

## Post Revised Bethesda System for Reporting Thyroid Cytology 2017 Experience: Lessons Learned

Imad A El Hag\*

Department of Pathology and Laboratory Medicine, Security Force Hospital, Riyadh, Saudi Arabia

\*Corresponding author: Dr. Imad A El Hag, Department of Pathology and Laboratory Medicine, Security Force Hospital, Riyadh, Saudi Arabia, E-mail: imadam12@gmail.com

Received: 10-Jun-2022, Manuscript No. DPO-22-66364; Editor assigned: 15-Jun-2022, PreQC No. DPO-22-66364 (PQ); Reviewed: 29-Jun-2022, QC No. DPO-22-66364; Revised: 06-Jul-2022, Manuscript No. DPO-22-66364 (R); Published: 13-Jul-2022, DOI: 10.4172/2476-2024.7.3.205

Copyright: © 2022 Hag IAEI. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Abstract

**Back ground:** BSFRTC is widely adopted in the management of the thyroid nodules. Post implementation experiences and the lessons learned from them is the subject of this short review.

**Methods:** Published experiences and appraisals post BSFRTC reviewed, focusing on the variation in reported ROM, state of Atypia of Undetermined Significance (AUS) category and the cytological diagnosis of Non-Invasive Follicular Neoplasm with Papillary like Nuclear Features (NIFTP) as well as the role of two key indicators, namely Risk of Neoplasia (RON) and surgical rates.

**Results:** The differences in ROM between the reported values and values expected by BSFRTC are due to variation in practice among institutions and it does not affect the performance of the schema. Using the total number of cases in the category as a dominator in the calculation of ROM shall reduce or even abolish the differences. Opinion about AUS varied between abolishing and merging it with SFN or maintaining it with further sub-categorization. The consensus post BSFRTC 2017 is to keep the AUS category and stratified into different subgroups based on the type and degree of atypia. RON and the surgical follow-up rates are essential quality indicators. RON/ROM ratio could be utilized to determine the appropriate management for each diagnostic category on an institutional basis. A RON/ROM ratio close to unity in indeterminate categories is indicative for surgical triaging. The rate of surgery in different Bethesda group is an indication of how far clinicians accept the schema. NIFTP cannot be easily separable from other follicular lesions on cytology alone or combined with available molecular tests. Considering NIFTP benign, improve the stratification in ROM between indeterminate groups.

**Conclusion:** BSFRTC is an effective schema, well accepted by clinician, and has improved greatly the management of thyroid nodules. Maintaining AUS with sub categorization, considering NIFTP benign and calculating RON and surgical rates shall improve the performance of the schema.

**Keywords** Revised Bethesda System for Reporting Thyroid Cytology (BSFRTC); Risk of Malignancy (ROM); AUS; Non-Invasive Follicular Tumor with Papillary like Nuclear Features (NIFTP)

### Introduction

The published data post the first edition of BSFRTC 2007 and the new development in thyroid pathology and management, specifically the introduction of the indolent tumor Non-Invasive Follicular Tumor with Papillary like Nuclear Features (NIFTP) have mandated a revision to the first edition of BSFRTC of 2007, which was completed in 2017 [1,2]. The revised format has retained the six reporting categories, recalculated the Risk of Malignancy (ROM) when NIFTP is counted as malignant or benign, and incorporated molecular testing in the final work as optional. It also allowed an overlap in the estimated ROM between AUS and SFN in the presence of NIFTP as a malignant lesion and recommended sub- classification of AUS into five subgroups based on the type of atypia.

The implementation of the BSFRTC has dramatically improved the management of thyroid nodules. It removed the confusion and ambiguity created by descriptive, non-categorized reporting and

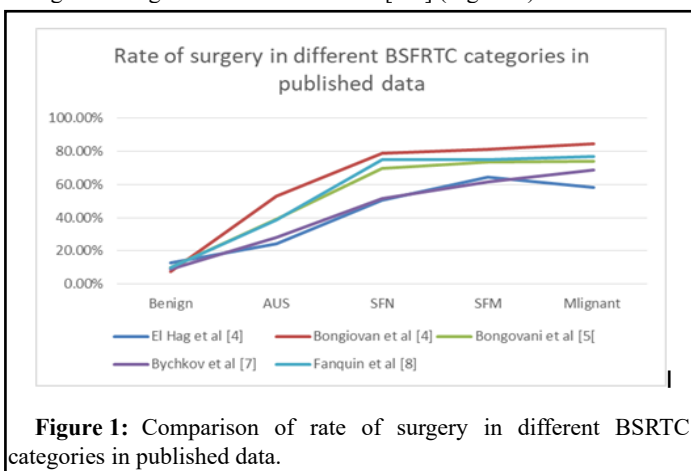
enhanced the communication between pathologists and clinicians [3]. Additionally, comparisons among different institutions can easily be obtained and most importantly, the BSFRTC has reduced unnecessary surgeries among benign nodules [4,5]. However, the implementation of BSFRTC has been associated with certain issues, which have been the subject of post implementation appraisals and discussions, among them are; the variation in the reported ROM and its deviation to the higher side of the estimated values by BSFRTC, the state of the debatable AUS group and the separation of NIFTP from other follicular tumors in cytology material. These issues are the main subjects of this short review.

### Literature Review and Comment

#### Reported vs. estimated ROM

Certainly there is variation in the calculated ROM between the published studies and the estimated ones by BSFRTC [6]. The reported values tended to be towards the higher side. The reported ROM ranges 31.2%-43.9% for AUS, 33.2%-57.5% for suspicious for SFN and 82.6%-93.6% for the SFM is higher than the estimated

ranges of 10%-30%; 10%-40%; and 50%-74%, respectively. The variation in the calculated ROM between studies and their deviation from the predicted ranges by the BSFRTC is not necessarily a flaw to the schema nor inherent to it, but most likely due to differences in practice between institutions. Pathologist experience and training, the number of pathologists involved in reporting thyroid FNAs, types of stains, palpation vs. US-guided and on-site evaluation, and the bias in triaging patients for surgery, are all factors that inevitably impacted the results. Despite the variation in the reported ROM in various diagnostic categories between institutions, the results revealed the same trend of increasing ROM from benign, through AUS, SFN, and SFM to malignant. The variation between the studies was mainly in the slope of the increase of ROM, but not the trend. This was the main reason that both the rate of surgical follow-up and Risk of Malignancy increased steadily from benign through intermediate to frank malignant categories in several studies [4-8] (Figure 1).



**Figure 1:** Comparison of rate of surgery in different BSRTC categories in published data.

The consistent results transpired despite the variation in their ROM. There is direct proportional relationship between category ROM and rate of surgical follow-up in many studies [6], which is a good indicator of the degree of perception and acceptance of the terminologies by clinicians. The reported ROM for the AUS category generally been over estimated, as many AUS nodules are just observed with no surgical interference or “enriched” due to referral or selection bias [9]. To avoid the over inflation of ROM produced by referral selection and bias, ROM has to be calculated from the total number of cases in the category and not only from those with surgical pathology follow up, especially in the intermediate categories. From our experience, this will produce a ROM falling closer to the estimated values by the BSFRTC [6].

### AUS screening vs. diagnostic category

The consensus during the NCI art of science conference on thyroid cytopathology held in 2007 was to split the indeterminate category into two separate categories, AUS and SFN, which remained the same in the second edition of the BSFRTC. Cibas and Ali [1,2] have enlisted nine diverse scenarios for an AUS category diagnosis; these include architectural atypia, cytological atypia and technically compromised specimens, making it the least defined among the BSFRTC categories. Because of its heterogeneity, the AUS category has the highest inter-observer variability among the BSFRTC categories [10,11]. The post implementation opinions about AUS divided between merging it with SFN in a five or four category schema, which has certainly improved the stratification between

diagnostic categories, and subcategorizing it into different morphologically distinct subgroups with different probabilities of malignancy [5,12-16]. The new recommendation in the latest BSFRTC edition was to sub-classify AUS into five subgroups based on the type of atypia [2]. Looking at the data post second edition of the BSFRTC, a true overlap in the ROM between AUS and SFN observed in most of the published studies [6], particularly when NIFTP is included among malignant tumors. Interestingly, removal of NIFTP from the malignant tumors, slightly improved the stratification in the ROM between AUS and SFN in those studies. The second edition of the BSFRTC has astutely anticipated this. The calculation of RON has also amplified the stratification between the two indeterminate groups, AUS and SFN [6]. In our latest study, while the RON and ROM remained comparable in SFN, SFM and malignant categories, ROM was significantly lower than RON in the AUS category, especially when NIFTP was not counted among the malignant tumors [6]. This is because AUS tended to include more benign and indolent neoplasms, such as FA, NIFTP than malignant ones. Sensitivity of a test is defined as the tests ability to correctly classify an individual as diseased, while specificity is defined as the test ability to correctly classify an individual as disease-free. This is crucial for a screening test, which should have a well-balanced sensitivity and specificity; a higher sensitivity might lead to overtreatment and unnecessarily increased patient’s anxiety, while high specificity might lead to loss of diseased individuals for routine follow-up. When sensitivity and specificity, PPV and NPV values were compared between the three intermediate groups, AUS, SFN and SFM, it was observed that AUS had a well-balanced sensitivity and specificity (58.1% vs. 81%) with a low PPV (34%), suiting a screening rather than diagnostic category (6). The sensitivity, specificity and PPV in SFN (63.9% vs. 89.5% and 57.5%) and SFM (59.4% vs. 97.3% and 82.6%) favored diagnostic categories with higher probabilities for malignancy. The findings are consistent with the view that AUS should remain separate from SFN and its presence, especially with the new tumor, NIFTP, makes patient management safer. Treating AUS as a screening category, including more benign and indolent neoplasms rather than malignant ones, appears to be the best approach; therefore, it has to be managed conservatively. This is very important for patient safety. From statistical perspectives, fewer categories in a diagnostic system would reduce the intra and inter-observer variability. However, the main objective of any schema is to improve patient care and as such, reducing diagnostic variability should not be a major incentive to minimize a working and clinically useful scheme. In a working histopathological continuum model, Morris [17] has shown that reducing the number of category leads to an increase in inter-observer agreement, as measured by kappa statistics, but also leads to a decrease in the information transmitted. Because the main role of a cytopathology’s is to communicate information obtained from visual patterns to clinicians, more categories, rather than fewer, might prove more clinically effective. Therefore, not only keeping AUS was good, but subcategorizing it into several morphological groups with a spectrum of ROM, ranging from low to high might reduce false negative and false positive results in favor of patient safety. Several sub-categorization schemes 2, 3, 4 and 5-tired have been proposed, based mainly on cytological and architectural atypia [18]. These schemes all shared the highest ROM in the cytological atypia group, rather than the architectural atypia group.

## The complementary role of calculating RON

The calculation of RON beside ROM can shed a light on the appropriateness of the management recommendation within each diagnostic category, and provides a good guide and checks for safe practice on an individual institutional basis. Our personal experience supported the current BSFRTC recommendation for management, with a significantly higher RON compared to ROM in AUS, favoring a conservative approach for management [6], in contrast to the comparable RON and ROM in SFN and SFM categories. Wu and Chen's [19,20] experiences were quite different. They reported higher RON in SFN and even SFM, almost double and triple the ROM values in these categories, suggesting a more conservative approach to management (Table 1).

Study	AUS	SFN	SFM
El Hag et al [6]	1.56	1.03	1.05
Chen et al [19]	3.55	2.95	1.95
Wu et al [20]	2	2.45	1.15
Mahajan et al [16]	1.32	2.78	1.33

**Note:** AUS: Atypia of Undetermined Significance; SFN: Suspicious for Follicular Neoplasm; SFM: Suspicious for Malignancy

**Table 1:** The ration RON/ROM in indeterminate BSFRTC diagnostic categories in different published studies.

This is a good example of the variation between institutions, which could be overlooked without the calculation of RON. Moreover, the ROM by itself might be reductive, as it excludes benign nodules like hyperplastic nodules, FA and indolent tumors like NIFTP. These lesions and tumors are of low risk; however, they might represent precursors or premalignant lesions. A morphological spectrum for NIFTP [6,21-23], ranging from minute tumoral foci in a goitrous background, to a pure solid tumor is noted. When this is taken together with the common RAS mutation between NIFTP and FVPC, the former might be a precursor for the invasive counterpart FVPC in a second pathway. The same is true for FA and FC. Partyka, et al [24] tested three molecular platforms (Afirma Thyroid FNA Analysis, RosettaGX Reveal, and Interpace ThyGenX/ThyraMIR) on thyroid FNA samples from indeterminate categories AUS, SFN and SFM, to rule in/rule out malignancy. The test was found to be more predictive for neoplasm rather than malignancy. They concluded that, the risk of neoplasm is a good index for surgery, and have endorsed conservative treatment with lobectomy for cases that are indeterminate by FNA but suspicious in molecular testing. Their results were more consistent with those of Wu and Chen et al [19,20]. The lessons learned here is that management recommendation of BSFRTC should not be taken rigidly, but rather be determined depending on the RON and ROM values on institutional basis. A RON/ROM ratio close to unity in indeterminate categories is a marker indicative for surgical triaging.

## NIFTP

The IV WHO classification of tumors of the endocrine organs in 2017 abolished the old diagnostic category of non-invasive encapsulated follicular variant of papillary thyroid carcinomas, and replaced it with the new terminology Non-Invasive Follicular Thyroid Neoplasms with Papillary-like Nuclear Features, NIFTP [25], which remained the same in update of 2022 [26]. The cyto-morphological

features of NIFTP are intermediate between those of FA and FVPC and the possibility of NIFTP can only be raised in cytology; a definitive diagnosis of NIFTP is confirmed in surgical specimens [27]. The tumor is indolent with low risk of adverse outcome, and adequately treated by simple lobectomy [21]. Therefore, this tumor is similar to FA and should be treated as such. Since the recommended treatment is surgical lobectomy, its diagnosis as AUS or SFN in the preoperative cytology is relatively harmless, because lobectomy is the recommendation for their management. The concern is to diagnose the NIFTP preoperatively as SFM or malignant. This will lead to surgical overtreatment with total thyroidectomy. Published studies with the distribution of NIFTP between different BSFRTC diagnostic categories were variable, with most NIFTP reported in AUS and SFN [7,8,23,28]. Nevertheless, an appreciable number of NIFTP were also reported in SFM and malignant categories that led to overtreatment. To avoid overcalling NIFTP in SFM and malignant categories, molecular tests on cytology material, which might have appeared to be a good practice, do not, in fact, seem very promising at least for the time being. In addition to the high cost and availability of the tests, none of the common molecular alterations associated with NIFTP (RAS mutation, BRAFK601E mutation PAX8-PARG rearrangement and THADA fusion) is pathognomic for the entity [22]. There is a great overlap in the molecular alteration between NIFTP and other follicular and non-follicular carcinomas, which make its triage difficult [21,22]. On the other hand, up to 21.1% of NIFTP at least in our little experience might be lost preoperatively in the benign category, here molecular testing could be of help and can be performed on the base of suspicious US findings.

Another issue with molecular testing is result interpretation by the clinician. Hang et al [29] found that more total thyroidectomies performed based on a suspicious Afirma GEC result in their study group. Clinicians might have equated a suspicious molecular test with malignancy, leading to overtreatment. In our opinion, the best solution for the time being is to improve the performance of the BSFRTC. Considering NIFTP benign like FA and omitting it from the calculation of ROM, introducing RON as a quality indicator, sub-categorization of AUS and abiding to diagnostic criteria could reduce the risk of reporting NIFTP in the SFM and the malignant categories. Our cytological analysis of NIFTP cases revealed clearly the similarities between it and benign follicular nodules and follicular neoplasms [6]. Of notice here is that nuclear grooves seen in all cases of NIFTP, however, small in number and never crossed the suggested significant cut off level of 20%, which characterize PC and it's variant [30].

## Conclusion

Exclusion of NIFTP from malignant lesions has led to the reduction in the calculated ROM and improved the stratification between the AUS and SFN categories. Availability, costs, specificity and interpretation issues are limiting factors for including molecular tests on a routine basis to triage for NIFTP. A simple solution is to consider NIFTP a benign tumor that should be managed similarly to FA. The ratio RON/ROM might be used to determine the appropriate approach to management, conservative vs surgical, within each indeterminate category, with a ratio close to unity indicative for the need to immediate surgical management. The reported high ROM in intermediate categories compared to the estimated values in BSFRTC is probably due to referral bias.

## References

1. Cibas ES, Ali SZ (2009) The Bethesda System for reporting thyroid cytopathology. *Am J Clin Pathol* 19:658-665.
2. Cibas ES, Ali SZ (2017) The 2017 Bethesda system for reporting thyroid cytopathology. *Thyroid* 27:1341-1346.
3. Redman R, Yoder BJ, Massol NA (2006) Perception of diagnostic terminology and cytopathologic reporting of fine-needle aspiration of the thyroid nodules. A survey of clinicians and pathologists. *Thyroid* 16:1003-1008.
4. Bongiovanni M, Spitale A, Faquin WC, Mazzuchelli L, Baloch ZW (2012) The Bethesda system for reporting thyroid cytopathology: A meta-analysis. *Acta Cytologica* 56:333-339.
5. Bongiovanni M, Crippa S, Baloch ZW, Piana S, Spitale A, et al. (2012) Comparison of 5-tiered and 6-tiered diagnostic system for reporting of thyroid cytopathology. *Cancer Cytopathol* 120:117-125.
6. El Hag IA, Johnston J, Alessa E, Shammari M (2021) Revised Bethesda system for reporting thyroid cytology: lessons learned from 5 years' experience in a central hospital. *Cytopathology* 32:482-492.
7. Bychkov A, Keelawat S, Agarawal S, Jain D, Jung CK, et al. (2018) Impact of non-invasive follicular thyroid neoplasm with papillary-like nuclear features on the Bethesda system for reporting thyroid cytopathology: a multi-institutional study in five Asian countries. *Pathology* 50: 411-417.
8. Faquin WC, Wong LQ, Afrogheh AH, Ali SZ, Bishop JA, et al. (2016) Impact of Reclassifying Noninvasive Follicular Variant of Papillary Thyroid Carcinoma on the Risk of Malignancy in the Bethesda System for Reporting Thyroid Cytopathology. *Cancer Cytopathol* 124:181-187.
9. Ho AS, Sarti EE, Jain KS, Wang H, Nixon IJ, et al. (2014) Malignancy rate in thyroid nodules classified as Bethesda category III (AUS/FLUS). *Thyroid* 24:832-839.
10. Bhasen TS, Mannan R, Manjari M, Mehra M, Sekhon AK, et al. (2013) Reproducibility of the 'Bethesda System for Reporting Thyroid Cytopathology: A Multicenter study with review of literature. *JCDR* 7: 1051-1054.
11. Krauss EA, Mahon M, Fede JM, Zhang L (2016) Application of Bethesda Classification for Thyroid Fine-Needle Aspiration: Institutional experience and meta-analysis. *Arch Path Lab Med.* 140:1211-1213.
12. Renshaw AA (2010) Should atypical follicular cells in thyroid fine-needle aspirates be sub classified. *Cancer Cytopathol* 118:186-189.
13. Singh RS, Wang HH (2011) Eliminating the "Atypia of Undetermined Significance/Follicular Lesion of Undetermined Significance" category from the Bethesda System for reporting Thyroid Cytopathology. *Am J Clin Pathol.* 139: 896-902.
14. Kocjan G, Cochand-Priollet B, de Agustin PP, Bourgain C, Chandra A, et al. (2010). Diagnostic terminology for reporting fine needle aspiration cytology: European of Cytology Societies Thyroid Working Party Symposium, Lisbon 2009. *Cytopathology* 21: 86-92.
15. Baloch ZW, Mandel SJ, LiVolsi VA (2013) Are we ready to modify the Bethesda Thyroid Fine-Needle Aspiration Classification Scheme?. *Cancer Cytopathol* 121:171-174.
16. Mahajan S, Srinivasan R, Rajwanshi A, Radotra B, Panda N, et al. (2017) Risk of Malignancy and Risk of Neoplasia in the Bethesda Indeterminate Categories: Study on 4,532 Thyroid Fine-Needle Aspirations from a Single Institution in India. *Acta Cytologica* 61:103-110.
17. Morris JA (1994) Information and observer disagreement in histopathology. *Histopathology* 25:123-128.
18. Furness PN, Taub N (2006) Interobserver reproducibility and application of the ISN/RPS classification of lupus nephritis-a UK-wide study. *Am J Surg pathol.* 30:1030-1035.
19. Chen HY, Partyka KL, Dougherty R, Cramer HM, Wu HH (2020) The importance of risk of neoplasm as an outcome in cytologic-histologic correlation studies on thyroid fine needle aspiration. *Diagn Cytopathol* 48:1237-1243.
20. Wu HH, Rose C, Elsheikh TM (2012) The Bethesda system for reporting thyroid cytopathology: An experience of 1,381 cases in a community practice setting with the implication for risk of neoplasia and risk of malignancy. *Diagnostic cytopathol.* 40:399-403.
21. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, et al. (2016) Nomenclature Revision for Encapsulated Follicular Variant of Papillary Thyroid Carcinoma: A Paradigm Shift to Reduce Overtreatment of Indolent Tumors. *JAMA Oncol.* 2: 1023-1029.
22. Pusztaszeri M, Bongiovanni M (2019) The impact of non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) on the diagnosis of thyroid nodules. *Gland Surg.* 8:86-97.
23. Canini V, Leni D, Pincelli AI, Scardilli M, Garancini M, et al. (2019) Clinico-pathological issues in thyroid pathology: study on the routine application of NIFTP diagnostic criteria. *Sci Rep* 9:1-8.
24. Partyka KL, Trevino K, Rnaldolph MI, Cramer H, Wu HH (2019) Risk of malignancy and neoplasia predicted by three molecular testing platform in indeterminate thyroid nodule on fine needle aspiration. *Diagn Cytopathol* 47:853-862.
25. Lloyd RV, Osamura RY, Klöppel G, Rosai J (2017) WHO Classification of Tumours of Endocrine Organs. IARC.
26. Baloch ZW, Asa SL, Barletta JA, Ghossein RA, Juhlin CC, et al. (2022) Overview of the WHO classification of thyroid neoplasm. *Endocrine Pathol* 33:27-63.
27. Brandler TC, Zou F, Liu CZ, Cho M, Lau RP, et al. (2017) Can non-invasive follicular neoplasm with papillary-like nuclear feature be distinguish from classic papillary thyroid carcinoma by fine needle aspiration cytology. *Cancer cytopathol* 125:378-388.
28. Strickland KC, Howitt BE, Marqusee E, Alexander EK, Cibas ES, et al. (2015) The Impact of Noninvasive Follicular Variant of Papillary Thyroid Carcinoma on Rates of Malignancy for Fine-Needle Aspiration Diagnostic Categories. *Thyroid* 25:987-992.
29. Hang JF, Westra WH, Cooper DS, Ali SZ (2017) The impact of non-invasive follicular thyroid neoplasm with papillary-like nuclear features on the performance of the afirma gene expression classifier. *Cancer Cytopathol* 125:683-691.
30. Bhat AS, Varma L, Fernandes H, Jayaprakash CS (2019) Role of nuclear groove in the cytological diagnosis of papillary carcinoma thyroid. *Thyroid Res. Pract* 16:105-107.