To operate or not to operate? The value of fine needle aspiration cytology in the assessment of thyroid swellings

N Leonard, D H Melcher

Abstract

Aims—To evaluate all thyroid fine needle aspirations (FNA) done over a six year period to assess the accuracy and value of the technique.

Methods—There were 335 FNAs of which 184 had subsequent histology and 49 others had clinical follow up, providing 233 patients for analysis. All cytology and histology was reviewed with no significant alterations in diagnosis. The FNAs were classified into three groups: benign, suspicious (recommend excision), and malignant. The histology and medical records were reviewed to determine whether the cytology was accurate.

Results—There were 130 benign FNAs, 126 had non-malignant histology or normal clinical follow up, and four had malignancies on histology (two lymphomas, one follicular carcinoma, and one carcinoma not otherwise specified). There were 45 suspicious FNAs. Of these five had either follicular or papillary carcinoma, 14 had follicular adenomas, and 26 had colloid nodular goitres or normal clinical follow up. Of the 21 malignant FNAs, 11 had carcinoma and 10 had either a non-malignant histology or normal follow up. There were 126 true negatives, 30 true positives, 4 false negatives, and 36 false positives. This gives a sensitivity of 88%, a specificity of 78%, a positive predictive value of 46%, a negative predictive value of 97%, and an accuracy of 80%.

Conclusions—FNA cytology of the thyroid has a high negative predictive value, which is useful to reassure the majority of patients presenting with thyroid enlargement. However, a negative FNA should never exclude malignancy if there is a strong clinical suspicion. If this rule is adhered to a large number of patients will be spared unnecessary surgery and no malignant nodule will go untreated.

Keywords: fine needle aspiration; thyroid nodules; predictive value

Thyroid nodules are common, especially in women, and occur spontaneously in 4–7% of the adult population.1 Between 0.1% and 20% of nodules that are surgically resected contain carcinoma. Few people die from thyroid carcinoma and thyroid surgery has a number of risks, therefore, an accurate screening test is needed to select those nodules that are more likely to need an operation.

Fine needle aspiration (FNA) cytology was first described in the 1930s by Martin and Ellis.1 It is a standard procedure used in the assessment of patients with a thyroid swelling and it can answer two clinical questions: whether the nodule needs surgical removal, and, if malignant, whether it is amenable to surgery as tumours such as anaplastic carcinomas and lymphomas may require other treatments. In one centre the introduction of thyroid FNA cytology decreased the number of surgical biopsies by 50% and doubled the tumour detection rate.2

The aim of this study was to audit FNA cytology in a large centre to calculate its sensitivity and specificity. This study has large numbers with histological verification, which tends to be low in many previously published studies. We evaluated 233 FNA cytologies performed to assess a thyroid swelling. Histology was available in 184, the outcome in the remaining 49 was assessed by reviewing medical records.

Methods

Three hundred and seven patients (33 male, 274 female) had 335 FNA cytology of thyroid over a six year period (1989–95); 184 subsequently had surgical biopsy or excision. Of the 123 patients with no subsequent histology, medical records were found for 49 and these were reviewed. The overall follow up time was 2.7 years.

FNA cytology was performed by the same operator (a radiologist) using a 19 G needle on a 20 ml syringe and applying negative pressure. All cytology and histology was reviewed in a blind fashion. None of the original diagnoses were changed as a result of this review.

CYTOLOGY

A standard FNA cytology technique was used. All smears were stained with May–Grunwald–Giemsa stain. Smears containing only blood or very few follicular cells without colloid were judged to be inadequate for diagnosis. The cytological diagnosis was classified as: benign—smears with colloid and containing relatively few follicular cells; suspicious—smears with a relatively large amount of follicular cells compared to colloid (these are follicular lesions and this category recognises that the distinction between a follicular adenoma and carcinoma is not possible on
Table 1 Results of FNA cytology and subsequent histology

<table>
<thead>
<tr>
<th>FNA result</th>
<th>Malignant</th>
<th>Colloid nodular goitre</th>
<th>Adenoma</th>
<th>Thyroiditis</th>
<th>Negative clinical follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>21 malignant</td>
<td>11</td>
<td>7</td>
<td>2</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>45 suspicious, excise</td>
<td>5</td>
<td>19</td>
<td>14</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>130 benign</td>
<td>4</td>
<td>78</td>
<td>3</td>
<td>1</td>
<td>34 plus 10 inadequate biopsies*</td>
</tr>
</tbody>
</table>

*All 10 patients were well on clinical follow up.

cytology; surgical excision was advised); or malignant—smears with definite features of papillary carcinoma, anaplastic carcinoma or other malignancies.

The above classification was used as it relates to intention to treat. The cytology reports would also include descriptions and further subclassification of lesions if possible. Other audits of thyroid FNA cytology have used this three tiered division4 as it is clearer than listing each separate histological entity.

HISTOLOGY
Histology included biopsy specimens and surgical excisions. These were classified as: benign (intervention not needed)—colloid nodular goitre, follicular adenoma, Hashimoto's thyroiditis; or malignant—follicular carcinoma, papillary carcinoma, anaplastic carcinoma, lymphoma or metastatic carcinoma.

STATISTICAL ANALYSIS
The results were analysed on an intention to treat principle. It is impossible to distinguish between follicular adenomas and follicular carcinomas on FNA cytology, therefore, classifying a follicular adenoma as suspicious was not regarded as a false positive.

Sensitivity, specificity, positive predictive value, and negative predictive value were calculated. True positives (TP) were defined as cases deemed either suspicious or definite for malignancy on cytology that were malignant or follicular adenomas on histology. False positives (FP) were those positive or suspicious of malignancy on cytology that were benign on histology (with the exception of follicular adenomas) or negative for malignancy on clinical follow up. True negatives (TN) were benign on cytology and either benign on histology or clinical follow up. False negatives (FN) were negative for malignancy on cytology but positive for further histological assessment. The diagnostic accuracy was calculated as (TP + TN) - (FP + FN + TN). Inadequate smears were not included in calculation of the above figures.

Results
Thirty seven smears were judged to be inadequate for diagnosis (18%); 35 of the patients with inadequate smears had subsequent surgery and four had carcinoma and six had follicular adenomas. They are not included in any further calculations.

One hundred and thirty FNA cytologies were classified as benign (66%), 45 were suspicious (23%), and 21 (11%) were malignant (table 1).

Of the 130 patients with benign FNA cytology, 96 had surgery and four malignancies were later found on histology: two lymphomas, one follicular carcinoma, and one carcinoma not otherwise classified. All of these false negatives were felt to be clinically suspicious and two had nuclear imaging tests suspicious for malignancy. One patient had four FNAs, two of these were inadequate, one was a colloid nodular goitre, and one was reported as thyroiditis. The tumour was a lymphoma on histology but even this diagnosis was difficult and an expert opinion was sought. There were 78 goitres, three adenomas, one Hashimoto's disease, and 10 insufficient biopsies. All of the insufficient biopsies had normal clinical follow up. As none of these needed surgery the designation as benign was appropriate. In total, there were 126 true negatives including 34 with negative clinical follow up and no biopsy.

Of the 45 patients with suspicious FNA cytology, 39 had subsequent surgery and there were three papillary carcinomas, two follicular carcinomas, and 14 follicular adenomas (true positives), and 20 benign diagnoses (false positives). Of the six lesions followed by clinical means alone, none experienced worsening symptoms or clinically malignant disease (false positives).

Of the 21 patients with malignant FNA cytology, 20 had subsequent surgery and there were 11 true positives (four papillary, three anaplastic, two metastatic, one follicular, and one unclassifiable). There were nine false positives: seven colloid nodular goitres and two adenomas. The single patient followed up clinically did not develop features of malignancy (another false positive).

Therefore, in this study there are 126 true negatives, 30 true positives, 4 false negatives, and 36 false positives. This gives the following values for FNA cytology of thyroid: sensitivity, 88%; specificity, 78%; positive predictive value, 46%; negative predictive value, 97%; accuracy, 80%.

Table 2 Results of previous studies

<table>
<thead>
<tr>
<th>Reference</th>
<th>No of FNAs</th>
<th>No with histology</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>False -ve (n)</th>
<th>False +ve (n)</th>
<th>Inadequate (%)</th>
<th>Accuracy (%)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hall et al7</td>
<td>795</td>
<td>72</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5</td>
<td>7</td>
<td>1</td>
<td>89</td>
<td>Different rates are given for each disease</td>
</tr>
<tr>
<td>Kendall8</td>
<td>113</td>
<td>34</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>7</td>
<td>11</td>
<td>-</td>
<td>No follow up given on non-histology patients</td>
</tr>
<tr>
<td>Godinho-Matos et al13</td>
<td>144</td>
<td>28</td>
<td>73</td>
<td>100</td>
<td>69</td>
<td>4</td>
<td>4</td>
<td>13</td>
<td>83</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Rosen et al9</td>
<td>59</td>
<td>41</td>
<td>80</td>
<td>100</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>12</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mandrek et al9</td>
<td>1992</td>
<td>238</td>
<td>-</td>
<td>-</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>65</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Holleman et al11</td>
<td>112</td>
<td>53</td>
<td>84</td>
<td>52</td>
<td>53</td>
<td>83</td>
<td>2</td>
<td>1</td>
<td>11</td>
<td>65</td>
<td>-</td>
</tr>
</tbody>
</table>

PPV, positive predictive value; NPV, negative predictive value.
Discussion

Thyroid nodules are very common occurring in 4% of the population aged between 30 and 60. Most are benign and only between 10% and 20% are malignant. Therefore, surgery as the initial intervention or investigation will have a very low yield. Very few diagnostic tests help differentiate between benign and malignant nodules. Thyroid ultrasound can distinguish solid from cystic lesions but not all cystic lesions are benign. Thyroid isotope scans using technetium-99 classify nodules as hot or cold. Hot nodules are functioning and should be biopsied whereas cold nodules are non-functioning and might be malignant. However, fewer than 20% of cold nodules are malignant leading to a high false positive rate. The addition of thallium scans is a newer development but again has high false positive rates. The poor accuracy and high cost of nuclear imaging plus the significant radiation burden it places on patients has led some to suggest that its routine use in investigating the solitary thyroid nodule should be abandoned, while recognising its value in follow up after a thyroid malignancy.

FNA cytology is inexpensive, can be performed in an outpatients clinic, and has few complications including abscess of tumour implants along the needle track. The percentage of inadequate biopsies in this study (15%) is not excessively high compared with the literature, which has ranges from 0–25%. Inadequate biopsies are said to be less common when the pathologist reporting the biopsy does the aspiration personally. It has also been attributed to insufficient numbers performed annually. Mazzaferri has estimated that an individual needs to do 20–35 aspirations annually to gain acceptable experience.

Surgical complications of thyroidectomy are common and were responsible for 4% of all general surgical claims settled by the Medical Defence Union. All of these involved recurrent laryngeal nerve damage. Other injuries include parathyroid damage, which causes permanent hypocalcaemia following 0.4–13.8% of total thyroidectomies and 0.2–1.9% of subtotal thyroidecтомies. Hypothyroidism is inevitable following a total thyroidectomy. Other complications include postoperative haemorrhage and wound complications.

Every technique will have false positives and negatives. As there are only three possible diagnostic categories in our study, we accept that false positives may be high. We feel that this is safer than having a high false negative rate. False negatives may occur because of missed sampling of a lesion or cytological error. In particular, a multinodular goitre with recent change can be difficult to sample adequately. Ultrasound guided aspiration may have a role. In this study there were four false negatives. However, these were all clinically significant, highlighting the point that a negative FNA should never exclude malignancy. If there is a strong clinical suspicion of malignancy this should always override a negative FNA. This will apply to a relatively small number of patients.

In this study seven of nine false positives were colloid nodular goitres on histology. False positives may occur if a hyperplastic nodule is sampled producing a high cell to colloid ratio. The cellulosity of an aspirate also depends on the pressure used by the aspirator and can vary from case to case.

Other wrong diagnoses include the classification of follicular carcinomas as papillary carcinomas and vice versa. This is understandable when one looks at the lack of concordance between pathologists even when looking at histological sections.

Previous studies, although variably reported, have shown similar levels of accuracy to this study. It can be difficult to compare studies as many different methods of classification are used and some authors do not calculate the sensitivity and specificity. This is one of the larger studies and has a high level of histological verification. It is to be hoped that our levels of histological verification will decrease as surgeons and pathologists appreciate the value of FNA of the thyroid. In particular, a high negative predictive value is the most important figure when investigating large numbers of benign nodules.

A positive result is of less predictive value, although the presence of papillary features on a lymphoma or anaplastic carcinoma may spare the patient unnecessary surgery.

CONCLUSION

Following a benign result of FNA cytology of a thyroid swelling, the physician can be 97% sure that the lesion is indeed benign. However, a negative FNA should never exclude malignancy if there is a strong clinical suspicion. If this is always adhered to no malignant thyroid will be overlooked and many benign nodules will be spared surgery.

We acknowledge the help of Mr Gumpert, Consultant Surgeon, Royal Sussex County Hospital and the Department of Clinical Audit, Royal Sussex County Hospital, Brighton.

To operate or not to operate? The value of fine needle aspiration cytology in the assessment of thyroid swellings.
N Leonard and D H Melcher

*J Clin Pathol* 1997 50: 941-943
doi: 10.1136/jcp.50.11.941

Updated information and services can be found at:
http://jcp.bmj.com/content/50/11/941

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/