The Bethesda System for Reporting Thyroid Cytopathology: A Four-Year Single Academic Institution Experience

David C Chhieng[1], Constantine Theoharis[1], Diane Kowalski[1], Tobias Carling[2], Robert Udelsman[2], Adebowale J. Adeniran[1]

Departments of [1]Pathology and [2]Surgery, Yale School of Medicine, New Haven, CT, USA

ABSTRACT

Background: The Bethesda System for Reporting Thyroid Cytopathology (BSRTC) is a standardized reporting system for classifying thyroid fine-needle aspiration results comprising of 6 diagnostic categories with unique risks of malignancy and recommendations for clinical management. Our objective was to report our past 4-year experience with the BSRTC, review its distribution and evaluate its efficiency based on the cytologic-histologic correlation.

Design: A total of 12,930 thyroid nodules undergoing FNA were examined at our institution between January 2008 and December 2011. All FNAs were classified prospectively into unsatisfactory, benign, FLUS/AUS (follicular lesions or atypia of undetermined significance), follicular neoplasm (FN), suspicious for malignancy, or positive for malignancy. The Cyto-histologic correlation was recorded.

Results: Table 1 summarizes the diagnostic frequencies. The false positive rate for a malignant and suspicious diagnosis was 0.4% and 9.9%, respectively. The specificity of diagnosing malignant thyroid nodules was 97.7% whereas the specificity as a screening test for all neoplasms was 89.6%.

Conclusions: The BSRTC shows excellent specificity in diagnosing malignant nodules and in screening for neoplasms. Each diagnostic category conveys specific risks of malignancy, which offers guidance for clinical management. In addition, the frequency distribution of the individual diagnostic categories remained relatively stable over time.

BACKGROUND

The Bethesda System for Reporting Thyroid Cytopathology (BSRTC) is a standardized reporting system for classifying thyroid fine-needle aspiration results comprising of 6 diagnostic categories with unique risks of malignancy and recommendations for clinical management. In 2009, the American Thyroid Association (ATA) incorporated the BSRTC terminology into the 2009 revised guidelines for management of patients with thyroid nodules. Initial studies by our group as well as others have shown that each diagnostic category of the BSRTC conveys a different level of risk of malignancy to the caregivers. However, others have not observed any significant impact of the BSRTC on the diagnostic accuracy and the false positive rates of thyroid FNA or the frequency of intraoperative consultation. The objective of this study was to report our past 4-year experience in using the BSRTC to review the distribution of diagnostic categories and to evaluate the efficiency of the system based on the cytologic-histologic correlation.

DESIGN

The study population included all consecutive thyroid FNAs that were evaluated at the Yale Pathology Laboratory from January 2008 to December 2011. The thyroid FNAs were performed by either the endocrinologists or radiologists under ultrasound guidance. On-site assessment of adequacy was performed in less than 5% of the cases. The study also included specimens submitted by outside laboratories for second opinion. At our institution, in service training was provided to all cytopathologists and cytotechnologists to explain the terms and implications of each diagnostic category before implementation of the BSRTC. To minimize biases in the application of the BSRTC, all cases that were initially interpreted as AUS/FLUS, follicular neoplasm, suspicious for malignancy, and positive for malignancy as well as a substantial number of negative and non-diagnostic cases were reviewed by 2 or more cytopathologists. Follow up cytology and histology were obtained from our laboratory information system. Histologic diagnoses were used as the gold standard for the correlation with cytologic interpretations. Categorical analysis was performed using both a log-linear model (likelihood ratio) and Chi-square model with symmetric measure of association. Statistical significance was set at p-value of 0.5 or less.

RESULTS

- A total of 12,930 thyroid FNAs from 9,901 patients were evaluated during the 4-year study period. The average age was 54.3 ± 15.0 years, and a female to male ratio of 4:1. There were no significant differences among the age and gender distribution in the calendar years during the study period.
- The overall distribution of the diagnostic categories is summarized in **Table 1**. **Table 2** highlights the percentage of cases that were found to have a malignant follow up on histology according to various diagnostic categories. **Table 3** summarizes the overall 4-year cytology-histology patient correlation.
- The false positive rate for a malignant diagnosis was 0.2%. The false negative rate was 16.1%. The specificity of diagnosing malignant thyroid nodules was 98.8%, whereas the sensitivity as a screening test for all neoplasm was 89.6%. The positive predictive value was 99.8%.
- The false negative rate of 16.1% represented 4.5% of all FNAs (99/2196) from 89 patients over the 4-year period.
  - 9 (0.4%) cases were due to interpretation errors.
    - 4 FVPTC, 2 PTC NOS, 2 follicular CA and 1 lymphoma
  - 90 (4.1%) cases were due to sampling errors.
    - Size of nodules: 0.8 ± 0.9 cm
    - 29 FV PTC, 53 PTC NOS, 4 PTC (other variants), 2 Hürthle cell CA, 2 medullary CA
    - 1/4 had a positive/suspicious diagnosis in the contra-lateral lobe.

RESULTS

Table 1. Overall Distribution of the Diagnostic Categories

Cytologic Category	2008	2009	2010	2011	AVERAGE 2008-2011
Unsatisfactory	11.2%	13.0%	8.7%	7.1%	10.0%
Benign/Negative for Malignancy	73.5%	73.9%	76.8%	76.7%	75.2%
FLUS/AUS	4.3%	4.4%	7.1%	5.8%	5.4%
Follicular /Hürthle Cell Neoplasm	5.3%	5.6%	4.2%	3.1%	4.6%
Suspicious for Malignancy	1.3%	1.3%	2.0%	1.9%	1.6%
Malignancy	5.4%	6.4%	5.0%	5.4%	5.6%
Total (n)	3,208	3,071	3,043	3,607	12,930

Table 2. Percentage of Cases that were Found to have a Malignant Follow Up on Histology According to Various Diagnostic Categories

Cytologic Category	2008	2009	2010	2011	Overall
Unsatisfactory	21.1%	25.5%	34.1%	16.7%	23.8%
Benign/Negative for Malignancy	15.2%	19.6%	14.8%	22.1%	16.1%
FLUS/AUS	54.3%	46.4%	64.4%	60.5%	56.8%
Follicular /Hürthle Cell Neoplasm	30.8%	29.8%	38.4%	33.8%	32.7%
Suspicious for Malignancy	89.2%	86.7%	94.0%	91.8%	91.0%
Malignancy	100.0%	98.8%	100.0%	100.0%	99.8%
Total	54.4%	49.9%	55.0%	55.6%	52.2%

Table 3. Cytologic-Histologic Correlation

	Unsatisfactory	Negative	FLUS/ AUS	FN	Suspicious	Positive	Total
Malignant	23.8%	16.1%	56.8%	32.7%	91.0%	99.8%	52.2%
Neoplastic	10.3%	6.6%	15.0%	41.3%	2.4%	0.0%	12.1%
Non-neoplastic	65.9%	77.3%	28.2%	26.0%	6.6%	0.2%	35.7%
Total n (%)	164 (13.4%)	621 (6.5%)	287 (43.2%)	400 (43.2%)	166 (77.6%)	561 (79.2%)	2,199 (17.0%)

Table 4. Statistical Analysis of the Cytologic Classification System

Comparisons Between Different Cytologic Categories	$X^{2\,a}$	Phi [b]	LR[c]	p value (2-sided)
All categories	999	0.891	244:1	<0.001
All categories excluding unsatisfactory category	999	0.908	243:1	<0.001
B vs I	155	0.42	7:1	<0.001
B vs FN	38	0.19	3:1	<0.001
B vs S	335	0.66	52:1	<0.001
B vs M	833	0.84	2929:!	<0.001
A vs FN	39	0.24	1:1	<0.001
A vs S	57	0.36	8:1	<0.001
A vs M	277	0.57	427:1	<0.001
FN vs S	158	0.53	21:1	<0.001
FN vs M	517	0.74	1152:1	<0.001
S vs M	43	0.25	55:1	<0.001

B: Benign/Negative for Malignancy, A: AUS/FLUS, FN: Follicular Neoplasm, S: Suspicious for Malignancy, M: Positive for Malignancy, NS: not statistically significant.
a: X^2 model for association between categorical data.
b: Phi correlation coefficient for measuring the strength of association between categorical data; values range from +1 to -1; +1denotes strong positive association, -1 strong negative association, and 0 no association.
c: Likelihood ratio (log linear model) for association between categorical data.

CONCLUSIONS

The BSRTC shows excellent specificity in diagnosing malignant nodules and in screening for neoplasms. Each diagnostic category conveys specific risks of malignancy, which offers guidance for clinical management. In addition, the frequency distribution of the individual diagnostic categories remained relatively stable over time.

REFERENCES

1. Theoharis CG, Schofield KM, Hammers L, Udelsman R, Chhieng DC 2009 The Bethesda thyroid fine-needle aspiration classification system: year 1 at an academic institution. Thyroid. **19**:1215-1223.
2. Baloch ZW, Cibas ES, Clark DP, Layfield LJ, Ljung BM, Pitman MB, Abati A 2008 The National Cancer Institute Thyroid fine needle aspiration state of the science conference: a summation. Cytojournal. **5**:6.
3. Crowe A, Linder A, Hameed O, Salih C, Roberson J, Gidley J, Eltoum IA 2011 The impact of implementation of the Bethesda System for Reporting Thyroid Cytopathology on the quality of reporting, "risk" of malignancy, surgical rate, and rate of frozen sections requested for thyroid lesions. Cancer Cytopathol. **119**:315-321.