The 2017 Bethesda System for Reporting Thyroid Cytopathology

Edmund S. Cibas, MD, Syed Z. Ali, MD

Department of Pathology, Brigham and Women’s Hospital and Harvard Medical School, Boston, Massachusetts
Department of Pathology, The Johns Hopkins Medical Institutions, Baltimore, Maryland

KEYWORDS
Thyroid; Cytopathology; Fine-needle aspiration; Terminology; Bethesda; Non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP); Molecular testing

The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) established a standardized, category-based reporting system for thyroid fine-needle aspiration (FNA) specimens. The 2017 revision re-affirms that every thyroid FNA report should begin with 1 of 6 diagnostic categories, the names of which remain unchanged since they were first introduced: (1) Nondiagnostic or Unsatisfactory; (2) Benign; (3) Atypia of Undetermined Significance (AUS) or Follicular Lesion of Undetermined Significance (FLUS); (4) Follicular Neoplasm or Suspicious for a Follicular Neoplasm; (5) Suspicious for Malignancy; and (6) Malignant. There is a choice of two different names for some of the categories: a laboratory should choose the one it prefers and use it exclusively for that category; synonymous terms (e.g., AUS and FLUS) should not be used to denote 2 distinct interpretations. Each category has an implied cancer risk that ranges from 0% to 3% for the “Benign” category to virtually 100% for the “Malignant” category, and, in the 2017 revision, the malignancy risks have been updated based on new (post 2010) data. As a function of their risk associations, each category is linked to updated, evidence-based clinical management recommendations. The recent reclassification of some thyroid neoplasms as non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) has implications for the risk of malignancy, and this is accounted for with regard to diagnostic criteria and optional notes. Such notes can be useful in helping guide surgical management.

© 2017 American Society of Cytopathology. Published by Elsevier Inc. All rights reserved.

Introduction

With its inception, The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) established a standardized reporting system with a limited number of diagnostic categories for thyroid fine-needle aspiration (FNA) specimens. Using TBSRTC, cytopathologists can communicate their interpretations to the referring physician in terms that are succinct, unambiguous, and clinically useful.1-3
TBSRTC has been widely adopted in the United States and in many places worldwide and has been endorsed by the American Thyroid Association.\(^4\) It has improved communication and provided a uniform template for sharing data among investigators. Since its acceptance in clinical practice, however, questions have arisen over the proper use of the diagnostic categories, the associated risks of malignancy, and the appropriate management. By 2016 the time had come to consider revisions. The 2017 revision described herein was inspired by new data and new developments in the field of thyroid pathology: revised guidelines for the management of patients with thyroid nodules,\(^4\) the introduction of molecular testing as an adjunct to cytopathologic examination, and the reclassification of the non-invasive follicular variant of papillary thyroid carcinoma as non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP).\(^5\) Much of the groundwork for this revision was laid down by a symposium entitled “The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC): Past, Present, and Future” at the 2016 International Congress of Cytology in Yokohama, Japan. Preparations for the symposium began 12 months earlier with the designation of a steering group and the appointment of an international panel of 16 cytopathologists and an endocrinologist whose task was to review and summarize the published literature in English since the introduction of TBSRTC.

The symposium, moderated by Drs. Syed Ali and Philippe Vielh, took place on May 30, 2016, and the discussions and recommendations from the symposium have been summarized in a publication by Pusztaszeri et al.\(^6\) Based on the panel’s recommendation, the 6 original general categories (“Nondiagnostic/Unsatisfactory [ND/UNS],” “Benign,” “Atypia of Undetermined Significance/Follicular Lesion of Undetermined Significance [AUS/FLUS],” “Follicular Neoplasm/Suspicious for a Follicular Neoplasm [FN/SFN],” “Suspicious for Malignancy [SUS],” and “Malignant”) have been retained in the 2017 revision, and a revised atlas is in press, with updated and expanded chapters devoted to these categories and refined definitions, morphologic criteria, and explanatory notes.\(^7\)

### Format of the report

For clarity of communication, the 2017 BSRTC continues to recommend that each report begin with a general diagnostic category. Because they are more ambiguous and less clearly descriptive, numerical designations alone (eg, “Bethesda III”) are discouraged for the purposes of cytologic reporting, although the numerical designations may be used in conjunction with the category name—for example, “Atypia of Undetermined Significance (Bethesda III).”

The 6 general diagnostic categories are unchanged and shown in upper case in Table 1. Some categories have two alternative names; a laboratory should choose the one it prefers and use it exclusively for that category; synonymous terms (eg, AUS and FLUS) should not be used to denote two distinct interpretations. Each of the categories has an implied cancer risk (ranging from 0% to 3% for the Benign category to virtually 100% for the Malignant category) that links it to an evidence-based clinical management guideline (Table 2).

For some of the general categories, some degree of subcategorization can be informative and is often appropriate (see Table 1). Additional descriptive comments (beyond such subcategorization) are optional and left to the discretion of the cytopathologist.

Notes and recommendations are not required but can be useful in certain circumstances, particularly if the cytomorphicic features raise the possibility of NIFTP. Some laboratories, for example, may wish to state the risk of malignancy associated with the general category, based on its own data or that found in the literature.

Table 2 shows revised risks of malignancy (ROMs) based on data since 2010. NIFTP has added a wrinkle in this regard by excluding the non-invasive follicular

---

### Table 1: The 2017 Bethesda System for Reporting Thyroid Cytopathology: recommended diagnostic categories.

<table>
<thead>
<tr>
<th>Category</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Nondiagnostic or Unsatisfactory</td>
<td>Cyst fluid only</td>
</tr>
<tr>
<td></td>
<td>Virtually acellular specimen</td>
</tr>
<tr>
<td></td>
<td>Other (observing blood, clotting artifact, etc)</td>
</tr>
<tr>
<td>II. Benign</td>
<td>Consistent with a benign follicular nodule (includes adenomatoid nodule, colloid nodule, etc)</td>
</tr>
<tr>
<td></td>
<td>Consistent with lymphocytic (Hashimoto) thyroiditis in the proper clinical context</td>
</tr>
<tr>
<td></td>
<td>Consistent with granulomatous (subacute) thyroiditis</td>
</tr>
<tr>
<td></td>
<td>Other</td>
</tr>
<tr>
<td>III. Atypia of undetermined significance or follicular lesion of undetermined significance</td>
<td></td>
</tr>
<tr>
<td>IV. Follicular neoplasm or suspicious for a follicular neoplasm</td>
<td>Specify if Hürthle cell (oncocytic) type</td>
</tr>
<tr>
<td>V. Suspicious for malignancy</td>
<td>Suspicious for papillary carcinoma</td>
</tr>
<tr>
<td></td>
<td>Suspicious for medullary carcinoma</td>
</tr>
<tr>
<td></td>
<td>Suspicious for metastatic carcinoma</td>
</tr>
<tr>
<td></td>
<td>Suspicious for lymphoma</td>
</tr>
<tr>
<td></td>
<td>Other</td>
</tr>
<tr>
<td>VI. Malignant</td>
<td>Papillary thyroid carcinoma</td>
</tr>
<tr>
<td></td>
<td>Poorly differentiated carcinoma</td>
</tr>
<tr>
<td></td>
<td>Medullary thyroid carcinoma</td>
</tr>
<tr>
<td></td>
<td>Undifferentiated (anaplastic) carcinoma</td>
</tr>
<tr>
<td></td>
<td>Squamous cell carcinoma</td>
</tr>
<tr>
<td></td>
<td>Carcinoma with mixed features (specify)</td>
</tr>
<tr>
<td></td>
<td>Metastatic carcinoma</td>
</tr>
<tr>
<td></td>
<td>Non-Hodgkin lymphoma</td>
</tr>
<tr>
<td></td>
<td>Other</td>
</tr>
</tbody>
</table>

Adapted from Ali and Cibas\(^7\) with permission of Springer.
variant of papillary thyroid carcinoma from the list of thyroid carcinomas. NIFTP is, nonetheless, a “surgical disease”—surgery is necessary for these nodules—and Table 2 shows ROMs calculated two ways: when NIFTP is not considered a malignancy, and when NIFTP is still included among the “carcinomas.” The higher risk estimates arguably have more clinical relevance because they are defined for surgical disease.

### Nondiagnostic or unsatisfactory

Every thyroid FNA should be evaluated for specimen adequacy. Inadequate samples are reported as Nondiagnostic (ND) or Unsatisfactory (UNS). Examples include specimens with obscuring blood, poor cell preservation, and an insufficient sample of follicular cells. For a thyroid FNA specimen to be satisfactory for evaluation (and benign), at least 6 groups of benign follicular cells are required, each group composed of at least 10 cells. The minimum requirement for group size allows one to determine (by the evenness of the nuclear spacing) whether or not it represents a fragment of a macrofollicle.

Given that the great majority of ND/UNS nodules prove to be benign, one may question whether the criteria for adequacy are too stringent. Lowering the required number of follicular cells would save many patients a repeat FNA. Preliminary data suggest that doing so would substantially reduce ND/UNS interpretations without significantly impacting the false-negative rate.8,9 There is no consensus on a lower number, however, and therefore the criteria have been retained, with the understanding that this is an evolving area that would benefit from more evidence.

The 2017 BSRTC reinforces several exceptions to the numerical requirement of benign follicular cells. Any specimen that contains abundant colloid is adequate (and benign), even if 6 groups of follicular cells are not identified; a sparsely cellular specimen with abundant colloid is, by implication, a predominantly macrofollicular nodule and therefore almost certainly benign. Whenever a specific diagnosis (e.g., lymphocytic thyroiditis) can be rendered, and whenever there is any significant atypia, the specimen is, by definition, adequate for evaluation.

Specimens that consist only of cyst contents (macrophages) are ND/UNS. The significance (and clinical value) of a ND/UNS “cyst contents only” result depends in large part on sonographic correlation. If the nodule is entirely cystic, with no worrisome sonographic features, an endocrinologist might proceed as if it were a benign result. On the other hand, it might be clinically unsatisfactory if the sonographic features are worrisome and the endocrinologist is not convinced that the sample is representative.

The risk of malignancy for an ND/UNS interpretation is difficult to calculate because most ND/UNS nodules are not resected. Among surgically excised nodules initially reported as ND/UNS, the malignancy rate is 9% to 32%. Surgically resected nodules, however, are a selected subset that were either repeatedly ND/UNS or had worrisome clinical/sonographic features, or both. Thus, surgically resected ND/UNS nodules over-represent malignancies compared with the entire cohort of ND nodules. A reasonable extrapolation of the overall risk of malignancy is 5% to 10% (Table 2).9

A repeat aspiration with ultrasound (US) guidance is recommended for cytologically ND/UNS nodules and is diagnostic in most cases, but some nodules remain persistently ND/UNS. Excision is considered for persistently ND/UNS nodules.

---

### Table 2 The 2017 Bethesda System for Reporting Thyroid Cytopathology: implied risk of malignancy and recommended clinical management.

<table>
<thead>
<tr>
<th>Diagnostic category</th>
<th>Risk of malignancy if NIFTP ≠ CA (%)</th>
<th>Risk of malignancy if NIFTP = CA (%)</th>
<th>Usual management&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nondiagnostic or Unsatisfactory</td>
<td>5-10</td>
<td>5-10</td>
<td>Repeat FNA with ultrasound guidance</td>
</tr>
<tr>
<td>Benign</td>
<td>0-3</td>
<td>0-3</td>
<td>Clinical and sonographic follow-up</td>
</tr>
<tr>
<td>Atypia of Undetermined Significance or Follicular Lesion of Undetermined Significance</td>
<td>6-18</td>
<td>~10-30</td>
<td>Repeat FNA, molecular testing, or lobectomy</td>
</tr>
<tr>
<td>Follicular Neoplasm or Suspicious for a Follicular Neoplasm</td>
<td>10-40</td>
<td>25-40</td>
<td>Molecular testing, lobectomy</td>
</tr>
<tr>
<td>Suspicious for Malignancy</td>
<td>45-60</td>
<td>50-75</td>
<td>Near-total thyroidectomy or lobectomy&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Malignant</td>
<td>94-96</td>
<td>97-99</td>
<td>Near-total thyroidectomy or lobectomy&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>Actual management may depend on other factors (e.g., clinical, sonographic) besides the FNA interpretation.

<sup>b</sup>Some studies have recommended molecular analysis to assess the type of surgical procedure (lobectomy versus total thyroidectomy).

<sup>c</sup>In the case of “Suspicious for metastatic tumor” or a “Malignant” interpretation indicating metastatic tumor rather than a primary thyroid malignancy, surgery may not be indicated.

Abbreviations: NIFTP, non-invasive follicular thyroid neoplasm with papillary-like nuclear features; CA, carcinoma; FNA, fine-needle aspiration.
In the past, it was often recommended that the patient with an ND/UNS cytology wait 3 months before a repeat FNA, but this delay often causes patient anxiety. It was reasoned that a transient follicular cell atypia induced by the inflammation that results from a recent FNA might confound interpretation, but a pair of studies does not support this assumption.\textsuperscript{10,11} The American Thyroid Association guidelines now state that there is no need to wait several months before repeating the FNA.\textsuperscript{4}

Unless specified as ND/UNS, the FNA is considered adequate for evaluation; an explicit statement of adequacy remains optional.

**Benign**

The 2017 BSRTC has made essentially no changes to the usage, definition, criteria, or usual management association for this category. Data continue to support a very low false-negative rate (<3%).

**Atypia of undetermined significance or follicular lesion of undetermined significance**

This category has two alternative names; a laboratory should choose the one it prefers and use it exclusively when criteria are fulfilled for this category. AUS and FLUS are therefore synonyms and should not be used to denote two distinct interpretations. It is worth pointing out that, of the two, AUS is more versatile; FLUS applies only to follicular lesions of undetermined significance and cannot be used if the cells are not clearly follicular in origin (lymphoid, parafollicular, parathyroid, etc).

AUS/FLUS has been studied extensively since the advent of TBSRTC, but calculating the ROM associated with this interpretation has been challenging. Because only a minority of AUS/FLUS cases undergo excision, estimating the ROM based on histologic follow-up alone overestimates the ROM due to selection bias: AUS/FLUS nodules (much like Benign and ND/UNS nodules) are usually resected only if there are worrisome clinical or sonographic features, an abnormal repeat aspiration result, and/or an abnormal molecular testing result. AUS/FLUS nodules with a benign repeat aspiration result, and/or an abnormal molecular testing result remain (appropriately) unresected. On the other hand, when calculated using the total number of AUS/FLUS specimens (regardless of surgical follow-up) as the denominator, assuming that unresected nodules are benign, the ROM is underestimated. The actual ROM is between the values obtained using these two different calculations and thus requires extrapolation. It is likely that the ROM of AUS/FLUS has been further overestimated because of publication bias (unexpected/discrepant results are more likely to be published than expected findings).\textsuperscript{12}

Although the overall low-risk nature of AUS/FLUS aspirates has been borne out, new (pre-NIFTP) data suggest that the ROM is higher than originally estimated and closer to 10% to 30% (Table 2). On the other hand, if the risk is recalculated by removing NIFTPs from the tally of malignancies, the risk diminishes to 6% to 18%, because early data suggest that NIFTP constitutes a substantial proportion of the “malignancies” hidden in this category.\textsuperscript{13,14}

The ROM differs according to the nature of the atypia. The 2017 BSRTC recommends subclassification of the atypia, even though this will not generally affect patient management. Descriptive language like “cytologic atypia” and “architectural atypia” is preferred (rather than “rule out papillary carcinoma,” etc) because of its less provoking nature, as follows:

1. **Cytologic atypia.** This may take one of several different forms: focal nuclear changes, extensive but mild nuclear changes; atypical cyst lining cells, or “histiocytoid” cells.\textsuperscript{5-17}

2. **Architectural atypia.** This is often a sparsely cellular sample but one that is composed mostly of microfollicles.

3. **Cytologic and architectural atypia.** Cytologic atypia and architectural atypia are not mutually exclusive.

4. **Hürthle cell AUS/FLUS.** This is often a sparsely cellular sample comprised exclusively of Hürthle cells. Alternatively, AUS/FLUS may be used for a moderately or markedly cellular sample composed exclusively (or almost exclusively) of Hürthle cells, yet the clinical setting suggests a benign Hürthle cell nodule, such as in chronic lymphocytic (Hashimoto) thyroiditis or a multinodular goiter when other sampled nodules show similar features.

5. **Atypia, not otherwise specified.**

It is good to think of AUS/FLUS as a category of last resort. The original TBSRTC recommended that an effort be made to limit its use to approximately 7% or fewer of all thyroid FNAs. This proved a difficult challenge for many laboratories, and a more realistic limit might be 10%.

The usual management now includes consideration of molecular testing.

**Follicular neoplasm or Suspicious for a follicular neoplasm**

This category likewise has two alternative names; a laboratory should choose the one it prefers and use it exclusively; FN and SFN are synonymous terms and should not be used to denote two distinct interpretations. SFN is preferred by some laboratories because a significant proportion of cases (up to 35%) prove not to be neoplasms but rather hyperplastic proliferations of follicular cells, most commonly those of multinodular goiter.\textsuperscript{16,22}

The 2017 BSRTC includes a modification to the definition and diagnostic criteria for this category in light of NIFTP. In the original BSRTC, cases that demonstrated the nuclear features of papillary thyroid carcinoma were
excluded from this category. The new definition reads as follows:

Follicular-patterned cases with mild nuclear changes (increased nuclear size, nuclear contour irregularity, and/or chromatin clearing) can be classified as FN/SFN so long as true papillae and intranuclear pseudoinclusions are absent; a note that some nuclear features raise the possibility of a follicular variant of papillary thyroid carcinoma (FVPTC) or NIFTP can be included.13,14

If the cytologic features raise the possibility of FVPTC or NIFTP (a predominance of microfollicles and only mild or focal nuclear changes), the following optional note (or something similar) may be useful:

Note: Although the architectural features suggest a follicular neoplasm, some nuclear features raise the possibility of an invasive follicular variant of papillary carcinoma or its recently described indolent counterpart, NIFTP; definitive distinction among these entities is not possible on cytologic material.

This note will apply only to a subset of FN/SFN cases—those with mild nuclear changes.

As with AUS/FLUS, if the ROM for FN/SFN is recalculated by removing NIFTPs from the tally of malignancies, the risk diminishes (see Table 2); early data suggest that NIFTP constitutes a substantial proportion of the “malignancies” hidden in this category as well.13,14

The recommended management of a patient with a diagnosis of FN/SFN is surgical excision of the lesion, most often a hemithyroidectomy or lobectomy, but molecular testing may be used to supplement risk assessment rather than proceeding directly to surgery.

**Suspicious for malignancy**

As with AUS/FLUS and FN/SFN, if the ROM for this category (SUS) is recalculated by removing NIFTPs from the tally of malignancies, the risk diminishes (see Table 2); early data suggest that NIFTP constitutes a substantial proportion of the “malignancies” hidden in this category as well.13,14

Some (not all) cases in this category raise the possibility of FVPTC or NIFTP. For this subset, the following optional note (or something similar) may be useful:

Note: The cytomorphicologic features are suspicious for a follicular variant of papillary thyroid carcinoma or its recently described indolent counterpart NIFTP.

This can be useful in guiding the clinical team in the direction of lobectomy rather than thyroidectomy for this subset of SUS cases.

**Malignant**

The general category “Malignant” is used whenever the cytomorphologic features are conclusive for malignancy. Descriptive comments that follow are used to subclassify the malignancy and summarize the results of special studies, if any.

Based on early studies, NIFTP constitutes only a very small fraction of cases that are interpreted as Malignant. Nevertheless, the 2017 BSRTC has modified the definition and criteria for cases of papillary thyroid carcinoma (PTC) that belong in the Malignant category. To avoid false-positives due to NIFTP, it suggests limiting use of the Malignant category to cases with “classical” features of PTC (true papillae, psammoma bodies, nuclear pseudoinclusions).5,23 Nevertheless, it is likely that a small number of Malignant cytologic interpretations will be followed by a histologic NIFTP diagnosis, and thus the following optional note may be used when the diagnosis “Malignant; papillary thyroid carcinoma” is made:

Note: A small proportion of cases (~3%-4%) diagnosed as malignant and compatible with papillary thyroid carcinoma may prove to be NIFTP on histopathologic examination.

**Highlights of the 2017 Bethesda System for Thyroid Cytopathology**

The original 6 categories remain unchanged, but a number of enhancements have been introduced with the 2017 BSRTC:

1. The risks of malignancy have been recalculated based on post-2010 data.
2. The risks of malignancy are shown two ways (see Table 2): (1) when NIFTP is not considered a malignancy; and (2) when NIFTP is still included among the “carcinomas.” The higher risk estimates may have more clinical relevance because they are defined for surgical disease.
3. The “usual management” of AUS/FLUS and FN/SFN now incorporates the option of molecular testing.
4. The definition and diagnostic criteria for FN/SFN have been revised in light of NIFTP. Cases that demonstrate mild nuclear changes associated with papillary thyroid carcinoma are now included.
5. The definition and diagnostic criteria for the papillary thyroid carcinoma subset of the Malignant category have been modified to suggest limiting use to cases with “classical” features of PTC.
6. Optional educational notes may be used for the subsets of FN/SFN and SUS with cytomorphologic features suggestive of FVPTC or NIFTP.
7. An optional educational note may be used for “Malignant; papillary thyroid carcinoma” cases to acknowledge that a small proportion may prove to be NIFTP.

It is our hope that the 2017 BSRTC will continue to stimulate interest in the improvement of thyroid
cytopathologic diagnosis and the betterment of patients with thyroid nodular disease. Subsequent experience, it is expected, will lead to further refinements to this terminology framework.

Acknowledgments

The authors thank Dr. Erik Alexander for his review of the manuscript and helpful comments.

Funding sources

No specific funding was disclosed.

Conflict of interest disclosures

The authors made no disclosures.

References