Better Thyroid Cytopathology Reporting System May Increase the Clinical Management and Patients Outcome

Mulazim Hussain Bukhari1*, Asad Aslam Khan1, Shahida Niazi2, Madiha Arshad1, Zahid Mahmood Akhtar2 and Khalid Ahmed AL-Sindi2

1Department of Pathology, King Edward Medical University, Lahore, Pakistan
2Department of Pathology, King Hamad University Hospital, Bahrain

Abstract

Objective: The Bethesda System is better for Thyroid Cytopathology reporting and may increase the clinical management of Thyroid Disease.

Study design: Analytical Cross sectional.

Material and methods: A cross sectional study was conducted on 120 smears of thyroid swellings which were referred for fine needle aspiration to the Department of Pathology in 2011 and the Bethesda System for reporting Thyroid Cytopathology was followed in comparison to old conventional reporting system running from the last 30 years. Histopathology was used as a gold standard to compare the sensitivity of both systems. Three groups of Histopathologists were assigned three reporting systems without knowing the results of each other. The comparison was made in a meeting of 3 groups after histopathology biopsy reports were made available. Group A was assigned to report with the older system of thyroid aspiration with 7 categories, Group B reported with five classes of an older reporting system and Group C reported with the latest Bethesda Reporting System. Screening test was applied to compare the results.

Results: When the results of these three systems were compared Bethesda adapted method was found to be more superior as compared to the others. Sensitivity of Group B and C is significantly high 0.051 as compared to group A (p=0.051 and 0.000)’ Sensitivity of Group C is also significantly high as compared to Group B (p=0.000). Specificity of Group B is not significant (0.326) as compared to group A. Specificity of group C is also significantly high as compared to Group A and Group B (p=0.009 and 0.002). Our findings are consistent with others who used the Bethesda Cytopathology Reporting System.

Conclusions: Bethesda Cytopathology Reporting system can help with a better patient’s outcome due to proper clinical management of thyroid swellings and saves patients from unnecessary thyroid surgery.

Keywords: Thyroid gland swellings; FNAC; Thyroid cytopathology classification systems; Inadequate samples; Follicular lesions

Introduction

Palpable thyroid nodules may be found in 4–7% of the general population, and this prevalence may approach 60% when high-resolution Ultrasonography (USG) is used [1-3]. Fine needle aspiration cytology is a well-established technique for preoperative investigation of thyroid nodules. The technique is a noninvasive, cost-effective, and efficient specific and an excellent cost-effective method in the investigation of solitary thyroid nodules [4,5]. Indeterminate results, like suspicious for malignancy and follicular neoplasm or lesion [6], variability in reporting systems [7], and inadequate specimens limit the utility of FNAC and may complicate the management of thyroid nodules [8,9].

The history of Thyroid Fine Needle Aspiration Cytology (FNAC) at our institution is over 30 years old and reporting was modified with the passage of time [10]. The recent developments in the reporting system of Thyroid Fine Needle Aspiration Cytology (FNAC) is due to the need of the day because thyroid nodules are becoming more common day by day, however thyroid cancer is still comparatively rare [11].

“The literature shows wide variations in the criteria for inadequate thyroid FNAC and study of inclusion or exclusion criteria. In-clinic assessment of specimen adequacy and in-clinic reporting of thyroid FNAC has become popular although the costs and resource implications of in-clinic thyroid FNAC assessment and reporting are substantial” [12].

Standardized categorical systems for FNAC reporting can make results easier to understand for clinicians and give clear indications for therapeutic action [12,13].

Materials and Methods

A cross sectional study was conducted on 120 smears of thyroid swellings which were sent for fine needle aspiration to the Department of Pathology in 2011 and the Bethesda System for reporting thyroid cytopathology was followed in comparison to the old conventional reporting system running for the last 30 years. Histopathology was used as a gold standard to compare the sensitivity of both systems.

Three groups of Histopathologists were assigned three reporting systems without knowing the results of each other. The comparison
was made in a meeting of the three groups after biopsy reports were
made available.

**Inclusion criteria**

1. Age 10 to 70 Years.
2. Both genders.
3. Patients presenting with thyroid swelling in any lobe of thyroid
selected by clinical palpation (multinodular, solitary nodules,
diffuse goiter etc).
4. Patients with recurrent thyroid swellings after a previous
thyroid surgery.

**Exclusion criteria**

1. Patients with already diagnosed thyroid lesions.
2. All toxic goiters confirmed by clinical evaluation and laboratory
parameters.

All patients presenting with solitary thyroid nodules in the
OPD and fulfilling the inclusion criteria were included in this study.
Informed consent from all the patients included in the study was taken.
All included patients were recorded for their demographic features,
that is, age, sex, and address and telephone contacts (for follow up).
History of present illness with regard to symptoms and duration was
recorded. They were examined for signs related to the solitary thyroid
swelling. All routine investigations and serum T3, T4, and TSH levels
were performed by Radioimmunooassay (RIA), (normal range of T3,
2.5–5.8 nmol/L, T4, 11.5–23.0 nmol/L, and TSH, 0.2–4.0 mIU/L). Patients
with thyroid swellings also underwent a thyroid scan. Thyroid
swellings were marked through by the nuclear department and then
FNAC was performed [12].

**Group A**

Was assigned to report with the conventional system according to
which; Cytological diagnosis was categorized into the following 7
groups [10,12].

**Nondiagnostic or Unsatisfactory:** (when smears are hemorrhagic
or containing less than six groups of well-preserved follicular cells on
each of at least two slides.

- **Colloid goiter:** When smears contained follicular cells with
  abundant thick colloid in the background.

- **Colloid cysts:** When follicular cells, thin or thick colloid in the
  background and hemosiderin laden macrophages were seen in the
  smears.

- **Follicular lesions/Neoplasm:** When smears contain many
  follicular cells without or scanty colloid in the background or when
  smears contain predominant population of Hurthle cells, the
differential diagnosis would include hyperplastic adenomatoid nodule
  with Hurthle cell change, Hurthle cell adenoma, and Hurthle cell
  carcinoma.

- **Indeterminate smears:** When smears containing cells with findings
  that were not clearly benign but were not diagnostic of a neoplasm or
  malignant lesions.

- **Suspicious for malignancy:** Suspicious when aspirates suggest a
  follicular neoplasm, i.e., hypercellular sample with scant colloid and
  a significant proportion of microfollicules, trabeculae, or crowded
  overlapping clusters of follicular cells (also includes lesions consisting
  of oncocytic [Hurthle cell] neoplasms).

- **Malignant lesions**
  - Papillary Carcinoma
  - Medullary carcinoma
  - Anaplastic Carcinoma
  - Lymphoma
  - Metastatic

**Group B**

Was assigned to report with the following conventional system;
Cytological diagnosis was categorized into 5 categories:

**Unsatisfactory Smears** (Same as described in group) A

- **Benign or Negative for malignancy:** This group included thyroid
cysts, colloid goiters, thyroiditis and hyperplasia, benign when aspirates
were hypo cellular to moderately cellular with moderate to abundant
colloid and follicular cells with round nuclei of uniform size.

- **Follicular lesions:** Aspirates of Follicular patterned lesions other
  than follicular variant of papillary carcinoma; (Aspirates of processes
  which cannot be fully classified by FNAC as they require histological
  assessment for actual classification).

- **Indeterminate:** Aspirates showing some but not all features of
  malignancy, e.g., cell clusters with enlarged and grooved nuclei without
  true pseudo inclusions; and .The indeterminate group included
  follicular neoplasm’s, Hurthle cell neoplasms, and suspicious thyroid
  carcinoma.

- **Positive for malignancy:** Any Malignant category.

**Group C**

Was assigned to report thyroid FNAC by the Bethesda System
having the following six categories [14].

- **Nondiagnostic or Unsatisfactory:** Cyst fluid only virtually a
  cellular specimen other (obscuring blood, clotting artifact, etc).

- **Benign:** Consistent with a benign follicular nodule (includes
  adenomatoid nodule, colloid nodule, etc) Consistent with lymphocytic
  (Hashimoto) thyroiditis in the proper clinical context. Consistent with
  granulomatous (sub acute) thyroiditis others (Figure 1).

- **Atypical of Undetermined Significance (AUS) or Follicular Lesion
  of Undetermined Significance (FLUS)** (Figures 2 and 3).

- **Follicular neoplasm or Suspicious for a follicular neoplasm**
  Specify if Hürthle cell (oncocytic) type (Figure 3).

- **Suspicious for malignancy:** Suspicious for papillary carcinoma,
  Suspicious for medullar carcinoma, Suspicious for metastatic
  carcinoma, Suspicious for lymphoma (Figure 4).

- **Malignant:** Papillary thyroid carcinoma, Poorly differentiated
  carcinoma, Medullar thyroid carcinoma, Undifferentiated (anaplastic)
carcinoma, Squamous cell carcinoma, Carcinoma with mixed features (specify), Metastatic carcinoma, Non-Hodgkin lymphoma or others (Figure 4).

The cases were operated and evaluated for histopathological changes

Histopathological diagnoses of patients who had undergone surgery were used as the gold standard for correlation with the cytological interpretations. In the event where more than one nodule underwent biopsy on the same patient, the most abnormal FNA result was used for analysis. The specimens with discrepant cytological and histological diagnoses were reviewed to determine the plausible explanations of these discrepancies.

Statistical analysis

The sensitivity, specificity and diagnostic accuracy were calculated considering thyroid FNA as a 'screening test'; FNA specimens interpreted as benign were considered to be true negative samples and the remaining categories were considered to be true-positive samples because they led to a recommendation of surgery. The false-positive category included cases that were diagnosed as follicular and malignant.
but which were confirmed as benign on histopathological evaluation. The false-negative cases included those diagnosed as benign on FNA but confirmed as malignant upon surgical excision.

**Results**

Of the 120 specimens 13 samples (10.8%) were unsatisfactory for diagnosis as evaluated by the three groups. In Group A the old system was adopted where cytological categories were made as Colloid Goiter, Colloid cyst, Thyroiditis, Follicular/ Lesions Neoplasm, Indeterminate, Suspicious for Malignancy and frankly malignant smears. In this group false positive and false negative rate was much higher as compared to group B and C. The sensitivity and specificity was calculated as: Sensitivity: 77% and Specificity: 69% (Table 1 and 2).

In group B, only four categories were suggested like Unsatisfactory, negative for malignancy, Indeterminate and malignant smears. In this group false positive results improved to a certain extent but false negative results remained the same when compared with histopathology. The sensitivity was 85% and specificity was 65% (Table 3 and 4).

In group C latest system for evaluation of thyroid smears were performed which gave no false negative result with highest sensitivity, while false positive results also improved with higher specificity outcome as compared to the other two groups (Table 5 and 6).

When the sensitivities and specificity of two systems (Group A and B) were compared with the Bethesda Method (Group C), far better results were achieved (Table 7).

**Discussion**

Thyroid nodules are a common clinical problem and are noted much more frequently on imaging examinations than are apparent by palpation and clinical observation. Fine-Needle Aspiration Biopsy (FNAB), which yields a cytology specimen for analysis, is the standard test to determine whether surgical removal of a detected nodule is recommended or not. Fortunately, the vast majority of nodules are benign, but when they are discovered, an assessment regarding the need to exclude malignancy using FNA must be performed [14].

Various reporting systems for thyroid cytology have been adopted at our institution for the last three decades but none of them was related to the prognosis of disease and patients outcome[10,12]. These reporting schedules were least informative due to variability of sensitivity and least reproducibility. The introduction of new the simplified Bethesda Thyroid Reporting System into six categories logically relates to the prognosis of thyroid diseases and may increase the reproducibility of diagnosis [11].

<table>
<thead>
<tr>
<th>Group A Lesions ( on FNA REPORTING)</th>
<th>Numbers</th>
<th>Histopathology</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Benign</td>
<td>Malignant</td>
</tr>
<tr>
<td>Non-diagnostic or Unsatisfactory</td>
<td>13</td>
<td>Excluded</td>
<td>Excluded</td>
</tr>
<tr>
<td>Colloid Goiter</td>
<td>50</td>
<td>47</td>
<td>3</td>
</tr>
<tr>
<td>Colloid cyst</td>
<td>16</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Thyroiditis</td>
<td>5</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Follicular/ Lesions Neoplasm</td>
<td>18</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>Indeterminate</td>
<td>5</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Suspicious for Malignancy</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Malignant Lesions</td>
<td>8</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Totals</td>
<td>120</td>
<td>90</td>
<td>17</td>
</tr>
</tbody>
</table>

**Table 1:** Distribution of different lesions diagnosed by groups A of Histopathologists.

<table>
<thead>
<tr>
<th>Cytopathology</th>
<th>Histopathology</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>FNAC +ve</td>
<td>17</td>
<td>30</td>
</tr>
<tr>
<td>FNAC –VE</td>
<td>05</td>
<td>68</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>98</td>
</tr>
</tbody>
</table>

Sensitivity: 77%; Specificity: 69%, Positive predictive value:37%, Negative predictive value 93%

<table>
<thead>
<tr>
<th>Group B Lesions (FNA-cytology)</th>
<th>Numbers</th>
<th>Histopathology</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory Smears</td>
<td>13</td>
<td>Nil</td>
<td>Excluded</td>
</tr>
<tr>
<td>Negative for malignancy</td>
<td>69</td>
<td>66</td>
<td>3</td>
</tr>
<tr>
<td>Follicular Lesions</td>
<td>20</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>Indeterminate</td>
<td>8</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Malignant Lesions</td>
<td>10</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Totals</td>
<td>120</td>
<td>90</td>
<td>107</td>
</tr>
</tbody>
</table>

**Table 2:** Comparison of FNAC of Group A with histopathology.

<table>
<thead>
<tr>
<th>Cytopathology</th>
<th>Histopathology</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>FNAC +ve</td>
<td>17</td>
<td>35</td>
</tr>
<tr>
<td>FNAC –VE</td>
<td>03</td>
<td>65</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>100</td>
</tr>
</tbody>
</table>

Sensitivity: 85%, Specificity: 65%, Positive predictive value:32%, Negative predictive value 95.5%

<table>
<thead>
<tr>
<th>Group C CYTOPATHOLOGY</th>
<th>Histopathology</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Non-diagnostic or Unsat</td>
<td>NIL</td>
<td>13</td>
</tr>
<tr>
<td>II Benign</td>
<td>72</td>
<td>0</td>
</tr>
<tr>
<td>III Follicular Neoplasm or Suspicious for a Follicular Neoplasm</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>IV Atyopia of Undetermined Significance or Follicular Lesion of Undetermined Significance</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>V Suspicious for Malignancy</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>VI Malignant</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>90</td>
<td>107</td>
</tr>
</tbody>
</table>

**Table 3:** Distribution of different lesions diagnosed by groups B of Histopathologists.

<table>
<thead>
<tr>
<th>Cytopathology</th>
<th>Histopathology</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>FNAC +ve</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>FNAC –VE</td>
<td>0</td>
<td>85</td>
</tr>
<tr>
<td>Total</td>
<td>17</td>
<td>103</td>
</tr>
</tbody>
</table>

Sensitivity: 100%, Specificity: 82.5%, Positive predictive value:45%, Negative predictive value 100%

**Table 4:** Comparison of FNAC of Group B with histopathology.

FNAC of the thyroid is the key preoperative investigation of thyroid lesions. There are overlaps in the criteria for diagnosis of certain lesions, particularly important regarding those reported as follicular neoplasm’s. Thyroid FNAC allows binary triaging for surgery, to enable a decision to operate or not, albeit with some subtleties as some cases may require re-aspiration or reassessment after a period of time. The reporting of thyroid FNAC, however, is in many cases not binary. Various diagnostic category systems for reporting FNAC have been reviewed recently by Dr Helen Wang [15,16].

These data demonstrate that the recently introduced Bethesda classification system is excellent for reporting thyroid FNAs. Each
for Reporting Thyroid Cytopathology is a standardized reporting system that is based upon numbers or stepwise descriptions. The Bethesda System for Thyroid Cytopathology is the new system for thyroid FNAC. This would allow for better assessment of adequacy and an acceptable rate of inadequate diagnoses.

We experienced difficulty in reporting two categories of Bethesda system for thyroid cytopathology reporting namely Atypia of undetermined significance (IV) and Follicular Neoplasm or Suspicious for a Follicular Neoplasm (V). When results of these three systems were compared, Bethesda adapted method was found to be more superior as compared to others. Sensitivity of Group B and C is significantly higher (p=0.051 and 0.000) as compared to Group A (p=0.051 and 0.000). Specificity of C is also significantly high as compared to B (p=0.000) Specificity of Group B is not significant (0.326) as compared to group A. Specificity of C is also significantly high as compared to A and B (p=0.009 and 0.002).

<table>
<thead>
<tr>
<th>Classes</th>
<th>Description</th>
<th>Risk of malignancy (Percentage)</th>
<th>Usual management</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Non-diagnostic or Unsatisfactory</td>
<td>1-4</td>
<td>Repeat</td>
</tr>
<tr>
<td>II</td>
<td>Benign</td>
<td>0-3</td>
<td>Follow up</td>
</tr>
<tr>
<td>III</td>
<td>Follicular Neoplasm or Suspicious for a Follicular Neoplasm</td>
<td>5-15</td>
<td>surgical lobectomy</td>
</tr>
<tr>
<td>IV</td>
<td>Atypia of Undetermined Significance or Follicular Lesion of Undetermined Significance</td>
<td>15-30</td>
<td>surgical lobectomy</td>
</tr>
<tr>
<td>V</td>
<td>Suspicious for Malignancy</td>
<td>60-75</td>
<td>near-total thyroidectomy or surgical lobectomy</td>
</tr>
<tr>
<td>VI</td>
<td>Malignant</td>
<td>100</td>
<td>near-total thyroidectomy or surgical lobectomy</td>
</tr>
</tbody>
</table>

Table 7: Comparison of Sensitivity and specificity of all systems Group A, B, and C.

We improved our system a decade back and results were published in medical journals. Our journey was to make it simple for the understanding of consultants to propose their line of management. A simplified reporting scheme would, undoubtedly, have to address the assessment of adequacy and an acceptable rate of inadequate diagnoses could then be determined [10]. At the present time there is no accepted definition of what constitutes an inadequate thyroid FNAC, or how cystic thyroid lesions should be classified. In the old system the false positive and false negative results were also higher as compared to the newly adopted Bethesda system. The sensitivity was 85% and specificity was 65% in Group B where the old system was adopted.

The vast array of diagnostic nomenclature currently in use can usually be made to fit into these systems and thus easily explained to clinicians. There is now a need for a more unified approach to the reporting of thyroid FNAC. This would allow for better assessment of how FNAC diagnoses relates to therapy and outcome and for the development of truly evidence-based treatment recommendations. In the newly adopted system we used the six tire technique based on Bethesda Reporting system for Thyroid Cytopathology. The reporting is based upon numbers or stepwise descriptions. The Bethesda System for Reporting Thyroid Cytopathology is a standardized reporting system for classifying thyroid fine-needle aspiration results comprising of 6 diagnostic categories with unique risks of malignancy and recommendations for clinical management. Like (I) Nondiagnostic (II) Benign (III) Aspirates of atypia/follicular lesion of undetermined significance (IV) Follicular neoplasm/suspicion for a follicular neoplasm (V) Suspiciously malignant aspirates and (VI) Malignant aspirates [11-13,19]. The sensitivity and negative predictive values proved to be 100% with no false positive results and high (82.5%) specificity with low false negative values. The findings are consistent with the study of Wong et al. [19].

When results of these three systems were compared, Bethesda adapted method was found to be more superior as compared to others. Sensitivity of Group B and C is significantly higher (p=0.051 and 0.000) as compared to Group A (p=0.051 and 0.000). Specificity of C is also significantly high as compared to B (p=0.000). Specificity of Group B is not significant (0.326) as compared to group A. Specificity of C is also significantly high as compared to A and B (p=0.009 and 0.002).

<table>
<thead>
<tr>
<th>Categories</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Significance values</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>77</td>
<td>69</td>
<td>Sensitivity of Group B and C is significantly higher 0.051 as compared to group A (p=0.051) and 0.000)</td>
</tr>
<tr>
<td>B</td>
<td>85</td>
<td>65</td>
<td>Sensitivity of Group B and C is also significantly high as compared to B (p=0.000) Specificity of Group B is not significant (0.326) as compared to group A. Specificity of C is also significantly high as compared to A and B (p=0.009 and 0.002)</td>
</tr>
<tr>
<td>C</td>
<td>100</td>
<td>82.5</td>
<td></td>
</tr>
</tbody>
</table>

Note: While specificity of three groups is non significant 0.278, 0.236 and 0.316 respectively.

Table 8: Diagnostic categories, associated risk of malignancy and clinical management [20].

We used different methods of reporting since 1990, and we tried to improve our reporting technique because we found that our surgeons had difficulty in understanding the complexities of our thyroid cytopathology reporting. We also understand that this problem may lead to difficulty in the management of patients with follicular neoplasm, follicular lesions etc. In Pakistan, Agha Khan University Hospital and Shaukat Khanum Research Centre have also developed their own in house reporting systems for thyroid FNAC [12,18].

When we compared our different reporting systems, we concluded that the journey is still long to approach a final destination. The false positive reports were commoner in our previous two decades older techniques, where the screening tests show 77% sensitivity and 69% specificity. The older system did not help in the prognosis of disease and patients outcome especially in Follicular lesions.

Conclusion

By Adapting the Bethesda Cytopathology Reporting system a high sensitivity and high negative predictive values can help to determine a better patient outcome due to proper clinical management of thyroid swellings.

Problems experienced in description of Bethesda reporting systems and suggestions

We experienced difficulty in reporting two categories of Bethesda system for thyroid cytopathology reporting namely Atypia of undetermined significance or follicular lesions of undetermined significance and Follicular Neoplasm or Suspicious for a Follicular Neoplasm specify if hurthle cell (oncocytic) type, therefore we suggest both classes should be submerged in one category and a new name should be proposed. This will reduce the false positive results if both categories are reported under one new class. Our suggestion is to name...
it as Atypica of undetermined Significance with Follicular Neoplasm. Both benign and malignant follicular lesions should be reported under this heading along with the hurthle cell lesions.

References