Thyroid Cytopathology:
Bethesda and Beyond

Syed Z. Ali
Division of Cytopathology, Department of Pathology and Radiology, The Johns Hopkins Hospital, Baltimore, Md., USA

Key Words
Ancillary tests · Bethesda · BRAF · Cancer · Cytopathology · Fine-needle aspiration · Terminology · Thyroid

Abstract
Thyroid nodules are commonly encountered in clinical practice. Although the overwhelming majority of them turn out to be benign, the small subset of cancerous nodules needs to be accurately identified for optimal and timely surgical management. Fine-needle aspiration has proven to be the most valuable diagnostic modality for pre-operative distinction of benign from malignant nodules. The recently introduced and much anticipated Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) has standardized our diagnostic approach to reporting and cytomorphological criteria. TBSRTC has well-defined and rational management algorithms with implicit risk of malignancy in each of the 6 diagnostic categories. Recently published data supports the clinical utility and wide acceptance of TBSRTC by both practicing pathologists and clinicians. The problematic category of ‘indeterminate’ cytopathologic diagnoses has led to the discovery and development of unique and useful molecular markers, such as BRAF, which have displayed promising potential in recently published studies. As a result of the publication of TBSRTC, in 2009 the American Thyroid Association revised its clinical guidelines for the management of patients with thyroid disease and TBSRTC offers a useful source of information for the pathologist as well.

Historical Background
Thyroid nodules are a common clinical occurrence. More than 50% of the world’s population harbors at least 1 thyroid nodule and the frequency of nodular thyroid disease increases with age. It is therefore no surprise that thyroid fine-needle aspiration (FNA) is one of the most commonly practiced areas in non-gynecologic cytopathology. Despite the frequent occurrence of thyroid nodules, the majority (~95%) of them are benign. The critical issue in the management of patients with thyroid disease, therefore, is to find a way to distinguish pre-operatively, benign nodules (the overwhelming majority) from cancers (the minority of cases). Radiological imaging, serological and molecular studies have made major advances in the last decade in the diagnosis and management of patients with thyroid disease. However, to date, there is no single appropriate non-invasive diagnostic test in clinical medicine that can accurately, and in a timely and cost-effective fashion, distinguish benign (mostly approached non-surgically) from malignant (mostly managed surgically) nodules, except for FNA. Numerous clinical studies have proven over time the tremendous impact of thyroid FNA in care of patients with thyroid disease [1, 2]. The entire current management approach for patients with thyroid disease hinges upon a cytopathologic diagnosis. As illustrated in the 2 studies referenced above, the incidence of encountering cancer after thyroid surgery has risen from 14% to more than 50% in
a period of just over 2 decades, something that is totally credited to an increasing and optimum use of thyroid FNA and better knowledge of the cytomorphological characteristics of thyroid disease. In other words, triaging of patients with thyroid disease who are to be surgically managed has become more refined and accurate now as compared with 20–30 years ago.

No other area in diagnostic cytopathology demands more of a multidisciplinary approach than thyroid FNA. An accurate FNA diagnosis leading to timely and optimal patient management is highly dependent on a close and friendly interaction amongst the pathologist, radiologist and medical endocrinologist/head and neck surgeon. The cytopathologist should render the diagnosis based not only on cytomorphological characteristics of the nodule but also on the clinical history/findings and radiological features. On the other hand endocrinologists/surgeons cannot adequately manage thyroid patients unless they get an FNA report which is unambiguous, clear, succinct and clinically relevant. Historically we have all been guilty of using a highly variable thyroid cytopathology reporting system. In some instances the reporting systems were regional and in others they were more or less institutionally developed. In the United States alone, more than a dozen reporting systems have been used by cytopathologists as recently as a few years ago. There was so much laboratory to laboratory variation in reporting that patients and clinicians had started to question the validity of cytopathologic diagnoses of thyroid FNAs.

There was a dire need to standardize the reporting system and an excellent model already existed in the form of the Bethesda System for Reporting Cervical Cytology. Additionally, it was felt that a new reporting system of thyroid cytopathology should be more reproducible with robust diagnostic criteria and should be clinically relevant and informative. These requirements thus needed input and endorsement of all major groups involved in the care of thyroid patients, i.e. radiologists, endocrinologists and endocrine surgeons.

Clinical management plans and guidelines for thyroid disease, particularly neoplasia, have always existed, such as the American Thyroid Association and National Comprehensive Cancer Network; however, there was never an effective linkage between these clinical management plans and thyroid FNA reporting, thus undermining the clinical utility of our cytopathologic diagnoses.

With this background, Dr. Andrea Abati, the then Director of Cytopathology at the National Cancer Institute, took the initiative and formed a multidisciplinary steering committee in 2006, which ultimately led to the formation of 8 specific committees/forums which were given specifically defined tasks and assignments, such as exploring the indications/pre-FNA requirements, training and credentialing, FNA technique, terminology and morphologic criteria, ancillary testing, publications and website/print atlas.

After 18 months of hard work, which included reviewing and summarizing a huge amount of published data, 6 ’review and conclusions’ documents were drafted. These documents underwent extensive discussion and appropriate modifications after 2 runs of ‘bulletin board’ discussions at the NCI website (http://thyroidfna.cancer.gov).

Finally, it was time for a face-to-face discussion and the NCI hosted the ‘State of the Science’ conference for 2 days in Bethesda on October 22–23, 2007. More than 150 experts registered and most of them took an active part in the deliberations of the conference, which was moderated by Drs. Edmund Cibas and Susan Mandel. The conference, at which most major pathology and endocrine professional organizations were represented, resulted in important recommendations ranging from indications and technique of thyroid FNA to post-FNA testing and management plans. Perhaps the most anticipated recommendations came from the ‘terminology and morphologic criteria’ committee chaired by Dr. Zubair Baloch [3]. The conclusions of that committee’s work formed the framework for the subsequently published Bethesda thyroid monograph, an initiative taken by Drs. Ali and Cibas as the chairs of the atlas and publication committees, respectively. With the untiring efforts of over 40 well-known experts from various disciplines from around the globe, the Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) was finally published in January 2010. The book, which is approximately 200 pages with over 200 color images, has an easy-to-read format and includes definitions, diagnostic/morphologic criteria, explanatory notes and a brief management plan for each diagnostic category. It is already on its second print run and has been successfully translated into Chinese, while a Japanese translation is in the works.

The Bethesda System

TBSRTC [4] has 6 distinct diagnostic categories (Table 1) and is constructed on the concept of the probabilistic approach, i.e. the probability of finding malignancy in...
Each diagnostic category or, in other words, each diagnostic group carries an assigned positive predictive value which varies from generally <1.0% in 'benign' to as high as 99% for 'malignant'. This approach is highly useful and led the way for well-defined clinical management algorithms linked to each diagnostic category. Thus, for first time in a non-gynecologic setting, a multidisciplinary approach led to immensely useful linking of primary cytopathologic diagnoses to patients’ clinical management plan. In essence, potentially all cytopathologic interpretations of thyroid FNAs will belong to 1 of these 6 diagnostic groups. TBSRTC has clarity of communication which was desperately needed before. The 6 diagnostic categories are well defined, morphologically distinct and ensure a better cytohistologic correlation and inter-observer reproducibility for better and more meaningful exchange of data across national and international boundaries for optimum patient care and basic and translational research. A brief discussion of each diagnostic category follows.

### Non-Diagnostic or Unsatisfactory

These are specimens that have an inadequate number of follicular cells or the preparations are compromised due to obscuring blood, overly thick smears, air-drying artifacts, etc. It is recommended that to be satisfactory for evaluation when contemplating a benign diagnosis a specimen should display at least 6 groups of follicular cells, with each group composed of at least 10 cells. Some exceptions to the numerical requirement may include abundant colloid, background suggestive of chronic lymphocytic thyroiditis or when cytologic atypia is encountered. Non-diagnostic interpretations occur in 2–20% of cases (should be limited to 10% of thyroid FNAs in any individual laboratory). A special subset of non-diagnostic cases consisting only of cystic fluid and macrophages are problematic. TBSRTC recommends that cyst-fluid-only cases should technically be considered a special subset of non-diagnostic cases. However, an optional note can be added that if the nodule is almost entirely cystic, without any suspicious ultrasound/clinical findings or family history, the results can be considered adequate and benign. It should be remembered that in cyst-fluid-only cases, the

### Table 1. The Bethesda System for Reporting Thyroid Cytopathology

<table>
<thead>
<tr>
<th>Diagnostic category</th>
<th>ROM, %</th>
<th>Management plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-diagnostic/unsatisfactory</td>
<td>1–4</td>
<td>Repeat FNA with ultrasound guidance</td>
</tr>
<tr>
<td>Limited cellularity or acellular</td>
<td>0–3</td>
<td>Clinical follow-up</td>
</tr>
<tr>
<td>Technically compromised</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyst fluid only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benign</td>
<td>20–25 (for repeat AUS)</td>
<td>Repeat FNA</td>
</tr>
<tr>
<td>Adenomatoid or colloid nodule</td>
<td>15–30</td>
<td>Surgical lobectomy</td>
</tr>
<tr>
<td>Chronic lymphocytic thyroiditis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atypia of undetermined significance (AUS)/follicular lesion of undetermined significance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspicious for a follicular neoplasm/follicular neoplasm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspicious for malignancy</td>
<td>60–77</td>
<td>Surgical lobectomy or near total thyroidectomy</td>
</tr>
<tr>
<td>Papillary CA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medullary CA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lymphoma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metastatic neoplasm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malignant</td>
<td>97–99</td>
<td>Near total thyroidectomy (radiation/chemotherapy for some)</td>
</tr>
</tbody>
</table>

ROM = Risk of malignancy; CA = carcinoma.
risk of malignancy (mostly cystic papillary carcinomas) is estimated to be around 4%. The clinical management plan for non-diagnostic cases is a repeat FNA with ultrasound guidance, which is mostly diagnostic. However, for persistently non-diagnostic cases, a limited excision is recommended.

**Benign**
This category includes all adequately cellular aspirates comprised of varying proportions of colloid and benign follicular cells mostly displaying a macrofollicle architecture. Other benign diagnoses include chronic lymphocytic (Hashimoto) thyroiditis, granulomatous (subacute) thyroiditis, Graves’ disease and other. These patients are followed with repeat examination by palpation or ultrasound at 6- to 18-month intervals.

**Atypia of Undetermined Significance or Follicular Lesion of Undetermined Significance**
This is a heterogeneous group of cases that have cellular features which are not easily classifiable into the benign, suspicious, or malignant categories (fig. 1–4). Atypia of undetermined significance (AUS)/follicular lesion of undetermined significance (FLUS) represents a minority of thyroid FNAs (3–6%) and this diagnosis should not exceed 7% in an individual laboratory. TBSRTC illustrates 8 specific scenarios for AUS/FLUS:
1. Predominance of microfollicles in a sparsely cellular aspirate with scant colloid.
2. Predominance of Hürthle cells in a sparsely cellular aspirate with scant colloid.
3. Follicular cell atypia hindered by sample preparation artifact (air-drying artifact or clotting artifact).
4. Cellular aspirate with virtual all Hürthle cells, but in a clinical setting suggesting chronic lymphocytic thyroiditis or multinodular goiter.
5. Focal features of papillary carcinoma (especially focal nuclear features) in a benign aspirate with features suggestive of chronic lymphocytic thyroiditis.
6. Cyst-lining cells which may display atypia such as nuclear grooves, elongated nuclei and/or intranuclear inclusions in an otherwise predominantly benign aspirate.
7. Focal follicular cell atypia (nuclear enlargement, prominent nucleoli, etc.) in patients with history of radioactive iodine, etc., or repair due to involutional changes in a cystic lesion.
8. Atypical lymphoid infiltrate in which flow cytometric phenotyping could not be performed and a repeat FRNA is needed.

The recommended clinical management plan is a repeat FNA at an appropriate interval, which often results in a more definitive interpretation with only approximately 20% cases remaining atypical. The risk of malignancy for repeated AUS/FLUS cases is 20–25% of patients, therefore justifying the clinical approach of surgical lobectomy.

**Suspicious for a Follicular Neoplasm/Follicular Neoplasm**
FNA diagnosis attempts to play the role of a ‘screening test’ by identifying cases that might turn out to be the biologically significant lesion of follicular carcinoma (fig. 5). Although the cytomorphic features do not allow distinction of a follicular adenoma from a follicular carcinoma, the diagnosis usually leads to a surgical lobectomy. The term ‘suspicious for a follicular neoplasm’ is preferred over ‘follicular neoplasm’ for this category because studies have shown that up to 35% of these cases on histologic follow-up turn out to be adenomatoid nodules. Cytomorphology often displays high cellularity with scant or absent colloid. The diagnostic feature is a predominant microfollicular pattern with equisized rosette-like structures. Management is a surgical lobectomy, which may be immediate or near total thyroidec- tomy, based on the histological interpretation.

**Suspicious for Malignancy**
These are cases that either have sufficient cellular atypia but lack the quantitative and/or qualitative features for a definitive cancer diagnosis, or are sparsely cellular but a malignant diagnosis cannot be made with certainty (fig. 6, 7). ‘Suspicious for papillary carcinoma’ is the most common member of this category and, due to its high positive predictive value, is resected, either by lobectomy or near total thyroidectomy. Most (60–75%) turn out to be papillary carcinomas. The same basic approach is applied to other thyroid cancers such as medullary carcinoma, anaplastic carcinoma, or lymphoma.

**Malignant**
This category is employed in cases where the cytomorphic features are diagnostic for cancer (fig. 8). TBSRTC recommends that descriptive comments should be added to subclassify the cancer type. A malignant diagnosis is followed by a total thyroidectomy and/or chemotherapy/radiation depending on the specific tumor type. The positive predictive value of a malignant FNA diagnosis is 97–99%.
Although long-term follow-up studies are still needed, retrospective application of TBSRTC or limited-volume prospective analysis have shown promising results. Post-TBSRTC published data clearly highlights the ease of use of the new standardized nomenclature for much improved interinstitutional reproducibility and wider acceptance by the clinicians for better and consistent management approaches [5]. One of the expected contentious areas is the newly defined diagnostic category of AUS/FLUS. TBSRTC recommends a 7% cutoff for such diag-

**Post-Bethesda Experience**

Fig. 1. AUS/FLUS. The aspirate shows abundant clotted blood trapping fragments of follicular epithelium which display artifactual nuclear elongation and grooves. Follow-up was benign. Papanicolaou stain. ×200. Direct smear.

Fig. 2. AUS/FLUS. Cyst-lining cells with atypical nuclear forms (elongated with focal palisading and grooves). Follow-up was benign. Diff Quik stain. ×200. Direct smear.

Fig. 3. AUS/FLUS. Metaplastic-appearing cells, one showing a small intranuclear inclusion (right of center). Follow-up was benign. Diff Quik stain. ×400. Direct smear.

Fig. 4. Suspicious for a follicular neoplasm. Cellular smear with microfollicular proliferation. Follow-up was a follicular adenoma. Diff Quik stain. ×200. Direct smear.
noses by individual laboratories, based on some initially published data [2, 6]. However, others have found a much more variable incidence ranging from 2.5 to 28.6% for individual pathologists or 3.3–14.9% for among various institutions [7]. Most of these studies were published before the Bethesda monograph was introduced. A dedicated chapter on AUS/FLUS adequately illustrates 8 specific scenarios where cytomorphological characteristics will enable this particular diagnosis to be made. It is anticipated that post-publication of the monograph will result in better fine tuning of the cytomorphologic criteria resulting in fewer than 7% of these ‘undetermined’ diag-

Fig. 5. Suspicious for papillary thyroid carcinoma. This was the only cellular fragment with nuclear features of papillary carcinoma in an aspirate with predominantly colloid. Follow-up was papillary carcinoma. Diff Quik stain. ×200. Direct smear.

Fig. 6. Malignant (papillary thyroid carcinoma). Note oval-shaped nuclei and abundant intranuclear inclusions. Diff Quik stain. ×200. Direct smear.

Fig. 7. Malignant (medullary thyroid carcinoma). Loosely dispersed cells with round to oval and eccentric nuclei and fragile, somewhat granular cytoplasm. Diff Quik stain. ×200. Direct smear.

Fig. 8. Malignant (anaplastic thyroid carcinoma). A mixture of high-grade epithelioid, spindled and multinucleated giant cells. Diff Quik stain. ×400. Direct smear.
noses on FNA. Some have observed that different types of 'AUS cells' have significantly different risks of malignancy and they suggest that this disparity of risk be communicated to the treating physician [8].

**Future Trends**

One of the highlights of the Bethesda thyroid conference was to revisit and explore further potential ancillary tests, immunomarkers and molecular tools applied on FNA samples (smears, cell blocks) for better and timely identification of thyroid cancers [9]. The role of immunomarkers such as thyroglobulin, TTF-1, calcitonin, CEA, PTH and various selected cytokeratins is well known. Flow cytometry for lymphoproliferative disorders, PTH assays in needle rinse fluid for parathyroid lesions, and thyroglobulin for suspected lymph node metastasis are also well-recognized. The major issue is the so-called indeterminate FNA diagnosis and the potential use of some of the newer molecular markers (genetic mutations – BRAF and RAS) or chromosomal rearrangements (RET/PTC, PAX8/PPARG) in such cases [10]. The prevalence of these mutations is highly variable (BRAF = 45% and RET/PTC = 20% in papillary carcinoma, RAS = 45% and PAX8-PPAR = 35% in follicular carcinoma, and TP53 = 70% in anaplastic carcinoma, etc.) [11]. These analyses are highly variable, being performed on extracted DNA or RNA and using a range of methodologies such as PCR, RT-PCR and FISH. Although a combination of marker analyses appears more promising for primary diagnosis and prognostication [11], there is no single molecular test that can solve the problematic issue of the indeterminate/suspicious diagnosis.

Of all the markers described to date, the molecular detection of BRAF mutation appears to be most useful [11, 12]. BRAF mutation can be detected by various molecular techniques which involve DNA extracted from fresh or fixed FNA material. Although still not ready for ‘prime time’ clinical use due to limited validation studies, numerous studies have proved it to be the single most specific marker for papillary thyroid carcinoma (PTC), particularly for highlighting the more aggressive forms of the cancer (such as the tall cell variant, which is mutated in 90% of cases, as opposed to follicular variant which is mutated in 9% of cases). This makes a strong case for its use in prognostication for optimum clinical management plans (extent of surgery etc.) by the surgeon in BRAF-positive PTCs on FNA. It has to be kept in mind, however, that most of these analyses will require dedicated aspirates and special processing of the cytologic samples.

Studies have also shown that the sensitivity of BRAF for PTC is further enhanced when combined with ultrasound examination. A recent study showed BRAF mutation was detected in 67% of ‘ultrasound-positive’ thyroid nodules as opposed to 14% of ‘ultrasound negative’ lesions [13]. The value of performing BRAF mutational analysis assumes a more important role when done in a BRAFV600E-prevalent population, where it has been clearly shown to increase the diagnostic accuracy of FNA [14]. Several other publications, including a recent one summarizing a large volume of published data, stated that the utility of using BRAF mutation testing for nodules with indeterminate cytology is limited since many of those nodules (benign and malignant) do not harbor BRAF mutations. The authors concluded that when the pathologist sees cytomorphological characteristics suspicious for PTC, BRAFV600E mutation analysis may enhance the assessment of preoperative risks for PTC, thus directing a more aggressive initial surgical management when appropriate [15].

**Clinical Perspective**

The American Thyroid Association guidelines were originally published in 1996 and then in 2006, and were endorsed by many American, British and other European endocrine societies. However, due to the rapid growth of the literature on this topic, the American Thyroid Association came out with revised guidelines for management of patients with thyroid nodules and differentiated cancer in 2009 [16–18]. The revisions were based on an extensive review of all published data through December 2008 by a task force. The revised guidelines include recommendations regarding initial evaluation, clinical and ultrasound criteria for FNA, interpretation of FNA results, and management of thyroid nodules. These are all presumably evidence-based recommendations with a goal of providing optimal care for patients with thyroid disease. The guidelines also took into consideration TBSRTC.

The following are extracts from these published guidelines, more relevant to our practice as pathologists.

The guidelines regarding FNA clearly state its role as ‘the most accurate and cost-effective method for evaluating thyroid nodules’. The role of ultrasound-guided aspirations has been highlighted with fewer non-diagnostic and false negative cases. For non-palpable nodules, nodules with a higher likelihood of a non-diagnostic aspirate...
Certain characteristics of thyroid nodules on ultrasound have been shown to be associated with a higher probability of encountering malignancy. These include: hypoechogenicity compared to surrounding thyroid tissue, increased vascularity of the nodule, irregular margins, microcalcifications, an absent halo, and when the lesion is taller than its width in the transverse dimension. However, there is no single or combination of features on ultrasound to adequately identify a malignant nodule. Microcalcifications are highly specific for PTC but often can be difficult to distinguish from colloid. Follicular carcinoma is more often iso- to hyperechoic with a thick irregular halo. A pure cystic nodule or aggregation of multiple microcystic components ‘spongiform’ signify benign nodule.

Routine FNA is not recommended for sub-centimeter nodules unless there is a solid hypoechogenic nodule with microcalcifications, family history of PTC or previous history of external beam radiation, exposure to ionizing radiation, history of thyroid cancer in the other lobe or an FDG-PET-positive nodule. Detection of cervical lymphadenopathy should lead to FNA of the enlarged node. For non-diagnostic cytology, a repeat FNA with ultrasound is the recommended procedure and results in a diagnostic cytology in 75% of solid and 50% of cystic nodules. About 7% of nodules are repeatedly non-diagnostic and may be cancerous on resection. For a benign FNA diagnosis but with a growth of 20% in nodule diameter or >50% change in nodule volume, a repeat aspiration, preferably under ultrasound guidance, is recommended.

FNA plays a pivotal role in management of patients with thyroid cancer. The goal of thyroid surgery in these cases is the removal of the primary cancer in the thyroid, staging and patient preparation for radioactive ablation and serum thyroglobulin monitoring. For patients with indeterminate or suspicious cytology, a limited lobectomy is the standard approach unless the tumor is large (>4 cm) or is ‘suspicious for PTC’ or there is a family history of thyroid cancer or prior history of radiation exposure when total thryoidectomy is performed. For all biopsy diagnostically positive cases, total or near total thyroidectomy is indicated if the tumor is >1 cm. For smaller cancers, a lobectomy may be sufficient if the patient is lymph node negative. Cervical lymph nodes can be involved in as many as 20–50% of patients (or in up to 90% with PTC) at the time of presentation with preoperative ultrasound identifying 20–31% of these cases. Ultrasound-guided FNA confirms the metastases and is often done with thyroglobulin level of the needle rinse. After surgery, in patients with total or near total thyroidectomy a serum thyroglobulin assay combined with neck ultrasonography should be performed. Rising thyroglobulin levels over time are considered suspicious for recurrence.

References


Thyroid Cytopathology: Bethesda and Beyond

Acta Cytologica 2011;55:4–12


