

# Incidence and Malignancy Rates of Diagnoses in the Bethesda System for Reporting Thyroid Aspiration Cytology: An Institutional Experience

Ji Hye Park · Sun Och Yoon  
Eun Ju Son<sup>1</sup> · Hye Min Kim  
Ji Hae Nahm · SoonWon Hong

Departments of Pathology and <sup>1</sup>Radiology,  
Gangnam Severance Hospital, Yonsei University  
College of Medicine, Seoul, Korea

Received: February 17, 2014

Revised: March 29, 2014

Accepted: March 31, 2014

## Corresponding Author

SoonWon Hong, M.D.

Department of Pathology, Gangnam Severance  
Hospital, Yonsei University College of Medicine,  
211 Eonju-ro, Gangnam-gu, Seoul 135-720, Korea  
Tel: +82-2-2019-3543  
Fax: +82-2-3463-2103  
E-mail: soonwonh@yuhs.ac

**Background:** The Bethesda System for Reporting Thyroid Cytopathology (BSRTC) uses six diagnostic categories to standardize communication of thyroid fine-needle aspiration (FNA) interpretations between clinicians and cytopathologists. Since several studies have questioned the diagnostic accuracy of this system, we examined its accuracy in our hospital. **Methods:** We calculated the incidences and malignancy rates of each diagnostic category in the BSRTC for 1,730 FNAs that were interpreted by four cytopathologists in Gangnam Severance Hospital between October 1, 2011, and December 31, 2011. **Results:** The diagnostic incidences of categories I-VI were as follows: 13.3%, 40.6%, 9.1%, 0.4%, 19.3%, and 17.3%, respectively. Similarly, the malignancy rates of these categories were as follows: 35.3%, 5.6%, 69.0%, 50.0%, 98.7%, and 98.9%, respectively. In categories II, V, and VI, there were no statistically significant differences in the ranges of the malignancy rates among the four cytopathologists. However, there were significant differences in the ranges for categories I and III. **Conclusions:** Our findings suggest that institutions that use the BSRTC should regularly update their diagnostic criteria. We also propose that institutions issue an annual report of incidences and malignancy rates to help other clinicians improve the case management of patients with thyroid nodules.

**Key Words:** Incidence; Biopsy, fine-needle; Thyroid; Pathology; Terminology

The Bethesda System for Reporting Thyroid Cytopathology (BSRTC) was developed in 2008 to facilitate more accurate communication of thyroid fine-needle aspiration (FNA) interpretations between clinicians and cytopathologists.<sup>1</sup> This system, which we adopted in 2010, classifies FNA results into six general diagnostic categories, namely, I) nondiagnostic or unsatisfactory, II) benign, III) atypia of undetermined significance or follicular lesion of undetermined significance, IV) follicular neoplasm or suspicious for a follicular neoplasm, V) suspicious for malignancy, and VI) malignant.<sup>1</sup> Each of these categories is associated with a risk of malignancy as follows: I) 1-4%, II) 0-3%, III) 5-15%, IV) 15-30%, V) 60-75%, and VI) 97-99%.<sup>1</sup> Although this system is useful, some studies that have investigated these risks have had controversial results, especially for category III. Furthermore, other studies have reported differences in malignancy rates among different institutions and have offered various explanations.

In this study, we determined the distribution of FNA results and malignancy rates in each diagnostic category of the BSRTC

in our hospital to determine whether our cytopathologists are using this system properly. Specifically, by analyzing data from individual cytopathologists, we hoped to ascertain whether our hospital has consistent FNA results in each diagnostic category. Finally, we suggest ways to improve the accuracy of classifying FNA diagnoses.

## MATERIALS AND METHODS

### Thyroid FNA cytology cases

We retrospectively analyzed 1,538 patients who had thyroid nodules that were diagnosed by FNA between October 1, 2011, and December 31, 2011, in Gangnam Severance Hospital in Korea. This study met criteria for exemption from review from the institutional review board. Each FNA diagnosis was made independently by one of four cytopathologists. Each thyroid aspiration sample was analyzed by using liquid-based preparation or conventional smear. Some FNAs were originally performed by other hospitals; however, in these cases, the slides were re-

evaluated by our cytopathologists. If a case was diagnosed by multiple FNA procedures, then we only considered the last diagnosis in the period of 2011. If a patient had more than two thyroid nodules, then we considered each FNA diagnosis as a separate case.

### Follow-up cases

Among 1,538 patients, we included 1,383 patients who had follow-up data in the same hospital after diagnosis of their thyroid nodules by FNA. The criteria for considering the follow-up data were at least one additional FNA, sonography of the thyroid or thyroid surgery from the date of initial FNA to December 31, 2013. For follow-up cases that were classified as category IV, V, or VI, we only considered the histologic diagnosis of surgical cases. However, for follow-up cases that were classified as category I, II, or III, we considered both the histologic diagnosis of surgical cases (if applicable) and the most recent FNA diagnosis during the follow-up period. The most recent FNA diagnoses that were classified as category I, II, or III were considered to be benign.

### Statistical analysis

To determine the sensitivity, specificity, false negative rate, false positive rate, positive predictive value, and negative predictive value of the malignancy rate of individual cytopathologists, we divided the BSRTC categories into two groups: 1) surgery is not recommended (categories I, II, and III) because the suggested malignancy risks of these categories are low, and 2) surgery is recommended (categories IV, V, and VI) because the suggested malignancy risks of these categories are high. The cut-off value for a malignant diagnosis by each cytopathologist was calculated by using a receiver operating characteristic (ROC) curve. ROC curve analysis was performed with MedCalc Statistical Software ver. 12.7.5 (MedCalc Software, Ostend, Belgium). p-values less than .05 were considered to be statistically significant.

## RESULTS

### Patients and distribution of diagnostic categories in all cases

In this study, we examined 1,538 patients who were 14 to 86 years old (mean, 50 years). The ratio of females to males was 3.8. Among these patients, 201 patients had more than two thyroid nodules, which resulted in a total of 1,730 FNA cases.

As shown in Table 1, the distribution of all cases in the six BSRTC diagnostic categories were as follows: 230 cases (13.3%) of category I, 702 cases (40.6%) of category II, 157 cases (9.1%) of category III, 7 cases (0.4%) of category IV, 335 cases (19.3%) of category V, and 299 cases (17.3%) of category VI. The total number and distributions of all cases that were analyzed by each cytopathologist is shown in Table 1.

### Patients and distribution of outcomes in follow-up cases

The patients who met the criteria for follow-up was 1,383. Among these 1,383 patients, 125 had more than two thyroid nodules, rendering a total of 1,547 cases which met the follow-up criteria. Of these 1,547 cases, 213 cases were examined by at least one additional FNA, 485 cases were examined by sonography and the remaining 849 cases were noted by surgery after the initial FNA. The ages of these patients ranged from 14 to 86 years old (mean, 49 years) and the ratio of females to males was 4.0. Their follow-up periods ranged from 2 days to 2 years and 2 months (median, 150 days).

The distributions of follow-up diagnoses for each initial BSRTC diagnostic classification are shown in Table 2. Specifically, category I diagnoses (116 cases) remained benign in 75 cases (64.7%), but were histologically confirmed as papillary carcinoma in 41 cases (35.3%). Category II diagnoses (702 cases) remained benign in 663 cases (94.5%), but were histologically confirmed as papillary carcinoma in 36 cases (5.1%), follicular carcinoma was present in two cases (0.3%), and poorly differentiated carcinoma occurred in one case (0.1%). Category III diagnoses (126 cases) remained benign in 39 cases (30.9%), but were histologically confirmed as papillary carcinoma in 84 cases

**Table 1.** Distribution of the diagnostic categories in all cases by each cytopathologist (%)

Diagnostic category of Bethesda system	Cytopathologists				Total (n=1,730)
	A (n=348)	B (n=330)	C (n=382)	D (n=670)	
I	28 (8.0)	50 (15.1)	63 (16.5)	89 (13.3)	230 (13.3)
II	128 (36.8)	151 (45.8)	122 (31.9)	301 (44.9)	702 (40.6)
III	41 (11.8)	18 (5.5)	40 (10.5)	58 (8.7)	157 (9.1)
IV	5 (1.4)	0 (0)	1 (0.3)	1 (0.1)	7 (0.4)
V	57 (16.4)	66 (20.0)	75 (19.6)	137 (20.5)	335 (19.3)
VI	89 (25.6)	45 (13.6)	81 (21.2)	84 (12.5)	299 (17.3)

**Table 2.** Distribution of the outcomes and malignancy rate in the follow-up cases (%)

Diagnostic category of Bethesda system	Follow-up diagnosis					Malignancy
	BN	PTC	MTC	FC	PD	
I (n = 116)	75 (64.7)	41 (35.3)	0 (0)	0 (0)	0 (0)	41 (35.3)
II (n = 702)	663 (94.5)	36 (5.1)	0 (0)	2 (0.3)	1 (0.1)	39 (5.6)
III (n = 126)	39 (30.9)	84 (66.7)	1 (0.8)	2 (1.6)	0 (0)	87 (69.0)
IV (n = 4)	2 (50.0)	2 (50.0)	0 (0)	0 (0)	0 (0)	2 (50.0)
V (n = 314)	4 (1.3)	306 (97.4)	3 (1.0)	0 (0)	1 (0.3)	310 (98.7)
VI (n = 285)	3 (1.1)	282 (98.9)	0 (0)	0 (0)	0 (0)	282 (98.9)

BN, benign; PTC, papillary carcinoma; MTC, medullary carcinoma; FC, follicular carcinoma; PD, poorly differentiated carcinoma.

**Table 3.** Distribution of the malignancy rate in follow-up cases by each cytopathologist (%)

Diagnostic category of Bethesda system	Cytopathologists			
	A (M/T)	B (M/T)	C (M/T)	D (M/T)
I	41.2 (7/17)	48.4 (15/31)	28.0 (7/25)	27.9 (12/43)
II	4.7 (6/128)	4.0 (6/151)	7.4 (9/122)	6.0 (18/301)
III	63.6 (21/33)	73.3 (11/15)	70.0 (21/30)	70.8 (34/48)
IV	33.3 (1/3)	0 (0/0)	0 (0/0)	100.0 (1/1)
V	100.0 (51/51)	95.4 (62/65)	100.0 (68/68)	99.2 (129/130)
VI	98.8 (80/81)	100.0 (43/43)	98.7 (78/79)	98.8 (81/82)
Total	53.0 (166/313)	44.9 (137/305)	56.5 (183/324)	45.5 (275/605)

M, case number of malignancy; T, total case number.

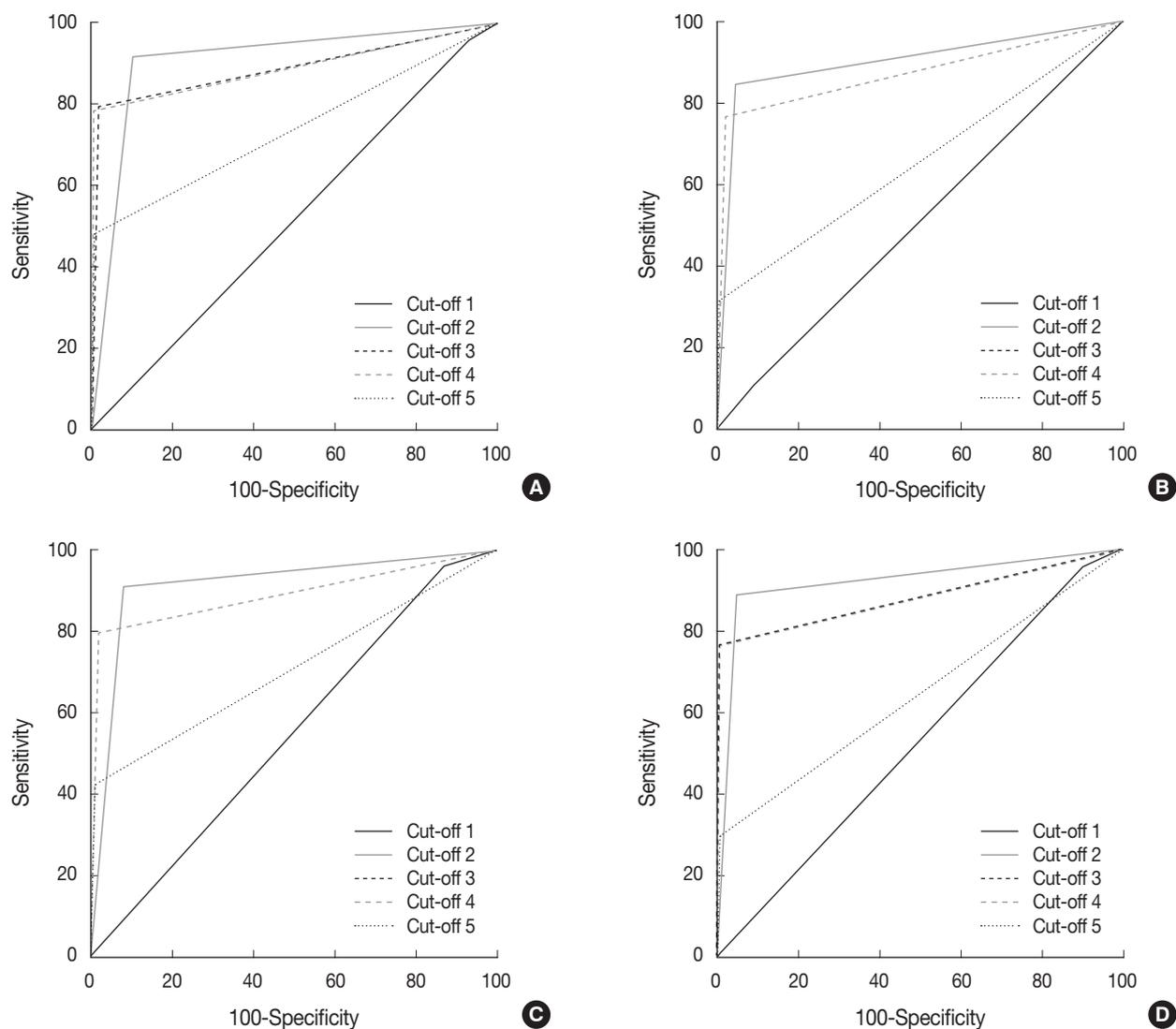
(66.7%), follicular carcinoma in two cases (1.6%), and medullary carcinoma in one case (0.8%). Category IV diagnoses (4 cases) were histologically confirmed as follicular carcinoma in two cases (50%) and were histologically confirmed as benign in the other two cases (50%). Category V diagnoses (314 cases) were histologically confirmed as papillary carcinoma in 306 cases (97.4%), medullary carcinoma in three cases (1.0%), and poorly differentiated thyroid carcinoma in one case (0.3%), but were histologically confirmed as benign in three cases (1.1%). Finally, category VI diagnoses (285 cases) were histologically confirmed as papillary carcinoma in 282 cases (98.9%), but were histologically confirmed as benign in three cases (1.1%). The histologically confirmed as benign cases included 53 cases (6.7%) of adenomatous hyperplasia, 13 cases (1.7%) of lymphocytic thyroiditis, nine cases (1.1%) of follicular adenoma, one case (0.1%) of a hyalinizing trabecular tumor, and three cases (0.4%) of a fibrocalcific nodule. The total number and distributions of follow-up cases that were analyzed by each cytopathologist are shown in Table 3. We did not find any significant differences in the malignancy rates for categories II, V, and VI among four cytopathologists, but found considerable difference in the malignancy rates for categories I and III among the cytopathologists. However, the statistical significance of the differences could not be calculated because each FNA diagnosis was made independently by one of four cytopathologists.

Instead, the cut-off value of malignant diagnosis was analyzed using ROC curves to see if there is a difference in the malignancy-suggesting category among the cytopathologists, and consequently to see if there is a difference in the diagnoses among the cytopathologists. As shown in Fig. 1, we found that all cytopathologists used category III as the cut-off category for differentiating malignant and benign cases ( $p < .0001$ ).

The sensitivity, specificity, false negative rate, false positive rate, positive predictive value, and negative predictive value of these malignancy rates for each cytopathologist are shown in Table 4.

## DISCUSSION

We compared our findings with those reported in 11 previous studies<sup>2-13</sup> and found several differences, which we attempt to explain below. The percentage of FNA diagnoses in the BSRTC categories V and VI in our study (19.3% and 17.3%, respectively) was higher than that in other studies (mean, 4.6% and 7.9%, respectively) (Table 5). In addition, the percentage of FNA diagnoses in category II in our study (40.6%) was lower than that in previous studies (mean, 62.7%) (Table 5). One possible reason for these differences may be that our hospital is a referral hospital for thyroid surgery, so many patients who are suspected of having thyroid cancer in other hospitals come here to



**Fig. 1.** Receiver operating characteristic curve showing the cut-off value of the diagnostic category of the Bethesda system for the diagnosis of histologically-confirmed malignant thyroid by cytopathologists A, B, C, and D. The cut-off values have the following values derived from the area under the curve: (A) cytopathologist A: cut-off value > 1, 0.513; cut-off value > 2, 0.910; cut-off value > 3, 0.887; cut-off value > 4, 0.891; cut-off value > 5, 0.738; (B) cytopathologist B: cut-off value > 1, 0.507; cut-off value > 2, 0.900; cut-off value > 3, 0.871; cut-off value > 4, 0.871; cut-off value > 5, 0.654; (C) cytopathologist C: cut-off value > 1, 0.544; cut-off value > 2, 0.914; cut-off value > 3, 0.888; cut-off value > 4, 0.888; cut-off value > 5, 0.706; (D) cytopathologist D: cut-off value > 1, 0.528; cut-off value > 2, 0.921; cut-off value > 3, 0.881; cut-off value > 4, 0.879; cut-off value > 5, 0.646.

**Table 4.** Sensitivity, specificity, false negative rate, false positive rate, positive predictive value, and negative predictive value of the malignancy rate by each cytopathologist (%)

	Cytopathologists			
	A	B	C	D
Sensitivity	79.5	76.6	79.8	76.7
Specificity	98.0	98.2	99.3	99.4
False negative rate	20.5	23.4	20.2	23.3
False positive rate	2.0	1.8	0.7	0.6
Positive predictive value	97.8	97.2	99.3	99.1
Negative predictive value	80.9	83.8	79.1	83.7

have their FNA diagnoses confirmed and if possible undergo surgery. Baloch *et al.*<sup>10</sup> also reported a high percentage of FNA diagnoses in BSRTC categories V and VI (19.1% and 21.3%, respectively) in a referral hospital, which is similar to our results. Similarly, Lee *et al.*<sup>13</sup> found that 13.0% of FNA diagnoses in a referral hospital were classified as category VI.

We also found differences in the malignancy rates of some BSRTC categories in our hospital and those reported in previous studies. For example, the malignancy rates in categories I, III, IV, and V in our study (35.3%, 69.0%, 50.0%, and 98.7%,

**Table 5.** Comparison of the distribution of the diagnostic categories of the present report with that of other reports (%)

Diagnostic category of Bethesda system	Present study	Mondal <i>et al.</i> <sup>3</sup>	Jo <i>et al.</i> <sup>4</sup>	Yassa <i>et al.</i> <sup>5</sup>	Yang <i>et al.</i> <sup>6</sup>	Nayar and Ivanovic <sup>7</sup>	Theoharis <i>et al.</i> <sup>8</sup>	Mufti and Molah <sup>9</sup>	Baloch <i>et al.</i> <sup>10</sup>	Stamataki <i>et al.</i> <sup>11</sup>	Wu <i>et al.</i> <sup>12</sup>	Lee <i>et al.</i> <sup>13</sup>
I	13.3	1.2	18.6	7.0	10.4	5.0	11.1	11.6	17.5	4.0	20.1	10.0
II	40.6	87.5	59.0	66.0	64.6	64.0	73.8	77.6	15.8	74.6	39.0	67.7
III	9.1	1.0	3.4	4.0	3.2	18.0	3.0	0.8	9.8	6.7	27.2	3.1
IV	0.4	4.2	9.7	9.0	11.6	6.0	5.5	4.0	16.5	-	8.4	0.6
V	19.3	1.4	2.3	9.0	2.6	2.0	1.4	2.4	19.1	2.8	2.6	5.1
VI	17.3	4.7	7.0	5.0	7.6	5.0	5.2	3.6	21.3	11.9	2.7	13.0

**Table 6.** Comparison of the distribution of the malignancy rate of the present report with that of other reports (%)

Diagnostic category of Bethesda system	Present study	Mondal <i>et al.</i> <sup>3</sup>	Jo <i>et al.</i> <sup>4</sup>	Yassa <i>et al.</i> <sup>5</sup>	Yang <i>et al.</i> <sup>6</sup>	Nayar and Ivanovic <sup>7</sup>	Theoharis <i>et al.</i> <sup>8</sup>	Mufti and Molah <sup>9</sup>	Baloch <i>et al.</i> <sup>10</sup>	Stamataki <i>et al.</i> <sup>11</sup>	Wu <i>et al.</i> <sup>12</sup>	Lee <i>et al.</i> <sup>13</sup>
I	35.3	0	8.9	10.0	10.7	9.0	-	20.0	64.0	30.0	12.0	77.8
II	5.6	4.5	11.0	0.3	0.7	2.0	9.8	3.1	11.0	6.0	8.0	0
III	69.0	20.0	17.0	24.0	19.2	6.0	48.0	50.0	50.0	33.0	27.0	79.0
IV	50.0	30.6	25.4	28.0	32.2	14.0	34.0	20.0	63.0	67.0	33.0	61.5
V	98.7	75.0	70.0	60.0	64.8	53.0	87.0	80.0	71.0	-	68.0	97.6
VI	98.9	97.8	98.1	97.0	98.4	97.0	100.0	100.0	100.0	100.0	100.0	100.0

respectively) were higher than those reported in the original BSRTC guidelines (1-4%, 5-15%, 15-30%, and 60-75%, respectively).<sup>1</sup> In addition, the malignancy rates in these categories in our study were higher than those in the other studies that we examined (mean, 24.2%, 33.9%, 37.2%, and 72.6%, respectively) (Table 6).<sup>2-13</sup> There are two possible reasons for these differences. First, although the BSRTC guidelines recommend that patients with categories I or III diagnoses have a repeat FNA, in Korea, patients who have thyroid nodules that are strongly suspicious for malignancy in a clinical aspect undergo surgery without a repeat FNA, but a frozen section examination may be performed. Second, Korean patients tend to be more concerned about false positive results than false negative results, which may pressure cytopathologists to underdiagnose FNA cases to avoid making false positive diagnoses. Lee *et al.*<sup>13</sup> also reported high malignancy rates for categories III and V (79.0% and 97.6%, respectively), which was similar to our result. However, their explanation was different from ours; they suspected that their cytopathologists did not properly apply the BSRTC classification criteria, which were still relatively new at the time of their study. If our explanation that Korean cytopathologists have a tendency to underdiagnose FNA cases is considerable, then clinicians may need to rely on intraoperative diagnoses in those cases. As a result, we need to refine and adapt to the BSRTC classification criteria or the malignancy rates of the BSRTC diagnostic categories may need to be modified. In addition, we recommend providing clinicians with current institutional data about malignancy rates in these categories to

help them improve their management of thyroid nodule cases.

In this study, we did not find any significant differences in the malignancy rates for categories II, V, and VI among four cytopathologists, but found considerable difference in the malignancy rates for categories I and III among the cytopathologists. However, the large number of patients in these categories who did not undergo surgery may have biased these results. Several previous studies have noted that malignancy rates in the BSRTC categories can differ among cytopathologists both intra- and inter-institutionally. For example, Layfield *et al.*<sup>14</sup> reported that there is wide variation in the malignancy rate for category III diagnoses between different institutions and among cytopathologists within same institution, depending on whether they had received cytopathology fellowship training. Similarly, Wu *et al.*<sup>12</sup> found that differences in thyroid cytopathology diagnoses may arise due to differences in the amount of experience or training that cytopathologists have. Furthermore, Cibas and Ali,<sup>1</sup> who wrote the original BSRTC paper, later admitted that "category III may never have good interobserver reproducibility, even after pathologists familiarize themselves with the criteria in the atlas,"<sup>15</sup> but they argued that this category was still useful because it provides clinically important distinctions for some cases. Since category III diagnoses may differ significantly among cytopathologists, we recommend that institutions prepare annual reports about the malignancy rates of diagnoses in categories III-VI by individual cytopathologists, so clinicians can improve their case management of patients with thyroid nodules. We also recommend that clinicians consider refining category III in

order to better characterize the threshold for differentiating malignant cases from benign ones.

Our comparison of malignancy rate was difficult because category I, II, and III usually does not require surgery, and distribution of category IV is too small to compare. Other reports also commented on the difficulty in comparing malignancy rates arising from the same cause.<sup>16-18</sup>

Our comparisons of the malignancy rates of individual cytopathologists may be limited in two ways. First, there are some methodological differences between our study and previous studies that have compared the FNA diagnoses of individual cytopathologists. Specifically, other studies made comparisons among cytopathologists who had all diagnosed the same slides.<sup>19-24</sup> However, this method was not applicable to our study, because we performed a retrospective data analysis. In addition, such quality control methods are often not practical on a daily basis in hospitals. Second, there were no statistically significant differences in the cut-off value of malignant diagnosis by each cytopathologist. We found that all cytopathologists used category III as the cut-off category for differentiating malignant and benign cases ( $p < .0001$ ). Therefore, we suggested that there was no difference in the malignancy rates of the diagnoses made by individual cytopathologists. However, an improvement in cut-off value is needed because category III is not appropriate for determining malignancy.

The sensitivity, specificity, false negative rate, false positive rate, positive predictive value, or negative predictive value of the malignancy rates of individual cytopathologists were not significantly different among cytopathologists. There is some limitation due to the large number of cases that were not treated with surgery.

## Conclusion

In conclusion, the distribution of diagnostic categories in this institute is shifted towards category V and VI, and the malignancy rates of category I, III, and V were higher than those in other reports. Thus, our findings regarding the distribution of FNA diagnoses in the BSRTC diagnostic categories and their malignancy rates in our hospital suggest the need for future improvements in the BSRTC. Specifically, the determination of the malignancy rate needs to be modified to reduce additional diagnostic procedures, such as intraoperative diagnosis. In addition, we propose that institutions prepare, communicate, and use annual reports of malignancy rates of their cytopathologists' diagnoses to help clinicians practice better case management of patients with thyroid nodules.

## Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

## REFERENCES

1. Cibas ES, Ali SZ; NCI Thyroid FNA State of the Science Conference. The Bethesda System For Reporting Thyroid Cytopathology. *Am J Clin Pathol* 2009; 132: 658-65.
2. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda System for Reporting Thyroid Cytopathology: a meta-analysis. *Acta Cytol* 2012; 56: 333-9.
3. Mondal SK, Sinha S, Basak B, Roy DN, Sinha SK. The Bethesda system for reporting thyroid fine needle aspirates: a cytologic study with histologic follow-up. *J Cytol* 2013; 30: 94-9.
4. Jo VY, Stelow EB, Dustin SM, Hanley KZ. Malignancy risk for fine-needle aspiration of thyroid lesions according to the Bethesda System for Reporting Thyroid Cytopathology. *Am J Clin Pathol* 2010; 134: 450-6.
5. Yassa L, Cibas ES, Benson CB, *et al.* Long-term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. *Cancer* 2007; 111: 508-16.
6. Yang J, Schnadig V, Logrono R, Wasserman PG. Fine-needle aspiration of thyroid nodules: a study of 4,703 patients with histologic and clinical correlations. *Cancer* 2007; 111: 306-15.
7. Nayar R, Ivanovic M. The indeterminate thyroid fine-needle aspiration: experience from an academic center using terminology similar to that proposed in the 2007 National Cancer Institute Thyroid Fine Needle Aspiration State of the Science Conference. *Cancer* 2009; 117: 195-202.
8. Theoharis CG, Schofield KM, Hammers L, Udelsman R, Chhieng DC. The Bethesda thyroid fine-needle aspiration classification system: year 1 at an academic institution. *Thyroid* 2009; 19: 1215-23.
9. Mufti ST, Molah R. The Bethesda System for Reporting Thyroid Cytopathology: a five-year retrospective review of one center experience. *Int J Health Sci (Qassim)* 2012; 6: 159-73.
10. Baloch ZW, Hendreen S, Gupta PK, *et al.* Interinstitutional review of thyroid fine-needle aspirations: impact on clinical management of thyroid nodules. *Diagn Cytopathol* 2001; 25: 231-4.
11. Stamataki M, Anninos D, Brountzos E, *et al.* The role of liquid-based cytology in the investigation of thyroid lesions. *Cytopathology* 2008; 19: 11-8.
12. Wu HH, Rose C, Elsheikh TM. The Bethesda System for Reporting Thyroid Cytopathology: an experience of 1,382 cases in a community practice setting with the implication for risk of neoplasm and

- risk of malignancy. *Diagn Cytopathol* 2012; 40: 399-403.
13. Lee K, Jung CK, Lee KY, Bae JS, Lim DJ, Jung SL. Application of Bethesda System for Reporting Thyroid Aspiration Cytology. *Korean J Pathol* 2010; 44: 521-7.
  14. Layfield LJ, Morton MJ, Cramer HM, Hirschowitz S. Implications of the proposed thyroid fine-needle aspiration category of "follicular lesion of undetermined significance": a five-year multi-institutional analysis. *Diagn Cytopathol* 2009; 37: 710-4.
  15. Crippa S, Mazzucchelli L, Cibas ES, Ali SZ. The Bethesda system for reporting thyroid fine-needle aspiration specimens. *Am J Clin Pathol* 2010; 134: 343-4.
  16. Crowe A, Linder A, Hameed O, *et al.* 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* 2011; 119: 315-21.
  17. Marchevsky AM, Walts AE, Bose S, *et al.* Evidence-based evaluation of the risks of malignancy predicted by thyroid fine-needle aspiration biopsies. *Diagn Cytopathol* 2010; 38: 252-9.
  18. Renshaw A. An estimate of risk of malignancy for a benign diagnosis in thyroid fine-needle aspirates. *Cancer Cytopathol* 2010; 118: 190-5.
  19. Gerhard R, da Cunha Santos G. Inter- and intraobserver reproducibility of thyroid fine needle aspiration cytology: an analysis of discrepant cases. *Cytopathology* 2007; 18: 105-11.
  20. Redman R, Yoder BJ, Massoll NA. Perceptions of diagnostic terminology and cytopathologic reporting of fine-needle aspiration biopsies of thyroid nodules: a survey of clinicians and pathologists. *Thyroid* 2006; 16: 1003-8.
  21. Bomeisl PE Jr, Alam S, Wakely PE Jr. Interinstitutional consultation in fine-needle aspiration cytopathology: a study of 742 cases. *Cancer* 2009; 117: 237-46.
  22. Kocjan G, Chandra A, Cross PA, *et al.* The interobserver reproducibility of thyroid fine-needle aspiration using the UK Royal College of Pathologists' classification system. *Am J Clin Pathol* 2011; 135: 852-9.
  23. Clary KM, Condel JL, Liu Y, Johnson DR, Grzybicki DM, Raab SS. Interobserver variability in the fine needle aspiration biopsy diagnosis of follicular lesions of the thyroid gland. *Acta Cytol* 2005; 49: 378-82.
  24. Stelow EB, Bardales RH, Crary GS, *et al.* Interobserver variability in thyroid fine-needle aspiration interpretation of lesions showing predominantly colloid and follicular groups. *Am J Clin Pathol* 2005; 124: 239-44.