Original Article
Yonsei Med J 2026 Apr;67(4):314-321
https://doi.org/10.3349/ymj.2025.0257
Yonsei Medical Journal
YMJ
pISSN: 0513-5796 · eISSN: 1976-2437

Adequacy Criteria for Thyroid Fine-Needle Aspiration in the Era of the Bethesda Reporting System

Jin Young Kwak[1]*, Sangwoo Cho[2]*, Hye Sun Lee[3], Jung Hyun Yoon[1], Syed Z. Ali[4], and Soon Won Hong[5]

[1]Department of Radiology and Research Institute of Radiological Science, Yonsei University College of Medicine, Seoul, Korea;
[2]Yonsei University College of Medicine, Seoul, Korea;
[3]Biostatistics Collaboration Unit, Yonsei University College of Medicine, Seoul, Korea;
[4]Department of Pathology, The Johns Hopkins Medical Institutions, Baltimore, MD, USA;
[5]Department of Pathology, Yonsei University College of Medicine, Seoul, Korea.

Purpose: There is a lack of consensus and data validating lower cell counts for sample adequacy of thyroid fine-needle aspiration (FNA). We investigated less stringent adequacy thresholds under the Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) and evaluated malignancy risks for "nondiagnostic" nodules by ultrasound features.

Materials and Methods: A total of 2459 nodules with initial FNA were included. We built 11 new adequacy criteria based on the number of cell groups and total follicular epithelial cells. Diagnostic performance of each criterion was compared with the original criterion using the general estimating equation. Nondiagnostic nodules under each criterion were categorized by the American College of Radiology Thyroid Imaging Reporting and Data System (ACR TIRADS), and malignancy rates of each ACR TIRADS category were compared.

Results: Malignancy rates of nondiagnostic nodules were under 3.5% across all criteria, and showed no significant differences compared with the original. More than 40 cells, regardless of cell group, or more than three cell groups showed no significant difference in diagnostic accuracy and false negative rates compared with the original criterion. Malignancy rates of nondiagnostic nodules were above 28.6% for ACR TIRADS 5, and below 5% for ACR TIRADS 1 to 4, in all criteria.

Conclusion: Less stringent thresholds for sample adequacy can show comparable diagnostic performances to the original criterion of the TBSRTC. Given their markedly higher malignancy rates, nondiagnostic ACR TIRADS 5 nodules may warrant more active management than lower category nodules.

Key Words: Thyroid nodule; biopsy, fine-needle; ultrasonography; cytology

INTRODUCTION

Fine-needle aspiration (FNA) has been widely used as a diagnostic tool for thyroid nodules.[1] However, one inherent limitation of FNA is that it is not always possible to obtain sufficient materials for definitive characterization.[1] Although ultrasonographic (US) guidance can enable a physician to accurately and safely target a thyroid nodule during FNA, several factors affect sample adequacy, such as the technique of the performer, the number of nodules aspirated, and the size and type of nodules being biopsied.[2-5]

For sample adequacy, the Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) recommends a minimum of six groups of well-visualized follicular epithelial cells with at least 10 cells per group, based on criteria developed at the Mayo Clinic.[6,7] Although the 2017 TBSRTC commented that it may be possible to reduce the number of follicular cells needed for sample adequacy without significantly increasing the false-negative rate (FNR), the 2023 TBSRTC reaffirmed the original adequacy criteria.[8] This decision reflects concerns about the importance of maintaining false-negative rates, the lack of con-

Received: July 30, 2025 **Revised:** October 29, 2025
Accepted: November 4, 2025 **Published online:** February 12, 2026
Corresponding author: Soon Won Hong, MD, PhD, Department of Pathology, Yonsei University College of Medicine, 211 Eonju-ro, Gangnam-gu, Seoul 06273, Korea. E-mail: SOONWONH@yuhs.ac

*Jin Young Kwak and Sangwoo Cho contributed equally to this work.
•The authors have no potential conflicts of interest to disclose.

sensus and no robust data validating lower cell counts, and inconsistent follow-up across different studies and practice settings.[9] Still, considering that most nondiagnostic nodules ultimately prove to be benign, reducing the required quantity of follicular cells needed to define "nondiagnostic" cytology could potentially help safely reduce unnecessary procedures.[1,10] Consequently, it is imperative to thoroughly investigate the appropriateness of this standard.

The risk of malignancy was higher in thyroid nodules that had suspicious US features even with the same cytologic results.[11-13] Several US risk stratification systems have been introduced to evaluate thyroid nodules.[2,14-16] When thyroid nodules with nondiagnostic cytology were evaluated using the Thyroid Imaging Reporting and Data System (TIRADS), their malignancy risk varied according to their TIRADS categories.[14,17] In this study, we investigated whether it was feasible to lower the thresholds of sample adequacy for nondiagnostic nodules with the current incorporation of the TBSRTC and evaluated the risk of malignancy for nondiagnostic nodules according to the American College of Radiology (ACR) TIRADS.[16]

MATERIALS AND METHODS

Study design

This study is of a retrospective design and was approved by our Institutional Review Board (IRB institution: Severance Hospital; No. 4-2024-1038). Neither patient approval nor informed consent was required to review medical records or US images.

The TBSRTC has been consistently utilized for thyroid cytopathology reports in our institution (a referral center) since its implementation in December 2009.[6] From March 2016 to February 2019, 5898 focal thyroid nodules in 5741 consecutive patients aged 19 years or older underwent initial US-guided thyroid FNA. In 4035 patients, 4096 focal thyroid nodules were 1 cm or greater along their maximum diameter on US. Nodules were included if they met one of the following criteria: 1) underwent surgery (n=1048); 2) definitive benign (n=167) or malignant cytology (n=1) on follow-up US-FNA (n=154) or US-guided core needle biopsy (n=14); 3) decreased (n=425) or no change in size on follow-up US performed at least 1 year after the initial FNA (n=763); or 4) malignant on the initial FNA (n=55). Finally, 2459 thyroid nodules in 2431 patients were included in this study. Fig. 1 shows the diagnostic flowchart of included nondiagnostic nodules.

US examinations and image analysis

US examinations of both thyroid glands and neck areas were performed using a 5–12 MHz linear array transducer (iU22 or EPIQ5, Philips Healthcare, Amsterdam, Netherlands). Real-time US and subsequent US-guided FNA were performed by one of 17 radiologists with 1–25 years of experience in thyroid imaging (5 faculties and 12 fellows). US-guided FNA was performed on thyroid nodules showing suspicious US features or on the largest mass when none of the multiple thyroid nodules observed showed any US features suspicious for cancer.[14]

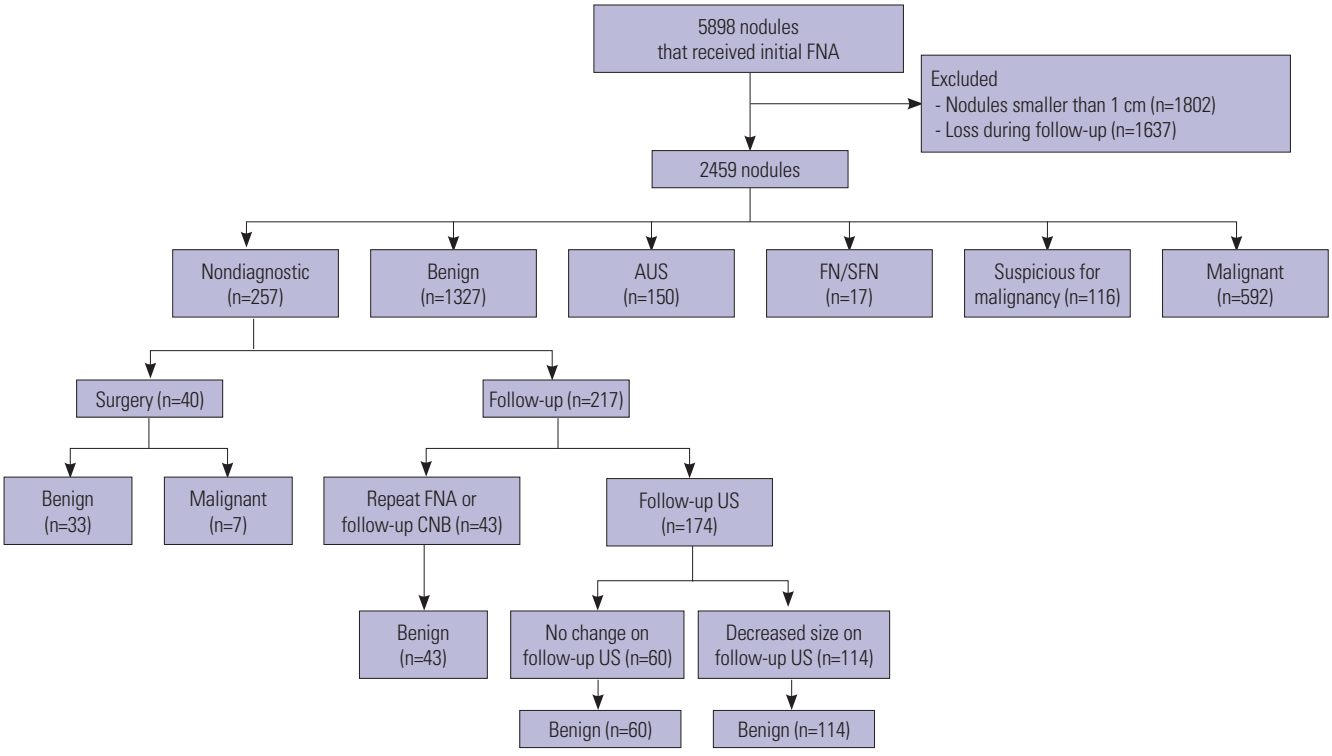

Fig. 1. Diagnostic flowchart of the included nondiagnostic thyroid nodules. FNA, fine-needle aspiration; AUS, atypia of undetermined significance; FN, follicular neoplasm; SFN, suspicious for a follicular neoplasm; CNB, core-needle biopsy; US, ultrasonography.

Transverse and longitudinal gray-scale US images were obtained from nodules. The radiologist who conducted the US and US-FNA procedures reported the US features of each thyroid nodule and prospectively entered them into our institutional database. The recorded US features were composition, echogenicity, margin, calcifications, and shape. From June 2012 to the present, our database has been utilizing these descriptors.[14] Using the imaging features recorded in our institutional database, focal thyroid nodules on US were retrospectively categorized based on the ACR TIRADS,[16] assigned by an experienced radiologist (J.Y.K., with 22 years of experience in thyroid imaging) who was blinded to the final diagnoses of the nodules.

US-FNA and cytologic interpretation

US-guided FNA was performed on each thyroid nodule using a 23-gauge needle attached to a 2-mL disposable syringe without an aspirator. Local anesthesia was not routinely administered, and onsite evaluation was not routinely performed. The aspirated material was expelled on two glass slides and immediately placed in 95% ethanol for Papanicolaou staining. The remaining material in the syringe was collected in a CytoLyt transport medium (Hologic Co., Marlborough, MA, USA) or CytoRich^{R} Red preservative fluid (BD, Franklin Lakes, NJ, USA).[18] A single Papanicolaou-stained liquid-based cytology slide was routinely prepared for each case. Cell counts and cluster assessments were determined from the combined evaluation of all prepared slides. All cytopathologic slides were reviewed by one of the 15 cytopathologists in our institution according to the pathology department's internal review schedule. Cytologic results were reported according to the TBSRTC.[9]

For this study, the cytologic slides of nodules with nondiagnostic results were reviewed. Six nondiagnostic nodules were excluded because their cytologic slides were not available. Seven nondiagnostic nodules were reclassified as AUS after their cytologic slides were reviewed by an experienced cytopathologist (S.W.H.). The cytopathologist retrospectively reviewed the slides of 257 nodules that were diagnosed with nondiagnostic cytology according to the following characteristics: follicular cell count, number of clusters with at least 10 well-visualized follicular epithelial cells, presence of nuclear and cytoarchitectural atypia, histiocytoid cells, giant cells, lymphocytes, colloid, hemosiderin-laden macrophages, cyst-lining cells, and artifacts. The cytopathologist did not know the demographic data, US findings, and final diagnoses of the patients.

Statistical analysis

We evaluated the diagnostic performances of 12 different criteria, including 11 criteria newly built for sample adequacy (Fig. 2).[9] Criterion 1 is the current criterion recommended by the TBSRTC, with a minimum of six groups of follicular epithelial cells needed with at least 10 cells per group for a sample to be considered adequate for a cytology examination.[6] The newly added 11 criteria for sample adequacy were built by modifying criterion 1 to lower the thresholds for specimen adequacy. Criteria 2 to 6 were defined by the number of groups of follicular epithelial cells with at least 10 cells per group. More specifically, criterion 2 was set as a minimum of five groups; criterion 3 as a minimum of four groups; criterion 4 as a minimum of three groups; criterion 5 as a minimum of two groups; and criterion 6 as a minimum of one group. Criteria 7 to 12 were defined by the total follicular epithelial cell count, regardless of the number of cell clusters. More specifically, criterion 7 was set as having at least 60 follicular epithelial cells total; criterion 8 as at least 50 cells total; criterion 9 as at least 40 cells total; criterion 10 at least 30 cells total; criterion 11 as at least 20 cells total; and criterion 12 as at least 10 cells total. Nodules were categorized as nondiagnostic regardless of follicular epithelial cell count when slides showed poor quality, such as dry artifacts, bloody obscuring, or squeezed artifact.

Categorical variables were compared between benign and malignant nodules with the χ^2 test. For patient based-analysis, patients with at least one malignant nodule were categorized as malignant. Continuous variables were tested for normality using the Kolmogorov–Smirnov test and Shapiro–Wilk test. Both patient age and nodule size were not normally distributed and were evaluated using the Mann–Whitney U test. Malignancy

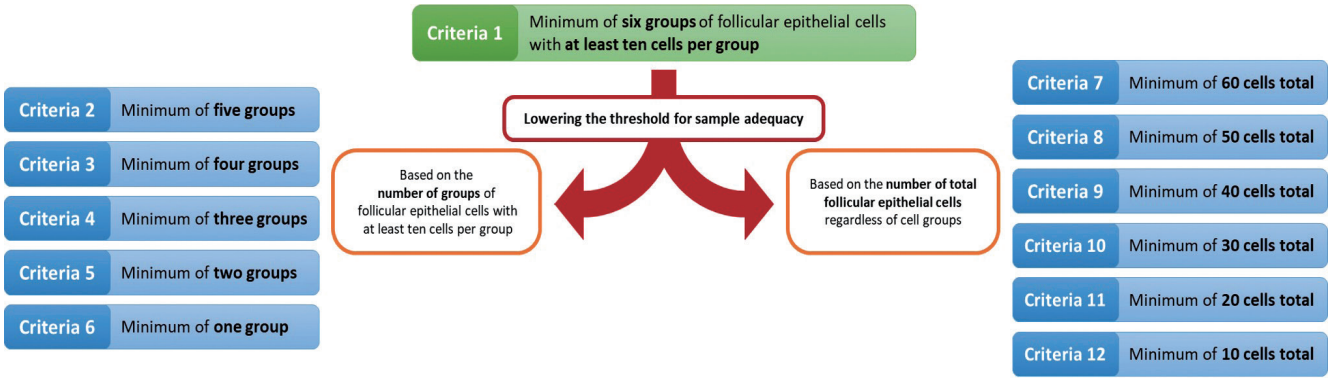

Fig. 2. Definition of the 12 criteria for sample adequacy. Current criteria for nondiagnostic cytology results, recommended by the Bethesda System for Reporting Thyroid Cytopathology.

https://doi.org/10.3349/ymj.2025.0257

rates of nodules with nondiagnostic or benign cytology according to each criterion were evaluated using the generalized estimating equation (GEE) to account for cases where multiple nodules were included from the same patient.

The primary outcome of this study was diagnostic performances according to each criterion. We defined "benign" cytology as test-negative, and "suspicious for malignancy" and "malignant" cytology as test-positive. We evaluated diagnostic performances for the detection of thyroid cancer including sensitivity, specificity, accuracy, positive predictive value, negative predictive value, and false negative rate (FNR) according to each criterion along with 95% confidence intervals. The diagnostic performances of criteria 2 to 12 were compared with those of the original criterion (criterion 1) using the GEEs. As repeat FNA is recommended for nodules with nondiagnostic cytology according to the TBSRTC management guidelines, we evaluated the decrease in repeat FNA after each criterion was applied. The secondary outcome was defined as the malignancy rates of each ACR TIRADS category in nondiagnostic nodules according to each criterion. Nondiagnostic nodules, according to each criterion, were categorized according to the ACR TIRADS. Malignancy rates of each ACR TIRADS category were compared using GEEs.

Differences with a p-value<0.05 were considered statistically significant. All statistical analyses were conducted with SAS version 9.4 (SAS Institute, Cary, NC, USA).

RESULTS

Baseline characteristics of the patients and nodules are summarized in Table 1. Of the 2459 nodules enrolled in this study, 810 (32.9%) were malignant and 1649 (67.1%) were benign. Patients with benign nodules were significantly older than patients with malignant nodules. Malignant nodules were significantly smaller than benign nodules. Gender distribution was significantly different between patients with malignant and benign nodules (p<0.001). There were also significantly different malignancy rates according to the ACR TIRADS categories (p<0.001).

The malignancy rates for nodules with nondiagnostic and benign cytology according to each criterion are summarized in Table 2. Criterion 1, the original criterion for sample adequacy according to the TBSRTC, showed malignancy rates of 2.7% for nondiagnostic nodules and 2.6% for nodules with benign cytology. Malignancy rates of nondiagnostic nodules were under 3.5% for all criteria, and there were no significant differences between the malignancy rates of criterion 1 and any of the other criteria. Malignancy rates of nodules with benign cytology were under 3% for all criteria, and there were no significant differences between the malignancy rates of criterion 1 and any of the other criteria.

The diagnostic performances of each criterion are listed in Supplementary Table 1 (only online). The diagnostic performance of criterion 1 is as follows: FNR, 4.7%; sensitivity, 95.3%;

Table 1. Baseline Characteristics of the Included Patients and Their Nodules

All patients	Total (n=2431)	Malignant (n=802)	Benign (n=1629)	*p*
Age (yr)	50 (19–92)	44 (19–88)	52 (19–92)	<0.001
Gender				<0.001
Men	528 (21.7)	213 (26.6)	315 (19.3)	
Women	1903 (78.3)	589 (73.4)	1314 (80.7)	
All nodules	**Total (n=2459)**	**Malignant (n=810)**	**Benign (n=1649)**	***p***
Size (mm)	19 (10–100)	15 (10–80)	22 (10–100)	<0.001
Cytology*				<0.001
Nondiagnostic	257 (10.5)	7 (0.9)	250 (15.2)	
Benign	1327 (54.0)	34 (4.2)	1293 (78.4)	
AUS/FLUS	150 (6.1)	67 (8.3)	83 (5.0)	
FN/SFN	17 (0.7)	6 (0.7)	11 (0.7)	
Suspicious for malignancy	116 (4.7)	106 (13.1)	10 (0.6)	
Malignant	592 (24.1)	590 (72.8)	2 (0.1)	
ACR TIRADS				<0.001
TR 1	27 (1.1)	0 (0.0)	27 (1.6)	
TR 2	618 (25.1)	19 (2.4)	599 (36.3)	
TR 3	593 (24.1)	24 (3.0)	569 (34.5)	
TR 4	585 (23.8)	162 (20.0)	423 (25.7)	
TR 5	636 (25.9)	605 (74.7)	31 (1.9)	

AUS, atypia of undetermined significance; FLUS, follicular lesion of undetermined significance; FN, follicular neoplasm; SFN, suspicious for a follicular neoplasm; ACR TIRADS, American College of Radiology Thyroid Imaging Reporting and Data System; TR, TIRADS category.
Data are presented as median (range) or n (%).
*Categorized based on the Bethesda System for Reporting Thyroid Cytopathology.

specificity, 99.1%; and accuracy, 97.7%. Compared with criterion 1, criteria 2, 3, 4, 7, 8, and 9 did not show significantly different FNRs. There were no significant differences between the diagnostic accuracies of criterion 1 and the other criteria. Specificity was significantly higher with all other criteria compared to criterion 1 (p<0.05 for all) (Supplementary Table 2, only online). When the original criterion recommended by the TBSRTC was applied, the number of repeat FNA performed on nodules with nondiagnostic cytology was 257, followed by 211 for criterion 2; 201 for criterion 3; 189 for criterion 4; 166 for criterion 5; 126 for criterion 6; 184 for criterion 7; 170 for criterion 8; 160 for criterion 9; 143 for criterion 10; 124 for criterion 11; 108 for criterion 12 (Fig. 3). The number of repeat FNA for nodules with nondiagnostic cytology showed a decreasing trend from criterion 1 to criterion 6 (p<0.001), and from criterion 7 to criterion 12 (p<0.001).

Of nondiagnostic nodules according to criterion 1, the malignancy rates of those categorized as ACR TIRADS 1 to 5 were 0% (0 of 15), 1% (1 of 103), 3.2% (2 of 63), 2.9% (2 of 68), and 28.6% (2 of 7), respectively. The malignancy rates of nondiagnostic nodules categorized as ACR TIRADS 5 were all above 28.6%, and of those assessed as ACR TIRADS 1 to 4 were below 5% for all adjusted criteria (Table 3). Most of the criteria did not show significantly different malignancy rates between the ACR TIRADS categories, except for criterion 6 (p=0.046), criterion 11 (p=0.046), and criterion 12 (p=0.046).

DISCUSSION

In this era of widespread thyroid nodule detection on US, it is crucial to identify clinically significant thyroid cancers while minimizing the need for invasive procedures whenever possible. Although FNA is considered a reliable method for diagnosing thyroid nodules, nondiagnostic cytology may necessitate additional invasive procedures.[2,19] However, many nodules with nondiagnostic cytology are ultimately proven to be benign.[6,8,9,20] In light of the current situation, reducing the requirements for sample adequacy and thus reducing the number of FNAs performed on nodules with nondiagnostic results would benefit the clinical practice.[1,8,10] However, errors associated with sampling and preparation techniques may result in inadequate samples that are categorized as nondiagnostic, leading to false-negative diagnoses that could directly affect the actual outcome of patients.[21,22] Therefore, we need to ensure a dependable diagnostic performance with FNA, especially regarding the FNR. In this study, among criteria with less stringent thresholds for sample adequacy, criteria 2, 3, 4, 7, 8, and 9 demonstrated FNRs ranging from 4.8% to 5.0% that did not significantly change from the FNR of the original criterion (criterion 1). All criteria showed

Table 2. Malignancy Rates of Nondiagnostic and Benign Nodules according to Different Criteria for Sample Adequacy

Malignancy rates	Criteria											
	1	2	3	4	5	6	7	8	9	10	11	12
Nondiagnostic	2.7% (7/257)	2.8% (6/211)	3.0% (6/201)	2.7% (5/189)	1.8% (3/166)	1.6% (2/126)	3.3% (6/184)	2.9% (5/170)	2.5% (4/160)	2.1% (3/143)	1.6% (2/124)	1.9% (2/108)
Benign	2.56% (34/1327)	2.55% (35/1373)	2.53% (35/1383)	2.58% (36/1395)	2.68% (38/1418)	2.67% (39/1458)	2.50% (35/1400)	2.55% (36/1414)	2.60% (37/1424)	2.64% (38/1441)	2.67% (39/1460)	2.64% (39/1476)

There was no statistically significant difference in malignancy rates between criterion 1 and any of the other criteria (p>0.05 for all).

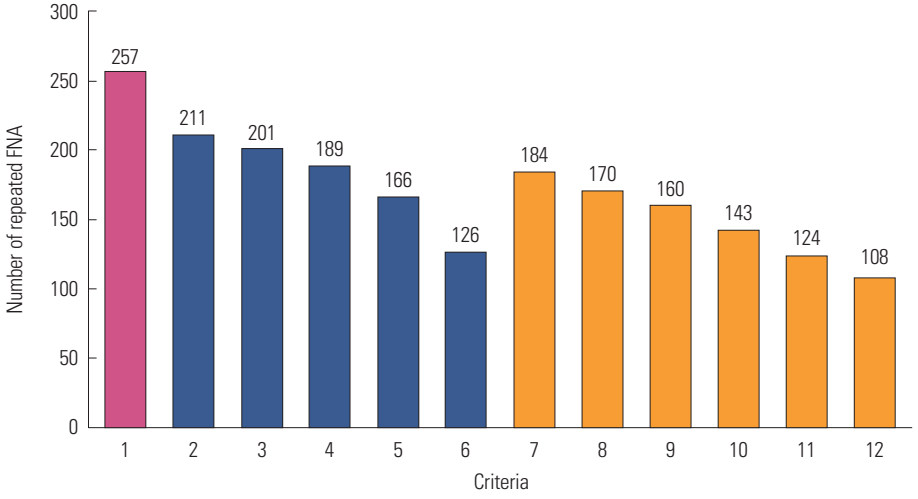

Fig. 3. Number of repeat FNA according to each criterion. Criterion 1 is the original criterion recommended by the TBSRTC. Criteria 2 to 6 were defined as a minimum of five, four, three, two, and one cell groups, respectively. Criteria 7 to 12 were defined as a minimum of 60, 50, 40, 30, 20, and 10 cells, regardless of cell groups, respectively. FNA, fine-needle aspiration; TBSRTC, the Bethesda System for Reporting Thyroid Cytopathology.

https://doi.org/10.3349/ymj.2025.0257

Table 3. Distribution and Malignancy Rates of Nondiagnostic Nodules Categorized by ACR TIRADS according to Different Criteria for Sample Adequacy

Criteria	ACR category	Nondiagnostic nodules			
		Benign (%)	Malignant (%)	Estimated malignancy rate (SE)	*p*
Criterion 1	2	102 (40.8)	1 (14.3)	1.0 (1.0)	0.094
	3	61 (24.4)	2 (28.6)	3.2 (2.2)	
	4	67 (26.8)	2 (28.6)	2.9 (2.0)	
	5	5 (2.0)	2 (28.6)	28.6 (17.1)	
Criterion 2	2	86 (42.0)	1 (16.7)	1.2 (1.1)	0.140
	3	46 (22.4)	2 (33.3)	4.2 (2.9)	
	4	53 (25.9)	1 (16.7)	1.9 (1.8)	
	5	5 (2.4)	2 (33.3)	28.6 (17.1)	
Criterion 3	2	81 (41.5)	1 (16.7)	1.2 (1.2)	0.130
	3	46 (23.6)	2 (33.3)	4.2 (2.9)	
	4	49 (25.1)	1 (16.7)	2.0 (2.0)	
	5	4 (2.1)	2 (33.3)	33.3 (19.3)	
Criterion 4	2	75 (40.8)	1 (20.0)	1.3 (1.3)	0.172
	3	43 (23.4)	1 (20.0)	2.3 (2.3)	
	4	48 (26.1)	1 (20.0)	2.0 (2.0)	
	5	3 (1.6)	2 (40.0)	40.0 (21.9)	
Criterion 5	2	65 (39.9)	1 (33.3)	1.5 (1.5)	0.082
	3	39 (23.9)	0 (0.0)	0.0 (0.0)	
	4	45 (27.6)	0 (0.0)	0.0 (0.0)	
	5	2 (1.2)	2 (66.7)	50.0 (25.0)	
Criterion 6	2	51 (41.1)	0 (0.0)	0.0 (0.0)	0.046
	3	30 (24.2)	0 (0.0)	0.0 (0.0)	
	4	30 (24.2)	0 (0.0)	0.0 (0.0)	
	5	2 (1.6)	2 (100.0)	50.0 (25.0)	
Criterion 7	2	70 (39.3)	1 (16.7)	1.4 (1.4)	0.114
	3	43 (24.2)	2 (33.3)	4.4 (3.1)	
	4	49 (27.5)	1 (16.7)	2.0 (2.0)	
	5	3 (1.7)	2 (33.3)	40.0 (21.9)	
Criterion 8	2	66 (40.0)	1 (20.0)	1.5 (1.5)	0.133
	3	40 (24.2)	1 (20.0)	2.4 (2.4)	
	4	45 (27.3)	1 (20.0)	2.2 (2.2)	
	5	2 (1.2)	2 (40.0)	50.0 (25.0)	
Criterion 9	2	61 (39.1)	1 (25.0)	1.6 (1.6)	0.110
	3	38 (24.4)	0 (0.0)	0.0 (0.0)	
	4	44 (28.2)	1 (25.0)	2.2 (2.2)	
	5	2 (1.3)	2 (50.0)	50.0 (25.0)	
Criterion 10	2	55 (39.3)	0 (0.0)	0.0 (0.0)	0.081
	3	35 (25.0)	0 (0.0)	0.0 (0.0)	
	4	37 (26.4)	1 (33.3)	2.6 (2.6)	
	5	2 (1.4)	2 (66.7)	50.0 (25.0)	
Criterion 11	2	52 (42.6)	0 (0.0)	0.0 (0.0)	0.046
	3	29 (23.8)	0 (0.0)	0.0 (0.0)	
	4	29 (23.8)	0 (0.0)	0.0 (0.0)	
	5	2 (1.6)	2 (100.0)	50.0 (25.0)	
Criterion 12	2	45 (42.5)	0 (0.0)	0.0 (0.0)	0.046
	3	27 (25.5)	0 (0.0)	0.0 (0.0)	
	4	22 (20.8)	0 (0.0)	0.0 (0.0)	
	5	2 (1.9)	2 (100.0)	50.0 (25.0)	

ACR TIRADS, American College of Radiology Thyroid Imaging Reporting and Data System; SE, standard error.
In all criteria, the estimated malignancy rate for ACR TIRADS category 1 was 0.0% and therefore was excluded from the table.

diagnostic accuracy above 97% and were not significantly different from criterion 1. To satisfy the current TBSRTC standards for sample adequacy, the biopsied sample must meet the minimum requirements for both clusters and cells. From the first edition of the TBSRTC in 2009 to the recent third edition of the TBSRTC in 2023, the nondiagnostic category has not included atypical cells, suggesting that adequacy thresholds can be lowered without noticeably reducing sensitivity.[6,8-10,23]

To the best of our knowledge, no research has used cluster numbers to determine sample adequacy; nevertheless, a few studies have used cell counts to decide whether a sample is sufficient for diagnosis.[1,10,23] When using a threshold of at least 10 cells for sample adequacy in a previous study, specificity was significantly higher, while sensitivity and malignancy rate did not show significant difference compared to when a threshold of 60 cells was applied (specificity, sensitivity, and malignancy rate being 55%, 96%, and 17% for 10 follicular cells vs. 42%, 97%, and 13% for 60 follicular cells, respectively).[10] However, the study did not compare the diagnostic performances of the other criteria to the original standard criteria. When 146 nondiagnostic nodules were reclassified based on criteria for sample adequacy using cell counts ranging from 0 to 60 cells in 10-cell increments, regardless of cell clusters, sensitivity, specificity, and FNR showed minimal changes in nondiagnostic nodules even without any benign follicular cells (sensitivity, specificity, specificity, and FNR being 93%, 58%, and 7.7% according to the original criterion vs. 92%, 60%, and 7.7% with no follicular cells, respectively),[1] compared with the performances seen with the original criterion.[7] Although these studies demonstrated the possibility of lowering the threshold for sample adequacy, only surgically resected nodules were included, which may have introduced a selection bias.[1,10] In our study, we defined new criteria based on cell counts regardless of clusters and those based on the number of cell clusters with more than 10 cells to evaluate potential alternate thresholds for sample adequacy. More than 40 cells, or a minimum of 3 cell groups encompassing more than 10 cells each, showed diagnostic accuracies of 97.7%–97.8% and FNRs of 4.8%–5.0%, which were not significantly dif-

ferent from those seen with the original criterion. When applying a lower threshold for sample adequacy, repeat FNA could be reduced from 257 to 189 (26.5%) when using criterion 4, which was defined with a minimum of three cell groups of more than 10 cells each, and could be reduced from 257 to 160 (37.7%) when using criterion 9, which requires a minimum of 40 cells regardless of cell cluster.

US features can influence the probability of malignancy, even in cases where nodules have the same cytology.[24] In nondiagnostic nodules, malignancy rates were 5.1% for those with high suspicion US features, in contrast to less than 3% for nodules without high suspicion US features based on the 2015 American Thyroid Association Management Guidelines.[2,24] When applying the Kwak TIRADS, nondiagnostic nodules with more than two suspicious US features (marked hypoechogenicity, noncircumscribed margins, microcalcifications or mixed calcifications, and nonparallel shape) showed malignancy rates higher than 7%, compared to nodules with no more than one suspicious US feature, showing a malignancy rate of 3% or less.[11,14] In our study, malignancy rates of nondiagnostic nodules categorized as ACR TIRADS 5 were all above 28.6% and those categorized as ACR TIRADS 1 to 4 were below 5% for all criteria. This suggests that US stratification systems may serve as complementary tools to aid in the risk stratification of nodules with nondiagnostic cytology, potentially assisting decisions regarding follow-up US or repeat FNA.[11,17,24] Given their markedly higher malignancy rates, nondiagnostic nodules classified as ACR TIRADS 5 may warrant closer surveillance or more proactive diagnostic and therapeutic intervention compared to those in lower TIRADS categories. Importantly, a nondiagnostic cytology result alone should not be regarded as an automatic indication for repeat FNA in all cases. Management ought to be individualized, ideally within a multidisciplinary framework that integrates pathology, imaging, and patient risk factors. In this context, US classification could provide significant guidance.

There are some limitations to our study. First, this study was of retrospective design, which may have caused a selection bias. Second, we included nodules with nondiagnostic cytology that did not undergo follow-up US-guided FNA or surgery. While only nodules that remained stable or decreased in size 1 year or more following the initial FNA were classified as benign, variability in follow-up duration and methods may still pose a limitation. A subsequent FNA or surgery, along with longer or more consistent observation, could have revealed changes leading to a different histopathologic outcome. Third, nondiagnostic cytologic slides were reevaluated by a single experienced cytologist. Although the implementation of TBSRTC has shown a moderate-to-good level of interobserver concordance among pathologists of various levels of experience, the lack of interobserver assessment in our study could have introduced variability in diagnostic accuracy, especially for nodules with AUS cytology.[25,26] Fourth, ACR TIRADS categorization was retrospectively evaluated based on prospectively recorded US features by 17 radiologists with varying levels of experience, suggesting that interobserver variability does need to be considered.[27,28] Fifth, all criteria showed significantly higher specificity compared with criterion 1, although the actual differences in values were in the decimals. This is due to limitations in statistical testing with extreme values, and these differences may not have clinical significance in real patient care. Sixth, five cases of noninvasive follicular thyroid neoplasm with papillary-like nuclear features were classified into the benign group for analysis, which may not fully reflect the intermediate nature of these neoplasms in the WHO classification. Finally, the single-institution setting, along with variability in operator technique, experience, and ultrasound guidance, may limit the generalizability of these findings to wider, real-world clinical practice.[3,29] Further multicenter studies are needed to validate whether less stringent thresholds can be safely applied across diverse practice environments, alongside efforts to standardize procedures and explore additional strategies—such as on-site adequacy assessment—to improve overall specimen quality.[30]

In conclusion, less stringent thresholds for sample adequacy can show comparable diagnostic performances to the original criterion of the TBSRTC. Moreover, US stratification systems may still help guide management of nondiagnostic nodules, even with lower thresholds determining sample adequacy.

AUTHOR CONTRIBUTIONS

Conceptualization: Jin Young Kwak, Sangwoo Cho, Syed Z. Ali, and Soon Won Hong. Data curation: Jin Young Kwak, Sangwoo Cho, Jung Hyun Yoon, and Soon Won Hong. Formal analysis: Jin Young Kwak, Sangwoo Cho, and Hye Sun Lee. Funding acquisition: Jin Young Kwak. Investigation: Jin Young Kwak and Sangwoo Cho. Methodology: Jin Young Kwak, Sangwoo Cho, and Hye Sun Lee. Project administration: Jin Young Kwak. Resources: Jin Young Kwak. Software: Sangwoo Cho and Hye Sun Lee. Supervision: Syed Z. Ali and Soon Won Hong. Validation: Sangwoo Cho and Hye Sun Lee. Visualization: Jin Young Kwak and Sangwoo Cho. Writing—original draft: Jin Young Kwak and Sangwoo Cho. Writing—review & editing: Jin Young Kwak and Sangwoo Cho. Approval of final manuscript: all authors.

ORCID iDs

Jin Young Kwak https://orcid.org/0000-0002-6212-1495
Sangwoo Cho https://orcid.org/0000-0002-3840-7725
Hye Sun Lee https://orcid.org/0000-0001-6328-6948
Jung Hyun Yoon https://orcid.org/0000-0002-2100-3513
Syed Z. Ali https://orcid.org/0000-0002-9069-8427
Soon Won Hong https://orcid.org/0000-0002-0324-2414

REFERENCES

1. Vivero M, Renshaw AA, Krane JF. Adequacy criteria for thyroid FNA evaluated by ThinPrep slides only. Cancer Cytopathol 2017;125:534-43.
2. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiat-

https://doi.org/10.3349/ymj.2025.0257

ed thyroid cancer: the American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer. Thyroid 2016;26:1-133.
3. Ghofrani M, Beckman D, Rimm DL. The value of onsite adequacy assessment of thyroid fine-needle aspirations is a function of operator experience. Cancer 2006;108:110-3.
4. Degirmenci B, Haktanir A, Albayrak R, Acar M, Sahin DA, Sahin O, et al. Sonographically guided fine-needle biopsy of thyroid nodules: the effects of nodule characteristics, sampling technique, and needle size on the adequacy of cytological material. Clin Radiol 2007;62:798-803.
5. Alexander EK, Heering JP, Benson CB, Frates MC, Doubilet PM, Cibas ES, et al. Assessment of nondiagnostic ultrasound-guided fine needle aspirations of thyroid nodules. J Clin Endocrinol Metab 2002;87:4924-7.
6. Cibas ES, Ali SZ. The Bethesda system for reporting thyroid cytopathology. Am J Clin Pathol 2009;132:658-65.
7. Goellner JR, Gharib H, Grant CS, Johnson DA. Fine needle aspiration cytology of the thyroid, 1980 to 1986. Acta Cytol 1987;31:587-90.
8. Cibas ES, Ali SZ. The 2017 Bethesda system for reporting thyroid cytopathology. Thyroid 2017;27:1341-6.
9. Ali SZ, Baloch ZW, Cochand-Priollet B, Schmitt FC, Vielh P, VanderLaan PA. The 2023 Bethesda system for reporting thyroid cytopathology. Thyroid 2023;33:1039-44.
10. Renshaw AA. Histologic follow-up of nondiagnostic thyroid fine needle aspirations: implications for adequacy criteria. Diagn Cytopathol 2012;40(Suppl 1):E13-5.
11. Yoon JH, Lee HS, Kim EK, Moon HJ, Kwak JY. Thyroid nodules: nondiagnostic cytologic results according to thyroid imaging reporting and data system before and after application of the Bethesda system. Radiology 2015;276:579-87.
12. Yoon JH, Lee HS, Kim EK, Moon HJ, Park VY, Kwak JY. Follow-up strategies for thyroid nodules with benign cytology on ultrasound-guided fine needle aspiration: malignancy rates of management guidelines using ultrasound before and after the era of the Bethesda system. Thyroid 2019;29:1227-36.
13. Lee YB, Kim JY, Cho H, Hahn SY, Shin JH, Lee SE, et al. Modified Bethesda system informing cytopathologic adequacy improves malignancy risk stratification in nodules considered benign or atypia (follicular lesion) of undetermined significance. Sci Rep 2018;8:13503.
14. Kwak JY, Han KH, Yoon JH, Moon HJ, Son EJ, Park SH, et al. Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk. Radiology 2011;260:892-9.
15. Russ G, Bonnema SJ, Erdogan MF, Durante C, Ngu R, Leenhardt L. European Thyroid Association guidelines for ultrasound malignancy risk stratification of thyroid nodules in adults: the EU-TIRADS. Eur Thyroid J 2017;6:225-37.
16. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, et al. ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS committee. J Am Coll Radiol 2017;14:587-95.
17. Moon HJ, Kim EK, Yoon JH, Kwak JY. Malignancy risk stratification in thyroid nodules with nondiagnostic results at cytologic examination: combination of thyroid imaging reporting and data system and the Bethesda system. Radiology 2015;274:287-95.
18. Cha H, Pyo JY, Hong SW. The usefulness of immunocytochemistry of CD56 in determining malignancy from indeterminate thyroid fine-needle aspiration cytology. J Pathol Transl Med 2018;52:404-10.
19. Richards ML, Bohnenblust E, Sirinek K, Bingener J. Nondiagnostic thyroid fine-needle aspiration biopsies are no longer a dilemma. Am J Surg 2008;196:398-402.
20. Anderson TJ, Atalay MK, Grand DJ, Baird GL, Cronan JJ, Beland MD. Management of nodules with initially nondiagnostic results of thyroid fine-needle aspiration: can we avoid repeat biopsy? Radiology 2014;272:777-84.
21. Haider AS, Rakha EA, Dunkley C, Zaitoun AM. The impact of using defined criteria for adequacy of fine needle aspiration cytology of the thyroid in routine practice. Diagn Cytopathol 2011;39:81-6.
22. Yeh MW, Demircan O, Ituarte P, Clark OH. False-negative fine-needle aspiration cytology results delay treatment and adversely affect outcome in patients with thyroid carcinoma. Thyroid 2004;14:207-15.
23. Renshaw AA. Evidence-based criteria for adequacy in thyroid fine-needle aspiration. Am J Clin Pathol 2002;118:518-21.
24. Park CJ, Kim EK, Moon HJ, Yoon JH, Park VY, Kwak JY. Thyroid nodules with nondiagnostic cytologic results: follow-up management using ultrasound patterns based on the 2015 American Thyroid Association guidelines. AJR Am J Roentgenol 2018;210:412-7.
25. Pathak P, Srivastava R, Singh N, Arora VK, Bhatia A. Implementation of the Bethesda system for reporting thyroid cytopathology: interobserver concordance and reclassification of previously inconclusive aspirates. Diagn Cytopathol 2014;42:944-9.
26. Sripodok S, Benjakul N. Interobserver variability in inconclusive diagnostic categories of thyroid fine needle aspiration cytology: an urban-based tertiary hospital experience. Ann Diagn Pathol 2023;63:152083.
27. Hoang JK, Middleton WD, Farjat AE, Teefey SA, Abinanti N, Boschini FJ, et al. Interobserver variability of sonographic features used in the American College of Radiology thyroid imaging reporting and data system. AJR Am J Roentgenol 2018;211:162-7.
28. Park CS, Kim SH, Jung SL, Kang BJ, Kim JY, Choi JJ, et al. Observer variability in the sonographic evaluation of thyroid nodules. J Clin Ultrasound 2010;38:287-93.
29. Gunes P, Canberk S, Onenerk M, Erkan M, Gursan N, Kilinc E, et al. A different perspective on evaluating the malignancy rate of the non-diagnostic category of the Bethesda system for reporting thyroid cytopathology: a single institute experience and review of the literature. PLoS One 2016;11:e0162745.
30. Issa PP, McCarthy C, Hussein M, Albuck AL, Emad E, Shama M, et al. Assessing adequacy: a meta-analysis of rapid onsite evaluation of thyroid nodules. J Surg Res 2024;296:523-31.