Diagnostic accuracy of fine needle aspiration cytology of thyroid and evaluation of discordant cases

Chetna Sharma *

Department of Pathology, PSG Institute of Medical Sciences, Coimbatore, India

Received 23 April 2015; revised 30 May 2015; accepted 1 June 2015
Available online 14 July 2015

Abstract Introduction: The main role of fine needle aspiration cytology (FNAC) lies in differentiating between a malignant and benign thyroid nodule. It greatly influences the treatment decision. The current study was undertaken to evaluate the cytology–histopathology correlation and to analyze the cause of diagnostic errors with an eventual aim to improve diagnostic accuracy.

Materials and Methods: This is a retrospective study comparing cytology and corresponding histopathology report in 724 thyroid cases. The statistical analysis included false positive rate, false negative rate, sensitivity, specificity, positive predictive value, negative predictive value and accuracy.

Results: On cytological examination, 635/724 were reported as benign, 68 malignant and 21 suspect. On histopathological examination, 626/635 cases were confirmed as benign but there were 9 discordant cases. Among the other cases histopathology diagnosis of malignancy matched in 66/68 and 11/21 cases. Diagnosis correlated in 703/724 cases (97%) \(p<0.001\). False positive and false negative rates were 1.9% and 10.5%, respectively. The sensitivity and specificity were 89.5% and 98%, respectively. The positive predictive value was 84.6% and negative predictive value was 98.6%. Accuracy of FNA was 97%.

Conclusion: In spite of high accuracy of FNAC in differentiating between a benign and malignant lesion, certain pitfalls should be kept in mind. The common false negative diagnoses were follicular pattern cases which constitute a ‘gray zone’, cystic papillary thyroid carcinoma (PTC) and papillary microcarcinoma. The reason for false positive diagnoses was the occurrence of nuclear features characteristic of PTC in other thyroid lesions. Awareness of pathologist regarding these pitfalls can minimize false negative/positive diagnoses.

© 2015 The Author. Production and hosting by Elsevier B.V. on behalf of National Cancer Institute, Cairo University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Today it remains the mainstay of diagnostic workup for thyroid pathologies. Thyroid FNA is a safe and relatively simple test that is widely recognized as the critical, primary diagnostic procedure of first choice for the evaluation of patients presenting with thyroid nodules [2–7]. One of the major advantages is that FNAC can be done as an out-patient procedure.

Thyroid nodules are a common clinical finding and have a reported prevalence of 4–7% in the general population [2,8,9]. The vast majority of these nodules are non-neoplastic lesions or benign neoplasms. FNAC is relied upon to distinguish benign from neoplastic or malignant thyroid nodules and therefore has lead to a dramatic decrease in thyroid surgeries [10]. Nevertheless, Fine needle Aspiration has some limitations like specimen inadequacy, sampling techniques and WHAFFT changes [11–14]. Cytopathologist should be aware of these potential limitations and pitfalls of FNA interpretation.

The present study was performed with the aim to evaluate the value of FNA in differentiating benign and malignant lesions of thyroid. Our another aim was to analyze the false positive and false negative diagnosis so as to highlight the pitfalls of FNA and probable reasons for the same are discussed. We hope that a better understanding of these pitfalls would help avoiding them in future and will contribute to better patient care.

**Materials and methods**

This is a retrospective review of all thyroid FNAs done at our institution between January 2010 and December 2014. Data were retrieved from the institutional database and analyzed. There were 945 thyroid FNAs done during this period. Ultrasound guided FNAs at our institution are very infrequent and most of the FNAs in the present study were without ultrasound guidance. This cytological study was based on evaluation of FNA smears. In patients with multiple nodules, even if one nodule was malignant and the other benign it was reported as malignant in cytology and considered in the malignant category.

Cytological diagnosis of thyroid lesions at our institution is influenced by the guidelines of Royal college of Pathologists [15]. The cytological results were categorized into 5 categories: unsatisfactory; benign; follicular pattern lesions; suspicious (includes atypical) and malignant.

Samples were considered “unsatisfactory” if there was insufficient cellularity; cellularity was obscured by blood, poor quality smears or delayed/inadequate fixation. For cystic lesions, smears with insufficient cellularity i.e. if there were less than six groups of thyroid follicular epithelial cells across all the submitted slides, each with at least 10 well visualized epithelial cells and which contain mostly macrophages but without abundant colloid were included in the unsatisfactory category [15].

Cystic lesions which were not unsatisfactory by the above definition were considered satisfactory and considered for this correlation study. As per the Royal College guidelines [15], these include:

- Cystic lesion fluid samples which have sufficient thyroid follicle cells to achieve the adequacy criterion, irrespective of any possible colloid and/or macrophage content
- Cystic lesion specimens which consist predominantly of colloid and macrophages, even if too few follicular epithelial cells are present to meet the adequacy criteria outlined above, can be considered to be ‘consistent with a colloid cyst’ in the appropriate clinical setting.

At our institute, all the cystic lesions with ‘atypia’ are reported as “suspicious” of malignancy. As this paper was specifically aimed to find out the efficacy of FNA in ruling out malignancy; to keep the comparisons clear we have included all the cysts without atypia in the benign category and those reported as “suspicious” in malignant category.

‘Follicular pattern lesion’ forms a ‘gray zone’ in cytology and it is difficult to rule out malignancy on cytology in these lesions [16,17] We report these lesions as ‘follicular pattern lesions’ and advice excision of the lesion followed by HPE for confirmation.

If atypical cells are seen it is mentioned additionally in the report to provide a clear picture to the surgeon. For the purpose of this study the follicular pattern lesions with atypical cells are classified in ‘suspicious category’ as they are managed the same way i.e. surgical excision and HPE. In follicular pattern lesions without atypia clinical judgment should play a large role in deciding to watch, repeat FNA or even excise in certain circumstances. Since there was no evidence of malignancy on cytology in these cases we have grouped them in benign category for this study.

As the present study aims to correlate the accuracy of the FNA in detecting the malignant/benign lesions on cytology the follicular lesions without atypia were categorized as benign and follicular lesions with atypia as ‘suspicious’. Suspicious lesions have been clubbed with malignant ones as both are treated as same as far as treatment is concerned. Such lesions are advised surgery followed by histopathological examination for definite diagnosis.

The corresponding histopathology report was available for 724 cases. The patients who underwent either of the two tests in any other institution were excluded from the study. The FNA samples which were inadequate for evaluation were also excluded. A total of 724 cases who had both cytology and histopathology reports available formed the study group to analyze the value of FNA in thyroid pathologies. The statistical analysis included false positive rate, false negative rate, sensitivity, specificity, positive predictive value, negative predictive value and accuracy.

**Results**

945 thyroid FNAs were done in our institute between 2010 and 2014. Most of the patients were in the age group 30–50 years (Fig. 1). There were 82% female and 18% male patients (Fig. 2). Out of these 945 cases, the FNA sample was adequate in 870 cases. Histopathology correlation was available for 724 cases, which formed the study population. Among the 724 cases 635 (87.7%) were reported as benign on cytology and the diagnosis was as follows: 324 (44.8%) cases were nodular goiter, 232 (32.0%) were Hashimoto’s thyroiditis, 48 (6.6%) were follicular lesions and 44 (6.0%) were cystic lesions of thyroid (Fig. 3). 68 cases out of 724 were diagnosed as malignant on cytology. Of these, 62 (8.6%) cases were reported as PTC and 6 (0.8%) as medullary carcinoma (Fig. 3). 21 (2.90%)
out of 724 cases were in the suspicious/atypical cells category, where a possibility of malignancy was suggested and surgical excision was advised for a definite diagnosis (Fig. 3).

HPE of 635 benign cytology cases revealed 626 cases as benign and 9 cases as malignant (Table 1). The HP diagnosis for these 9 cases was as follows: 3 FVPTC, 2 conventional PTC, 2 cystic PTC, 1 follicular carcinoma and 1 Hashimoto’s thyroiditis with papillary microcarcinoma (Table 2). Out of 62 cases with a FNA diagnosis of PTC, 60 proved to be the same on HPE but the HP diagnosis differed in 2 cases and was that of nodular goiter (Tables 1 and 2). The cytology diagnosis correlated with HP diagnosis for all medullary carcinomas. Out of the 21 suspicious/atypical cases almost half of the cases i.e. 11 cases were malignant on HPE, but, 10 cases came out to be benign contrary to the cytological diagnosis (Table 1). 5 out of the 11 malignant cases were conventional PTC and 6 were follicular variant of PTC. Among the 10 benign cases, 7 were nodular goiter or adenomatous goiter, 2 were follicular adenomas and 1 case was Hashimoto’s thyroiditis (Table 2).

---

**Table 1**  Cyto-histopathology correlation of all the cases.

<table>
<thead>
<tr>
<th>Cytological diagnosis</th>
<th>Histopathology diagnosis</th>
<th>Discordant cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign (N = 635)</td>
<td>Benign (N = 626)</td>
<td>9</td>
</tr>
<tr>
<td>[TN]</td>
<td>[FN]</td>
<td></td>
</tr>
<tr>
<td>Malignant (N = 68)</td>
<td>Malignant (N = 66)</td>
<td>2</td>
</tr>
<tr>
<td>[TP]</td>
<td>[FP]</td>
<td></td>
</tr>
<tr>
<td>Suspicious (N = 21)</td>
<td>Malignant (N = 11)</td>
<td>10*</td>
</tr>
<tr>
<td>[TP]</td>
<td>[FP]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Benign (N = 10)</td>
<td></td>
</tr>
</tbody>
</table>

TN – True Negative; FN – False Negative; TP – True Positive; FP – False Positive.
* The suspicious cases were included in the malignant category for statistical analysis as such cases are usually treated by clinicians as malignancies.

**Table 2**  Details of the discordant cases.

<table>
<thead>
<tr>
<th>Discordant cases (No.)</th>
<th>Cytological diagnosis</th>
<th>Histopathology diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Benign</td>
<td>Malignant (3-FVPTC; 2-Conventional PTC; 2-Cystic PTC; 1-Follicular Carcinoma; 1-Hashimoto’s with papillary microcarcinoma)</td>
</tr>
<tr>
<td>2</td>
<td>Malignant (PTC)</td>
<td>Benign (2-Nodular Goiter)</td>
</tr>
<tr>
<td>10</td>
<td>Suspicious</td>
<td>Benign (7-Nodular Goiter; 2-Follicular adenoma; 1-Hashimoto’s thyroiditis)</td>
</tr>
</tbody>
</table>

Total = 21

Statistical analysis was done to find out the value of FNA in detecting malignancy in thyroid swellings.

**Statistical analysis included**

- Sensitivity (probability that a test result will be positive when the disease is present – ‘true positive rate’Sensitivity = True positive/(True positive + False negative) = 77/(77 + 9) = 89.5%)
- Specificity (probability that a test result will be negative when the disease is not present – ‘true negative rate’Specificity = True negative/(True negative + False positive) = 626/(626 + 12) = 98%)
- Positive predictive value (probability that the disease is present when the test is positive)Positive predictive value (PPV) = True positive/(True positive + False positive) = 66/(66 + 12) = 84.6%
- Negative predictive value (probability that the disease is not present when the test is negative)Negative predictive value (NPV) = True negative/(True negative + False negative) = 626/(626 + 9) = 98.6%
- False positive rate = False positive/(False positive + True negative) = 12/(12 + 626) = 1.9%
- False negative rate = False negative/(False negative + True positive) = 9/(9 + 77) = 10.5%
- Accuracy is the portion of the correct results, true positive and true negative in relation to all cases

**Figure 1**  Age distribution in our study group.

**Figure 2**  Sex distribution in our study group.

**Figure 3**  Distribution of FNA diagnosis in 724 cases.
studied. Accuracy = True positive + True negative/Total number of cases = 77 + 626/724 = 97%.

To summarize the results: False positive rate was 1.9% and false negative rate was 10.5%. The sensitivity and specificity were 89.5% and 98% respectively. The positive predictive value was 84.6% and negative predictive value was 98.6%. Accuracy of FNA in differentiating benign from malignant thyroid lesion was 97%.

Discussion

Fine needle aspiration cytology (FNAC) is regarded as a gold standard in the initial diagnosis of thyroid nodules. It is simple, reliable, time saving, minimally invasive and cost effective [1–7]. However there are some limitations of FNAC which a pathologist must be aware of [11–14]. The main information one wants from FNAC is to distinguish a malignant lesion from a benign one. This distinction has dramatically reduced the surgery rates in thyroid pathologies [10]. Though the accuracy rates as high as 95–98% have been reported in best hands FNAC has some limitations, and false positive and false negative results do occur [3,5].

The reporting systems for thyroid cytology vary among institutions and include 4 category system [2], 5 category system [15] or 6 category systems [18]. The most widely used ones are the Bethesda system [18] (6 category) and the Royal College of Pathologist [15] (5 category) system.

The Royal College of Pathologists thyroid fine-needle aspiration diagnostic classification is a robust tool for the clinical management of abnormal thyroid nodules with established reproducibility [19,20]. Cytological diagnosis of thyroid lesions at our institution is influenced by the guidelines of Royal college of Pathologists. Use of terminologies like – undetermined significance, atypical leads to confusion for the clinicians and are avoided at our institution. We categorize lesions as, unsatisfactory; benign; follicular pattern lesions; suspicious (includes atypical) and malignant.

The current study was undertaken to evaluate the correlation between cytology and histology in thyroid FNA in our institute and to analyze the cause of the diagnostic errors with an eventual aim to improve accuracy. It is of utmost importance because the treatment is greatly influenced by the FNAC report. The surgical rate for benign nodules in our study is much higher because our institution is a surgical referral center and most of the patients are referred to the hospital for surgical intervention. Most of the patients would have had initial workup at the primary hospital and referred for surgical intervention to our surgical team and FNAs are done to reassess the diagnosis at our institution before surgical intervention.

As FNAC is mainly aimed to rule out malignancy it should have a low false-negative rate, acceptable sensitivity and specificity for detection of malignancy and high negative predictive value. The reported sensitivity of thyroid FNA ranges from 65% to 99% and its specificity from 72% to 100% [10,21–25]. In addition, there are some studies in the literature reporting sensitivity as low as 55.3% and false negative as low as 44.7% [26]. The likely reasons reported for lower sensitivity in series of Mistry et al. were combination of operator variability, low number (56 cases) and diagnostic difficulty of using FNAC in certain thyroid pathologies [26]. This shows that there is significant disparity in reported thyroid FNAC statistics [27]. The determinant factors for wide range of difference could be number of cases, the included diagnostic categories and how the cytopathologist classifies the suspicious lesions [28].

There is disparity in the literature in the way the cystic lesions are categorized; some categorize all the cystic lesions in unsatisfactory category whereas others do not [2,15]. In our institution, hypocellular aspirates in the presence of cystic fluid are not considered to be non-diagnostic because in these cases the presence of no or a few cells is correlated with the structure of the nodule itself and not with an inadequate sampling technique. We agree with the Royal College guidelines [15] and follow them in our institution because for such lesions if they are labeled as ‘unsatisfactory’ it could imply that the material is not appropriate. In such lesions even a repeat sample would be ‘unsatisfactory’. Hence, it is more appropriate to not include them in ‘unsatisfactory’ category.

All the cystic lesions with ‘atypia’ were reported as ‘suspicious’ of malignancy. This is in accordance with the work published by Ravetto where specimens were diagnosed as ‘suspicious’ when cells showed atypical features suggestive of, but not diagnostic for, malignancy [7]. As our study specifically aimed to find out the efficacy of FNA in ruling out malignancy; we have included all the cysts without atypia in the benign category and those reported as ‘suspicious’ in malignant category. Distribution of the cysts into either a benign or those with ‘atypia’ in malignant category allowed easier comparison and clearer outcome.

In the present study the cases reported as ‘suspicious’ on cytology were included in the malignant category for statistical analysis because both lead to surgical management as far as the treatment is concerned [27]. This allows for an easier comparison and clearer final results.

In our study the sensitivity of FNAC was 89.5% and specificity was 98%. This indicated that ability of FNA to detect malignancy in our series was quite high. The comparable studies in literature report sensitivity in the range of 82–93.4% and specificity of 74.9–96% [2,7,11,12,21,28,29].

The positive predictive value in our series was 84.6% compared to 85.7–98.6% in other studies [11,28,29]. The negative predictive value in our series was 98.6% as compared to 91.8–94% in similar studies in literature [12,28,29]. The accuracy of FNA in detecting malignancy in thyroid lesions in our study was 97% which reinforces that FNAC can be used as a reliable tool to detect thyroid malignancy. Other studies in the literature report accuracy ranging from 83.6% to 93.6% and support our results [12,28,29].

The reported false-negative rate ranges from 1% to 19% [10,22,25,28,30–32]. However, it is difficult to know the true frequency of false-negative results because only a small percentage (approximately 10%) of patients with benign cytologic findings undergo surgery [10,25]. It has been postulated that the true false-negative rate is below 5% if all patients with thyroid FNA also have a histologic examination [25]. In our study false negative rate was 10.5%. Our false-negative rate is in the higher side of reported range in literature probably because of the same fact. Also, this reflects the limitation of FNAC in accurately diagnosing follicular pattern lesions, cystic papillary thyroid carcinoma (PTC) and papillary microcarcinoma.
The follicular pattern lesions include – hyperplastic nodule (nodular goiter), follicular neoplasm (follicular adenoma and carcinoma) and follicular variant of PTC. Use of various terms, including “atypical,” indeterminate, favor, cannot exclude, possible, or probable and suspicious by pathologists often creates confusion for the clinicians [33]. It is truly a gray zone in thyroid cytology [16,17]. We report these lesions as ‘follicular pattern lesions’ and advice excision of the lesion followed by HPE for confirmation.

As the present study is to correlate the accuracy of the FNA in detecting the malignant/benign lesions on cytology the follicular lesions without atypia were categorized as benign and follicular lesions with atypia as ‘suspicious’. Suspicious lesions have been clubbed with malignant ones as both are treated as same as far as treatment is concerned.

Our practice regarding use of the category “atypical” is similar to MMC (Montefiore Medical Center (MMC) [33] and includes all potentially neoplastic follicular lesions that do not show overt papillary features with the expectation that these patients will be offered a hemithyroidectomy in the appropriate clinical setting. These lesions in our study are included in the ‘suspicious category’ as this indicates the need for surgery and confirmation by HPE.

Studies in the past have highlighted the fact that in follicular variant of papillary thyroid carcinoma (FVPTC), the nuclear features essential for a cytologic diagnosis of PTC are infrequent [34–37] and sensitivity of FNA in establishing a diagnosis of FVPTC is low [38]. These concerns in the literature are further affirmed in our observation study where FVPTC cases accounted for 3 of the 9 discordant cases with benign cytology. Even the presence of infrequent nuclear features of PTC should alarm the pathologist to consider the diagnosis of FVPTC which is a major diagnostic challenge. As FNA is considered a screening procedure, particular attention should be given to minimizing false-negative diagnoses, even at the expense of accepting false-positive diagnoses [39].

False negative diagnosis can also occur due to inadequate sampling or due to aspiration of fluid from cystic lesions of thyroid with an underlying malignancy. This could be the reason for another two false negative cases (out of 9 discordant cases with benign cytology) in our series (Table 2). Among all the thyroid cancers, PTC tends to undergo marked hemorrhagic degenerative changes. Sampling of this hemorrhagic fluid with sparse tumor cells may result in false interpretation as a benign cyst [40]. In these two cases, suggestion for surgical excision had been given by the pathologist in order to rule out an underlying malignancy, which led to timely surgery and saved the patient of undue complications of PTC.

Although, fine-needle biopsy is the best predictor of malignancy in either cystic or solid thyroid lesions, it is slightly less reliable when a thyroid lesion is fluid filled rather than solid [41]. A close follow up or thyroid lobectomy for diagnosis should be strongly considered in these patients even when FNA cytologic finding is interpreted as benign [42,43]. Careful attention to details of the procurement process will have a significant impact on both inadequacy and false-negative rates. Meticulous examination of all the smears is of paramount importance in reducing discrepant cases [44]. Any recurrent cystic lesion should raise a strong suspicion for malignancy and should be treated so.

One of the other false negative cases (out of total 9) was reported as Hashimoto’s thyroiditis on cytology. This case on HPE was found to be papillary microcarcinoma of thyroid. The term papillary micr0carcinoma is used when the papillary carcinoma is an incidental finding and measures less than one centimeter in diameter. Papillary microcarcinoma arising in Hashimoto’s thyroiditis or in any other benign lesion has high chances of being missed because FNAC is mostly done as a blind procedure and as a result of which the aspirated sample may not be fully representative of the existing lesion. Papillary microcarcinomas of the thyroid (i.e. papillary carcinomas <1 cm in size) are frequent; autopsy surveys have revealed prevalence of up to 30% in high risk groups [45]. However, their clinical impact is limited in the vast majority of cases and do not cause a significant reduction in survival [7].

One another false negative case which was reported as follicular lesion on FNA turned out to be follicular carcinoma. This is one of the pitfalls of FNA, as FNA cannot distinguish benign follicular nodules from follicular carcinomas because the criteria to distinguish between them are based upon histologic evidence of transcapsular or vascular invasion which cannot be assessed on cytology [46].

The false positive rate reported in the literature range from 6% to 8% [10,27,28]. In our study the rate is very low (1.9%). Out of our 12 cases of false positive diagnosis, six were diagnosed as PTC or suspicious of PTC on FNA. The HP diagnosis of these cases was adenomatous goiter in 5 cases and Hashimoto’s thyroiditis in one case. The reason accountable for these false positive diagnoses in our study and also in the literature was the presence of features diagnostic of PTC even in benign conditions of thyroid. There is evidence in the literature that the characteristic features of PTC like nuclear grooves, nuclear pseudo-inclusions and many more features can be seen in benign conditions of thyroid like adenomatous goiter, Hashimoto’s thyroiditis, nodular goiter, follicular neoplasm, etc [18,47]. However, studies have stated that occurrence of these features in benign lesions is very less as compared to PTC [48]. Hence, awareness of pathologists in this context and strict adherence to adequacy criterion can reduce the false positive rate.

The other six cases (out of 12 false positive cases) had a cytological diagnosis of suspicious follicular lesion, of these four turned out to be adenomatous goiter and two as follicular adenoma on HPE. The diagnosis of follicular patterned lesions can be challenging because of overlapping features between benign and malignant lesions and is considered as a ‘gray zone’ [16,33,49].

Conclusion

FNAC is a highly reliable and accurate tool to differentiate a malignant lesion from a benign one with accuracy as high as 97%. However, certain pitfalls of FNA should be kept in mind while reporting. The most common false negative diagnosis was in cases with follicular pattern which constitute a ‘gray zone’. Diagnosis of FVPTC requires meticulous examination of smears in order to identify the nuclear features. Follicular neoplasms i.e. follicular adenoma and carcinoma cannot be distinguished on FNA and excision is mandatory for a definite diagnosis, which should be indicated in FNA report. Cystic lesions of thyroid should undergo surgical excision or close follow up as these can harbor PTC, which can be missed as FNA smears may not be representative. Owing to their small size.
papillary microcarcinomas also can get missed in FNA. The commonest reason for false positive diagnosis is the occurrence of nuclear features characteristic of PTC in other thyroid lesions like nodular goiter, Hashimoto’s thyroiditis, follicular neoplasms, etc. However, occurrence of these features in benign lesions is very less as compared to PTC and a pathologist should be aware of this fact. Nevertheless, as FNA is considered a screening procedure, particular attention should be given to minimizing false-negative diagnoses.

Conflict of interest

We have no conflict of interest to declare.

References


