

Content available at: <https://www.ipinnovative.com/open-access-journals>

Indian Journal of Pathology and Oncology

Journal homepage: www.ijpo.co.in

Original Research Article

Reproducibility of the Bethesda system for reporting thyroid cytopathology (TBSRTC): An observational study of 100 patients

Reena Naik¹, Kishori Moni Panda^{2*}, Mohd Ghouse Mohiddin²¹Government Medical College, Raigarh, Chhattishgarh, India²Dept. of Pathology, TRR Institute of Medical Sciences, Hyderabad, Telangana, India

ARTICLE INFO

Article history:

Received 22-10-2024

Accepted 02-12-2024

Available online 12-12-2024

Keywords:

Thyroid

FNAC

TBSRTC

Reproducibility

ABSTRACT

Background: FNAC is the most commonly used screening test to classify patients with thyroid nodules into six Bethesda categories (category I-VI) and for therapeutic decision making. The 2023 edition of TBSRTC, further simplifies these categories, using single name and describing more clear diagnostic criteria for each of the categories. The aim of this study is to investigate the interobserver reproducibility and usefulness of this classification scheme.

Materials and Methods: FNAC smears from 100 patients with thyroid nodules from october 2022 to September 2024, over a period of 2 years were analysed retrospectively, which were reviewed and reclassified by two pathologists independently according to 2023 TBSRTC. The interobserver reproducibility was statistically assessed using Cohen's kappa.

Results: Interobserver agreement was seen in 95 cases (95%) and disagreement was seen in 5 cases (5%). The concordance rate was 100% in the Bethesda I and Bethesda VI, 95% (76 cases) in Bethesda II, 66% (02 cases) in Bethesda III, 66% (08 cases) in Bethesda IV and 66% (02 cases) in Bethesda V. So, interobserver variation was seen among Bethesda II, III, IV and V. Overall Cohen's kappa value showed excellent agreement.

Conclusion: TBSRTC should be encouraged in our country because it reduces interobserver variability. Although there is excellent interobserver agreement seen in this study, disagreements were seen in the Bethesda categories II, III, IV and V, which corroborated with the findings of the studies done elsewhere.

This is an Open Access (OA) journal, and articles are distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License](https://creativecommons.org/licenses/by-nc-sa/4.0/), which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprint@ipinnovative.com

1. Introduction

"The Bethesda System for Reporting Thyroid Cytopathology" (TBSRTC) classifies patients with thyroid nodules into six Bethesda categories (category I-VI) according to risk of malignancy (ROM) and for therapeutic decision making (e.g. active surveillance, surgery, radioactive iodine, targeted therapy). TBSRTC not only gives the definitions, cytomorphologic criteria, explanatory notes, but also mentions clinical management plan for each of the six diagnostic categories including

the role of molecular testing.^{1,2} The 2023 edition of the TBSRTC, has clarified the diagnostic criteria and uses a single name for each of the six categories, i.e. non-diagnostic (ND); benign (B); atypia of undetermined significance (AUS); follicular neoplasm (FN), suspicious for malignancy (SM), and malignant (M).³ Statistical validity of FNAC in patients with thyroid nodules has been very high including diagnostic accuracy, which reaches almost near to 100% in thyroid malignancy.⁴⁻⁶ However, the indeterminate categories (categories III-V) may require molecular testing to decide the mode of treatment, i.e. surgery versus conservative. But some of the new borderline 'low-risk' entities are 'molecular

* Corresponding author.

E-mail address: drkishoripanda@gmail.com (K. M. Panda).

indeterminate' and molecular testing is not available at many institutions.⁷ So, morphologic evaluation of FNAC smears continue to remain as the gold standard screening test for assessment of patients with thyroid nodules. However, significant interobserver variabilities exist for indeterminate categories due to the subjective morphological interpretation. The authors aim to investigate the interobserver reproducibility and usefulness of 2023 TBSRTC classification scheme, which can help in the daily routine diagnosis and management of thyroid lesions.

2. Materials and Methods

FNAC smears from patients with thyroid nodules from October 2022 to September 2024, over a period of 2 years were analysed retrospectively in this observational study. Wright Giemsa and Leishman-stained smears and relevant clinical and radiological details of 100 patients were retrieved from cytopathology records. Then two pathologists reviewed and reclassified the cases according to 2023 edition of TBSRTC. FNAC smears with well-defined cellular morphology were included in the study, where as those smears with extensive cellular overlapping, unclear morphology and degenerative changes were excluded.

The reclassification of FNAC smears was done based on the diagnostic criteria described in TBSRTC, 2023 and categorized using a single name, i.e. non-diagnostic (ND); benign (B); atypia of undetermined significance (AUS); follicular neoplasm (FN), suspicious for malignancy (SM), and malignant (M) (Figure 1 A-F).³ A category was assigned to each case independently by each reviewer and interobserver variability was calculated as the percentage of agreement between the results of the two pathologists. Cross tabulation and Cohen's kappa was used to estimate the degree of reproducibility and to confirm statistically significant agreement.⁸ The interpretation of Cohen's kappa value i.e. percent agreement as follows: 0–0.2 shows poor agreement, 0.3–0.4 shows fair agreement, 0.5–0.6 shows moderate agreement, 0.7–0.8 shows strong agreement, and >0.8 shows excellent agreement.

As it was an observational study, the patient's identity was not disclosed and institutional ethical clearance was not obtained.

3. Results

Demographic evaluation of 100 patients with thyroid nodules revealed 82 females (82%) and 18 males (18%), with a female: male ratio of 4.5:1. The age distribution ranged from 5 to 70 years, with a mean of 38.5 years. The most commonly affected age group was 21–50 years.

In the present study, interobserver agreement was seen in 95 out of 100 cases (95%) and disagreement was seen in 5 cases (5%). The concordance rate was 100% in the ND and 'M' category, where as in the benign category (Figure 1

A, B) interobserver agreement was 95% (76 cases). The agreement rate was lower in the AUS (02 cases: 66%), FN (08 cases: 66%), and SM categories (02 cases; 66%). So, interobserver variation was seen among Bethesda category II, III, IV and V. (Table 1). There were four cytologically discordant cases in the 'B' category, which were previously diagnosed as hyperplastic nodule in nodular goiter by first observer and were put into FN category by the second observer. Disagreement was also seen in one case in the AUS category by the first observer, which was put into SM category by the second observer. In this case, there was higher suspicion of malignancy than AUS but lower suspicion than malignant. In Bethesda category VI (M), cytomorphologic features in all 06 cases were consistent with papillary thyroid carcinoma (PTC). (Figure 1E, F) Percent agreement using Cohen's kappa was calculated to confirm an agreement beyond chance, which revealed "excellent agreement" (86.7%; 0.867).

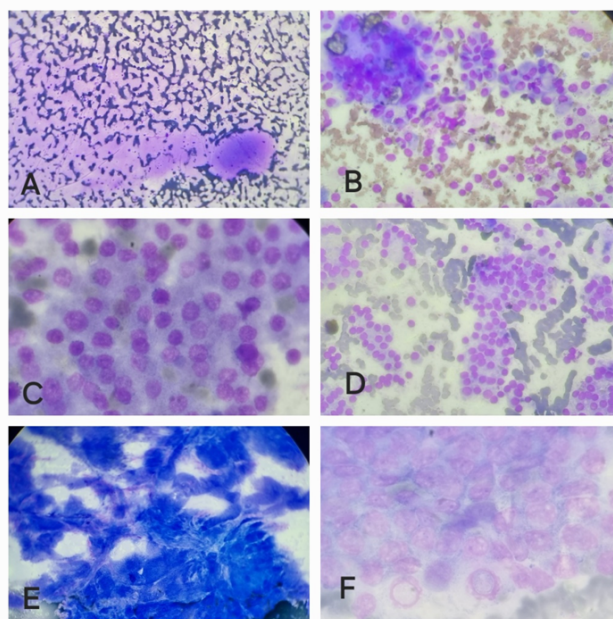


Figure 1: A): 10x, paucicellular smear with abundant colloid c/w colloid goiter; B): 10x, Cellular smear with thyroid epithelial cell clusters including Hurthle cells, being infiltrated by lymphocytes c/w Hashimoto's thyroiditis; C): 40x, Thyroid follicular cells showing mild nuclear atypia c/w AUS; D): 10x, Cellular smear with predominant microfollicular pattern c/w FN; E): 40x, Multiple papillae with nuclear atypia c/w PTC; F): 100x, Intracellular inclusion of PTC

In summary, there was interobserver agreement in 95 out of 100 cases (95%). The interobserver agreement was highest in benign (B) category (76 out of 80 cases; 95%), followed by 08 cases in follicular neoplasm (FN), 06 cases in malignant (M), 02 cases each in atypia of undetermined significance (AUS) and suspicious for malignancy (SM).

Table 1: Comparison of reclassified lesions according to TBSRTC between observers as agreement versus disagreement (observer 1- horizontal and observer 2-vertical)

Category	ND	B	AUS	FN	SM	M	Total
ND	01	-	-	-	-	-	01
B	-	76	-	04	-	-	80
AUS	-	-	02	-	01	-	03
FN	-	-	-	08	-	-	08
SM	-	-	-	-	02	-	02
M	-	-	-	-	-	06	06
Total	01	76	02	12	03	06	100

ND: Nondiagnostic; B: Benign; AUS: Atypia of undetermined significance; FN: Follicular neoplasia; SM: Suspicious of malignancy; M: Malignant

4. Discussion

The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) is an international reporting system, which not only standardizes the reporting of thyroid FNA but also improves communication between pathologists and surgeons. In order to match with the 2022 WHO Thyroid Tumor Classification, the 2023 TBSRTC diagnostic criteria is simplified and additional terminology is added. Instead of multiple names, which is confusing, each category is assigned a single name as follows: Bethesda I-nondiagnostic (ND); Bethesda II- benign (B); Bethesda III-atypia of undetermined significance (AUS); Bethesda IV- follicular neoplasm (FN); Bethesda V- suspicious for malignancy (SM); and Bethesda VI- malignant (M). The implied risk of malignancy (ROM) for each category has been updated, which is based on data reported after 2017 edition of TBSRTC (2017).⁹ However, indeterminate categories continue to exist in the 2023 edition of the reporting system.³ In this context, excellent interobserver reproducibility with Cohen's kappa (0.867), observed in this study is particularly valuable as clinical relevance of any categorization scheme requires diagnostic accuracy along with excellent interobserver reproducibility. So, "excellent agreement" (86.7%; 0.867) seen in this study is clinically relevant and in corroboration with findings of other studies in the literature.^{8,10–13} Cohen's kappa value range in these studies was 0.61-0.99, which is comparable to our kappa value (Table 2).

However, diagnostic disparity was found mostly in the Bethesda II category with four discrepant cases, where first observer diagnosed it as 'B' category and the second observer diagnosed it as FN category because of the presence of high cellularity and the microfollicular pattern (Figure 1 D). Similar disagreement has been observed in earlier studies.⁸ So, a clear distinction between hyperplastic nodule in multinodular goiter and follicular adenoma is not possible by FNA and does not have much clinical significance.^{14,15}

Table 2: Comparison of interobserver reproducibility among various studies

Study	Interobserver reproducibility
Awasthi et al. ⁸	Good agreement (Cohen's kappa score 0.613)
Ahmed et al. ¹¹	Strong agreement (Cohen's kappa score 0.735)
Anand et al. ¹³	Excellent agreement (Cohen's kappa score 0.99)
Our study	Excellent agreement (Cohen's kappa score 0.867)

Interobserver disagreement was also seen in one case of Bethesda III, where first observer diagnosed it as AUS with nuclear atypia and second observer diagnosed as SM category. This was due to low cellularity and focal papillary-like nuclear features. (Figure 1 C). Cytologic features were strongly suspicious of malignancy but were not sufficient for a conclusive diagnosis. The purpose of separating Bethesda V from Bethesda VI is to preserve the very high positive predictive value of the malignant category without compromising the overall sensitivity of FNAC. However, the mode of treatment for both Bethesda V and Bethesda VI is near total thyroidectomy. So, the distinction between Bethesda V and Bethesda VI does not have much clinical significance.⁴

TBSRTC has proven to be a highly efficient diagnostic tool for reporting thyroid lesions. The six tiers of TBSRTC is not only contributes to a solid stratification and management, but also is helpful for pathologists faced with challenging cases, allowing them to approach the case considering the clinical outcome first, rather than focusing solely on labeling a diagnosis.⁷

The 2023 edition of TBSRTC, compared to previous editions, results in reducing the number of indeterminate categories, which is probably due to the introduction of more strict diagnostic criteria. In addition, TBSRTC provides a common language for communication and clarity, reduces the number of inappropriate surgery in benign cases, and enables to perform the operation on time in patients with malignant lesions. It also provides a simple and reliable exchange of data between different institutions throughout the world.^{16,17}

Although the interobserver reliability observed in this study is excellent, a few pitfalls were noted in the present study. The overlapping cytological features of hyperplastic nodule and follicular adenoma led to interobserver variability. Lastly, for the AUS and SM category, the focal nuclear atypia may be easily missed by inexperienced pathologist. However, differentiation of nuclear atypia in AUS and SM categories could be improved by molecular testing.

5. Conclusion

To conclude, TBSRTC should be encouraged in our country because it reduces interobserver variability. Although there is excellent interobserver agreement seen in this study, disagreements were seen in the Bethesda categories II, III, IV and V, which corroborated with the findings of the studies done elsewhere. By incorporating novel molecular data into cytology classification scheme could help in improving morphologic diagnosis in indeterminate categories and preventing unnecessary surgery

6. Source of Funding

Nil.

7. Conflicts of Interest

There are no conflicts of interest.

References

1. Ali SZ, Cibas ES. The Bethesda system for reporting thyroid cytopathology: definitions, criteria and explanatory notes. New York: Springer Science & Business Media; 2010.
2. Cibas ES, Ali SZ. The 2017 Bethesda System for reporting thyroid cytopathology. *Thyroid*. 2017;27:1341–6.
3. Ali SZ, Vanderlaan PA. The Bethesda System for Reporting Thyroid Cytopathology. Definitions, Criteria, and Explanatory Notes. 3rd ed. Switzerland: Springer; 2023.
4. Lebruna L, Salmons I. Pathology and new insights in thyroid neoplasms in the 2022 WHO classification. *Curr Opin Oncol*. 2024;36(1):13–21.
5. Leonard N, Melcher DH. To operate or not to operate? The value of fine needle aspiration cytology in the assessment of thyroid swellings. *J Clin Pathol*. 1997;50:941–3.
6. Bakhos R, Selvaggi SM, DeJong S, Gordon DL, Pitale SU, Herrmann M, et al. Fine-needle aspiration of the thyroid: rate and causes of cytohistopathologic discordance. *Diagn Cytopathol*. 2000;23(4):233–7.
7. Luis R, Thirunavukkarasu B, Jain D, Canberk S. Welcoming the new, revisiting the old: A brief glance at cytopathology reporting systems for lung, pancreas, and thyroid. *J Pathol Transl Med*. 2024;58(4):165–73.
8. Awasthi P, Goel G, Khurana U, Joshi D, Majumdar K, Kapoor N. Reproducibility of "The Bethesda System for Reporting Thyroid Cytopathology": A Retrospective Analysis of 107 Patients. *J Cytol*. 2018;35(1):33–6.
9. Ali SZ, Baloch ZW, Cochand-Priollet B, Schmitt FC, Vielh P, Vanderlaan PA. The 2023 Bethesda System for reporting thyroid cytopathology. *J Am Soc Cytopathol*. 2023;12(5):319–25.
10. Bhasin TS, Mannan R, Manjari M, Mehra M, Sekhon AKG, Chandey M, et al. Reproducibility of 'The Bethesda System for reporting Thyroid Cytopathology': A Multi Center Study with Review of the Literature. *J Clin Diagn Res*. 2013;7(6):1051–4.
11. Ahmed S, Ahmad M, Khan MA, Kazi F, Noreen F, Nawaz S, et al. The interobserver reproducibility of thyroid cytopathology using Bethesda Reporting System: Analysis of 200 cases. *J Pak Med Assoc*. 2013;63(10):1252–4.
12. Cochand-Priollet B, Schmitt FC, Totsch M, Vielh P. The Bethesda terminology for reporting thyroid cytopathology: From theory to practice in Europe. *Acta Cytol*. 2011;55:507–11.
13. Anand B, Ramdas A, Ambroise MM, Kumar NP. The Bethesda System for Reporting Thyroid Cytopathology: A Cytohistological Study. *J Thyroid Res*. 2020;2020:8095378.
14. Somma J, Schlecht NF, Fink D, Khader SN, Smith RV, Cajigas A. Thyroid fine needle aspiration cytology: follicular lesions and the gray zone. *Acta Cytol*. 2010;54(2):123–31.
15. Cibas ES, Ali SZ. The Bethesda System for reporting thyroid cytopathology. *Am J Clin Pathol*. 2009;132(5):658–65.
16. Guerreiro SC, Tastekin E, Mourao M, Loureiro I, Eusebio R, Marques HP, et al. Impact of the 3rd Edition of the Bethesda System for Reporting Thyroid Cytopathology on Grey Zone Categories. *Acta Cytol*. 2023;67(6):593–603.
17. Wesola M, Jeleń M. The Bethesda System in the evaluation of thyroid nodules: Review. *Adv Clin Exp Med*. 2017;26(1):177–82.

Author's biography

Reena Naik, Associate Professor

Kishori Moni Panda, Professor  <https://orcid.org/0000-0002-4476-3830>

Mohd Ghouse Mohiddin, Professor and Head

Cite this article: Naik R, Panda KM, Mohiddin MG. Reproducibility of the Bethesda system for reporting thyroid cytopathology (TBSRTC): An observational study of 100 patients. *Indian J Pathol Oncol* 2024;11(4):385–388.