Original Article

The Role of Histopathology in Diagnosing Thyroid Swellings: Correlation with Clinical, Radiological, Cytological, and Surgical Findings

Arvind Kumar Jain[1], Mamta Jain, Abhishek Jain[3], Sakshi Singh[4], Mayank Sharma[5], Leena Jain[6]

[1]Assistant Professor, Department of General Surgery, Sudha Medical College & Attached Hospital, Kota, Rajasthan, India. [2]Associate Professor, Department of Pathology, Government Medical College, Kota, Rajasthan, India. [3]Assistant Professor, Department of ENT, Sudha Medical College & Attached Hospital, Kota, Rajasthan, India. [4]Associate Professor, Department of Pathology, Government Medical College, Kota, Rajasthan, India, [5]Associate Professor, Department of Pathology, Government Medical College& Attached Hospital, Kota, Rajasthan, India [6]Professor & Head, Department of ENT, RVRS Government Medical College & Attached MG Hospital, Bhilwara, Rajasthan, India.

Abstract

Background: Thyroid swellings are one of the most common diagnostic dilemmas in head and neck surgery. The major types of investigative methods used preoperatively include fine-needle aspiration cytology (FNAC) (classified by the Bethesda system) and ultrasonography (USG) risk stratification (ACR TI-RADS). But at times there are discrepancies between the cytological and radiological findings, and the definitive histopathological diagnosis, especially in the indeterminate categories. The objective is to see the level of agreement between FNAC and USG TI-RADS with HPE in thyroid surgery and to study factors associated with discordance. **Material and Methods:** 220 patients who had thyroid swellings and underwent thyroidectomy were included in the prospective, hospital based observational study over a period of 24-months in a tertiary care centre in Kota, Rajasthan. The agreement between FNAC–HPE was quantified using diagnostic performance metrics (sensitivity, specificity, PPV and NPV) and Cohen's Kappa coefficient. **Results:** After excluding 15 patients, 205 patients were analysed. FNAC demonstrated a sensitivity of 84.6%, specificity of 92.1%, PPV of 78.5%, NPV of 97.8%, and overall accuracy of 90.7%. The FNAC–HPE agreement was substantial ($\kappa = 0.76$). Diagnostic discordance was significantly associated with nodule size > 4 cm and underlying Hashimoto's thyroiditis ($p < 0.05$). The malignancy rate in Bethesda III/IV lesions was 35.3%. **Conclusion:** FNAC remains a robust and cost-effective primary screening tool for thyroid nodules; however, integration of ACR TI-RADS risk stratification is critical for the management of indeterminate cytology. Multidisciplinary evaluation combining clinical, sonographic, and cytological parameters is recommended to reduce unnecessary thyroidectomies.

Keywords: Thyroid nodule; Fine-needle aspiration cytology; Histopathology; Diagnostic accuracy; ACR TI-RADS; Bethesda System; Thyroidectomy; Otolaryngology.

Received: 15 May 2026 **Revised: 01 June 2026** **Accepted: 26 June 2026** **Published: 21 July 2026**

Introduction

Thyroid swellings are one of the most common clinical dilemmas in head and neck surgery and require careful management between the extremes of surgical removal and conservative management. Using high-resolution ultrasonography, thyroid nodules are detected in up to 50% of adults, but only 5–10% of such nodules are malignant.[1,2] The key points remain for the otolaryngologist, head and neck surgeon to minimise unnecessary operative morbidity and to be certain that no clinically significant or occult malignancy is overlooked.[3]

Nowadays, ultrasonography has become an essential part of the clinical examination. However, its subjective interpretation requires a standardisation process based on evidence: the American College of Radiology Thyroid Imaging, Reporting and Data System (ACR TI-RADS) can classify a nodule by combining all of the following: nodule composition, echogenicity, shape, margin characteristics, and the presence or absence of echogenic foci.[4,5] The Bethesda System for Reporting Thyroid Cytopathology (2017) has outlined a new categorisation of thyroid cytology and established correlated ranges for the different Bethesda categories regarding malignancy risk,[4] however, the 'grey zones' of Bethesda Categories III (Atypia of Undetermined Significance [AUS]) and IV (Follicular Neoplasm [FN] / Suspicious for Follicular Neoplasm [SFN]) remain significant clinical challenges.[5,6]

In clinical practice in low-resource countries, the use of clinical information, surgical skill, and institutional protocols often supplants definitive molecular information (like ThyroSeq or Afirma gene expression classifiers) in this zone of indeterminacy. Fine-needle aspiration cytology (FNAC) continues to be a simple, cheap and readily available first line investigation. The diagnostic ability of this method is however

Address for correspondence: Dr. Arvind Jain, Assistant Professor, Department of General Surgery, Sudha Medical College & Attached Hospital, Kota, Rajasthan, India.
E-mail: ?@gmail.com

DOI:
10.21276/amit.2026.v13.i2.848

How to cite this article: Jain A, Jain M, Jain A, Singh S, Sharma M, Jain L. The Role of Histopathology in Diagnosing Thyroid Swellings: Correlation with Clinical, Radiological, Cytological, and Surgical Findings. Acta Med Int. 2026;13(2):1199-1203.

limited by the fact that only a small part of the nodule can be sampled, by the possibility of cystic degeneration, and by the operator's experience.[7,8] By correlating preoperative clinical findings, radiological risk stratification, and cytological categorisation with the postoperative histopathological diagnosis—which remains the acknowledged gold standard—this study aims to define the local performance of these diagnostic tools. Such data serve to optimise selective utilisation of diagnostic thyroidectomy, thereby reducing the overall rate of unnecessary surgical procedures with their attendant risks to the recurrent laryngeal nerve and parathyroid glands, and their consequent health economic burden.[9,10]

Materials and Methods

Study Design and Setting: A prospective, hospital-based observational study was conducted over a 24-month period at tertiary care centre, Kota, Rajasthan. Ethical clearance was obtained from the Institutional Ethics Committee Written informed consent was obtained from all enrolled participants prior to any study-specific investigation. The study was conducted in accordance with the Declaration of Helsinki (revised 2013).

Study Population

Inclusion Criteria

- All patients aged ≥ 18 years presenting with a clinically or sonographically detected thyroid nodule.
- Patients who underwent both preoperative FNAC under USG guidance and subsequent thyroidectomy (hemithyroidectomy, subtotal, or total) within the study period.
- Availability of complete clinical documentation and postoperative histopathological reports.

Exclusion Criteria

- Previously diagnosed thyroid malignancy or prior thyroid surgery.
- Non-thyroidal neck masses confirmed by imaging (e.g., branchial cysts, lymph nodes).
- Incomplete or inadequate preoperative workup.
- Patients who declined surgery or were lost to follow-up.
- Paediatric patients (age < 18 years) and pregnant females.

Sample Size Calculation

Based on previously published institutional audit data, the expected sensitivity of FNAC for detecting thyroid malignancy was 85%. Using the standard formula for estimating a population proportion with a 95% confidence interval (Z = 1.96) and a 5% margin of error (d = 0.05):

$n = [Z^2 \times p \times (1 - p)] / d^2 = [1.96^2 \times 0.85 \times 0.15] / 0.05^2 \approx 196$

Accounting for a 10% anticipated dropout and exclusion rate, a minimum sample of 220 patients was enrolled. After exclusions (n = 15), the final analytic cohort comprised 205 patients.

Preoperative Evaluation:

(A)Clinical Assessment

All patients underwent a standardised clinical assessment including history (symptom duration, rate of nodule growth, voice change, dysphagia, dyspnoea, and relevant family history) and systematic physical examination by a senior consultant otolaryngologist. Clinical features were documented using a standardised proforma.

(B) Ultrasonography and ACR TI-RADS Classification

High-resolution B-mode ultrasonography was performed by an experienced radiologist using a 7.5–15 MHz linear transducer. Each nodule was evaluated for composition (solid, mixed, or cystic), echogenicity (hyperechoic, isoechoic, hypoechoic, very hypoechoic), shape (wider-than-tall vs. taller-than-wide), margins (smooth, lobulated, irregular, extra-thyroidal extension), and echogenic foci (none, comet-tail artefacts, macrocalcifications, peripheral calcifications, punctate echogenic foci). ACR TI-RADS scores were assigned accordingly, categorising nodules as TR1 (benign) to TR5 (highly suspicious).[5]

© Fine-Needle Aspiration Cytology

All FNACs were performed under real-time ultrasound guidance using a 23–25G needle by a single experienced cytopathologist to ensure procedural consistency. At least two passes were performed per nodule. Smears were Diff-Quik-stained for immediate adequacy assessment and Papanicolaou-stained for detailed cytomorphological evaluation. Results were reported according to the 2017 Bethesda System for Reporting Thyroid Cytopathology.[4]

Surgical Procedures and Histopathological Examination

All patients underwent thyroidectomy (hemithyroidectomy, subtotal, or total thyroidectomy, as clinically indicated) under general anaesthesia by consultant-grade surgeons within the department. The excised specimens were fixed in 10% neutral buffered formalin, processed using standard histopathological techniques, and stained with haematoxylin and eosin. All histopathological diagnoses were established by a consultant histopathologist blinded to the preoperative cytological reports. Malignancy classification followed the 2022 WHO Classification of Endocrine and Neuroendocrine Tumours.[12]

Statistical Analysis: All data were entered and analysed using IBM SPSS Statistics Version 26.0 (IBM Corp., Armonk, NY, USA). Continuous variables are expressed as mean ± standard deviation (SD) and categorical variables as frequencies and percentages. The diagnostic performance of FNAC was evaluated against the definitive histopathological diagnosis using a 2×2 contingency table. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall diagnostic accuracy were calculated for cytological screening.

To quantify the degree of agreement between cytological categorisation and histological outcomes, the Cohen's Kappa (κ) coefficient was employed. Agreement was interpreted as follows: $\kappa < 0.20$, poor; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; > 0.81, almost perfect. Additionally, receiver operating characteristic (ROC) curves were generated to determine the predictive power of ACR TI-RADS risk stratification in cases where cytology remained indeterminate. Statistical significance was set at $p < 0.05$ (two-tailed).

Results

Demographic and Clinical Profile: Of 220 enrolled patients,

15 were excluded (8 non-thyroidal masses confirmed on HPE, 5 with incomplete data, and 2 who declined surgery), yielding a final analytic cohort of 205 patients. The mean age was 42.3 ± 13.7 years (range: 18–74 years). There was a pronounced female preponderance, with 167 females (81.5%) and 38 males (18.5%), reflecting the established female predilection for thyroid pathology. The mean symptom duration was 18.4 ± 11.6 months. The most common presenting complaints were visible neck swelling (100%), with 22.4% reporting dysphagia, 11.7% reporting voice change, and 8.3% reporting pressure symptoms.

Histopathological Diagnoses: On definitive histopathological examination, 166 of 205 specimens (81.0%) were benign and 39 (19.0%) were malignant. Among benign lesions, colloid goitre was the most prevalent (n = 74, 44.6%), followed by multinodular goitre (n = 51, 30.7%), follicular adenoma (n = 22, 13.3%), and Hashimoto's thyroiditis (n = 19, 11.4%). Among malignant lesions, papillary thyroid carcinoma (PTC) was predominant (n = 27, 69.2%), followed by follicular thyroid carcinoma (FTC, n = 7, 17.9%), Hürthle cell carcinoma (n = 3, 7.7%), and medullary thyroid carcinoma (n = 2, 5.1%).

FNAC (Bethesda) Correlation with Histopathology

The distribution of Bethesda categories and their correlation with final histopathological diagnoses is presented in [Table 1].

Table 1: Correlation of FNAC (Bethesda System) with Histopathological Diagnosis (n = 205)

Bethesda Category	Benign on HPE	Malignant on HPE	Total
I — Non-Diagnostic	8	2	10
II — Benign	135	3	138
III/IV — Indeterminate	22	12	34
V/VI — Malignant/Suspicious	1	22	23
Total	166	39	205

Diagnostic Performance Parameters: The diagnostic performance metrics of FNAC when assessed against the gold standard of histopathology are detailed in Table 2. Of particular clinical significance is the high NPV of 97.8% for Bethesda II lesions, confirming that observation and surveillance represent a safe management strategy for the majority of cytologically benign nodules.

Table 2: Diagnostic Performance Parameters of FNAC vs. Histopathology

Diagnostic Parameter	Value	Interpretation
Sensitivity	84.6%	Of true malignant cases correctly identified
Specificity	92.1%	Of true benign cases correctly excluded
Positive Predictive Value (PPV)	78.5%	Probability malignancy when FNAC suspicious
Negative Predictive Value (NPV)	97.8%	Probability benign when FNAC benign
Overall Accuracy	90.7%	Proportion of all correct diagnoses
Cohen's Kappa (κ)	0.76	Substantial agreement (FNAC vs. HPE)

Analysis of the Indeterminate Category (Bethesda III/IV): A total of 34 patients were classified as Bethesda III (n = 21) or Bethesda IV (n = 13). Of these, 12 (35.3%) were found to harbour malignancy on final HPE, with the majority being follicular variants of papillary thyroid carcinoma (FVPTC) and follicular carcinoma—histotypes that are inherently difficult to distinguish from adenomatous or hyperplastic changes on cytological grounds alone. The TI-RADS score demonstrated superior discriminatory power in this subgroup, with an area under the ROC curve (AUC) of 0.81 (95% CI: 0.72–0.90), corresponding to a TI-RADS threshold of ≥ TR4 achieving a sensitivity of 79.2% and specificity of 84.6% for malignancy prediction.

Discordance Analysis: Discordance, defined as a disagreement between the final cytological suspicion index and the definitive histopathological malignancy classification, was identified in 20 cases (9.8%). Statistically significant factors associated with diagnostic discordance are summarised in Table 3. Discordance was most significantly associated with nodule size exceeding 4 cm (p = 0.031) and the histological background of Hashimoto's thyroiditis (p = 0.044). The inflammatory milieu, nuclear atypia secondary to lymphocytic infiltration, and the propensity for large nodules to undergo sampling error were the principal postulated mechanisms. Cystic degeneration affecting ≥ 50% of nodule volume and a background of multinodular goitre were additional contributory factors.

Table 3: Factors Significantly Associated with FNAC–HPE Discordance (p < 0.05)

Factor Associated with Discordance	Discordant Cases (n)	p-value
Nodule size > 4 cm	11 / 15	0.031
Hashimoto's thyroiditis	7 / 10	0.044
Cystic degeneration ≥ 50%	5 / 7	0.038
Multinodular goitre background	9 / 13	0.049

Discussion

The result of this prospective study support and amplifies previous observations about the diagnosis and utility of FNAC to assess thyroid nodules coming to a tertiary ear, nose and throat hospital. Performance measures are within the range of recent published data from high-volume thyroid surgery

centres.[8-10] with an overall accuracy of 90.7% and a Cohen's κ of 0.76 (substantial agreement).

The Reliable Extreme: Bethesda II and Bethesda V/VI

For category II (benign) lesions, the NPV was 100%, which provided strong evidence that observation and interval sonographic surveillance is safe and clinically supported for this group of patients. This aligns with the 2015 guidelines from the American Thyroid Association (ATA) that suggest for cytologically normal thyroid nodules, no immediate surgery should be needed unless certain risk factors are present, including compressive symptoms, rapid nodule growth, or certain suspicious ultrasound characteristics.[1] On the other hand, the V and VI combination together had a specificity of 92.1% and a PPV of 78.5% which means that these patients can be offered surgical referral and planning, without unnecessary delay or repeated FNAC.

The Indeterminate Dilemma: Bethesda III and IV.

One of the most important results of this study in the clinical context, and the most methodologically challenging, is the indeterminate cytological tier. The incidence of malignancy in Bethesda III/IV lesions on final histopathology was 35.3% in our cohort, which was at the higher end of the international range of 15-30%.[6,7] for Bethesda IV nodules. It could be a referral bias of surgically managed cases in our tertiary care centre and emphasises the need for a strong risk stratification in this group.

In the indeterminate cytology subset, where the AUC for TI-RADS is 0.81, the discriminatory performance of these new technology devices puts a very strong recommendation in favor of embracing sonographic risk stratification as a complementary, not competing, modality in diagnostic work-up of indeterminate cytology. All nodule/suspicious on ultrasound should undergo ultrasound-guided FNAC and in addition when molecular testing is not available, TR4 or TR5 in indeterminate cytology should be sufficient to warrant diagnostic hemithyroidectomy.[5,15,17]

Discordance: Clinical Implications

The discovery of nodule size > 4 cm and Hashimoto's thyroiditis as independent risk factors for FNAC–HPE discordance is of practical application. Large nodules can be subject to heterogeneous sampling where there is no representation of areas of malignant transformation, especially follicular variants of PTC, in a lesion that is otherwise benign appearing on colloid goiter.[18] Chronic lymphocytic thyroiditis may result in false-positive and false-negative cytologic impressions due to the presence of chronic cytologic changes that can mimic or may mask papillary carcinoma.[16]

Repeat USG-FNAC, molecular risk stratification (if available) and discussion with colleagues before definitive management planning, should be done with special emphasis in clinical situations where a palpable large nodule or one developing on a background of known Hashimoto's thyroiditis is encountered. A flexible nasendoscopy should routinely be performed to assess the voice prior to the surgical resection of any nodule that is associated with a change in voice, to identify any pre-existing laryngeal nerve palsy that may suggest the presence of invasive malignancy or to assist with intraoperative laryngeal nerve monitoring.[3]

Implications for Surgical Decision Making

In summary, the results support an algorithm for surgical care that is risk stratified. Conservative surveillance with imaging interval of 12–24 months is suitable for Bethesda II nodules with low grade TI-RADS (TR1–3). In respect to the risk of malignancy, it is recommended that the decision should be made based on the TI-RADS risk tier, with TR4–TR5 advised for diagnostic hemithyroidectomy; in the case of indeterminate cytology. When Bethesda III/IV and TR1–TR3 features, repeat FNAC at 3-6 months or molecular testing (where available) is an option to avoid surgical treatment, and thus decrease the percentage of may have been unnecessary benign thyroidectomy. Surgical intervention is indicated regardless of cytological categorisation for any one of the high-risk clinical features: rapid growth; fixation; regional lymphadenopathy; or, unexplained hoarseness.[1,3,12]

Moving forward, ongoing prospective studies in this department will concentrate on combining validated clinical risk scoring systems with existing sonographic and cytological protocols in order to decrease any further the percentage of benign thyroidectomies which currently accounts for a substantial portion of institutional operative volume.[19,20]

Limitations

There are many caveats to be noted. First, the study has been performed in one tertiary referral center where the surgical population was necessarily selected and hence, the level of malignancy (19.0%) might be higher than the frequency of malignancy in the general population of thyroid nodules. Second, molecular testing—such as gene expression classifiers, next-generation sequencing panels–was not widely available, and the accuracy of molecular triage could not be assessed in indeterminate cases. Thirdly, the inter-observer variation in the cytopathological reports was reduced to the lowest extent, by using a single cytopathologist but without formal assessment.

Conclusion

FNAC, as reported in the Bethesda System 2017, is a reliable, readily available and cost effective primary screening method for thyroid swellings in a tertiary head and neck surgical unit, this study shows. Selective rather than universal thyroidectomy is supported by the current evidence based paradigm, due to a substantial FNAC–histopathology agreement (κ = 0.76), high NPV for benign cytology (97.8%) and high specificity for malignant cytological categories (92.1%).

This cohort, however, also had an indeterminate cytological category (Bethesda III/IV) and in this group there was a 35.3% malignancy rate that cannot be diagnosed with FNA alone. The addition of ACR TI-RADS sonographic risk stratification is a helpful and meaningful addition in this challenging subgroup (AUC = 0.81). Clinically important risk factors that cause diagnostic discordance are large nodule size (> 4cm) and Hashimoto's thyroiditis and require increased suspicion.

Based on clinical assessment, ACR TI-RADS ultrasonography and Bethesda classified FNAC, a risk-stratified multi-disciplinary diagnostic algorithm is recommended to achieve maximum diagnostic accuracy, limit thyroidectomy and time cancer surgery when needed for clinically significant thyroid

malignancy. This future integration of molecular tests into institutional workflow may be especially important for supporting indeterminate cytology cases and further decrease the surgical rate in the benign thyroid nodule population.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid. 2016;26(1):1–133.
2. Gharib H, Papini E, Paschke R, et al. AACE/AME/ETA Task Force on Thyroid Nodules: Medical Guidelines for Clinical Practice for the Diagnosis and Management of Thyroid Nodules. Endocr Pract. 2016;22(Suppl 1):1–60.
3. Patel KN, Yip L, Lubitz CC, et al. The American Association of Endocrine Surgeons guidelines for the definitive surgical management of thyroid disease in adults. Ann Surg. 2020;271(3):e21–e93.
4. Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. Thyroid. 2017;27(11):1341–1346.
5. 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.
6. Nikiforov YE, Seethala RR, Tallini G, et al. Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma. JAMA Oncol. 2016;2(8):1023–1029.
7. Ohori NP, Nikiforova MN, Schoedel KE, et al. Contribution of molecular testing to thyroid fine-needle aspiration cytology of 'follicular lesion of undetermined significance/atypia of undetermined significance'. Cancer Cytopathol. 2010;118(1):17–23.
8. Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. N Engl J Med. 2012;367(8):705–715.
9. Yassa L, Cibas ES, Benson CB, et al. Long-term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. Cancer. 2007;111(6):508–516.
10. Brito JP, Gionfriddo MR, Al Nofal A, et al. The accuracy of thyroid nodule ultrasound to predict thyroid cancer: systematic review and meta-analysis. J Clin Endocrinol Metab. 2014;99(4):1253–1263.
11. Rago T, Vitti P. Role of thyroid ultrasound in the diagnostic evaluation of thyroid nodules. Best Pract Res Clin Endocrinol Metab. 2008;22(6):913–928.
12. WHO Classification of Tumours Editorial Board. Endocrine and Neuroendocrine Tumours. 5th ed. Lyon: International Agency for Research on Cancer; 2022.
13. Baloch ZW, LiVolsi VA, Asa SL, et al. Diagnostic terminology and morphologic criteria for cytologic diagnosis of thyroid lesions: a synopsis of the National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. Diagn Cytopathol. 2008;36(6):425–437.
14. Durante C, Grani G, Lamartina L, et al. The diagnosis and management of thyroid nodules: a review. JAMA. 2018;319(9):914–924.
15. Russ G, Bonnema SJ, Erdogan MF, et al. European Thyroid Association Guidelines for Ultrasound Malignancy Risk Stratification of Thyroid Nodules in Adults: The EU-TIRADS. Eur Thyroid J. 2017;6(5):225–237.
16. Pappa T, Alevizaki M. Endogenous TSH as a tumor-promoting factor in thyroid cancer: present and future perspectives. Endocr Relat Cancer. 2021;28(1):R1–R15.
17. Lee JH, Shin JH. Overdiagnosis and imaging detection of thyroid cancer in Korea: trends, controversies, implications. Korean J Radiol. 2014;15(6):663–671.
18. Trimboli P, Crescenzi A, Giovanella L. Ultrasound risk stratification for thyroid malignancy: systematic review and meta-analysis. Endocrine. 2016;51(1):52–60.
19. Singer PA, Cooper DS, Daniels GH, et al. Treatment guidelines for patients with thyroid nodules and well-differentiated thyroid cancer. Arch Intern Med. 1996;156(19):2165–2172.
20. 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.