The Bethesda System for Reporting Thyroid Cytopathology: Proposed Modifications and Updates for the Second Edition from an International Panel

Marc Pusztaszeri, Esther Diana Rossi, Manon Auger, Zubair Baloch, Justin Bishop, Massimo Bongiovanni, Ashish Chandra, Beatrix Cochand-Priollet, Guido Fadda, Mitsuyoshi Hirokawa, SoonWon Hong, Kennichi Kakudo, Jeffrey F. Krane, Ritu Nayar, Sareh Parangi, Fernando Schmitt, William C. Faquin

Department of Pathology, Geneva University Hospitals, Geneva, Switzerland; Division of Anatomic Pathology and Histology, Agostino Gemelli School of Medicine, Università Cattolica del Sacro Cuore, Rome, Italy; Department of Pathology, McGill University and McGill University Health Center, Montreal, QC, Canada; Department of Pathology and Laboratory Medicine, Hospital of the University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA, and Department of Pathology, Johns Hopkins School of Medicine, Baltimore, MD, USA; Institute of Pathology, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland; Department of Cellular Pathology, Guy’s & St Thomas’ NHS Foundation Trust, London, UK; Department of Pathology, Cochin Hospital, Paris, France; Department of Diagnostic Pathology, Kuma Hospital, Kobe, Japan; Department of Pathology, Gang Nam Severance Hospital, Yonsei University, Seoul, South Korea; Department of Pathology, Nara Hospital, Kinki University Faculty of Medicine, Nara, Japan; Department of Pathology, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA; Department of Pathology, Feinberg School of Medicine, Northwestern University and Northwestern Medicine, Chicago, IL, and Department of Surgery, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA; Department of Medicine, Laboratoire National de Santé, Dudelange, Luxembourg; Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Key Words
Bethesda system · Thyroid lesion · Classification systems · Diagnostic categories · NIFTP

Abstract
The Bethesda System for Reporting Thyroid Cytology (TBSRTC) was proposed in 2007 at the National Cancer Institute Thyroid Fine Needle Aspiration State of the Art and Science Conference held in Bethesda, Maryland. The aim was to address the inconsistent and sometimes confusing reporting terminologies used for thyroid FNA throughout the world. The TBSRTC consists of 6 diagnostic categories, each associated with an implied risk of malignancy that translates directly into a clinical management algorithm. Since the publication of the TBSRTC Atlas in January 2010, considerable experience has been gained regarding its application in cytology practice, clinical impact, and limita-

This article is being jointly published by the Journal of the American Society of Cytopathology and Acta Cytologica.
tnations. In conjunction with the International Academy of Cytology (IAC), an international panel composed of sixteen cytopathologists and an endocrinologist with special interest in thyroid cytology, including several co-authors of the 2010 TBSRTC Atlas, was created to: (1) analyze the current worldwide impact of TBSRTC, (2) report on the current state of TBSRTC based upon a review of the published literature, and (3) provide possible recommendations for a future update of TBSRTC. Herein, we summarize the panel’s deliberations and key recommendations that our panel hopes will be useful during the preparation of the second edition of TBSRTC.

© 2016 S. Karger AG, Basel

Introduction

The introduction of The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC), following the “National Cancer Institute Thyroid Fine Needle Aspiration State of the Science Conference” held in Bethesda, MD, in 2007, offered the opportunity to establish a uniform six-tiered reporting system for thyroid FNA [1]. The TBSRTC not only included the diagnostic designations for the commonly encountered benign and malignant thyroid lesions in FNA specimens but also for those that are often diagnosed as “indeterminate for malignancy (the so-called ‘grey zone’)” [1]. The latter were sub-classified into the diagnostic categories of: (i) Atypia of undetermined significance (AUS)/Follicular lesion of undetermined significance (FLUS); (ii) Follicular neoplasm (FN)/Suspicious for a follicular neoplasm (SFN) and (iii) Suspicous for malignancy (SM). For each of the diagnostic categories, TBSRTC defined an implied risk of malignancy (ROM) based upon the data from the available literature; and clinical management scenario(s) [1].

Over the past seven years, TBSRTC has found widespread international acceptance, and has contributed significantly to the management of thyroid nodules by increasing the quality and reproducibility of thyroid cytology reporting [2, 3]. Moreover, the use of TBSRTC has been endorsed by the American Thyroid Association (ATA) as part of the revised 2015 ATA guidelines [4].

In light of the recent advances and developments in the field of thyroid disease, a second edition of TBSRTC, scheduled to be released in 2018, seems appropriate. In anticipation of the latter, Drs. Syed Ali, Edmund Cibas and Philippe Viehl proposed the formation of an international panel of cytopathologists, pathologists and an endocrine surgeon to assess the available literature and propose useful modifications to TBSRTC. The topics to be addressed would include cytomorphological criteria for FNA classification, reporting terminology, implied ROM for each diagnostic category, and the anticipated changes due to the recently described non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) [5, 6]. The summary of this effort and the panel recommendations were presented in a Symposium entitled “TBSRTC: Past, Present, and Future” moderated by Drs. Ali and Vielh at the 2016 ICC meeting held in Yokohama, Japan. Herein, we summarize the panel’s deliberations and key recommendations that our panel hopes will be useful during the preparation of the second edition of TBSRTC.

Selected General Issues of TBSRTC

Reporting of Specific Uncommon Non-Thyroid Entities in TBSRTC

TBSRTC was designed specifically for the reporting of thyroid follicular and C-cell-derived lesions. On occasion, an FNA will be performed of a non-thyroid entity presenting as a thyroid nodule, the most common being an enlarged parathyroid secondary to hyperplasia or neoplasia. Other potential lesions include intrathyroidal paragangliomas, Langerhans cell histiocytosis, solitary fibrous tumor, and rare mesenchymal tumors. For such cases where specific classification is possible, our panel does not recommend creation of a separate diagnostic category for these uncommon lesions of non-thyroid origin. Instead, the lesion can be reported using traditional cytologic reporting categories.

Implied Risk of Malignancy (ROM)

Since the introduction of TBSRTC into clinical practice, numerous studies have contributed to refinement of the implied ROM attributed to each of the 6 diagnostic categories [2–4]. Some studies, however, have reported inflated ROMs particularly in the indeterminate diagnostic categories, that are influenced by case selection, partial verification, and publication biases [7]. Our panel recommends adjustments to the implied ROM in TBSRTC based upon a selected group of studies that included large cohorts of cases or meta-analyses [2, 3]. Such modifications could be provided as an update to Table 1.2 in the Atlas.

Non Invasive Follicular Thyroid Neoplasm with Papillary-Like Nuclear Features (NIFTP)

The introduction of NIFTP [5, 6], presents a challenge for thyroid cytology. Recent studies indicate that a major-
The reporting of thyroid FNA using TBSRTC offers the possibility for laboratory monitoring of key parameters, in particular, the over-utilization of the AUS/FLUS category [2, 3]. This can be avoided by intra-laboratory monitoring at regular intervals of the rate of AUS/FLUS. One study proposed using the AUS/M ratio as a quality control measure to accomplish this [18]. In addition, the panel encourages continuous cytologic-histologic correlation for thyroid FNA.

Specific Issues Pertaining to the Different Diagnostic Categories and Chapters of TBSRTC

Non-Diagnostic/Unsatisfactory
A specimen is defined as Non-Diagnostic (ND) in TBSRTC if it fails to meet the adequacy criteria (i.e., at least 6 groups, each with at least 10 benign-appearing, well-visualized follicular cells), with exceptions for thyroiditis, “abundant” colloid, or any degree of perceived atypia [1]. It is important to acknowledge that these adequacy criteria are based upon consensus opinion and are not evidence-based. Clarification of the specific adequacy criteria for laboratories using liquid based preparations (LBP) (ThinPrep and Surepath) alone or in combination with smear is recommended. While some controversy has existed concerning the reporting of cystic lesions, the panel recommends that “Cyst fluid (macrophages) only” cases should still be reported as ND with an explanatory note, as exemplified in the sample reports of TBSRTC Atlas.

The management recommendation for ND samples including cystic lesions have been updated in the 2015 revised ATA guidelines [4]. In large series of thyroid FNA cases, of which 7–26% were classified based upon TBSRTC, ND samples constituted 2–16% of all FNA samples, of which 7–26% were only one term would be assigned to each diagnostic category; however, it is understood that most laboratories have already established the use of one of two terms for selected diagnostic categories in their clinical practice. Thus, for practical reasons, it would be counterproductive to eliminate this flexibility in nomenclature. The application of TBSRTC terminology, with reasonably acceptable variations can be maintained as long as there is consistency at a given laboratory and good communication between the different members of the health care team. However, it should not open the doors to major alterations of the terminology which would defeat the purpose of TBSRTC.

Quality Control
The reporting of thyroid FNA using TBSRTC offers the possibility for laboratory monitoring of key parameters, in particular, the over-utilization of the AUS/FLUS category [2, 3]. This can be avoided by intra-laboratory monitoring at regular intervals of the rate of AUS/FLUS. One study proposed using the AUS/M ratio as a quality control measure to accomplish this [18]. In addition, the panel encourages continuous cytologic-histologic correlation for thyroid FNA.

Specific Issues Pertaining to the Different Diagnostic Categories and Chapters of TBSRTC

Non-Diagnostic/Unsatisfactory
A specimen is defined as Non-Diagnostic (ND) in TBSRTC if it fails to meet the adequacy criteria (i.e., at least 6 groups, each with at least 10 benign-appearing, well-visualized follicular cells), with exceptions for thyroiditis, “abundant” colloid, or any degree of perceived atypia [1]. It is important to acknowledge that these adequacy criteria are based upon consensus opinion and are not evidence-based. Clarification of the specific adequacy criteria for laboratories using liquid based preparations (LBP) (ThinPrep and Surepath) alone or in combination with smear is recommended. While some controversy has existed concerning the reporting of cystic lesions, the panel recommends that “Cyst fluid (macrophages) only” cases should still be reported as ND with an explanatory note, as exemplified in the sample reports of TBSRTC Atlas.

The management recommendation for ND samples including cystic lesions have been updated in the 2015 revised ATA guidelines [4]. In large series of thyroid FNA cases, of which 7–26% were classified based upon TBSRTC, ND samples constituted 2–16% of all FNA samples, of which 7–26% were only one term would be assigned to each diagnostic category; however, it is understood that most laboratories have already established the use of one of two terms for selected diagnostic categories in their clinical practice. Thus, for practical reasons, it would be counterproductive to eliminate this flexibility in nomenclature. The application of TBSRTC terminology, with reasonably acceptable variations can be maintained as long as there is consistency at a given laboratory and good communication between the different members of the health care team. However, it should not open the doors to major alterations of the terminology which would defeat the purpose of TBSRTC.

Quality Control
The reporting of thyroid FNA using TBSRTC offers the possibility for laboratory monitoring of key parameters, in particular, the over-utilization of the AUS/FLUS category [2, 3]. This can be avoided by intra-laboratory monitoring at regular intervals of the rate of AUS/FLUS. One study proposed using the AUS/M ratio as a quality control measure to accomplish this [18]. In addition, the panel encourages continuous cytologic-histologic correlation for thyroid FNA.

Specific Issues Pertaining to the Different Diagnostic Categories and Chapters of TBSRTC

Non-Diagnostic/Unsatisfactory
A specimen is defined as Non-Diagnostic (ND) in TBSRTC if it fails to meet the adequacy criteria (i.e., at least 6 groups, each with at least 10 benign-appearing, well-visualized follicular cells), with exceptions for thyroiditis, “abundant” colloid, or any degree of perceived atypia [1]. It is important to acknowledge that these adequacy criteria are based upon consensus opinion and are not evidence-based. Clarification of the specific adequacy criteria for laboratories using liquid based preparations (LBP) (ThinPrep and Surepath) alone or in combination with smear is recommended. While some controversy has existed concerning the reporting of cystic lesions, the panel recommends that “Cyst fluid (macrophages) only” cases should still be reported as ND with an explanatory note, as exemplified in the sample reports of TBSRTC Atlas.

The management recommendation for ND samples including cystic lesions have been updated in the 2015 revised ATA guidelines [4]. In large series of thyroid FNA cases, of which 7–26% were classified based upon TBSRTC, ND samples constituted 2–16% of all FNA samples, of which 7–26% were only one term would be assigned to each diagnostic category; however, it is understood that most laboratories have already established the use of one of two terms for selected diagnostic categories in their clinical practice. Thus, for practical reasons, it would be counterproductive to eliminate this flexibility in nomenclature. The application of TBSRTC terminology, with reasonably acceptable variations can be maintained as long as there is consistency at a given laboratory and good communication between the different members of the health care team. However, it should not open the doors to major alterations of the terminology which would defeat the purpose of TBSRTC.
eventually resected [4]. The ROM was 2–4% among all initially ND samples and 9–32% among those ND samples that were resected [4]. While it was originally recommended to wait >3 months to perform a repeat FNA, evidence exists to support shorter intervals for repeating an FNA after a ND result [4, 19, 20]. However, it should be noted that reactive atypia and other cellular changes may be present if the time interval is shortened.

**Benign**

Several recent studies have confirmed that the ROM is very low for this category [4, 21, 22]. Risk stratification based upon ultrasound patterns (ATA 2015 revised guidelines) can be used to guide clinical follow-up of thyroid nodules with benign cytology [4]. We propose the inclusions of more LBP images as well as IgG4 thyroiditis for this category in the Second Edition TBSRTC Atlas.

**AUS/FLUS**

Several retrospective and prospective studies have evaluated the role of the AUS/FLUS category in the management of thyroid nodules. The panel endorses the AUS/FLUS category, as it is now widely accepted and is also included in the current ATA clinical management guidelines [4]. In addition, the AUS/FLUS category has been shown to maintain the sensitivity for detecting thyroid neoplasms, with decreased false-negative and false-positive rates, making thyroid FNA a more effective screening test. While our panel would favor the use of a single term such as AUS/FLUS for this category, we recognize that the most practical approach would be to maintain current nomenclature practices. As such, only one term, either AUS or FLUS would be selected by a laboratory because they are synonymous.

Our panel recommends subclassification of AUS/FLUS, particularly with respect to the presence or absence of nuclear atypia. This could be accomplished using a concise note, Common patterns that would be candidates for subclassifiers include: architectural atypia, nuclear atypia, oncocytic features, and NOS. Several studies have demonstrated that subclassification within the AUS/FLUS category can be effective to better define the ROM, and that nuclear atypia is associated with a higher ROM than other AUS/FLUS patterns [23–28].

In order to limit the overall rate of AUS/FLUS, compromised samples lacking significant atypia should be classified as ND rather than AUS/FLUS. For the management of AUS/FLUS, all clinical, radiologic, pathologic, and molecular findings should be integrated for the most informed, accurate, and individualized patient assessment (a.k.a. personalized approach). While repeat FNA is recommended in the current version of TBSRTC, the revised ATA guidelines recognize molecular testing as a valid alternative [4].

**Follicular Neoplasm/Suspicious for a Follicular Neoplasm**

As with two other diagnostic categories where a choice between two different nomenclature designations is available, our panel favors the use of one; however, we recognize that for practical reasons, the use of either term is acceptable. In addition, there is concern that neither of the two nomenclature options for this category is ideal because: (1) SFN can be mistaken for SM due to the term ‘suspicious’ that these two categories share, and (2) FN denotes an unjustified certainty about the presence of a neoplasm. The alternative use of the term “Suggestive of FN” has been proposed by some, as this would connote uncertainty regarding neoplasia while also being distinct from the SM category. In our opinion, current diagnostic criteria for FN/SFN could be further defined to more clearly distinguish lesions that would belong in this category from those that might be more appropriately placed in the AUS/FLUS and SM categories. As mentioned above, with the advent of NIFTP, one could consider classifying follicular patterned lesions with nuclear atypia as FN/SFN rather than SM. Diagnostic surgical excision has been the long-established standard of care for the management of FN/SFN; however, the ATA 2015 guidelines now provide the option of molecular testing [4].

**Follicular Neoplasm/Suspicious for a Follicular Neoplasm, Hürthle Cell Type**

The use of the term “oncocytic” instead of “Hürthle cell” throughout TBSRTC is recommended to align with current WHO classification terminology. A rare malignant subset of oncocytic lesions with abundant colloid has been described [29], and this pitfall should be mentioned in TBSRTC. In addition, the accuracy of molecular testing may be lower for FNAs of oncocytic lesions. Specifically, there is an increased rate of “Suspicious” results in benign oncocytic lesions using the Afirma gene-expression classifier [30–33].

**Suspicious for Malignancy**

Poorly differentiated thyroid carcinoma (PDTC) should be included among the differential diagnoses in the SM category. The utility of molecular testing employing panels with high positive predictive value is likely to
be beneficial in the SM category, notably for the detection of potential NIFTP cases (e.g., RAS, THADA or BRAF<sup>K601E</sup> positive and BRAF<sup>V600E</sup> negative). According to the 2015 ATA guidelines, total thyroidectomy is an option for patients with nodules that are cytologically SM, positive for known mutations specific for carcinoma (BRAF<sup>V600E</sup>, TERT-Promoter, p53), sonographically suspicious, or of large size (>4 cm) [4] with the understanding that a subset of these may represent NIFTP on surgical excision. A comment that NIFTP comprises a subset of SM cases is optional for this category.

**Papillary Thyroid Carcinoma**

Recent publications have highlighted certain differences in the cytological features of PTC in LBPs as compared to conventional cytologic preparations [34, 35]. In addition, the cytological features of several variants of PTC, such as hobnail, diffuse sclerosing, solid, and cribriform-morular variants should be more fully described and illustrated in the revised TBSRTC. Although definitive cytologic subclassification as a specific PTC variant is not necessary, increasing familiarity with the morphologic features of PTC variants will decrease the risk of misdiagnosis.

More detailed descriptions of the FVPTC as mentioned previously would be useful as it relates to NIFTP, and there may be a role for classifying FVPTC as SM to avoid a false positive diagnosis. Updated information pertaining to the clinical behavior of FVPTC, especially in light of recent TCGA data [36], the new NIFTP terminology [5, 6], and other clinico-pathological studies is needed [37–39]. Hyalinizing trabecular tumor is now recognized as a mimic of PTC rather than as a variant of PTC per se [40].

**Medullary Thyroid Carcinoma**

MTC accounts for only 1–2% of thyroid cancers in the USA, a much lower range than frequently cited (3–5, and 7% in TBSRTC), primarily due to the marked increase in the relative incidence of PTC [41]. As suggested in the 2015 ATA Guidelines, the assessment of calcitonin levels can be helpful both in the blood and in FNA washout fluid for suspected MTC cases, especially when IHC confirmation is not possible (e.g., limited cellularity or equivocal IHC results) [41, 42]. Additional descriptors pertaining to rare morphologic variants of MTC (e.g., small cell, giant cell) that can be diagnostically challenging would be useful [43, 44]. Other diagnostic pitfalls that could be addressed include: (1) Morphological overlap with oncocytic follicular lesions and PDTC [45]; and (2) Morphological and immunohistochemical overlap with other neuroendocrine tumors [46]. In addition, some aspects of the management of MTC have been updated in the 2015 revised ATA guidelines [41].

**Poorly Differentiated Thyroid Carcinoma (PDTC)**

As mentioned above, a subset of PDTC will be diagnosed as SM, and in some cases PDTC can be a diagnostic pitfall for MTC. Additionally, an oncocytic variant of PDTC has been recently described [47]. Some advances have been made in the molecular characterization of PDTC, including TERT mutations and microRNA signatures; however, these molecular features are not specific for PDTC and therefore have limited use in routine practice.

**Undifferentiated (Anaplastic) Carcinoma (UTC) and Squamous Cell Carcinoma of the Thyroid**

There is a potential role of ICC to confirm UTC. PAX-8 is a helpful immunocytochemical stain [48–50], since it is usually retained in UTC, while TTF-1 can rarely be focal and variably positive, and thyroglobulin is always negative. Some advances have been made in the molecular characterization of UTC, including high rates of MAPK pathway genes mutations, p53 mutations, and TERT–promoter mutations. However, the clinical management and the prognosis of UTC or primary squamous cell carcinoma have not changed significantly.

**Metastases and Lymphomas**

TBSRTC defines the diagnostic criteria and provides explanatory notes for some of the most common metastatic cancers; the latter could be expanded to include others such as metastatic neuroendocrine tumors. Several recent FNA studies on secondary thyroid malignancies have emphasized the important role of clinical history and the judicious use of immunomarkers (limited mostly to cases with adequate cellularity) to increase the sensitivity and accuracy of FNA for detecting metastases [51, 52]. An expanded array of useful immunocytochemical markers, including PAX-8 and GATA-3 are available. In addition, information on the clinical management of metastatic disease and lymphomas is needed.

**Conclusions**

The launch of TBSRTC in January 2010 proved to be very successful, and TBSRTC is now the most common classification worldwide for the reporting of thyroid FNA.
specimens. Since its inception, additional knowledge and experience pertaining to various aspects of thyroid FNA have been obtained. In view of this, and in anticipation of the second edition of TBSRTC in 2018, our panel has provided its suggestions which we hope will prove useful for the second edition covering various aspects of thyroid FNA including nomenclature, differential diagnosis, the potential impact of NIFTP on the indeterminate diagnostic categories, utility of molecular and IHC markers, and clinical management.

References

4 Haugen BR, Alexander E, Bible KC, et al; American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer: 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. Thyroid 2016; 26:1–133.

Pusztazeri et al.


