Anatomic Pathology / AUS and ND in TBS Are Inversely Related

Atypia of Undetermined Significance and Nondiagnostic Rates in The Bethesda System for Reporting Thyroid Cytopathology Are Inversely Related

Paul A. VanderLaan, MD, PhD,1 Andrew A. Renshaw, MD,2 and Jeffrey F. Krane, MD, PhD1

Key Words: Thyroid; Fine-needle aspiration; Atypia of undetermined significance; Nondiagnostic; Quality assurance

DOI: 10.1309/AJCPI41QOQUSKDGP

Abstract

There is substantial variation in the reported rates of different diagnostic categories in The Bethesda System for Reporting Thyroid Cytopathology (TBS). Specifically, the relationship between the nondiagnostic (ND) and atypia of undetermined significance (AUS) categories has not been closely examined previously. Data from published series in the literature and from 2 separate hospitals with more than 15,000 thyroid aspirates were reviewed. The AUS and ND rates were consistently negatively correlated when analyzed by year, aspirator, and cytologist. The strongest correlation was with cytologists (P < .0003). Absolute ND rates decreased by 1% for every 3.5% increase in AUS, implying the existence of a discrete population of cases that cytologists will classify as ND or AUS. As such, AUS and ND are not independent variables. Awareness of this relationship may be useful for laboratories and individual cytopathologists for refining the use of TBS.

The Bethesda System for Reporting Thyroid Cytopathology (TBS) introduced a standardized 6-tiered system to promote uniform practice and reporting of thyroid fine-needle aspiration specimens with associated defined risks of malignancy and clinical management algorithms.1 Previous studies have suggested that diagnostic categories are interrelated. For example, studies based on the literature before TBS suggest a negative correlation between nondiagnostic (ND) and malignant categories, implying that a benign nodule is more likely than a malignant nodule to yield an ND aspirate.2 Recent studies based on the literature using TBS have shown that there is a consistent relationship between the atypia of undetermined significance/follicular lesion of undetermined significance (AUS) category and the malignant category, suggesting that malignant lesions are more likely to result in a diagnosis of AUS than benign nodules.3 As a result of this relationship, it was proposed that the AUS/malignant ratio may be a useful quality assessment measure in thyroid cytology. It is interesting that in this publication, it was noted that at least in some cases,4 the AUS and ND categories seemed to be negatively correlated.

In the present study, we more closely examined the relationship between ND and AUS in the published literature within TBS framework. Inevitably, such literature-based analysis is limited by the heterogeneity of published studies, including variations in the application of diagnostic criteria, differing patient populations and specimen preparation techniques, and variability in the aspirators performing the fine-needle aspirations (FNAs) and cytologists interpreting the aspirates. Whether these same relationships between diagnostic categories can be identified in data sets with less variation is not known. To address this, we also reviewed a large series of thyroid FNA specimens from 2 different practice settings.
that used different preparatory techniques, but that were both composed of a relatively small set of cytologists with lengthy experience practicing together as a group.

**Materials and Methods**

We used the same published experiences to analyze the relationship between ND and AUS using TBS in the available literature. Statistical analyses and linear regression were performed using GraphPad (Graphpad Software, La Jolla, CA) and Microsoft Excel (Microsoft, Redmond, WA). Statistical significance was established as a \( P \) value of less than .05.

All thyroid FNA reports interpreted at Baptist Hospital (BH), Miami, FL, and Homestead Hospital (HH), Homestead, FL, from October 1996 through May 2011 were reviewed and the results correlated with the results of histologic follow-up. Cytologic cases were classified according to TBS. All aspirates were performed by clinicians. Approximately one third of the aspirates were performed in clinicians’ offices without imaging studies. Two thirds of the aspirates were performed in the radiology department of the BH, with the aid of ultrasound guidance and immediate evaluation. Between 2 and 8 passes were performed. Direct smears were made in all cases, and all were alcohol-fixed and stained with Papanicolaou or an H&E stain. If sufficient material was obtained, cell blocks were also made. Core needle biopsy was also available in approximately 700 cases.

Thyroid aspirates performed from January 2005 to June 2011 at the Brigham and Women’s Hospital (BWH), Boston, MA, were included. All thyroid FNAs during this period were performed under ultrasound guidance by an endocrinologist using a 25-gauge needle (typically 3 or 4 passes without routine on-site evaluation). The specimens were collected immediately in CytoLyt (Hologic, Marlborough, MA). Papanicolaou-stained ThinPrep slides were prepared using the ThinPrep 2000 (Hologic), and reported by 1 of 8 staff cytopathologists using a 6-tiered diagnostic system identical to TBS except for minor terminology differences. When adequate material was present, cell block preparations were made (<5% of the cases).

Data from each site were compiled differently; the cytologist interpreting the case was available for BWH but not BH/HH, whereas the aspirator information was available for the BH/HH but not the BWH data. For the aspirator data, clinical fellows from the same year were grouped together for this analysis.

**Results**

Review of published laboratory experiences with TBS terminology **Table 1** demonstrates a negative correlation between AUS and ND rates **Figure 1** that nearly approaches statistical significance (\( P = .053 \)). However, this trend is strongly influenced by the 2 outliers with atypia rates of more than 12%. When these are removed, the correlation remains negative but is very weak (not shown).

We next analyzed the relationship between AUS and ND in the more homogeneous settings of our individual laboratories. AUS and ND were consistently negatively correlated when analyzed by year, aspirator, and cytologist. Correlation by aspirator at BH/HH is shown in **Figure 2**, whereas correlation by cytologist at BWH is shown in **Figure 3**. Although the negative correlation between AUS and ND rates was consistent throughout, only the correlation by cytologist achieved statistical significance (\( P < .0003 \)). Using the correlation by cytologist, on average, the absolute ND rate decreases 1% for approximately every 3.5% increase in AUS.

**Discussion**

We set out to explore the relationship between AUS and ND based on a prior observation that a high AUS rate observed in 1 study was likely tied to preferential diagnosis.

**Table 1**

<table>
<thead>
<tr>
<th>Study</th>
<th>ND</th>
<th>Benign</th>
<th>AUS</th>
<th>SusF</th>
<th>SUS</th>
<th>Malignant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faquin et al</td>
<td>13.9</td>
<td>66.9</td>
<td>10.0</td>
<td>2.0</td>
<td>3.2</td>
<td>3.9</td>
</tr>
<tr>
<td>Kim et al</td>
<td>1.8</td>
<td>58.3</td>
<td>16.3</td>
<td>1.2</td>
<td>6.2</td>
<td>16.2</td>
</tr>
<tr>
<td>Nayar and Ivanovic</td>
<td>5.3</td>
<td>64.2</td>
<td>17.8</td>
<td>5.9</td>
<td>1.9</td>
<td>4.9</td>
</tr>
<tr>
<td>Theoharis et al</td>
<td>11.1</td>
<td>73.8</td>
<td>3.0</td>
<td>5.5</td>
<td>1.3</td>
<td>5.2</td>
</tr>
<tr>
<td>VanderLaan et al</td>
<td>12.5</td>
<td>62.7</td>
<td>10.9</td>
<td>4.2</td>
<td>4.5</td>
<td>5.2</td>
</tr>
<tr>
<td>Marchevsky et al</td>
<td>12.9</td>
<td>71.6</td>
<td>9.8</td>
<td>1.5</td>
<td>2.3</td>
<td>2.0</td>
</tr>
<tr>
<td>Jo et al</td>
<td>18.6</td>
<td>59.0</td>
<td>3.4</td>
<td>9.7</td>
<td>2.3</td>
<td>7.0</td>
</tr>
<tr>
<td>Renshaw et al</td>
<td>23.6</td>
<td>54.0</td>
<td>7.7</td>
<td>8.6</td>
<td>1.8</td>
<td>4.2</td>
</tr>
</tbody>
</table>

AUS, atypia of undetermined significance/follicular lesion of undetermined significance; ND, nondiagnostic; SUS, suspicious for malignancy; SusF, suspicious for a follicular or Hürthle cell neoplasm; TBS, The Bethesda System for Reporting Thyroid Cytopathology.
of potentially ND cases within TBS as AUS. Our closer examination of published studies with TBS confirms the inverse relationship between AUS and ND rates and occurs in a wide variety of settings, including laboratories with very low ND rates and high AUS rates.

We sought to better understand the relationship between AUS and ND by focusing on the more homogeneous setting of 2 laboratories with extensive experience with thyroid FNA, a relatively stable group of cytologists, and consistent methods for specimen acquisition. A consistent negative correlation between ND and AUS was confirmed in both laboratories, although this correlation was most strong when stratified by individual cytologist. This strongly suggests that there is a group of cases that some cytologists are classifying as AUS and other cytologists are classifying as ND. Using the data from the correlation with cytologists, this amounts to an absolute change in 1% ND for roughly every 3.5% change in AUS. Most cytologists will recognize the existence of cases that are scant and difficult to classify, which might be reasonably placed in either category.

Our data indicate a link between the AUS and ND categories that extends beyond inadequacies in the aspirate. By obtaining a better aspirate (through improved aspiration or preparatory techniques), one would anticipate that the AUS rate would decrease as the overall ND rate decreased (conceptually represented by a downward shift of the AUS vs ND trend line). In addition, it is conceivable that with improved specimen quality, the negative correlation between ND and AUS would diminish (graphically equivalent to a decreased slope of the AUS vs ND trend line). However, our data from these 2 laboratories using completely different preparatory techniques (liquid-based vs direct smears) yielded similar results, indicating that there is a fundamentally persistent relationship between the ND and AUS categories related to the nature of the nodules aspirated, rather than the cytologist, the aspirator, or the sample preparation technique. It has also been previously argued that the ND rate is only finitely reducible owing to inherent properties of aspirated nodules (including but not limited to cystic lesions). Thus, it should be anticipated that an irreducible limit to the ND rate exists and that even as one approaches this limit, some degree of negative correlation with AUS rates would persist.

Nevertheless, there remains a small set of laboratories that have achieved a very low ND rate and a very low AUS rate. How these laboratories are achieving this is not entirely clear because our data suggest there are limits to how low a combination of these 2 diagnostic categories can go. Nevertheless, it is clear that changes in technique can result in marked changes of ND rates. As discussed, it is possible that the relationship between ND and AUS is not linear when one reaches very low ND rates, and perhaps in this range, it is also possible to reduce the AUS rate. However, a more
likely way to reduce the AUS rate in the setting of a low ND rate is to simply diagnose these cases as benign.11 Since most malignancies missed in this manner are minimally invasive follicular carcinomas or follicular variants of papillary carcinoma, the prognosis for the patients is excellent; differences in sensitivity for malignancy would be unlikely to be found based on clinical progression alone.

The implications of shifting diagnoses between the ND and AUS categories will vary based on clinical behavior. In TBS, the standard recommendation for ND and AUS is repeat aspiration. Thus, there should be little difference in short-term clinical outcome between these diagnoses. However, there is substantial controversy about this subject. Some authors have suggested the risk of malignancy for patients with an atypical diagnosis followed by a benign diagnosis does not reach that of a single benign diagnosis.15 Other authors have pointed out that the risk of malignancy for a diagnosis of AUS is not that different from that of the follicular neoplasm category,8 although the diagnosis of AUS may be less reproducible. Indeed, in some settings, patients with an AUS diagnosis undergo resection rather than repeat aspiration. In these laboratories, AUS and ND are managed quite differently, so simply switching between these diagnoses will have greater immediate clinical consequences. For the patients who undergo repeat sampling, the implications of an AUS with the repeat biopsy are quite different following an initial AUS (almost certainly resulting in surgery) than after an initial ND (likely another repeat FNA).

Our data support the notion that the diagnosis of AUS is negatively correlated with ND rates. When evaluating cytopathologist performance, evaluation of outlier data in the context of the overall laboratory ND rate may prove useful in identifying the source of significant deviations from expected levels. Further study of outlier laboratories that have achieved very low ND and AUS rates is warranted. In addition, studies of the reproducibility of these diagnoses might further define features that are challenging in distinguishing between the ND and AUS categories. Ultimately, molecular testing may also prove valuable as a tool for clinical triage and for refining diagnostic criteria in the cases lying on the cusp between ND and AUS.

References

From the Departments of Pathology, 1Brigham and Women’s Hospital and Harvard Medical School, Boston, MA; and Baptist Hospital of Miami, Miami, FL.

Address reprint requests to Dr Krane: Dept of Pathology, Brigham & Women’s Hospital, 75 Francis St, Boston, MA 02115.