

Thyroid Fine-Needle Aspiration Cytology

Performance Data of Neoplastic and Malignant Cases as Identified From 1558 Responses in the ASCP Non-GYN Assessment Program: Thyroid Fine-Needle Performance Data

Stan G. Eilers, MD, FASCP¹; Paula LaPolice, CT(ASCP)²; Perkins Mukunyadzi, MD, FASCP³; Umesh Kapur, MD, FASCP⁴; Amy Wendel Spiczka, MS, SCT, MP, HTL(ASCP)CM⁵; Ajay Shah, MD, FASCP⁶; Husain Saleh, MD, MBA, FASCP⁷; Adebowale Adeniran, MD, FASCP⁸; Amberly Nunez, MD⁹; Indra Balachandran, PhD, SCT(ASCP), CFIAC,¹⁰; Jennifer J. Clark, SCT(ASCP)CM¹¹; and Larry Lemon, CT(ASCP)¹¹

BACKGROUND: Fine-needle aspiration of the thyroid is a common procedure, with an established role in reducing unnecessary thyroid surgery and identifying neoplasms and malignancies. **METHODS:** The study evaluated 1558 responses in the American Society for Clinical Pathology (ASCP) Non-GYN Assessment program of aspirates of thyroid neoplasms and malignancies and placed them into the following groups: group A (target or correct interpretation), group B (incorrect interpretation as a benign thyroid nodule), group C (incorrect interpretation malignant aspirate as thyroid neoplasm), and group D (malignant diagnosis with incorrect interpretation). In clinical practice, responses in groups A, C, and D would lead to surgical excision, whereas responses in group B would not. **RESULTS:** Of a total of 1558 responses, 78.5% of the responses were in group A, 8.5% in group B, 3.75% in group C, and 9.25% in group D. By individual diagnosis, the group rates were 86.5%, 0%, 11%, and 2.5% for anaplastic thyroid carcinoma; 83%, 5.5%, 4.25%, and 7.25% for papillary thyroid carcinoma; 79%, 7%, 6%, and 8% for medullary thyroid carcinoma; 83.5%, 6.75%, 0%, and 9.75% for Hürthle cell neoplasm; and 61%, 22%, 0%, and 17% for follicular neoplasm in groups A, B, C, and D respectively. **CONCLUSIONS:** Fine-needle aspiration was effective in diagnosing thyroid neoplasms and malignancies and in separating thyroid nodules into surgical and nonsurgical categories. Data from a large group of cytology professionals showed good performance; however, there is room for improvement, especially in making specific diagnoses. In particular, follicular neoplasm and follicular variant of papillary thyroid carcinoma were challenging diagnoses for participants. *Cancer (Cancer Cytopathol)* 2014;122:745-50. © 2014 The Authors. *Cancer Cytopathology* published by Wiley Periodicals, Inc. on behalf of *American Cancer Society*. This is an open access article under the terms of the Creative Commons Attribution NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

KEY WORDS: follicular thyroid neoplasm; Hurthle cell thyroid neoplasia; thyroid cancer; follicular; thyroid cancer, Hurthle cell; thyroid carcinoma, anaplastic; thyroid carcinoma, follicular; thyroid carcinoma, medullary; thyroid carcinoma, papillary.

*Corresponding author: Stan Eilers, MD, Mercy Medical Center, 701 10th Street, Cedar Rapids, Iowa 52403; Fax: (319) 369-4631; staneilers@aol.com

¹Mercy Medical Center, Cedar Rapids, Iowa; ²Baystate Medical Center, Springfield, Massachusetts; ³Arkansas Pathology Associates, Little Rock Arkansas; ⁴Loyola University Medical Center, Maywood, Illinois; ⁵Mayo Clinic, Scottsdale, Arizona; ⁶FNA Clinic, Toledo, Ohio; ⁷Detroit Medical Center, Detroit, Michigan; ⁸Yale University School of Medicine, New Haven, Connecticut; ⁹University of Alabama at Birmingham Health System, Birmingham, Alabama; ¹⁰Albany College of Pharmacy and Health Sciences, Albany, New York; ¹¹American Society for Clinical Pathology, Indianapolis, Indiana.

The American Society for Clinical Pathology developed the Non-GYN Assessment program as part of its Cytopathology educational program offerings.

Received: March 31, 2014; **Revised:** April 22, 2014; **Accepted:** April 24, 2014

Published online June 9, 2014 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cncy.21440, wileyonlinelibrary.com

TABLE 1. Categorization of Incorrect Responses and Potential Significance

Group	Definition	Significance
A	Correct (target) interpretation	N/A
B	Incorrect interpretation as benign thyroid nodule	• Potential false negative for neoplasms • False negative for malignancies
C	Incorrect interpretation as thyroid neoplasm	• N/A for neoplasms • Probable second surgical procedure for malignancies
D	Malignant diagnosis with incorrect interpretation	• Unnecessary thyroidectomy for benign neoplasms • Minimal for most malignancies
E	Most common incorrect interpretation	N/A

INTRODUCTION

Fine-needle aspiration (FNA) of the thyroid is a commonly performed and widely accepted procedure in evaluating thyroid nodules for potential surgical excision. There are, however, inherent limitations in interpreting thyroid aspirates because of overlap in cytologic patterns and features between nonneoplastic and neoplastic nodules, as well as between a variety of neoplasms and the multiple disease processes that may coexist with them.¹ There are numerous published studies on the effectiveness of FNA of thyroid lesions. In a review article, Gharib and Goellner² found sensitivity varying from 65% to 98% (mean, 83%) and specificity from 72% to 100% (mean, 92%). The false-negative rate varied from 1% to 11.5% (mean, 5.2%), and the false-positive rate varied from 0% to 7.7% (mean, 2.9%). Nguyen³ reported that the accuracy of thyroid FNA approached 95%.

Most of these studies involved a large series of individual cases interpreted by a limited number of observers. The purpose of the current study was to evaluate the performance of numerous observers in interpreting cytologic preparations from a limited number of aspirates of thyroid neoplasms and malignancies in the American Society for Clinical Pathology (ASCP) Non-GYN Assessment program. The data from this program was evaluated to determine the effectiveness of FNA in diagnosing thyroid neoplasms and malignancies from a large, diverse group of cytology professionals.

MATERIALS AND METHODS

The American Society for Clinical Pathology Non-GYN Assessment program is a glass-slide-based program developed with oversight from the ASCP Non-GYN Assessment Committee. Each annual program is composed of 10 FNAs including thyroid aspirates, and 10 Non-GYN specimens including fluids and brushings for a total of 20 patient cases divided into 4 quarterly shipments of 5 cases. Laboratory and individual peer comparison statistics for

each case as well as a corresponding educational case study with high-resolution digital photographs and pertinent discussion are provided after the event. Cumulative data from the Non-GYN Assessment program were reviewed to evaluate the performance of participants in the diagnosis and classification of thyroid neoplasms and malignancies in FNAs.

On review of the glass slides, participants respond by choosing a diagnostic interpretation chosen from a list of 5 choices representing the typical differential diagnoses based on the unique clinical history and cytologic features of each case.

Although selecting clinical treatments is not part of the Non-GYN program, in clinical practice, the diagnostic interpretation largely determines the appropriate clinical response. For follicular neoplasm and Hürthle cell neoplasm, the standard clinical response is lobectomy with isthmusectomy.^{4,5} Depending on the subsequent surgical pathology diagnosis, additional surgical and therapeutic measures may be required. However, in 75% to 80% of cases, no further therapy is indicated.⁶ For the thyroid malignancies, anaplastic thyroid carcinoma, papillary thyroid carcinoma, and medullary thyroid carcinoma, a definitive surgical procedure, most likely a total thyroidectomy, is the standard clinical response. In general, nonsurgical follow-up is the standard clinical response for all other interpretations (benign thyroid nodules), thus reducing unnecessary surgery for patients. These relationships are summarized in Table 1.⁷ The unique clinical features of each patient may alter the standard clinical response.

For the study, 5 groups of cases representing thyroid malignancies and thyroid neoplasms, with a total of 1558 responses, were selected for statistical review and for microscopic review when appropriate. The 5 target diagnoses were: anaplastic thyroid carcinoma (ATC; n = 163), papillary thyroid carcinoma (PTC; n = 599), and medullary thyroid carcinoma (MTC; n = 225), categorized in the program as positive for malignancy; and

TABLE 2. Thyroid FNA Neoplastic Case Performance From the ASCP Non-GYN Assessment Program

Target Diagnosis	A	B	C	D	E	Responses
Anaplastic carcinoma	141 (86.5%)	0 (0%)	18 (11%)	4 (2.5%)	Hürthle cell neoplasm	163
Hürthle cell neoplasm	223 (83.5%)	18 (6.75%)	0 (0%)	26 (9.75%)	Papillary carcinoma	267
Papillary carcinoma	497 (83%)	32 (5.5%)	25 (4.25%)	45 (7.25%)	Medullary carcinoma	599
Medullary carcinoma	177 (79%)	16 (7%)	14 (6%)	18 (8.00%)	Follicular neoplasm	225
Follicular neoplasm	185 (61%)	67 (22%)	0 (0%)	52 (17%)	Papillary carcinoma	304
Totals	1223 (78.50%)	133 (8.5%)	57 (3.75%)	145(9.25%)		1558

Hürthle cell neoplasm (HCN; n = 267) and follicular neoplasm (FN; n = 304), categorized as lesions of uncertain biologic potential (indeterminate aspirates).

The responses were evaluated and placed in 1 of 5 defined groups. Group A represented the correct or target interpretation. Group B was a compilation of all incorrect interpretations that were classified as benign thyroid nodules. In clinical practice, group B responses may represent potentially significant errors. Because the standard clinical treatment of patients in these diagnostic categories is non-surgical, misidentification could potentially cause a delay in diagnosis and treatment. Group C represented malignant lesions incorrectly interpreted as a thyroid neoplasm, FN, or HCN. Because this diagnostic category would typically generate a subsequent diagnostic surgical procedure, these errors are relatively minor, although for most patients, a second surgical procedure may be necessary. Group D represented malignant diagnoses with incorrect interpretations. In the cases of the thyroid malignancies (ATC, PTC, and MTC), this may be of little clinical significance, although in the cases of ATC and MTC, more extensive surgery may be recommended. In the cases of HCN and FN, misinterpretations in this category could potentially result in a definitive surgical procedure that may not have been indicated because 70% to 85% of these cases will be either nonneoplastic or benign on surgical resection. Group E simply represented the most common misinterpretation for each case. These data are summarized in Table 2.

RESULTS

Anaplastic Thyroid Carcinoma

Of 163 responses for cases with a target interpretation of ATC, 86.5% gave the intended response, group A. Of the incorrect responses, none were placed in group B, 11% were placed in group C, and 2.5% were placed in group

D. Hürthle cell neoplasm was the most common misinterpretation. Some of the cases had a plasmacytoid morphology, which may help to explain the misinterpretations. Anaplastic thyroid carcinoma was not a common misinterpretation for any other entity.

Papillary Thyroid Carcinoma

Of 599 responses for cases with a target interpretation of PTC, 83% of the respondents had the intended response, group A. Of the incorrect responses, 5.5% were placed in group B, 4.25% were placed in group C, and 7.25% were placed in group D. The most common misinterpretation was MTC. In addition, PTC was the most common misinterpretation for HCN and FN. It is worth noting that slides from follicular variants of PTC significantly underperformed the PTC category as a whole, at 68%, 28%, 6%, and 0%. Although the distribution of responses may have been an artifact of the program's design, nearly a third of the participants did not answer with the correct response, whereas PTC associated with Hashimoto's thyroiditis performed better, with 99% of the respondents submitting the correct interpretation.

Medullary Thyroid Carcinoma

Of 225 responses for cases with a target interpretation of MTC, 79% of the respondents had the intended response, group A. Of the incorrect responses, 7% were placed in group B, 6% were placed in group C, and 8% were placed in group D. Follicular neoplasm was the most common misinterpretation. In addition, MTC was one of the most common misinterpretations for FN.

Hürthle Cell Neoplasm

Of 267 responses for cases with a target interpretation of HCN, 83.5% had the intended response, group A. Of incorrect responses, 6.75% were placed in group B, group C was not applicable for HCN, and 9.75% were placed in group D. Papillary thyroid carcinoma was the most

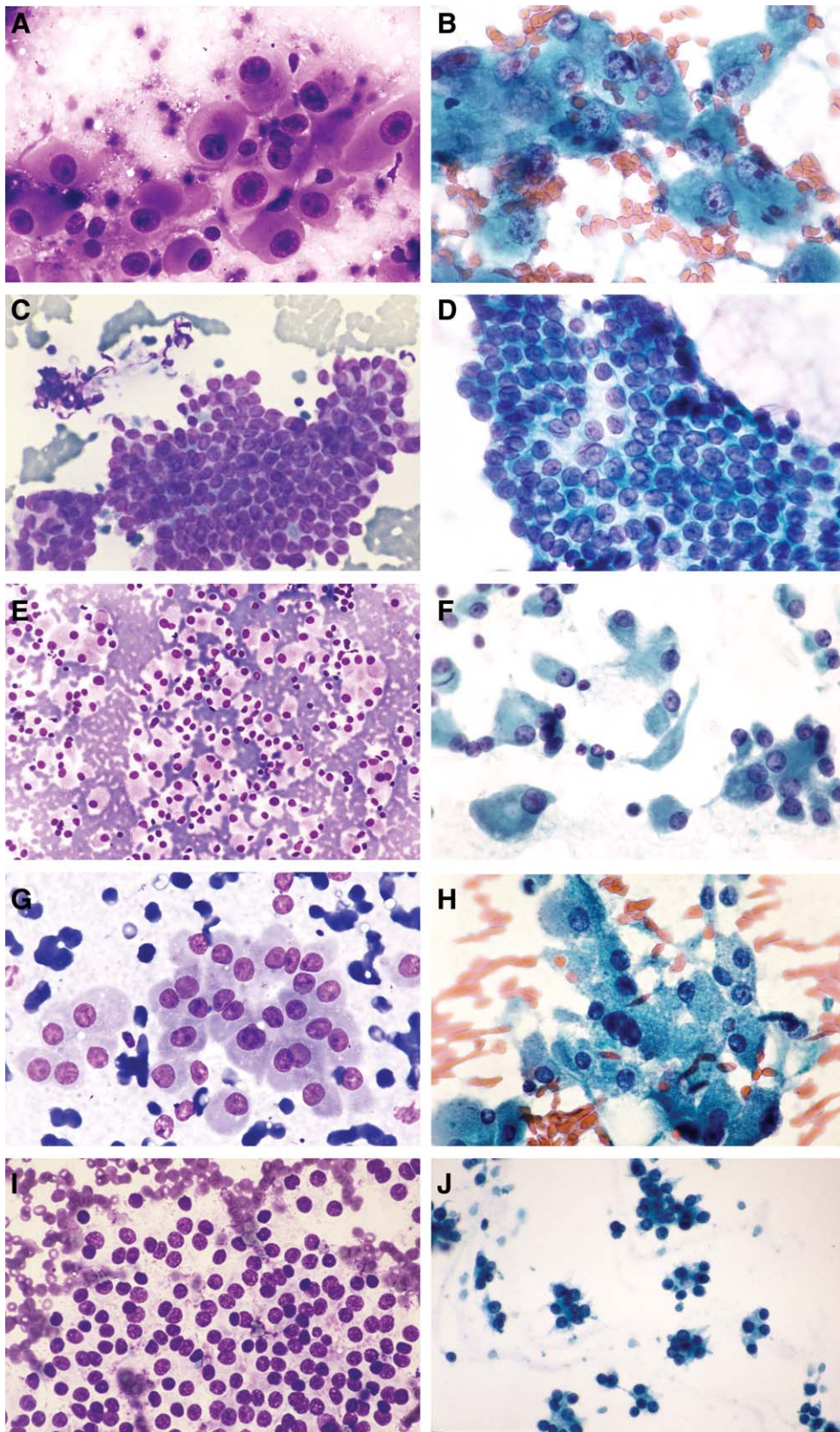


Figure 1.

common misinterpretation. In addition, HCN was the most common misinterpretation for ATC.

Follicular Neoplasm

Of 304 responses for cases with a target interpretation of FN, 61% had the intended response, group A. Of incorrect answers, 22% were placed in group B, group C was not applicable for FN, and 17% were placed in group D. Papillary thyroid carcinoma and MTC were the most common misinterpretations. In addition, FN was the most common misinterpretation for MTC.

Representative images of each diagnostic category are depicted in Figure 1.

DISCUSSION

As shown by these results, which were based on morphology alone, it is difficult to achieve reproducibility and consistency among a diverse group of cytology professionals for the diagnostic category of FN. In addition, this category also often includes the follicular variant of PTC, which was a problematic diagnosis for many program participants. The cytologic criteria for FN have historically been somewhat subjective and descriptive. In clinical practice, ultrasound findings, radionuclide findings, thyroid function testing, patient history, and previous biopsy results may aid in the correct diagnosis. Using the Bethesda System for Thyroid Cytology Reporting has been shown to improve consistency and reproducibility within the 6 diagnostic categories.^{8,9} In addition, the Bethesda System establishes a risk of malignancy for each diagnostic category. However, even when using the Bethesda System, a significant number of aspirates (20%-30%) fall into indeterminate categories, defined as either atypia of undetermined significance (AUS)/atypical follicular lesion of undetermined significance (FLUS) or suspicious for follicular (Hürthle) neoplasm (SFN).¹⁰ In the AUS/FLUS category, repeat FNA may result in up to 50% of these lesions being reclassified

as benign.¹¹ Although the data collection in this study predates the widespread adoption of the Bethesda System, the FN and HCN cases included in this study would be placed in the SFN category, which carries a malignancy rate of 10%-30%.

Historically, thyroid lobectomy has been necessary to evaluate these lesions for malignant features, primarily capsular or vascular invasion.¹² If a thyroid malignancy is diagnosed at lobectomy, a second surgery for completion thyroidectomy is generally required. Thus, for most patients with benign nodules, surgery turns out to be unnecessary, and for many patients with thyroid malignancies, a second procedure is necessary. In other words, at least in retrospect, thyroid lobectomy is rarely the ideal treatment for any patient with an aspirate of a thyroid nodule diagnosed in an indeterminate category.

Molecular testing in combination with the cytologic findings is helpful in determining which patients within the indeterminate categories have a high risk of malignancy and would most likely benefit from surgery. Certain point mutations have been shown to be strongly associated with thyroid cancers. For instance, PTC often has point mutations of the *BRAF* and *RAF* genes, as well as *RET*/PTC rearrangements that are very specific for PTC.^{13,14} Follicular carcinomas are associated with *RAS* mutations and *PAX8*/*PPAR* rearrangement. Medullary carcinoma often possesses point mutations in the *RET* gene. Thus, in indeterminate thyroid aspirates that have these mutations, surgery would be recommended (possibly a total thyroidectomy) with a strong possibility that these patients have a thyroid malignancy, especially those with a *BRAF* mutation. The presence of these mutations is strongly associated with malignancy, but their absence does not necessarily confer a benign process. Thus, patients lacking these mutations still present clinical management questions because they still have a risk, albeit relatively low, for thyroid carcinoma.

Figure 1. (A and B) Aspirates from anaplastic thyroid carcinoma show large pleomorphic cells with variable cytoplasm, marked nuclear pleomorphism, and prominent irregular nucleoli (panel A: Rapid Romanowsky, high power; panel B: Papanicolaou, high power). (C and D) Aspirates from papillary thyroid carcinoma show large flat sheets and 3-dimensional structures with crowded, overlapping nuclei. The nuclei are ovoid with clear chromatin, nuclear grooves, and inclusions (C, Rapid Romanowsky, low power; panel; D, Papanicolaou, medium power). (E and F) Aspirates from medullary thyroid carcinoma show a dyscohesive to loosely cohesive architecture with granular cytoplasm, eccentric nuclei, and nuclear pleomorphism (E, Rapid Romanowsky, low power; F, Papanicolaou, medium power). (G and H) Aspirates from Hürthle cell neoplasm show cells with abundant granular cytoplasm, low n/c ratios, and relatively uniform round nuclei and nucleoli (G, Rapid Romanowsky, medium power; H, Papanicolaou, medium power). (I and J) Aspirates from follicular neoplasm are cellular with follicular cells arranged in monotonous microfollicular or trabecular structures with little to absent colloid. The cells have uniform round nuclei and indistinct cytoplasm (I, Rapid Romanowsky, medium power; J, Papanicolaou, low power).

One commercially available gene expression classifier technology measures the gene expression of 142 genes and classifies indeterminate thyroid aspirates as benign or suspicious. In 1 report, the negative predictive value of cases classified as benign by the test was approximately 95%.¹⁵ The test had a sensitivity of 92% and a specificity of 52% in detecting thyroid malignancies in aspirates classified by the test as suspicious. This type of testing may be most useful as an adjunct in thyroid aspirates that fall into indeterminate categories to identify patients at low risk for thyroid malignancy and thus avoid unnecessary lobectomy.

Summary

In summary, participants in the ASCP Non-GYN Assessment program, which includes a broad spectrum of cytology professionals, performed well in identifying patients who would benefit from surgical resection of their thyroid nodules. There is room for improvement in making the correct diagnostic interpretation, even within the malignant category and in the diagnosis of the follicular variant of PTC. Review and adherence to established cytologic criteria could improve performance in this area. The cytologic diagnosis of FN was the most challenging for participants. Given that this category has historically been based on somewhat subjective and descriptive criteria and contains a heterogeneous group of benign, malignant, and nonneoplastic entities, achieving consensus and reproducibility on cytologic features alone, as shown by this study, will be difficult.

As the cytology profession moves toward providing more individualized and personalized medical care, using the Bethesda System for Thyroid Cytology Reporting, available clinical information, repeat FNA (in the AUS/FLUS category), and judicious use of molecular testing in the indeterminate categories will improve diagnostic accuracy over that achieved with morphologic features alone.

FUNDING SUPPORT

No specific funding was disclosed.

CONFLICT OF INTEREST DISCLOSURES

The authors made no disclosures.

REFERENCES

1. McCall A, Jarosz H, Lawrence AM, et al. The incidence of thyroid carcinoma in solitary cold nodules and multinodular goiter. *Surgery*. 1986;100:1128-1132.
2. Gharib H, Goellner JR. Fine-needle aspiration biopsy of the thyroid; an appraisal. *Ann Int Med*. 1993;118:282-289.
3. Nguyen GK, Ginsberg J, Crockford PM. Fine-needle aspiration biopsy cytology of the thyroid. Its value and limitations in the diagnosis and management of solitary thyroid nodules. *Pathol Annu*. 1991;25:63-91.
4. National Comprehensive Cancer Network. NCCN Clinical practice guidelines in oncology: thyroid carcinoma. Available at: http://www.nccn.org/professionals/physician_gls/PDF/thyroid.pdf. Accessed July 7, 2013.
5. Cooper DS, Doherty GM, Haugen BR, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2009;19:1167-214.
6. Cibas ES, Ali SZ. The Bethesda System for Reporting Thyroid Cytopathology. *Am J Clin Pathol*. 2009;132:658-665.
7. Kloos RT, Eng C, Evans DB, et al; American Thyroid Association Guidelines Task Force. Medullary thyroid cancer: management guidelines of the American Thyroid Association. *Thyroid*. 2009;19:565-612.
8. Song JY, Kim L, Park IS, Han JY, Kim JM. Reclassifying formerly indeterminate thyroid FNAs using the Bethesda System reduces the number of inconclusive cases. *Acta Cytol*. 2012;56:122-129.
9. Pathak P, Srivastava R, Singh N, Arora VK, Bhatia. Implementation of the Bethesda system for reporting thyroid cytopathology: Interobserver concordance and reclassification of previously inconclusive aspirates. *Diagn Cytopathol*. 2014 [Epub ahead of print].
10. Baloch ZW, Cibas ES, Clark DP, et al. The National Cancer Institute thyroid fine needle aspiration state of the science conference: a summation. *Cytojournal*. 2008;5:6-6.
11. Yassa L, Cibas ES, Benson CB, et al. Long term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. *Cancer*. 2007;506-516.
12. Gharib H, Papini E, Paschke R, et al. American Association of Clinical Endocrinologists, Associazione Medicie Endocrinologi, and European Thyroid Association medical guidelines for the diagnosis and management of thyroid nodules. *Endocrinol Invest*. 2010; 33(Suppl):51-56.
13. Nikiforov YE, Ohori NP, Hodak SP, et al. Impact of mutational testing on the diagnosis and management of patients with cytologically indeterminate thyroid nodules: a prospective analysis of 1056 FNA samples. *J Clin Endocrinol Metab*. 2011; 96:3390-3397.
14. Nikiforov YE, Steward DL, Robinson-Smith TM, et al. Molecular testing for mutations in improving the fine-needle aspiration diagnosis of thyroid nodules. *J Clin Endocrinol Metab*. 2011; 94:2092-2098.
15. Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med*. 2012;367:705-715.