



Interobserver Agreement in Interpreting Thyroid FNA with a Diagnosis of Atypia/ Follicular Lesion of Undetermined Significance (FLUS/AUS)



Adebowale J. Adeniran¹, Constanino Theoharis¹, Mitchell Wachtel⁷, Zubair Baloch³, Marilyn Bui⁶, Isam Eltoum⁵, Oscar Lin⁴, Aylin Simsir⁸, Ed Stelow², Guoping Cai¹, Malini Harigopal¹, Angelique Levi¹, Berrin Ustun¹, Meiklejohn McKenzie⁵, Diane Kowalski¹, David Chhieng¹

Departments of Pathology at: ¹Yale School of Medicine, New Haven, CT; ²University of Virginia, Charlottesville, VA; ³University of Pennsylvania, Philadelphia, PA; ⁴Memorial Sloan-Kettering Cancer Center, New York, NY; ⁵University of Alabama, Birmingham, AL; ⁶Moffit Cancer Center, Miami, FL; ⁷Texas Tech University, Lubbock, TX; and ⁸New York University, New York, NY

ABSTRACT

Background: The Bethesda System for Reporting Thyroid Cytopathology (BSRTC) was developed to refine definitions and improve clinical communication and management. However, the diagnostic category of atypia/follicular lesion of undetermined significance (FLUS/AUS) remains heterogeneous in terms of usage and clinical outcome. Because of the “gray zone” that exists in the interpretation of thyroid FNA demonstrating minor architectural and/or cytologic atypia, this study was undertaken to evaluate the degree of interobserver agreement in the evaluation of thyroid FNAs originally interpreted as FLUS/AUS.

Design: Twenty-three thyroid FNAs including 18 cases originally diagnosed as FLUS/AUS, 2 as negative for malignant cells, 2 as positive for malignancy, and 1 as follicular neoplasm were selected. Two representatives slides from each case were circulated to 13 board certified or eligible cytopathologists; all were from academic institutions with 6 were from the same institution. Each reviewer was asked to evaluate each cases using the BSRTC. The kappa statistics was calculated.

Results: Only 2 cases (22%) were in complete agreement: one originally interpreted as negative and one as positive for malignancy; both cases were confirmed on histology. There was a majority agreement (among 10 or more) in 6 cases: one originally diagnosed as positive, one negative, and 4 FLUS/AUS. Both positive and negative cases were confirmed on histology; among the 4 FLUS/AUS, 2 were found to be negative and 2 follicular adenoma on histology. For the entire group of reviewers, the mean kappa statistic was 0.34±0.13. The mean kappa statistic was 0.42±0.7 and 0.29±0.14 among reviewers of the same institution and among reviewers from different institutions. The difference was statistically significant.

Conclusions: The interobserver agreement for thyroid FNA cases originally classified as FLUS/AUS was fair among academic cytopathologists. It appeared that the interobserver agreement was better among cytopathologists who were from the same institution.

BACKGROUND

Interpretation of thyroid FNA is challenging because there is comparatively little difference in the morphologic features of the many non-neoplastic and neoplastic conditions of the thyroid and there is variability in FNA specimen preparation and Interpretation. The Bethesda System for Reporting Thyroid Cytopathology (BSRTC) was developed to refine definitions and improve clinical communication and management. However, the diagnostic category of atypia/follicular lesion of undetermined significance (FLUS/AUS) remains heterogeneous in terms of usage and clinical outcome. Because of the “gray zone” that exists in the interpretation of thyroid FNA demonstrating minor architectural and/or cytologic atypia, this study was undertaken to evaluate the degree of interobserver agreement in the evaluation of thyroid FNAs originally interpreted as FLUS/AUS. In the study we utilized kappa statistics to compare the interobserver diagnostic agreement obtained with evaluation of the thyroid FNAs.

DESIGN

Twenty-three thyroid FNAs including 18 cases originally diagnosed as FLUS/AUS, 2 as negative for malignancy, 2 as positive for malignancy, and 1 as follicular neoplasm were selected. Two representative slides from each case were reviewed independently by 13 cytopathologists with varying years of experience; all were from academic institutions with 6 of them from the same institution. Interobserver diagnostic agreement was evaluated by comparing the percent diagnostic agreement between the diagnoses rendered for each case, using the BSRTC. The diagnostic agreement data were analyzed across pathologists each reviewing all 23 cases with Fleiss’ generalized kappa statistics. Kappa values were interpreted using the criteria proposed by Fleiss where 0-0.4 = poor agreement; 0.4-0.75 = fair agreement; 0.75-1.0 = strong agreement.

RESULTS

The observer diagnoses using the BSRTC are shown in **Table 1**. Only 2 cases (22%) were in complete agreement: one originally interpreted as negative and one as positive for malignancy; both cases were confirmed on histology. **Table 2** shows the comparison of rates of FLUS/AUS diagnoses made by observers based on 18 cases with original diagnoses of FLUS/AUS. There was no 100% agreement of FLUS/AUS for any of the cases. There was a majority agreement (among 8 or more observers) in 6 cases: one originally diagnosed as positive, one negative, and 4 FLUS/AUS. Both positive and negative cases were confirmed on histology; among the 4 FLUS/AUS, 2 were found to be negative, 1 follicular adenoma and one classic variant of papillary thyroid carcinoma on histology. **Table 3** shows correlation of FLUS/AUS cases with final histologic diagnoses, while **Table 4** shows the final histologic diagnoses in the 4 FLUS/AUS cases with majority agreement among observers. For the entire group of reviewers, the mean kappa statistic was 0.34±0.13. The mean kappa statistic was 0.42±0.7 and 0.29±0.14 among reviewers of the same institution and among reviewers from different institutions, respectively. The difference was statistically significant.

RESULTS

Table 1. Observer diagnoses using the BSRTC

Case #1	Path1	Path 2	Path 3	Path 4	Path 5	Path 6	Path 7	Path 8	Path 9	Path 10	Path 11	Path 12	Path 13
1	NEG	NEG	NEG	NEG	NEG	NEG	NEG	NEG	NEG	NEG	NEG	NEG	NEG
2	POS	POS	POS	POS	POS	POS	POS	POS	POS	POS	POS	POS	POS
3	NEG	NEG	NEG	FLUS	SUS	NEG	NEG	NEG	FLUS	FLUS	NEG	FLUS	FLUS
4	SUS	SUS	FLUS	FLUS	NEG	POS	SUS	SUS	POS	SUS	SUS	SUS	POS
5	FLUS	FN	FN	FN	NEG	NEG	FN	FN	FN	FN	FN	FLUS	FN
6	FLUS	FN	FN	FN	FLUS	POS	NEG	FLUS	FN	FLUS	FLUS	NEG	FLUS
7	FLUS	POS	NEG	FN	SUS	POS	POS	POS	SUS	SUS	SUS	POS	POS
8	FLUS	FLUS	FLUS	FLUS	NEG	NEG	NEG	NEG	FN	FLUS	FN	FN	NEG
9	POS	SUS	POS	SUS	POS	POS	POS	POS	POS	POS	POS	POS	POS
10	FLUS	FLUS	NEG	NEG	NEG	FLUS	FLUS	FLUS	FLUS	FLUS	NEG	FLUS	SUS
11	POS	SUS	POS	SUS	SUS	SUS	SUS	POS	SUS	SUS	SUS	SUS	SUS
12	FLUS	FLUS	NEG	FN	NEG	NEG	NEG	NEG	NEG	FLUS	FLUS	NEG	FN
13	FN	FN	FN	FN	FN	SUS	FN	FLUS	FN	FN	FN	FN	FN
14	NEG	NEG	FLUS	FLUS	NEG	NEG	NEG	NEG	NEG	NEG	NEG	NEG	NEG
15	NEG	NEG	NEG	NEG	NEG	NEG	NEG	FN	NEG	NEG	FN	NEG	NEG
16	FLUS	FLUS	FLUS	NEG	FN	FLUS	FLUS	FLUS	FLUS	FLUS	FLUS	FN	FLUS
17	NEG	FN	NEG	FN	FN	NEG	FN	FLUS	FLUS	FN	NEG	NEG	FN
18	FLUS	NEG	NEG	NEG	FLUS	FLUS	NEG	FLUS	SUS	FLUS	FLUS	FLUS	FLUS
19	FLUS	FLUS	FLUS	NEG	FN	NEG	FN	FLUS	FN	NEG	NEG	FN	FN
20	FLUS	NEG	FLUS	FLUS	FLUS	FLUS	NEG	FLUS	FLUS	NEG	FLUS	NEG	FLUS
21	FLUS	NEG	NEG	FLUS	FLUS	FLUS	NEG	FLUS	SUS	NEG	FLUS	NEG	FLUS
22	NEG	NEG	NEG	NEG	FN	NEG	FLUS	NEG	NEG	NEG	NEG	NEG	NEG
23	FLUS	FN	SUS	FN	NEG	FLUS	NEG	FN	NEG	FN	SUS	NEG	NEG

Path, pathologist; NEG, negative for malignancy; FLUS, atypia/follicular lesion of undetermined significance; FN, follicular neoplasm; SUS, suspicious for papillary thyroid carcinoma; POS, positive for papillary thyroid carcinoma.

Table 2. Comparison of rates of FLUS/AUS diagnoses made by observers based on18 cases with original diagnoses of FLUS/AUS.

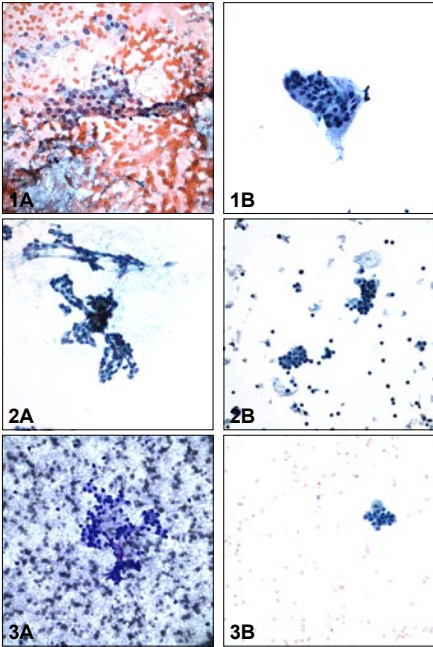
	Path 1	Path 2	Path 3	Path 4	Path 5	Path 6	Path 7	Path 8	Path 9	Path 10	Path 11	Path 12	Path 13
FLUS/AUS Rate (N=18)	10 (55.6%)	4 (22.2%)	5 (27.8%)	6 (33.3%)	4 (22.2%)	5 (27.8%)	3 (16.7%)	8 (44.4%)	5 (27.8%)	7 (38.9%)	6 (33.3%)	3 (16.7%)	5 (27.8%)
Years of Experience	14	8	16	2	7	15	15	10	7	4	5	6	1
YALE	NO	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES	YES

Table 3. Correlation of FLUS/AUS cases with final histologic diagnoses

Final histologic diagnosis	# of cases
PTC, classic variant	3
PTC, follicular variant	1
PTC, cystic	1
Medullary CA	1
Follicular adenoma, NOS	4
Hurthle cell adenoma	2
Hashimoto Thyroiditis	2
Goiter	4

Table 4. Final histologic diagnoses in the FLUS/AUS cases with majority agreement among observers

Case	Final histologic diagnosis
Case 10	Goiter
Case 16	PTC, classic variant
Case 18	Follicular adenoma
Case 20	Goiter



Case 10. Final histologic diagnosis: Goiter

Case 17. Final histologic diagnosis: Hurthle cell adenoma

Case 19. Final histologic diagnosis: Follicular adenoma

CONCLUSIONS

- The interobserver agreement for thyroid FNA cases originally classified as FLUS/AUS was poor among academic cytopathologists.
- It appeared that the interobserver agreement was better among cytopathologists who were from the same institution.
- The rate of diagnosis of FLUS/AUS was independent of the number of years of experience of the observers.
- The FLUS/AUS is a heterogeneous category of cases subject to diverse interpretation because of a lack of well-defined diagnostic criteria.
- Our study showed that this group is fraught with high interobserver variability.

REFERENCES

- Horne MJ, Chhieng DC, Theoharis C, et al. Thyroid follicular lesion of undetermined significance: Evaluation of the risk of malignancy using the two-tier sub-classification. Diagn Cytopathol. 2012;40:410-5.
- Tetikurt US, Oz Puyan F, Oz F, et al. Diagnostic value of liquid-based (Liqui-PREP) preparations and interobserver reproducibility in fine needle aspiration cytology of the nodular thyroid lesions Diagn Cytopathol. 2012;40:388-93.
- Duggal R, Rajwanshi A, Gupta N, Vasishta RK. Interobserver variability amongst cytopathologists and histopathologists in the diagnosis of neoplastic follicular patterned lesions of thyroid. Diagn Cytopathol. 2011;39:235-41.