Clinical Outcome for Atypia of Undetermined Significance in Thyroid Fine-Needle Aspirations

Should Repeated FNA Be the Preferred Initial Approach?

Paul A. VanderLaan, MD, PhD, Ellen Marqusee, MD, and Jeffrey F. Krane, MD, PhD

Key Words: Thyroid; Cytology; Fine-needle aspiration; Atypia of undetermined significance

Abstract

In the Bethesda System for reporting thyroid fine-needle aspirations (FNAs), atypia of undetermined significance (AUS) is a category with limited reported follow-up and outcome data. We report a retrospective analysis of our institution’s experience during nearly 4.5 years with a tiered classification scheme conforming to the Bethesda System in which repeated FNA was recommended for most patients with an initial AUS diagnosis. Of 4,691 thyroid FNAs, 512 (10.9%) had a diagnosis of AUS. Cytologic or histologic outcome data were available for 331 cases (64.6%), of which 240 (72.5%) were benign and 91 (27.5%) were malignant. Of patients with a surgical diagnosis, there was no statistically significant difference in malignancy rate among patients who went directly to surgery after a single AUS diagnosis (37/90 [41%]), patients having 2 successive AUS FNA diagnoses (22/51 [43%]), and patients with a benign aspirate after AUS (2/7 [29%]). Although AUS confers an intermediate risk of malignancy, guidelines recommending repeated FNA for most cases should be reevaluated.
performed at the Brigham and Women’s Hospital, Boston, MA, from January 2005 to May 2009. During this period, a total of 4,691 thyroid FNAs were performed using a 25-gauge needle by an endocrinologist under ultrasound guidance (typically 3 or 4 passes). The specimen was collected immediately in CytoLyt (Hologic, Marlborough, MA), and Papanicolaou-stained ThinPrep slides were prepared using the ThinPrep 2000 (Hologic). When adequate material was present, cell block preparations were made for 354 (7.5%) of the cases. On-site evaluation was not performed routinely.

All cases were reported by a staff cytopathologist using a 6-tiered diagnostic system with diagnostic criteria essentially identical to those of the 2007 National Cancer Institute Thyroid FNA State of the Science Conference guidelines1-3 and the Bethesda System for Reporting Thyroid Cytopathology.4 The FNA diagnostic categories were (with the corresponding Bethesda System designation following) as follows: nondiagnostic specimen (nondiagnostic), no malignant cells identified (benign), atypical cells of undetermined significance (AUS/follicular lesion of undetermined significance), suspicious for a Hürthle cell (oncocytic) neoplasm/suspicious for a follicular neoplasm (suspicious for a Hürthle cell/follicular neoplasm), suspicious for malignancy (suspicious for malignancy), and positive for malignant cells (malignant).

Based on prior institutional experience with thyroid FNA,14 including early experience with AUS, our standard approach to an initial AUS diagnosis has been repeated FNA. Surgical resection was recommended for patients with a repeated AUS diagnosis or a suspicious or malignant diagnosis on repeated FNA, while patients with a benign diagnosis on repeated FNA were followed up clinically. Deviation from this approach occurred in individual cases based on an overall assessment of clinical features, including such factors as nodule size and growth, patient preference, and/or imaging characteristics.

Clinical outcome for the aspirated thyroid nodule was categorized as benign or malignant. A benign outcome was defined as a benign cytologic diagnosis made on a follow-up FNA sample or a tissue diagnosis of multinodular hyperplasia, follicular adenoma, adenomatous nodule, Hashimoto thyroiditis, or papillary thyroid microcarcinoma (defined as papillary carcinoma ≤1 cm detected incidentally at surgery, not as the FNA-targeted mass) on histologic evaluation. Outcomes were classified as malignant when a histologic diagnosis of malignancy was made based on resection of the aspirated nodule. The target nodule was identified by comparing the cytology and surgical pathology reports and ultrasonographic findings, if needed, to resolve ambiguity. Outcome classification was solely based on the diagnosis made on the targeted nodule. All cases had at least 1 year of clinical follow-up after the initial thyroid FNA.

Data processing and statistical analysis were performed using Microsoft Excel (Microsoft, Redmond, WA) and GraphPad software (GraphPad Software, San Diego, CA). Categorical analysis was performed using a 2-tailed χ² test and a 2-tailed Fisher exact test when appropriate, with a predetermined level of significance set at a P value of .05.

Results

The distribution of case diagnoses during the study period is summarized in Table 1. A total of 512 (10.9%) cases had a diagnosis of AUS. This group included 99 males and 413 females (mean age, 53 years; range, 9-88 years). Cell block preparations were made for 21 (4.1%) of these cases. Of the 512 total AUS FNAs, 63 (12.3%) represented repeated AUS diagnoses for the same thyroid nodule, 72 (14.1%) had no further follow-up on record, and 46 (9.0%) had a repeated FNA giving an indeterminate diagnosis but without further follow-up on record. In total, a repeated FNA was performed on 287 (56.1%) of the nodules after the initial AUS diagnosis. Ultimately, 199 cases (38.9%) underwent surgical resection, with or without repeated FNA. Overall, a cytologic or histologic outcome for the thyroid nodule in question was achieved in 331 cases (64.6%).

Following a diagnosis of AUS on an initial thyroid FNA, a repeated FNA has been considered appropriate in most cases to provide more definitive risk stratification.1-3,9,14 The follow-up FNA results for this study population are provided in Table 2. In the majority of cases (67.9%) the diagnostic uncertainty resolved into low-risk (benign, 48.4%) or high-risk (suspicious or malignant, 19.5%) categories with repeated FNA. In 80 cases (27.9%), repeated FNA resulted in a second AUS diagnosis.

For the 331 cases with an initial diagnosis of AUS by FNA and outcome data, 240 (72.5%) were benign and 91 (27.5%) were malignant. These results are summarized in Table 3. Papillary thyroid carcinoma accounted for 89% of all malignant outcomes at surgical resection.
To assess the usefulness of repeated FNA after an initial AUS diagnosis in predicting outcome, the cases with surgical follow-up following an initial AUS diagnosis were evaluated (199 cases). The results are shown in Table 4. A diagnosis of suspicious or positive for malignancy on repeated FNA was significantly more likely to be malignant than a repeated AUS diagnosis (21/27 [78%] vs 22/51 [43%]; P < .001). Patients having surgery after a single AUS result had a similar rate of malignancy (37/90 [41%]) to patients who had 2 consecutive AUS results (22/51 [43%]). Seven patients with a benign diagnosis after an initial AUS diagnosis underwent surgery; 2 of these 7 (29%) had a malignant diagnosis. In patients with an excisional diagnosis, there was no statistical significance between the rate of malignancy in patients with a benign aspirate after initial AUS vs patients with no repeated FNA (P = .51) or a repeated AUS diagnosis (P = .81), although the number of cases in the first category was small.

To correct for clinical bias that may have led to surgery after a single AUS diagnosis rather than repeated FNA, we also examined follow-up data in which only patients who had a repeated FNA after an initial AUS diagnosis were considered. In patients who had a repeated FNA before surgery, 57 (52.3%) of 109 had a malignant outcome identified on surgical resection. Compared with the patients going directly to surgery (37/90 with a malignant diagnosis [41%]), there was no statistically significant difference in the rate of malignancy for patients who had undergone repeated FNA before surgery (P = .12). If the 132 patients whose only follow-up consisted of a repeat benign FNA are considered to represent a benign outcome, the overall malignancy rate for patients with a follow-up FNA was 23.7% (57/241).

Discussion

Our overall AUS diagnosis rate of about 11% is consistent with other recently published studies with rates between 2% and 18% Table 5. Given the inherent uncertainty associated with a diagnosis of AUS on thyroid FNA, the use of this category should be minimized and used only when absolutely necessary. An overall target laboratory use rate of 7% or less has been proposed,2-4 although this figure is based on minimal published laboratory experiences. Our experience, along with that reported in other published studies, suggests that the 7% figure may be an unrealistic goal for many practitioners at present. However, as additional experience with AUS is obtained and diagnostic criteria are further refined, it may be possible to define AUS more narrowly and use this category more sparingly.

Table 4
Repeated Fine-Needle Aspiration Diagnosis Following an Initial Diagnosis of Atypia of Undetermined Significance

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>No. (%) of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign</td>
<td>139 (48.4)</td>
</tr>
<tr>
<td>Atypia of undetermined significance</td>
<td>80 (27.9)</td>
</tr>
<tr>
<td>Suspicious for malignancy</td>
<td>26 (8.1)</td>
</tr>
<tr>
<td>Suspicious for a follicular/Hürthle cell neoplasm</td>
<td>25 (8.7)</td>
</tr>
<tr>
<td>Malignant</td>
<td>5 (1.7)</td>
</tr>
<tr>
<td>Nondiagnostic</td>
<td>12 (4.2)</td>
</tr>
<tr>
<td>Total</td>
<td>287 (100.0)</td>
</tr>
</tbody>
</table>

Table 5
Overall Outcome Following a Diagnosis of Atypia of Undetermined Significance by Initial Thyroid FNA in 331 Cases

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. (%) of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malignant</td>
<td>91 (27.5)</td>
</tr>
<tr>
<td>Papillary carcinoma</td>
<td>81 (89)</td>
</tr>
<tr>
<td>Follicular carcinoma</td>
<td>8 (9)</td>
</tr>
<tr>
<td>Anaplastic carcinoma</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Benign</td>
<td>232 (61.2)</td>
</tr>
<tr>
<td>Benign follicular cells on FNA</td>
<td>97 (40.4)</td>
</tr>
<tr>
<td>Adenoma/modular hyperplasia</td>
<td>12 (5.0)</td>
</tr>
<tr>
<td>Hashimoto thyroiditis</td>
<td>8 (3.3)</td>
</tr>
</tbody>
</table>

FNA, fine-needle aspiration.
Overall, our data and those of others support the usefulness of the AUS designation for risk stratification with a risk of malignancy intermediate between that of the benign (2.5%-3%) and suspicious for malignancy (60%-75%) categories. However, the overall risk of malignancy for a thyroid nodule with an initial diagnosis of AUS on FNA was relatively high at 45.7% (91/199) in patients who ultimately had surgical follow-up and 27.5% (91/331) if patients with benign cytologic follow-up who did not undergo surgery are included. This malignancy rate is higher than the rates in most other recently published studies using a Bethesda-like reporting system (Table 5).

One factor contributing to this difference may be heterogeneity in the interpretation of AUS in these initial published reports. Because some of these studies predate the Bethesda System and involve “retrofitting” of diagnoses into Bethesda categories, they may include cases that would be more suitably classified differently within the Bethesda framework, such as marginally cellular specimens. Although our data also partially predate the Bethesda guidelines, classification and diagnostic criteria for thyroid FNAs at our institution closely conformed to these guidelines throughout the study period, with only minor terminology differences from the Bethesda System. Dissemination of the uniform reporting guidelines of the Bethesda System, along with sample images, should promote more uniform application of the various diagnostic categories, including AUS, resulting in correspondingly more meaningful follow-up data in future studies.

Specimen preparation methods could also potentially contribute to differing characteristics of AUS across published studies. Many practitioners primarily perform direct air-dried or alcohol-fixed smears for thyroid aspirates, while our institution primarily uses liquid-based preparations. Confounding factors associated with direct smears, including obscuring blood or drying/fixation artifacts, contribute to the need to resort to the AUS diagnosis in some cases. Because these artifacts are removed from liquid-based preparation techniques, one might expect a lower percentage of AUS cases in liquid-based preparations, but this was not observed in our series. Alternatively, the absence of cases secondary to artifact might yield a different mix of atypical diagnoses in liquid-based preparations with correspondingly differing surgical outcomes. Our higher malignancy rate for AUS may be partly attributable to the elimination of AUS cases caused by specimen preparation artifact that would be unlikely to be associated with malignancy on follow-up. In addition, one could speculate that the improved fixation in liquid-based preparations may increase sensitivity for cytologic findings that might be seen in subtle cases of papillary carcinoma. Finally, other factors unrelated to the sample preparation method, such as the threshold for diagnosing papillary carcinoma in surgical pathology, could also have an impact on the malignancy rate in our population.

The proposed Bethesda classification algorithm for clinical follow-up of most patients with an initial AUS diagnosis is a repeated FNA in 3 to 6 months with subsequent surgical resection if the follow-up FNA is AUS or worse.1-4 Our data indicate that patients with a single AUS diagnosis who go directly to surgery have rates of malignancy similar to patients having 2 consecutive AUS diagnoses, with more than 40% of the nodules proving to be malignant. Patients going to surgery after a single AUS diagnosis may have a bias toward a higher risk of malignancy owing to clinical or radiologic factors that prompt immediate surgery rather than repeated FNA. However, eliminating patients in this group from consideration yielded comparable rates for malignancy in patients with surgical follow-up (52.3% [57/109] vs 45.7% [91/199] if these patients are included) and in patients with surgical or benign cytologic follow-up (23.7% [57/241] vs 27.5% [91/331] if these patients are included). These findings suggest that overall...
the factors leading patients to surgery after a single AUS FNA diagnosis do not identify a subset of patients at higher risk of malignancy. While the patients undergoing surgery after a repeated FNA had a higher malignancy rate than patients who went straight to surgery after an initial AUS diagnosis (52.3% [57/109] vs 41% [37/90]), this enrichment in the malignancy rate was not statistically significant, indicating minimal additional benefit to the repeated FNA.

One of the main benefits of repeated FNA after an initial AUS diagnosis is sparing patients with a benign repeated FNA result (48.4% in our series) from undergoing presumably unnecessary surgery. While patients with benign follow-up FNAs after an initial AUS diagnosis infrequently undergo surgery, 2 of our 7 patients who had surgery had a malignant diagnosis. Renshaw16 recently reported his experience with patients having benign and AUS interpretations of a single nodule on FNA and concluded that the risk of malignancy in such cases was intermediate (16%) between patients having a single benign aspirate (1.7%) and a single AUS aspirate (24.5%). Our small numbers in conjunction with Renshaw’s more extensive study of this question support the notion that the risk for malignancy in a patient with AUS followed by a benign FNA result remains higher than the extremely low risk for patients with a benign diagnosis alone. These findings therefore raise concern that treating patients with a benign diagnosis following an initial AUS diagnosis in a manner equivalent to patients having only a benign FNA result may not be justified.

A recent study10 indicated that AUS with features concerning for papillary carcinoma carries higher risk of malignancy than other types of AUS (38% on surgical resection), similar to the findings of earlier studies.17,18 In our study, AUS nodules ultimately found to be malignant largely correspond to papillary thyroid carcinomas, as reported by others.5,13 Taken together, these data suggest that repeated FNA may not be appropriate for AUS cases in which there is concern for the diagnosis of papillary carcinoma. Defining these cases in a reproducible manner that is distinguishable from the suspicious for malignancy diagnosis is challenging and in need of further clarification. In the interim, one manner of refining clinical management for AUS would be to limit the recommendation for repeated FNAs to cases not suggestive of papillary carcinoma, such as suboptimal specimens in which factors like obscuring blood, drying artifact, and/or sparse cellularity are major factors contributing to the AUS diagnosis.

The anticipated rate of malignancy for AUS within the Bethesda System was in the range of 5% to 15%.1-4 By including patients with benign cytologic follow-up after an initial AUS diagnosis, our malignancy rate of 24% to 27% is similar to that for patients with a diagnosis of suspicious for a follicular neoplasm (malignancy rate in the 20%-30% range) for whom surgery is typically the treatment of choice. Likewise, the relatively high risk of malignancy in our patients with AUS indicates that surgery rather than repeated FNA may be a more appropriate course for many patients. In fact, some authors have advocated eliminating the AUS category in favor of a combined AUS/follicular neoplasm category.12 Although the risks of malignancy are similar, we currently oppose this approach. Because of the diagnostic heterogeneity of AUS, the possibility remains that further study will identify subtypes of AUS that confer a lower risk of malignancy that would warrant more conservative management. In addition, overall, AUS confers risk most prominently for papillary carcinoma, while the follicular neoplasm category primarily identifies risk for follicular carcinoma. This information can influence clinical decision making (lobectomy vs total thyroidectomy) and would be lost by lumping these diagnoses together. Furthermore, clinical concern regarding the reproducibility of an AUS diagnosis might prompt some clinicians to prefer conservative management for AUS (such as repeated FNA or seeking an expert consultation) instead of recommending surgery as they would for the more well-defined follicular neoplasm classification, despite overall similar rates of malignancy.

AUS represents an evolving heterogeneous category with different cytologic scenarios potentially warranting this designation. Despite this heterogeneity, AUS successfully stratifies patient risk for malignancy. Our study indicates that as a group, AUS diagnoses carry a higher risk of malignancy than anticipated in the Bethesda System. Accordingly, the recommendations for repeated FNA for most patients following an initial AUS diagnosis should be reconsidered. Further refinement of this category and its associated clinical outcomes is desirable to define more clearly the appropriate use of AUS, identify patients with AUS who are at relatively higher risk of malignancy, and establish the most appropriate corresponding methods of clinical triage following an initial diagnosis of AUS.

References


