Indian J Otolaryngol Head Neck Surg
(October 2022) 74(Suppl 2):S2623–S2628; https://doi.org/10.1007/s12070-020-02335-5

ORIGINAL ARTICLE

Usefulness of The Bethesda System of Reporting Thyroid Cytopathology in Surgical Planning

Swetal Mahajan[1] · B. V. Manjula[1] · R. Vaishnavi[1] · Neena John[2] · Bhuvaneshwari Babu[1]

Received: 6 November 2020 / Accepted: 14 December 2020 / Published online: 7 January 2021

Abstract The Bethesda system for reporting thyroid cytopathology (TBSRTC) has attempted to standardize thyroid fine needle aspiration cytology (FNAC) reporting internationally into six diagnostic categories and help in clinical decision making. But there exists a significant variation in the reporting percentage and rates of malignancies in each category across the centres which complicates clinical decision making. To study the usefulness of TBSRTC in surgical planning of thyroid nodules and to correlate TBSRTC with final histopathology. 85 patients with thyroid nodules who underwent surgery were evaluated prospectively. Preoperatively FNAC and TBSRTC reporting was done and following surgery histopathology was correlated with cytology. Distribution of 85 patients amongst the six categories of TBSRTC was as follows: 2.35% in Category I, 68.23% II, 7.05% III, 16.47% IV, 2.35% V and 3.52% VI. In 93% (79) of patients TBSRTC correlated with histopathology whereas in 7% (6) it did not. Risk of malignancy calculated was 0% in II, 33.33% in III, 7.14% in IV and 100% in V and VI categories. Sensitivity, specificity, Positive Predictive Value and Negative Predictive Value of TBSRTC was 100% for V, VI categories, whereas it was 100%, 78%, 15% and 100% respectively for III, IV. The diagnostic accuracy in our study is 100% for Category V and VI whereas it is 79% for Category III and IV. TBSRTC proved to be a very good screening platform for triaging patients with thyroid nodules into benign and malignant groups, as it is directly related to risk of malignancy in each category. It has helped in appropriate surgical planning in 96.4% of our patients.

Keywords Thyroid nodule · FNAC · Bethesda · Malignancy · ACR-TIRADS · Thyroidectomy

✉ B. V. Manjula
drmanjubv@gmail.com

Swetal Mahajan
swetalmahajan85@gmail.com

R. Vaishnavi
ranganathan.vaishnavi@gmail.com

Neena John
neenajohndr2010@gmail.com

Bhuvaneshwari Babu
jillybabu@gmail.com

[1] Department of ENT and Head and Neck Surgery, Bangalore Baptist Hospital, Bellary Road, Hebbal, Bengaluru, Karnataka 560024, India

[2] Pathology, Bangalore Baptist Hospital, Bengaluru, Karnataka, India

Introduction

Prevalence of thyroid nodules ranges from 4 to 10% and 0.2 to 1.2% in adult and children population respectively [1]. A solitary thyroid nodule is a cause for concern because of high chance of malignancy ranging from 5 to 35% of all solitary nodules [1]. Thyroid Ultrasonography (USG) was commonly being used traditionally to decide on which nodules need further investigation including fine needle aspiration cytology (FNAC). The recent American College of Radiology Thyroid Imaging Reporting and Data System (ACR TI-RADS) scoring has high performance for selecting thyroid nodules for FNAC compared to other TIRADs and 2015 American Thyroid Association (ATA) guidelines risk stratification [2, 3]. ACR TI-RADS was primarily developed to triage nodules for FNAC and avoid unnecessary FNAC of benign nodules. But as clinicians we find TIRADs system cumbersome and time consuming.

Though ATA system is easier and quicker we can't rely on TIRADs and ATA for diagnostics as diagnosis of thyroid cancer involves patient, imaging and other tumor factors for risk stratification.

FNAC has proven to be a first-line tool to evaluate the thyroid lesions because of its cost effectiveness and high patient acceptance. In spite of acknowledged Cyto-diagnostic pitfalls, the use of FNAC in the preliminary evaluation of solitary nodule reduces the use of surgery by 1/3rd, doubles the proportions of malignancies in surgical resections and enables surgeons to take decision regarding mode of treatment [4]. In 2007 the National Cancer Institute (NCI) introduced The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) to standardize international terminology and to categorize morphological criteria in FNAC [5]. It includes 6 categories; I-nondiagnostic, II-Benign, III-Atypia of Undetermined Significance (AUS)/follicular lesion of undetermined significance (FLUS), IV-Follicular Neoplasm (FN)/suspicious for follicular neoplasm (SFN), V-suspicious for malignancy (SFM), and VI-Malignant and their risk of malignancy were 1–4%, 0–3%, 5–15%, 15–30%, 60–75%, and 97–99%, respectively [4]. With uniformity in reporting it has tried to bridge the gap between clinicians and pathologists.

Studies worldwide have proven that incorporation of TBSRTC in diagnostic algorithms for patients with thyroid nodules have reduced unnecessary thyroidectomies while also ensuring the quality of thyroid malignancy detection [6, 7]. Although Bethesda criteria appropriately stratify malignancy risk in thyroid nodules controversy exists regarding their accuracy and reliability in clinical decision-making [8–10]. This could be due to variability in the prevalence of cancer in the studied population, the inherent variability in the interpretation of cytology, and bias in the selection of patients for surgery. Category III includes AUS, FLUS or could be either architectural atypia or nuclear atypia or preparatory artifacts related atypia. Category IV includes FN where cytology does not show the capsular and or vascular invasion that can distinguish a follicular thyroid cancer from a benign follicular adenoma and also encompasses Hürthle cell neoplasm/suspicious for Hürthle cell neoplasm. So molecular testing for BRAF or 7 gene mutational panels for indeterminate nodules allows better risk stratification and reduce the need for diagnostic thyroid surgery [3]. However, they are neither freely available nor economic in Indian scenario. Studies have also compared TI-RADS scoring with Korean TIRADS and ATA guidelines to look at malignancy risk associated with each TI-RADS category and have found incremental increase in malignancy risk with higher point score. But the highest risk category still only predicted malignancy risk in 35%. A combined sonocytological BETH-TR score has also been proposed to triage these indeterminate thyroid nodules and have shown that a BETH-TR score ≥ 7 can correctly identify malignant nodules in 86% of cases [11]. With this conflicting evidence adoption of TI-RADS universally is a costly surveillance without clear benefit [12]. Therefore we planned this study to understand if TBSRTC alone can guide us in surgical planning.

Aims

The aim of the study was to evaluate the usefulness of TBSRTC in surgical planning of thyroid nodules with risk stratification in each category and to correlate TBSRTC with histopathology.

Materials and Methods

Study Design

This is a prospective observational study for 1 year duration conducted on 85 patients who underwent thyroidectomy at our institute after approval of our hospital's Ethics Committee.

Inclusion Criteria

1. Patients of age 18 years and above presenting with solitary thyroid nodule who underwent investigations and hemi/total thyroidectomy at our Institute.

Exclusion Criteria

1. Patients with recurrent thyroid swellings after previous surgery.
2. Patients with TBSRTC category II (benign nodules) those planned for observation and regular follow up.
3. Patients with multiple nodules or Multinodular goitre

Patients were explained about the study and informed consent was taken. FNAC was performed by the pathologist and application of TBSRTC was done to each FNAC result and divided into 6 diagnostic categories after which plan of treatment was decided by the surgeon based on the Bethesda category. Hemithyroidectomy was planned for Bethesda II, III, IV where as hemithyroidectomy and Frozen section for Bethesda V and Total thyroidectomy and Neck dissection for Bethesda VI. Thyroid specimen excised during surgery was sent for histopathological examination and final result was correlated with each Bethesda Category.

Statistical Analysis

Descriptive statistics were reported as mean (standard deviation [SD]), number and percentages. The reporting percentages of each Bethesda category and the percentage of malignancy in each category were calculated and compared with national studies. McNemar chi-square test was used to compare the proportion of positives between two diagnostic reports. The results were compared using the two-sample *t*-test between percents (Microsoft Excel). *p* value of <0.05 was considered statistically significant. Data entered in Microsoft excel 1997–2003 and analysed in SPSS version 25.0. Sensitivity, specificity, diagnostic accuracy, positive predictive (PPV) and negative predictive (NPV) values were calculated and usefulness of TBSRTC in deciding the treatment plan is reported in proportions.

Results

Patient Cohort

Clinicopathological features of 85 patients (13 male, 72 female) with M: F ratio = 1: 5.5, age range (18–80 years) with mean 45.5 years operated for solitary thyroid nodule at our institute were prospectively analyzed, Table 1.

Distribution of Cases According to Bethesda (Table 1)

Majority of patients, total 58 patients belonged to Bethesda category II (68.23%), whereas 2—Category I (2.35%), 6—category III (7.05%), 14—category IV (16.47%), 2—category V (2.35%) and 3 patients were reported as category VI (3.52%) respectively. 2 patients with Bethesda category I underwent repeat FNAC and were found to be II later.

All the patients belonging to category II were advised conservative management and follow up but in scenarios where a patient factor, like cosmetic reasons, or where presenting complaint was obvious large thyroid swelling and/or pressure symptoms, they were planned for hemithyroidectomy. Some anxious patients preferred surgery over follow up, as it becomes a quality of life issue for them, and they would be put at ease with an excision biopsy of the lesion. Out of 85 patients, 80 patients underwent hemithyroidectomy (44—right, 36—left) which included 60 patients of Category II, 6 Category III and 14 Category IV. 2 patients with Category V (SFM) underwent hemithyroidectomy followed by frozen section which of suggestive of papillary carcinoma, so they underwent total thyroidectomy and central compartment neck dissection. We did not plan for intra operative Frozen in Bethesda III and IV as it is well known fact that a careful and comprehensive evaluation of the tumor capsule is mandatory to visualize a capsular or a vascular invasion to confirm the diagnosis of malignancy. However, this assessment is possible only on permanent sections.

Table 1 Clinicopathological characteristics

Characteristics		No (%) (n = 85)
Gender	Male	13 (15.3%)
	Female	72 (84.7%)
Bethesda	I—Non diagnostic	2 (2.4%)
	II—Benign	58 (68.2%)
	III—AUS/FLUS	6 (7.1%)
	IV—FN/SFN	14 (16.5%)
	V—Suspicious for malignancy	2 (2.4%)
	VI—Malignant	3 (3.5%)
Plan	Left hemithyroidectomy	36 (42.4%)
	Right hemithyroidectomy	44 (51.8%)
	Hemithyroidectomy + Frozen and proceed	2 (2.4%)
	Total thyroidectomy + Neck Dissection	3 (3.5%)
Surgery done	Left hemithyroidectomy	36 (42.4%)
	Right hemithyroidectomy	44 (51.8%)
	Total + neck dissection	5 (5.9%)
Final HPR	Malignant	8 (9.4%)
	Benign	77 (90.6%)

Moreover, frozen section can jeopardize the detection of signs of capsular invasion on final pathologic examination. 3 patients with category VI (malignant disease) underwent upfront total thyroidectomy and neck dissection.

Distribution of Cases According to Final Histopathology (Table 2)

Maximum number of patients (38/85) had colloid goitre, 10 patients had colloid cyst and 9 colloid nodule. 14 patients had follicular adenoma, 5 patients had papillary carcinoma, 2 had follicular carcinoma and only 1 patient had medullary carcinoma of thyroid.

Cyto-Histopathological Correlation (Table 3)

Out of total 85 patients, FNAC of 93% (79) of patients matched with final histopathological report (HPR) where as in 7% (6) of cases, it didn't match. In these 6 patients, 2 belonging to Bethesda II (Benign) were found to have Follicular adenoma on final HPR, 2 patients of Bethesda III (AUS) had Follicular carcinoma and Papillary carcinoma respectively and 2 patients of Bethesda IV (SFN) had Follicular carcinoma and Colloid nodule respectively. In all patients belonging to category V and VI final HPR was consistent with malignancy.

Risk of malignancy (Table 4) calculated is 0% in Bethesda category I and II, 33.33% in category III, 7.14% in category IV and 100% in both category V and VI.

Sensitivity, specificity, positive predictive value, negative predictive value and diagnostic accuracy (Table 5) were calculated separately for Cat III, IV (indeterminate nodules) and Cat V, VI (malignant). Sensitivity, specificity, PPV and NPV of TBSRTC was 100% for V and VI whereas it was 100%, 78%, 15% and 100% respectively for III and IV categories. There was a statistically significant difference in the proportion of positives between HPR and Bethesda III, IV ($p < 0.0001$), however when malignancy (V, VI) was compared with HPR, there was no significant difference.

Discussion

TBSRTC has greatly improved uniformity in reporting thyroid FNACs and has tried to eliminate ambiguity by triaging nodules into benign and malignant and help in surgical planning. Though lot of literature exists on TBSRTC reproducibility and histopathological correlation there is need for consensus worldwide for clinical decision making. Therefore we planned this study to understand if TBSRTC can guide us in surgical planning.

In our study maximum cases were noted in category II (68%), which is consistent with other studies (64–66%) [13]. Distribution of cases in other 5 categories is also found to be similar to other studies. The risk of malignancy in our study for Category II was 0% which is consistent with literature (0–3%) [5]. The Category III (AUS) in our study yielded a malignancy rate of 33.33% which was higher than the international experience, but Category IV (SFM) yielded a malignancy rate of 7.14% in our study which was very less compared to other studies. This may be because of rigid adherence to all the diagnostic criteria by the pathologist and to keep the number of patients in AUS and FN to a minimum. Only 7.05% were in AUS (Category III) and 16.47% of our patients were in FN (Category IV). So at this point of time when molecular tests are uneconomical and at infancy stage in India either a repeat FNAC and regular follow up or diagnostic hemithyroidectomy which will relieve patient's anxiety can be offered to these patients. Category V and VI yielded a malignancy rate of 100% which is quite comparable with other studies [5].

Table 2 Cases distribution according to final histopathological examination

Histopathological diagnosis	No of cases	Percentage (%)
Colloid goitre	38	44.71
Colloid cyst	10	11.76
Colloid nodule	9	10.59
Hyperplastic thyroid lesion	1	1.176
Subacute granulomatous thyroiditis	1	1.176
Hashimoto's thyroiditis	4	4.706
Follicular adenoma	14	16.47
Follicular carcinoma	2	2.353
Medullary carcinoma	1	1.176
Papillary carcinoma	5	5.882
Total	85	100

Table 3 Cyto-histopathological correlation

TBSRTC (Cytopathology)		Histopathological diagnosis		
		Total	Benign	Malignant
II	Benign	60	60	0
III	Atypia of undetermined significance (AUS)	6	4	2
IV	Suspicious of follicular neoplasms(SFN/FN)	14	13	1
V	Suspicious of malignancy	2	0	2
VI	Malignant	3	0	3

Table 4 Risk of malignancy in each Bethesda category

Bethesda category	No. of cases	Malignancy rate No. of cases (%)
Benign	60	0 (0%)
Atypia of undetermined significance	6	2 (33.33%)
FN/Suspicious of follicular neoplasm	14	1 (7.14%)
Suspicious of malignancy	2	2 (100%)
Malignant	3	3 (100%)

Table 5 Diagnostic evaluation of Bethesda with histopathology

Bethesda	HPR		Sensitivity	Specificity	PPV	NPV	Diagnostic accuracy (%)	*p* value (McNemar test)
	Malignant	Benign						
Malignant (V and VI)	5	0	100 (47.8, 100)	100 (94.0, 100)	100 (56.0, 100)	100 (94.0, 100)	100	0.00
Indeterminate nodules (III and IV)	3	17	100 (29.0, 100)	78 (67.0, 86.6)	15 (10.4, 21.2)	100 (94.0, 100)	79	<0.0001
Benign	0	60						

In literature the actual risk of malignancy has been difficult to determine for category III and IV since confirmatory evidence is available only for that subset of patients who undergo surgery and is not available for those who are followed by repeat FNAC and observation according to Bethesda recommendations. There are no large prospective studies with histopathological correlation for this category. Also due to lack of access to molecular testing and financial constraints, in our study we offered hemithyroidectomy for all patients in Category III, IV to avoid selection bias. In present study, among 6 patients in Category III, 4 were benign and 2 malignant and in Category IV, among 14 patients 1 was malignant and remaining 13 benign and in these 3 (3.6%) patients with malignancy hemithyroidectomy was done. So TBSRTC helped in appropriate surgical planning in the rest 82 (96.4%) patients only.

A difference in the categorization of Bethesda III, IV and V has resulted in wide range of sensitivities and specificities. Some studies have categorized Bethesda IV as benign, while some studies have categorized it into malignant. If indeterminate nodules are considered positive, then sensitivity increases while specificity decreases. If they are excluded, then sensitivity for malignancy detection decreases and increases both false positive and false negative rates. So in the present study (Table 5) all parameters were calculated by grouping Category III, IV (Indeterminate nodules) separately and V, VI (SFM, malignant) separately into two groups. Sensitivity is 100% in both groups. In other studies, it was observed to be 76–94%. Specificity was 100% in Category V, VI where as it was 78% in Category III, IV which is comparable to 64–100% found in other studies. In our study PPV and NPV was 100% in Category V, VI; as compared to 15%

and 100% in category III, IV respectively whereas in literature it is varied [1, 14, 15]. Statistically significant difference ($p < 0.0001$) was noted when final HPR was compared to Category III, IV because of the very low PPV—15%. This shows that false negative rate is still high in this group of intermediate nodules.

Conclusion

Our study shows that in 93% of patients TBSRTC correlated with final histopathology where as only in 7% of cases, it didn't match. The diagnostic accuracy in our study is 100% for Category V, VI whereas it is only 79% for category III, IV. Our study is different from others as we have assessed TBSRTC accuracy separately.

TBSRTC is a simple, convenient, reproducible and cost effective tool but it still has false negative reports in Category III, IV. This can only be improved in future when molecular tests are widely available and economical. Also better diagnostic methodologies are necessary to obtain accurate diagnosis so as to avoid needless surgeries. Till then TBSRTC is the best available clinical tool for triaging thyroid nodules into benign and malignant. It has helped in appropriate surgical planning in 96.4% of our patients and it is important to utilize it. It is also important to adhere to the diagnostic criteria of TBSRTC and keep the number of patients with AUS and FN to a minimum so as to reduce false negatives in Bethesda III and IV categories.

Funding None.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Arul P, Akshatha C, Masilamani S (2015) A study of malignancy rates in different diagnostic categories of the Bethesda system for reporting thyroid cytopathology: An institutional experience. Biomed J 38(6):517–522
2. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA et al (2017) ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS Committee. J Am CollRadiol JACR 14(5):587–595
3. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE et al (2016) 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 26(1):1–133
4. Melo-Uribe MA, Sanabria Á, Romero-Rojas A, Pérez G, Vargas EJ, Abaúnza MC et al (2015) The Bethesda system for reporting thyroid cytopathology in Colombia: correlation with histopathological diagnoses in oncology and non-oncology institutions. J Cytol Indian AcadCytol 32(1):12–16
5. Cibas ES, Ali SZ (2009) NCI Thyroid FNA state of the science conference. The Bethesda System for reporting thyroid cytopathology. Am J ClinPathol 132(5):658–65
6. Cesur M, Corapcioglu D, Bulut S, Gursoy A, Yilmaz AE, Erdogan N et al (2006) Comparison of palpation-guided fine-needle aspiration biopsy to ultrasound-guided fine-needle aspiration biopsy in the evaluation of thyroid nodules. Thyroid Off J Am Thyroid Assoc 16(6):555–561
7. Anand B, Ramdas A, Ambroise MM, Kumar NP (2020) The bethesda system for reporting thyroid cytopathology: a cytohistological study. J Thyroid Res 16(2020):1–8
8. Yoon JH, Kwak JY, Moon HJ, Kim MJ, Kim E-K (2011) The diagnostic accuracy of ultrasound-guided fine-needle aspiration biopsy and the sonographic differences between benign and malignant thyroid nodules 3 cm or larger. Thyroid Off J Am Thyroid Assoc 21(9):993–1000
9. Stulak JM, Grant CS, Farley DR, Thompson GB, Heerden JAV, Hay ID et al (2006) Value of preoperative ultrasonography in the surgical management of initial and reoperative papillary thyroid cancer. Arch Surg 141(5):489–496
10. Solorzano CC, Carneiro DM, Ramirez M, Lee TM, Irvin GL (2004) Surgeon-performed ultrasound in the management of thyroid malignancy. Am Surg 70(7):576–580. (discussion 580–582)
11. Mehta S, Kannan S (2020) Approaching indeterminate thyroid nodules in the absence of molecular markers: "The BETH-TR score." Indian J EndocrinolMetab 24(2):170
12. Middleton WD, Teefey SA, Reading CC, Langer JE, Beland MD, Szabunio MM et al (2018) Comparison of performance characteristics of American College of Radiology TI-RADS, Korean Society of Thyroid Radiology TIRADS, and American Thyroid Association Guidelines. AJR Am J Roentgenol 210(5):1148–1154
13. Nayar R, Ivanovic M (2009) The indeterminate thyroid fine-needle aspiration: experience from an academic center using terminology similar to that proposed in the 2007 national cancer institute thyroid fine needle aspiration state of the science conference. Cancer Cytopathol 117(3):195–202
14. Muratli A, Erdogan N, Sevim S, Unal I, Akyuz S (2014) Diagnostic efficacy and importance of fine-needle aspiration cytology of thyroid nodules. J Cytol Indian AcadCytol 31(2):73–78
15. Basharat R, Bukhari MH, Saeed S, Hamid T (2011) Comparison of fine needle aspiration cytology and thyroid scan in solitary thyroid nodule. Pathol Res Int 2011:754041