



RESEARCH ARTICLE

Distribution of Bethesda Categories and Cytohistological Concordance in Thyroid Fine-Needle Aspiration: A Retrospective Study Design

Vivek Tukaram Bonde* 

¹Assistant Professor, Department of Pathology, IQ City Medical College, Durgapur

Abstract

Background: The Bethesda System for Reporting Thyroid Cytopathology assigns each thyroid aspirate to one of six categories, each with an implied risk of malignancy. Local audits of category distribution and histological outcome are needed to validate reporting practice. **Objectives:** To describe the distribution of Bethesda categories in thyroid fine-needle aspiration (FNA), to estimate the risk of malignancy for each category, and to assess the diagnostic accuracy of FNA against histopathology. **Materials and Methods:** This illustrative retrospective analysis uses a simulated dataset of 320 thyroid aspirates reported according to the 2017 Bethesda System. Histopathology was available for 96 operated cases. Risk of malignancy was calculated with Wilson 95% confidence intervals, and categories V and VI were treated as test-positive and category II as test-negative. **Results:** Women formed 81.9% of the cohort (female-to-male ratio 4.5:1). Category II (benign) was the most frequent (66.9%); the non-diagnostic rate was 6.9%, and indeterminate categories III–V accounted for 17.5%. The risk of malignancy rose from 3.3% in category II to 96.2% in category VI (trend $p < 0.001$). FNA showed a sensitivity of 97.1%, specificity of 87.9% and overall accuracy of 92.6%. **Conclusion:** In this simulated analysis, the Bethesda System stratified malignancy risk effectively and FNA showed high diagnostic accuracy. These estimates illustrate the method and require confirmation with real institutional data.

Keywords: Thyroid nodule, Fine-needle aspiration, Bethesda System, Cytopathology, Risk of malignancy, Histopathology

Copyright : © 2021 The Authors. Published by Publisher. This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

1 | INTRODUCTION

Thyroid nodules are common. They are palpable in a minority of adults but are detected far more often by high-resolution ultrasonography, and only a small proportion are malignant [1]. In India, thyroid disorders affect a large population and goitre remains an important clinical problem [2]. Thyroid cancer accounted for more than half a million new cases worldwide in 2018, with a markedly higher incidence in women [3]. The diagnostic challenge is to identify the minority of malignant nodules without operating on the large majority that are benign.

Fine-needle aspiration (FNA) cytology is the most accurate and cost-effective first-line test for this purpose and has reduced the number of unnecessary thyroidectomies [4]. Before 2007, however, terminology for thyroid cytology varied between laboratories, which hampered communication with clinicians and comparison of results. The National Cancer Institute Thyroid FNA State of the Science Conference produced uniform diagnostic terminology and morphological criteria [5]. These were published as the Bethesda System for Reporting Thyroid Cytopathology [6].

The Bethesda System places each aspirate in one of six categories: non-diagnostic (I), benign (II), atypia or follicular lesion of undetermined significance (III), follicular neoplasm or suspicious for a follicular neoplasm (IV), suspicious for malignancy (V), and malignant (VI). Each category carries an implied risk of malignancy and a recommended clinical action [6]. The 2017 revision updated these risk estimates and refined the terminology [7,8]. Management guidelines of the American Thyroid Association are built around the same categories [9], and a meta-analysis confirmed the system's value in risk stratification [10].

The published risks are reference values, and both the 2017 Bethesda revision and later studies recommend that each laboratory audit its own category frequencies and malignancy risks [7,8]. Such audits are especially relevant for the indeterminate categories, whose use and outcome vary between institutions.

This study design aims to describe the distribution of Bethesda categories in thyroid FNA, to estimate the risk of malignancy for each category by correlation with histopathology, and to assess the diagnostic accuracy of FNA.

2 | MATERIALS AND METHODS

Study design and setting. The protocol describes a retrospective record-based study in the cytopathology and histopathology sections of a department of pathology. The institution and study period are left as placeholders to be completed with verified information.

Study population. All thyroid FNAs reported during the study period would be included. Repeat aspirates from the same nodule would be counted once, using the most recent diagnostic report. Aspirates of non-thyroidal neck lesions and cases with incomplete records would be excluded. Histopathology would be available only for patients who underwent thyroid surgery.

Specimen collection and processing. FNA would be performed with a 23–25-gauge needle under palpation or ultrasound guidance, with two to four passes per nodule. Smears would be air-dried for May–Grünwald–Giemsa staining and fixed in 95% ethanol for Papanicolaou staining. Adequacy would require at least six groups of well-visualised follicular cells, each containing at least ten cells, with the exceptions defined in the Bethesda System [8].

Diagnostic classification. Aspirates would be reported according to the 2017 Bethesda System [7,8]. Thyroidectomy and lobectomy specimens would be processed routinely and diagnosed according to the 2017 World Health Organization classification of

Supplementary information The online version of this article (<https://doi.org/xx.xxx/xxx.xx>) contains supplementary material, which is available to authorized users.

Corresponding Author: *Vivek Tukaram Bonde*
Assistant Professor, Department of Pathology, IQ
City Medical College, Durgapur

endocrine tumours [11]. Non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) would be recorded separately [12]; none occurred in the simulated dataset.

Outcome variables and statistical analysis. The primary outcomes are the frequency of each category and the risk of malignancy, defined as the proportion of operated cases in a category with a malignant histological diagnosis. Proportions are given with Wilson 95% confidence intervals (CI). Associations were tested with the chi-square or Fisher exact test, and continuous variables were compared with the Mann–Whitney U test. The trend in risk of malignancy from category II to VI was tested with the Cochran–Armitage test. For diagnostic accuracy, categories V and VI were treated as positive and category II as negative; categories I, III and IV were excluded from this calculation. A two-sided $p < 0.05$ was considered significant. The analysis was performed in Python (SciPy).

Ethical considerations. No human data were used for this illustrative manuscript. Before any real study, approval or a waiver of consent from the Institutional Ethics Committee would be required, and records would be anonymised.

3 | RESULTS

The dataset included 320 thyroid aspirates from 262 women (81.9%) and 58 men (18.1%), a female-to-male ratio of 4.5:1. Table 1 summarizes the baseline characteristics. Mean age was 42.7 ± 12.3 years, and most patients were 21–60 years old. A solitary nodule was the commonest presentation (197; 61.6%). Median nodule size was 2.6 cm (interquartile range 1.8–3.3 cm). Age, presentation and nodule size did not differ significantly by sex.

Shows the distribution of Bethesda categories. Category II (benign) was the most frequent (214; 66.9%), followed by category VI (malignant; 8.8%) and category III (7.5%). The non-diagnostic rate was 6.9%. Indeterminate categories III, IV and V together accounted for 56 aspirates (17.5%). The ratio of category III to malignant categories (V + VI) was 0.57. Figure 1 illustrates the distribution.

CI, Wilson confidence interval; SD, standard deviation. Categories follow the 2017 Bethesda System for Reporting Thyroid Cytopathology [7].

A malignant cytological category (V or VI) was more frequent in solitary nodules (34/197; 17.3%) than in multinodular goitre (8/123; 6.5%; $\chi^2 = 6.77$, $p = 0.009$). It was more frequent in men (20.7%) than in women (11.5%), but this difference was not statistically significant (Fisher exact $p = 0.083$).

Histopathology was available for 96 cases (30.0%). The proportion operated rose with category, from 14.0% of benign aspirates to 92.9% of malignant aspirates (Table 3). Of the operated cases, 43 were malignant, most commonly papillary thyroid carcinoma (38). The risk of malignancy was 25.0% in category I, 3.3% in II, 30.0% in III, 28.6% in IV, 75.0% in V and 96.2% in VI, with a significant increasing trend from category II to VI (Cochran–Armitage $z = 7.27$, $p < 0.001$). These results are illustrated in Figure 2.

NH, nodular hyperplasia; FA, follicular adenoma; LT, lymphocytic thyroiditis; PTC, papillary thyroid carcinoma; FTC, follicular thyroid carcinoma; ROM, risk of malignancy; CI, Wilson confidence interval.

With categories V and VI as positive and category II as negative (68 cases), FNA gave 34 true-positive, 4 false-positive, 1 false-negative and 29 true-negative results. Sensitivity was 97.1% (95% CI 85.5–99.5), specificity 87.9% (72.7–95.2), positive predictive value 89.5% (75.9–95.8), negative predictive value 96.7% (83.3–99.4) and overall accuracy 92.6% (83.9–96.8).

4 | DISCUSSION

This illustrative analysis applies the 2017 Bethesda System to a simulated series of thyroid aspirates and correlates the categories with histopathology. The design reproduces the main elements of a laboratory audit: the frequency of each category, the risk of malignancy, and the diagnostic accuracy of FNA.

In the simulated dataset, benign aspirates formed two-thirds of all reports, in keeping with the predominantly benign nature of thyroid nodules [1]. The non-diagnostic rate of 6.9% and the category

DISTRIBUTION OF BETHESDA CATEGORIES AND CYTOHISTOLOGICAL CONCORDANCE IN THYROID FINE-NEEDLE ASPIRATION

TABLE 1: Baseline characteristics of the study population by sex (N = 320)

Characteristic	Category	Female (n = 262)	Male (n = 58)	Total (N = 320)	p value
Age, mean $\pm$ SD (years)		42.6 $\pm$ 12.4	43.0 $\pm$ 11.9	42.7 $\pm$ 12.3	0.668†
Age group (years)	≤ 20	8 (3.1)	1 (1.7)	9 (2.8)	0.396‡
	21–40	106 (40.5)	19 (32.8)	125 (39.1)	
	41–60	128 (48.9)	34 (58.6)	162 (50.6)	
	>60	20 (7.6)	4 (6.9)	24 (7.5)	
Clinical presentation	Solitary nodule	159 (60.7)	38 (65.5)	197 (61.6)	0.593§
	Multinodular goitre	103 (39.3)	20 (34.5)	123 (38.4)	
Nodule size on ultrasound (cm)	<2.0	73 (27.9)	16 (27.6)	89 (27.8)	0.975§
	2.0–4.0	173 (66.0)	38 (65.5)	211 (65.9)	
	>4.0	16 (6.1)	4 (6.9)	20 (6.2)	
Total		262 (81.9)	58 (18.1)	320 (100.0)	

Values are n (%) unless stated otherwise; percentages are column percentages. SD, standard deviation.

†Mann–Whitney U test. ‡Chi-square test, with the ≤ 20 and 21–40 year groups combined because of small expected counts. §Chi-square test.

TABLE 2: Distribution of thyroid aspirates by Bethesda category

Category	Diagnostic category	n (%)	95% CI (%)	Age, mean $\pm$ SD (years)	Female, n
I	Non-diagnostic or unsatisfactory	22 (6.9)	4.6–10.2	45.5 $\pm$ 14.3	18
II	Benign	214 (66.9)	61.5–71.8	42.2 $\pm$ 11.7	179
III	Atypia/follicular lesion of undetermined significance	24 (7.5)	5.1–10.9	40.9 $\pm$ 13.4	20
IV	Follicular neoplasm or suspicious for a follicular neoplasm	18 (5.6)	3.6–8.7	39.1 $\pm$ 12.6	15
V	Suspicious for malignancy	14 (4.4)	2.6–7.2	50.4 $\pm$ 8.8	11
VI	Malignant	28 (8.8)	6.1–12.4	44.0 $\pm$ 14.1	19
Total		320 (100.0)		42.7 $\pm$ 12.3	262

III rate of 7.5% fall within the performance targets of the 2017 Bethesda revision, which recommends that atypia of undetermined significance should not exceed about 10% of thyroid FNA interpretations [7,8]. Ultrasound-guided sampling within a multidisciplinary programme of nodule evaluation has been associated with low non-diagnostic rates [13].

The risk of malignancy increased in a stepwise manner from benign to malignant categories, consistent with the risk stratification on which the system is built [6,7]. A meta-analysis [10] and large institutional and multicentre series [14,15] have described the same gradient, and the simulated estimates for categories IV (28.6%) and V (75.0%) lie within the 2017 reference ranges, while those for categories II

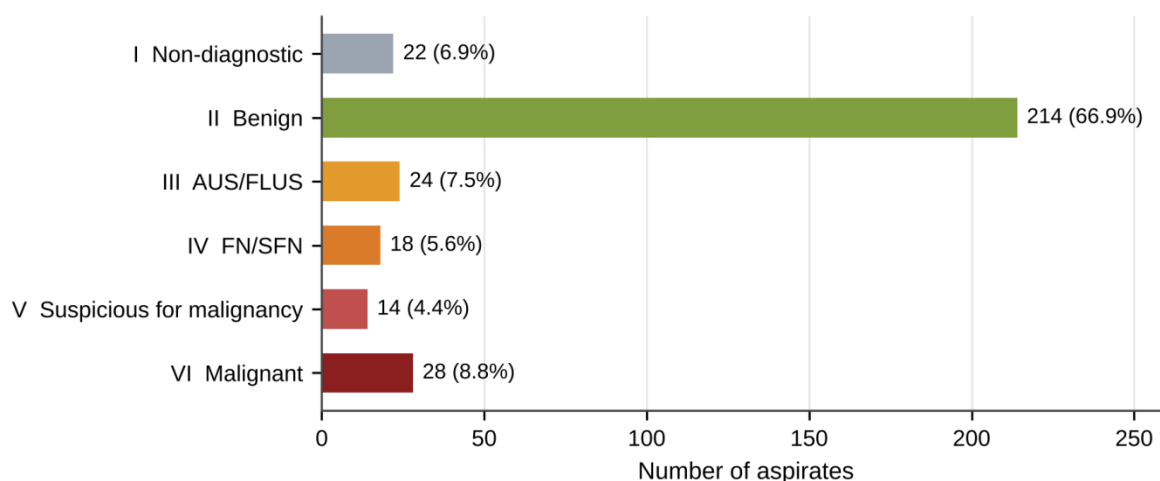


FIGURE 1: Distribution of thyroid fine-needle aspirates across Bethesda categories (N = 320).AUS/FLUS, atypia/follicular lesion of undetermined significance; FN/SFN, follicular neoplasm or suspicious for a follicular neoplasm.

TABLE 3: Cytohistological correlation and risk of malignancy by Bethesda category

Category	Operated, n (%)*	NH	FA	LT	PTC	FTC	Other†	Malignant, n	ROM, % (95% CI)
I	4 (18.2)	2	1	0	1	0	0	1	25.0 (4.6–69.9)
II	30 (14.0)	21	5	3	1	0	0	1	3.3 (0.6–16.7)
III	10 (41.7)	4	3	0	3	0	0	3	30.0 (10.8–60.3)
IV	14 (77.8)	2	8	0	1	3	0	4	28.6 (11.7–54.6)
V	12 (85.7)	1	1	1	9	0	0	9	75.0 (46.8–91.1)
VI	26 (92.9)	0	0	1	23	1	1	25	96.2 (81.1–99.3)
Total	96 (30.0)	30	18	5	38	4	1	43	44.8

*Percentage of all aspirates in the category. †Medullary thyroid carcinoma.

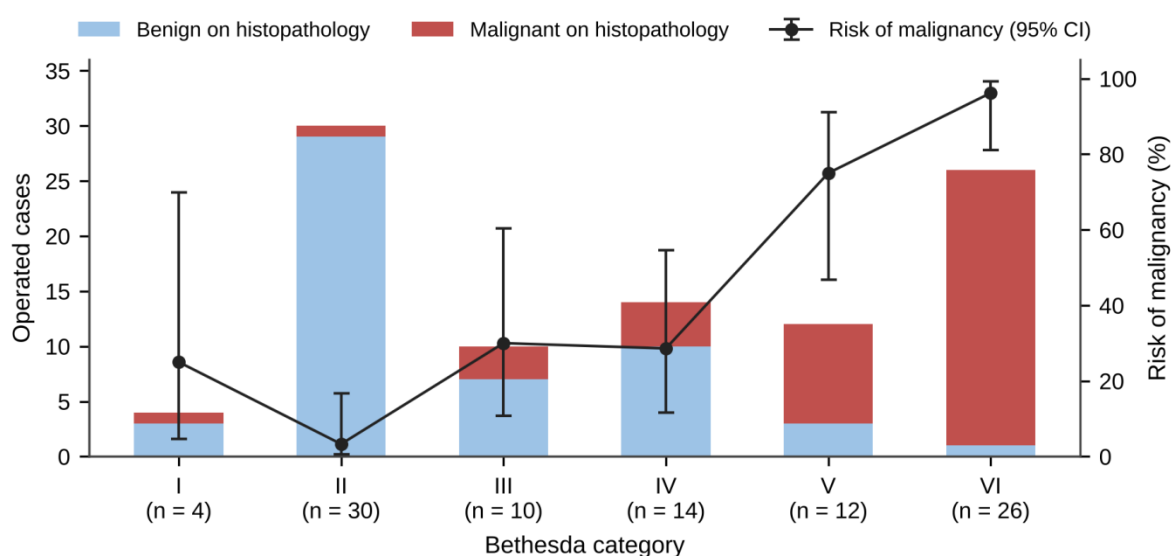


FIGURE 2: Histopathological outcome of operated cases (bars, left axis) and risk of malignancy with 95% confidence intervals (line, right axis) by Bethesda category (n = 96).

(3.3%) and VI (96.2%) lie just outside the reference ranges of 0–3% and 97–99% [7]. The high risk in category I (25.0%) rests on only 4 operated cases. The risk of 30.0% for category III is at the upper end of the published range. Higher-than-expected malignancy rates in this category have been reported, for example by Ho and colleagues [16], and are partly explained by selective surgery for nodules with suspicious clinical or ultrasound features. Category III therefore benefits most from local audit and, where available, molecular testing, as management guidelines recommend [9].

The high sensitivity (97.1%) and negative predictive value (96.7%) mirror the established role of FNA as a reliable rule-out test [4,10]. Specificity was lower (87.9%) because some category V aspirates proved benign on histology, a recognised consequence of overlapping nuclear features in follicular-patterned lesions. The reclassification of non-invasive encapsulated follicular variant tumours as NIFTP has shifted the expected risk in categories III to V [12]. A real study should report risk both with and without NIFTP counted as malignant [7]. Indian studies applying the Bethesda System have likewise shown female predominance and a large benign category, supporting its use in this setting [17].

Clinical implications. If confirmed with real data, these results would support routine use of the Bethesda System, conservative follow-up for category II, and surgical management for categories V and VI, with category III as the main target for repeat FNA, ultrasound risk assessment or molecular testing [9].

Strengths and limitations. The design uses standardized terminology, histological correlation and confidence intervals for every proportion. The principal limitation is that the data are simulated, so the estimates illustrate the method and are not evidence. A real retrospective study would also face verification bias, because only operated nodules have a histological diagnosis. This tends to inflate the risk of malignancy in categories where surgery is selective, particularly categories I to IV. Calculating the risk with all aspirates in the denominator, assuming unoperated cases are benign, would give a lower bound for comparison. Single-centre experience and inter-

observer variability in the indeterminate categories would further limit generalisation.

5 | CONCLUSION

This illustrative analysis outlines how a pathology department can audit thyroid FNA reporting under the 2017 Bethesda System. In the simulated dataset, benign aspirates predominated, the non-diagnostic and indeterminate rates fell within recommended targets, and the risk of malignancy increased progressively across categories. With categories V and VI as positive, FNA showed high sensitivity, negative predictive value and accuracy against histopathology. The expected clinical message is that the Bethesda System provides reliable risk stratification, with the indeterminate categories needing the closest attention. These findings must be confirmed with real, ethically approved institutional data before any conclusion is drawn. Future work should include risk estimates with and without NIFTP and should evaluate ultrasound risk scores and molecular markers in category III and IV nodules.

References

1. Dean DS, Gharib H. Epidemiology of thyroid nodules. *Best Pract Res Clin Endocrinol Metab.* 2008;22(6):901-11. doi:10.1016/j.beem.2008.09.019
2. Unnikrishnan AG, Menon UV. Thyroid disorders in India: an epidemiological perspective. *Indian J Endocrinol Metab.* 2011;15(Suppl 2):S78-81. doi:10.4103/2230-8210.83329
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394-424. doi:10.3322/caac.21492
4. Gharib H, Goellner JR. Fine-needle aspiration biopsy of the thyroid: an appraisal. *Ann Intern Med.* 1993;118(4):282-9. doi:10.7326/0003-4819-118-4-199302150-00007
5. Baloch ZW, LiVolsi VA, Asa SL, Rosai J, Merino MJ, Randolph G, et al. Diagnostic terminology and morphologic criteria for cytologic di-

agnosis of thyroid lesions: a synopsis of the National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. *Diagn Cytopathol.* 2008;36(6):425-37.

6. Cibas ES, Ali SZ. The Bethesda System for Reporting Thyroid Cytopathology. *Am J Clin Pathol.* 2009;132(5):658-65. doi:10.1309/AJCPHLWMI3JV4L

7. Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. *Thyroid.* 2017;27(11):1341-6. doi:10.1089/thy.2017.0500

8. Ali SZ, Cibas ES, editors. The Bethesda System for Reporting Thyroid Cytopathology: definitions, criteria, and explanatory notes. 2nd ed. Cham: Springer; 2018.

9. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid.* 2016;26(1):1-133. doi:10.1089/thy.2015.0020

10. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda System for Reporting Thyroid Cytopathology: a meta-analysis. *Acta Cytol.* 2012;56(4):333-9. doi:10.1159/000339959

11. Lloyd RV, Osamura RY, Klöppel G, Rosai J, editors. WHO classification of tumours of endocrine organs. 4th ed. Lyon: International Agency for Research on Cancer; 2017.

12. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LD, et al. Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma: a paradigm shift to reduce overtreatment of in-

dolent tumors. *JAMA Oncol.* 2016;2(8):1023-9. doi:10.1001/jamaoncol.2016.0386

13. Yassa L, Cibas ES, Benson CB, Frates MC, Doubilet PM, Gawande AA, et al. Long-term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. *Cancer.* 2007;111(6):508-16. doi:10.1002/cncr.23116

14. Jo VY, Stelow EB, Dustin SM, Hanley KZ. Malignancy risk for fine-needle aspiration of thyroid lesions according to the Bethesda System for Reporting Thyroid Cytopathology. *Am J Clin Pathol.* 2010;134(3):450-6. doi:10.1309/AJCP5N4MTHPAFXFB

15. Wang CC, Friedman L, Kennedy GC, Wang H, Kebebew E, Steward DL, et al. A large multicenter correlation study of thyroid nodule cytopathology and histopathology. *Thyroid.* 2011;21(3):243-51. doi:10.1089/thy.2010.0243

16. Ho AS, Sarti EE, Jain KS, Wang H, Nixon IJ, Shaha AR, et al. Malignancy rate in thyroid nodules classified as Bethesda category III (AUS/FLUS). *Thyroid.* 2014;24(5):832-9. doi:10.1089/thy.2013.0317

17. Mehra P, Verma AK. Thyroid cytopathology reporting by the Bethesda System: a two-year prospective study in an academic institution. *Patholog Res Int.* 2015;2015:240505. doi:10.1155/2015/240505

How to cite this article: Bonde V.T. **Distribution of Bethesda Categories and Cytohistological Concordance in Thyroid Fine-Needle Aspiration: A Retrospective Study Design.** *Innovative Journal of Medical and Health Science.* 2021;2312–2317. <https://doi.org/10.15520/ijmhs/11-4-15>