scoring, while 42.1% were diagnosed as malignant through histopathology. These histopathology findings are consistent with McClean et al.,[16] who reported 43.8%, while Weller et al. reported 34.0%, Amin et al. 21.93%, and Ain et al.[14] 38.0%. The results of this study showed a sensitivity of 90.16%, specificity of 95.24%, accuracy of 93.10%, a positive predictive value (PPV) of 93.23%, and a negative predictive value (NPV) of 93.01%. Persichetti et al.[9] reported a sensitivity of 74.0%, specificity of 92.0% and diagnostic accuracy of 89.0% of ultrasound scoring in predicting malignant nodules taking histopathology as gold standard. Amin et al.[13] reported a sensitivity of 88.0%, specificity of 97.8%, positive predictive value (PPV) of 91.7%, negative predictive value (NPV) of 96.7%, and a disease prevalence of 21.93%. Wahid et al.[15] demonstrated a sensitivity of 96.58%, specificity of 83.33%, PPV of 76.65%, NPV of 97.72%, diagnostic accuracy of 88.12%, and a disease prevalence of 21.93%. Ain et al.[14] found a sensitivity of 88.2%, specificity of 81.8%, PPV of 71.4%, NPV of 93.1%, and a disease prevalence of 38.0%. In the study by McClean et al.,[16] while diagnosing malignant thyroid nodules using the British Thyroid Association Ultrasound (BTA-U) scoring system, a sensitivity of 88.1%, specificity of 41.6%, positive predictive value (PPV) of 54.1%, negative predictive value (NPV) of 81.8%, and a disease prevalence of 43.8% were reported, taking histopathology as gold standard. Akhte et al.[17] reported a sensitivity of 80%, specificity of 34%, PPV of 100%, and NPV of 90% for ultrasonography in diagnosing malignant thyroid nodules.

CONCLUSION

Ultrasound scoring proved to be a reliable method for predicting malignant thyroid nodules, offering high diagnostic performance. The study showed no significant differences in malignancy prevalence across subgroups. These results suggest that ultrasound scoring is a valuable non-invasive tool for thyroid nodule evaluation.

LIMITATIONS & RECOMMENDATIONS

The strengths of this study include its large sample size, comparison of ultrasound scoring with histopathology as the gold standard, and the demonstration of high diagnostic accuracy of the BTA-U scoring system. However, limitations include the reliance on a single-center dataset and the potential variability in ultrasound interpretation. Future research should focus on multi-center studies, standardization of ultrasound practices, and further validation of the BTA-U score in diverse populations to enhance its clinical applicability and generalizability.

Conflict of Interest: None
Source of Funding: None

Authors Contribution

Author 1

Substantial contributions to study design, acquisition of data
Analysis & Interpretation of Data, Manuscript writing
Has given final approval of the version to be published
Agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Author 2

Substantial contributions to concept, study design
Data Analysis, Manuscript writing, Critical Review
Has given final approval of the version to be published
Agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

REFERENCES

1. Bakshi J, Patro SK, Kaur N, Panda NK, Budhiraja G. Understanding malignancies of the thyroid gland: institutional experience. Indian J Otolaryngol Head Neck Surg. 2018;70:482-9.
2. Rossi ED, Pantanowitz L, Raffaelli M, Fadda G. Overview of the ultrasound classification systems in the field of thyroid cytology. Cancers. 2021;13(13):3133.
3. Kitahara CM, Schneider AB. Epidemiology of thyroid cancer. Cancer Epidemiol Biomarkers Prev. 2022;31(7):1284-97.

4. Al-Hakami HA, Alqahtani R, Alahmadi A, Almutairi D, Algarni M, Alandejani T. Thyroid nodule size and prediction of cancer: a study at tertiary care hospital in Saudi Arabia. Cureus. 2020;12(3):1-7.
5. Ajani JA, D'Amico TA, Bentrem DJ, Chao J, Cooke D, Corvera C, et al. Gastric cancer, version 2.2022, NCCN clinical practice guidelines in oncology. J Natl Compr Cancer Netw. 2022;20(2):167-92.
6. Watkins L, O'Neill G, Young D, McArthur C. Comparison of British Thyroid Association, American College of Radiology TIRADS and Artificial Intelligence TIRADS with histological correlation: diagnostic performance for predicting thyroid malignancy and unnecessary fine needle aspiration rate. Br J Radiol. 2021;94(1123):20201444.
7. Alfuraih AM, Alotaibi AM, Alshammari AK, Alrashied BF, Mashhor YM, Mahmoud M, et al. Evaluating diagnostic accuracy and agreement of TI-RADS scoring in thyroid nodules: a comparative analysis between sonographers and radiologists. Plos One. 2024;19(10):e0312121.
8. 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.
9. Persichetti A, Di Stasio E, Guglielmi R, Bizzarri G, Taccogna S, Misischi I, Graziano F, et al. Predictive value of malignancy of thyroid nodule ultrasound classification systems: a prospective study. J Clin Endocrinol Metab. 2018;103(4):1359-68.
10. Foroughi AA, Mokhtari M, Heidari E, Nazeri M, Rastgouyan H, Babaei A. Concordance between TIRADS and cytology in thyroid nodule. Iran J Otorhinolaryngol. 2022;34(125):295-302.
11. Solomon C, Petea-Balea DR, Dudea SM, Bene I, Silaghi CA, Lenghel ML. Role of ultrasound elastography and contrast-enhanced ultrasound (CEUS) in diagnosis and management of malignant thyroid nodules—an update. Diagnostics. 2025;15(5):599.
12. 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. 2024;32(2):76-84.
13. Amin M, Zahira M, Siddiqui MA, Kamal R, Javed S, Maqsood A, Khan MM. Diagnostic accuracy of a new sonographic classification system in diagnosing malignant thyroid nodules. Pak J Med Health Sci. 2021;15(6):1865-7.
14. Ain SQ, Bibi S. Diagnostic accuracy of ultrasonography using u classification of thyroid nodules in detection of malignant thyroid. J Teknokes. 2023;16(1):1-6.
15. Wahid G, Tamkeen N, Maqsood F, Afsar M, Fahim S. diagnostic accuracy of ultrasound in detecting malignant thyroid nodules keeping histopathology as gold standard. J Postgrad Med Inst. 2024;38(3):165-9.
16. McClean S, Omakobia E, England RJ. Comparing ultrasound assessment of thyroid nodules using BTA U classification and ACR TIRADS measured against histopathological diagnosis. Clin Otolaryngol. 2021;46(6):1286-9.
17. Akhter T, Shahid K, Afzal U. Diagnostic accuracy of ultrasound u classification system of thyroid ultrasound in predicting thyroid malignancy by using histopathology as gold standard. Ann PIMS-Shaheed Zulfiqar Ali Bhutto Med Univ. 2018;14(3):222-6