Diagnostic Efficacy of Ultrasound-Guided Fine Needle Aspiration Combined with the Bethesda System of Reporting

K. Ajitha Kumari, Poonam D. Jadhav, Chaya Prasad, N. V. Smitha, Annie Jojo, V. D. Manjula
Department of Pathology, Believers Church Medical College, Thiruvalla, Department of Pathology, Amrita Institute of Medical Sciences and Research Centre, Kochi, Department of Community Medicine, Government Medical College, Idukki, Kerala, India

Abstract

Background: Image-guided fine needle aspiration cytology (FNAC) is emerging as an important diagnostic tool in the evaluation of thyroid swellings. Aim: This study aims to assess the efficacy of ultrasound (US)-guided FNAC combined with “The Bethesda system” of reporting as a primary screening test for all thyroid lesions. Settings and Design: A prospective cohort study was made and all the US-guided FNACs done were followed up to find out the histopathological diagnoses wherever surgery was done. Materials and Methods: In all, 1050 patients who underwent US-guided FNAC were studied during a period of 1 year. Age, sex, cytological features, and histological diagnoses were analyzed. Statistical analyses of all the findings were done to derive conclusions. Results: Of the 1050 patients, only 10.5% underwent surgery. Higher than expected rate (as per the Bethesda system) of malignancy was noted with the so-called grey zone lesions. The test results revealed a high level of sensitivity, specificity, and diagnostic accuracy. Conclusion: The study showed that The Bethesda System of Reporting Thyroid Cytology provides effective communication between clinician and pathologists thereby enabling clear management strategies. We also concluded that US-guided FNACs offer better results compared with palpation-guided FNACs.

Keywords: Cytopathology, The Bethesda System of Reporting Thyroid Cytology, thyroid lesions, ultrasound-guided fine needle aspiration

INTRODUCTION

There is a growing trend among institutions worldwide to do ultrasound (US)-guided fine needle aspiration along with implementation of The Bethesda System of Reporting Thyroid Cytology (TBSRTC) as an excellent, cost-effective and less invasive primary line of investigation in thyroid lesions. Pathologists were using different terminologies in cytodiagnosis of thyroid lesions till the introduction of an universal terminology called The Bethesda System of Reporting Thyroid Cytology in the year 2007. The most important clinical utility of TBSRTC is that the decisions for management strategy for these lesions can be taken appropriately based on the diagnostic category so that optimal and timely surgical management can be given to the patient. The combined impact of US guidance and TBSRTC is far superior to the standard palpation technique which is still practiced in most of the centers.

MATERIALS AND METHODS

A total of 1050 patients who underwent guided fine needle aspiration cytology (FNAC) for thyroid swellings were evaluated. All these cases were followed up to pick out those who underwent surgery in this hospital. Those who opted for surgery in other hospitals could not be traced. All the FNACs in this study were done by radiologist under US guidance. We received both wet-fixed and air-dried smears which were stained in the cytopathology laboratory using Papanicoloau’s stain and Giemsa stain. These smears were evaluated by a pathologist within 1–2 days. Wherever necessary, more than one pathologists were involved in the interpretation of the findings. Radiological findings were available in all the cases for correlation. Categorization was done strictly based on the guidelines and standard nomenclature for interpretation of thyroid nodules as suggested by TBSRTC. Histopathological diagnoses of all those who underwent surgery were followed...
up. Data were coded and entered in Microsoft Excel and analysis was done by SPSS version 16.0. Sensitivity, specificity, predictive value, and diagnostic accuracy were calculated with thyroid FNAC as screening test and histopathology as gold standard. Diagnostic categories IV, V, and VI which were histopathologically confirmed as malignant were defined as true-positive. True-negative included category II, which were histopathologically confirmed as benign. Category I was excluded from analysis. False-positive category included cases that were diagnosed as follicular neoplasm, suspicious for malignancy and as malignant that were confirmed as benign histopathologically. False-negative cases included those diagnosed as benign, but confirmed as malignant on histopathological examination. Sensitivity, specificity, predictive value, and accuracy were also calculated by including the atypia of undetermined significance/follicular lesion of undetermined significance into the test positive group.

**RESULTS**

A total of 1050 patients were included in the study, 916 of whom were females constituting 87.24% (males 12.76%). The mean age incidence was 47.5 (age range was 14–81 years).

The TBSRTC categories among the total population are obtained as shown in Table 1.

The biggest number of cases (870) in this study was in category II (benign), whereas the smallest number (10) was in category III. Categories IV, V, and VI together (neoplasm/suspicious of neoplasm, suspicious of malignancy, and frankly malignancy) formed a major group comprising 82 cases.

| TBSRTC category | Number of cases in each category | Percentage distribution |
|-----------------|---------------------------------|-------------------------|
| Category I      | 88                              | 8.38                    |
| Category II     | 870                             | 82.85                   |
| Category III    | 10                              | 0.95                    |
| Category IV     | 21                              | 2.00                    |
| Category V      | 27                              | 2.57                    |
| Category VI     | 34                              | 3.23                    |

TBSRTC: The Bethesda System of Reporting Thyroid Cytology

From the six different TBSRTC categories, 111 patients underwent surgery. In all, 67 of these cases (60%) were proved to be benign and 44 (40%) proved as malignant by histology. Papillary carcinoma was the most common histological type among malignancy comprising 39 cases (87%). All other histological types were far less common comprising three medullary carcinomas (9%), one poorly differentiated carcinoma (2%), and one Hurthle cell carcinoma (2%). There was a single case of anaplastic carcinoma which was referred to radiotherapy. Papillary microcarcinomas (incidentalomas) were considered as a separate category. There were a total of 14 cases in this category (11 from category II, 2 from category III, and one from category V). Table 2 shows the surgeries done in each category and the percentage distribution of benign and malignant cases. The calculated risk of malignancy in each category compared with the implied risk of malignancy as per TBSRTC is shown in Table 3. The overall sensitivity of US-guided FNAC was 97.6% and specificity was 83.9%.

**DISCUSSION**

Of the 1050 patients in this study, only 111 patients (10.5%) underwent surgery. Among them, 67 patients (60%) had a benign histopathology, whereas the remaining 44 patients (40%) were proved to be malignant. The percentage of patients undergoing surgery for thyroid swelling vary from 11.8% to 45.1% in different institutions and the reported rate of malignancy in thyroid lesions is only 5%.[6] The increased rate of malignancy in this study can be attributed to the appropriate selection of patients for surgery based on TBSRTC criteria.

In this study of 1050 patients, 88 (8.4%) were in the nondiagnostic category. Many of the smears in our study group were in fact classified as category I, not because of inadequacy of cellular material, but due to heavy blood contamination and clotted blood entrapping the cellular material wherein the morphology of the cells was obscured. Seven of 88 (8%) patients in this category underwent surgery due to clinical suspicion, of which one turned out to be malignant. Hence, the calculated risk of malignancy in this group comes to 14.3%, whereas the implied risk of malignancy as per TBSRTC is only 1%–4%.

The least percentage of surgery occurred with benign category II TBSRTC. In our study, of the 870 (82.9%) patients diagnosed

| Bethesda category | Number of cases in each category | Surgeries in each category | Percentage distribution | HPR benign (%) | HPR malignant (%) |
|-------------------|---------------------------------|---------------------------|-------------------------|----------------|------------------|
| Category I        | 88                              | 7                         | 8                       | 6 (86)         | 1 (14)           |
| Category II       | 870                             | 53                        | 6                       | 52 (98)        | 1 (2)            |
| Category III      | 10                              | 4                         | 40                      | 2 (50)         | 2 (50)           |
| Category IV       | 21                              | 10                        | 48                      | 3 (30)         | 7 (70)           |
| Category V        | 27                              | 16                        | 59                      | 4 (25)         | 12 (75)          |
| Category VI       | 34                              | 21                        | 62                      | 0              | 21 (100)         |
| Total             | 1050                             | 111                       | 10.6                    | 67 (60)        | 44 (40)          |

HPR: Histopathology report
The outstanding feature of this study is that category III and IV lesions, so-called “grey zone lesions,” was very few in number contributing only 1% and 2%, respectively, of the total 1050 patients. This may be considered as an advantage of image-guided FNACs over palpation-guided procedures where there is a chance of missing the target. TBSRTC explains clear-cut scenarios where a category III interpretation is appropriate. Significant variations were observed between different institutions and amidst pathologists in the diagnoses of category III lesions.[7,8] Bongiovanni et al. in their study observed the rate of category III lesions as ranging from 3% to 27% with a malignancy rate of 15.9%. According to TBSRTC, the malignancy risk for this category is 5%–15% only. However, recent reviews suggest higher rates of malignancy in this category.[10] Ho AS et al. have reported a malignancy rate of 26.6%–37.8%.[11]

Considering the findings of our study as well as other published data, the inclusion of category III in TBSRTC appears currently reasonable.[12] For effective clinical management strategies, further subclassification of category III lesions into high-risk and low-risk group can be very helpful.[13,14] We suggest that guidelines which recommend repeat FNA and observation for category III lesions need to be revised. High-risk patients may be treated surgically while a conservative approach maybe sufficient for the remaining low-risk patients. Moreover, the overall sensitivity of FNAC interpretation becomes more acceptable with the inclusion of category III in TBSRTC as shown in Table 3. The diagnostic accuracy of these indeterminate lesions can be improved further by using immunocytochemical and molecular methods.[15,16]

Lesions which were grouped under category IV TBSRTC included follicular neoplasm/suspicious for follicular neoplasm. TBSRTC gives definite criteria for putting a lesion under category IV which denotes a diagnosis of follicular neoplasm/suspicious of a follicular neoplasm. The term “suspicious for follicular neoplasm” is very advantageous for the cytopathologists because in fact some of these lesions which we consider as follicular neoplasms turn out to be hyperplastic nodules on histopathological examination. This happens due to overemphasis on microfollicular pattern, cellularity, and amount of colloid. The outstanding feature of category IV is that it includes potential follicular carcinoma

### Table 3: Calculated risk of malignancy in each category compared with the implied risk of malignancy as per TBSRTC categories

| FNAC diagnosis | Total number | Malignant after surgery | Risk of malignancy (%) | Implied risk of malignancy as per TBSRTC |
|----------------|--------------|-------------------------|------------------------|------------------------------------------|
| Category I     | 7            | 1                       | 14.3                   | 1-4                                      |
| Category II    | 53           | 1                       | 1.9                    | 0-3                                      |
| Category III   | 4            | 2                       | 50.0                   | 5-15                                     |
| Category IV    | 10           | 7                       | 70.0                   | 15-30                                    |
| Category V     | 16           | 11                      | 68.8                   | 60-75                                    |
| Category VI    | 21           | 21                      | 100.0                  | 97-99                                    |

FNAC: Fine needle aspiration cytology, TBSRTC: The Bethesda System of Reporting Thyroid Cytology

as category II, 53 (only 6%) underwent surgery due to clinical and radiological suspicion. The only one case which turned out to be malignancy was a follicular variant of papillary carcinoma which certainly has some overlapping cytological features with category II.

Of the total 1050 patients in this study, only 10 cases were categorized as category III. This is considerably lower (0.95%) and TBSRTC system advises to limit the use of this category to 7% or even less. Of these 10 cases, 4 patients underwent surgery due to strong clinical and radiological suspicion and other 6 cases were advised follow-up with repeat FNACs. Two of the operated cases proved to be papillary carcinoma, one classical variant and another follicular variant with a malignancy risk of 50% and two others had papillary microcarcinomas.

There were 21 cases in category IV (2%) and only 10 patients underwent surgery in the hospital. Histopathology reports revealed four cases as follicular variant of papillary carcinoma and three cases as classical papillary carcinoma together contributing to a malignancy risk of 70%, which is much higher compared with the implied risk as per TBSRTC which is 15%–30% only.

The number of patients in category V was 27 (2.57%). The number of patients who underwent surgery was 16, of which 11 were reported as malignant (papillary carcinoma) indicating a malignancy risk of 64.7% against a malignancy risk of 60%–75% as per TBSRTC.

In our study, there were 34 cases in the TBSRTC category VI (3.23%). A total of 21 cases in this category underwent surgery in this hospital and the remaining patients were lost to follow-up. Histopathological diagnoses of all the 21 cases were malignant giving a malignancy risk of 100% for category VI.

In cytology during the pre-TBSRTC period, the majority of the lesions of the thyroid were categorized together as “follicular lesions” without mentioning whether benign or malignant. The clinicians had no choice other than doing unwanted surgery and with a histopathology outcome of benign lesion in most of the cases. With the introduction of TBSRTC, all the frankly benign nonneoplastic lesions were grouped under category II so that these patients are followed up with repeated FNACs at appropriate interval with surgical option for only very few cases based on clinical suspicion.
Table 4: Overall sensitivity, specificity, accuracy, PPV, NPV, FPR, and FNR rates

| Main diagnostic indicators for six-tiered Bethesda system |
|----------------------------------------------------------|
| Without AUS | Sensitivity (%) | 92.9 | Specificity (%) | 87.1 | Accuracy (%) | 89.4 | PPV (%) | 83 | NPV (%) | 94.7 | FPR (%) | 12.9 | FNR (%) | 7.1 |
| With AUS | Sensitivity (%) | 97.6 | Specificity (%) | 83.9 | Accuracy (%) | 89.4 | PPV (%) | 80.4 | NPV (%) | 98.1 | FPR (%) | 16.1 | FNR (%) | 2.4 |

PPV: Positive predictive value; NPV: Negative predictive value; FPR: False-positive rate; FNR: False-negative rate; AUS: Atypia of undetermined significance

as well. We have to look forward to a panel of biomarkers that will help in delineating follicular carcinomas cytologically.

Inclusion of category V in TBSRTC is very helpful both for the clinician and the cytopathologist. At times with some of the FNAC samples, the pathologist may feel that it is malignant especially in the presence of positive clinical and radiological suggestion, but the available material may not be qualitatively or quantitatively sufficient to give a confident diagnosis of malignancy. Those cases can very well be placed under category V. In such a setting, rather than going in for ancillary tests, a meaningful discussion with the surgeon could be very effective so that he could opt for intraoperative frozen sections so as to go in for a radical surgery if indicated, rather than doing lobectomy.[17] TBSRTC suggested a target positive predictive value of 55%–85% with category V lesions and our study has very well reached a positive predictive value within this target.

Aspirates are put under category VI when the cytomorphological features are unequivocally malignant. Further subclassification as to the tumor type is always helpful for the clinician to decide on the management strategy. In our study, 34 of 1050 (3.23%) cases were reported as conclusively malignant. In all, 21 cases in this category underwent surgery and the remaining patients were lost to follow-up. Histopathological diagnoses of all the 21 cases were malignant giving a malignancy risk of 100% for category VI.

The sensitivity of US-guided FNAC (97.6%) in this study was remarkably higher and the specificity (83.9%) was very well within the acceptable range. Positive predictive value, negative predictive value, and diagnostic accuracy were also of high-level acceptability [Table 4].

With palpation-guided FNAC, various authors have reported sensitivity, specificity, and diagnostic accuracy considerably lower than what is obtained by US-guided FNAC. Sensitivity, specificity, and diagnostic accuracy of conventional palpation-guided FNAC compared with US-guided FNAC, respectively, were 91.8% versus 97.1%, 68.8% versus 70.9%, 72.6% versus 75.9%,[18] 54.54% versus 81.81%, 92.85% versus 92.85%, 76% versus 88%,[19] 30.7% versus 78.57%, 100% versus 91.67%, and 82.69% versus 86.2%.[20] The above studies show a wide variation in these indices with conventional FNACs. Sensitivity ranges from 30.7% to 91.8%, specificity ranges from 68.8% to 100%, and diagnostic accuracy varies from 72.6% to 82.69%, whereas with US-guided FNACs sensitivity ranges from 78.57% to 97.1%, specificity 70.9% to 92.85%, and diagnostic accuracy from 75.9% to 88%. In this study where all the FNACs were done under US guidance, the sensitivity, specificity, and diagnostic accuracy were 97.6%, 83.9%, and 89.4%, respectively.

Although this study provided useful information, there were a few limitations as well. With high incidence of papillary microcarcinoma, in the study group we found it difficult to designate TBSRTC category to this group of incidentalomas. In our study as shown by histopathology, there were a total of 11 cases of follicular variant of papillary carcinoma from all the six different Bethesda categories, but none we could pick up cytologically. Cytological diagnosis of follicular variant of papillary carcinoma is exclusively based predominantly on the presence of nuclear features, and hence always pose challenges to the cytopathologist and is often missed.[21] This emphasizes the need for repeat FNAC/surgery in category II lesions wherever it is suspicious clinically or radiologically and if no suspicion, these patients are to be followed up at 6–18 months interval.

Regarding technical limitations, if all these FNACs were done in the presence of cytopathologists, we could have made better utilization of the aspirated material and thereby effectively reduce the percentage of nondiagnostic category. This shows that there is a need for technical expertise in processing the material aspirated by the radiologist so as to assure the quality of the smears prepared.

**Conclusion**

Our study shows the utility of image guidance in fine needle aspiration of thyroid and the advantages of the implementation of TBSRTC. The overall increase in sensitivity, specificity, and diagnostic accuracy of this study compared to other published data[22,23] can be attributed to the efficacy of image guidance in thyroid aspiration. With this procedure, the chances of identifying frankly malignant lesions in the thyroid are very high as proved by the 100% correlation between category VI lesions and histopathology outcome. Another important conclusion is that the percentage of indeterminate lesions can be reduced by image-guided targeting of the lesion in the thyroid.

Hence, we recommend universal application of this procedure as the first step in the work-up of thyroid lesions.

**Financial support and sponsorship**
Nil.

**Conflicts of interest**
There are no conflicts of interest.

**References**
1. Robitschek J, Straub M, Wirtz E, Klem C, Sniezek J. Diagnostic
1. The prediction of malignant risk in thyroid nodules classified as Bethesda category III (AUS/FLUS). Thyroid 2014;24:832-9.

2. The Bethesda thyroid fine-needle aspiration classification system: Year at an academic institution. Thyroid 2009;19:1215-23.

3. Hyeon J, Ahn S, Shin JH, Oh YL. The prediction of malignant risk in thyroid nodules classified as Bethesda category III (AUS/FLUS). Thyroid 2014;24:832-9.

4. Danese D, Sciacchitano S, Farsetti A, Andreoli M, Pontecorvi A. Diagnostic accuracy of conventional versus sonography-guided fine-needle aspiration biopsy of thyroid nodules. Thyroid 1998;8:15-21.

5. Krishnappa P, Ramakrishnappa S, Kulkarni MH. Comparison of free hand versus ultrasound-guided fine needle aspiration of thyroid with histopathological correlation. J Environ Pathol Toxicol Oncol 2013;32:149-55.

6. Ramos CO, Mirasol RC. Ultrasound guidance improves the diagnostic yield and accuracy of fine needle aspiration biopsy of thyroid nodules. Thyroid 2017 for Best Published Article in Journal of Cytology

Dr. Panna Choudhury Memorial Award

1. This award shall be given to the best original article published in the Journal of Cytology during the past one year.

2. The award comprises of a gold plated silver medal and a certificate.

3. Only IAC members are eligible for this award.

4. The following criteria shall be followed in selection of the awardee:-
   (a) Publication should be in Journal of Cytology.
   (b) Publication should be during the past one year and shall comprise of papers from Issues 3 and 4 of previous year and Issues 1 and 2 of current year.
   (c) Ahead of print articles shall not be considered for the award.
   (d) Publication should be an original article.
   (e) The awardee shall be the first author of the published paper.