Comparison of the diagnostic performance of the modified Korean Thyroid Imaging Reporting and Data System for thyroid malignancy with three international guidelines

Eun Ju Ha[1], Jung Hee Shin[2], Dong Gyu Na[3], So Lyung Jung[4], Young Hen Lee[5],
Wooyul Paik[3], Min Ji Hong[6], Yeo Koon Kim[7], Chang Yoon Lee[8]

*Author affiliations appear at the end of this article.

Purpose: This study compared the diagnostic performance of the modified Korean Thyroid Imaging Reporting and Data System (K-TIRADS) for thyroid malignancy with three international guidelines.

Methods: From June to September 2015, 5,708 thyroid nodules (≥1.0 cm) in 5,081 consecutive patients who underwent thyroid ultrasound (US) at 26 institutions were evaluated. The US features of the thyroid nodules were retrospectively reviewed and classified according to all four guidelines. In the modified K-TIRADS, the biopsy size threshold was changed to 2.0 cm for K-TIRADS 3 and 1.0 or 1.5 cm for K-TIRADS 4 (K-TIRADS$_{1.0cm}$ and K-TIRADS$_{1.5cm}$, respectively). We compared the diagnostic performance and unnecessary fine-needle aspiration biopsy (FNAB) rates for thyroid malignancy between the modified K-TIRADS and three international guidelines.

Results: Of the 5,708 thyroid nodules, 4,597 (80.5%) were benign and 1,111 (19.5%) were malignant. The overall sensitivity was highest for the modified K-TIRADS$_{1.0cm}$ (91.0%), followed by the European (EU)-TIRADS (84.6%), American Association of Clinical Endocrinologists/American College of Endocrinology/Associazione Medici Endocrinologi (AACE/ACE/AME) (80.5%), American College of Radiology (ACR)-TIRADS (76.1%), and modified K-TIRADS$_{1.5cm}$ (76.1%). For large nodules (>2.0 cm), the sensitivity increased to 98.0% in both the modified K-TIRADS$_{1.0cm}$ and K-TIRADS$_{1.5cm}$. For small nodules (≤2.0 cm), the unnecessary FNAB rate was lowest with the modified K-TIRADS$_{1.5cm}$ (17.6%), followed by the ACR-TIRADS (18.6%), AACE/ACE/AME (19.3%), EU-TIRADS (28.1%), and modified K-TIRADS$_{1.0cm}$ (31.2%).

Conclusion: The modified K-TIRADS$_{1.5cm}$ can reduce the unnecessary FNAB rate for small nodules (1.0–2.0 cm), while maintaining high sensitivity for detecting malignancies >2.0 cm.

Keywords: Fine-needle aspiration; Thyroid neoplasms; Thyroid Imaging Reporting and Data System; Thyroid nodules; Ultrasonography

Introduction

Ultrasonography (US) is an essential tool for assessing the risk of malignancy of thyroid nodules and deciding whether to perform fine-needle aspiration biopsy (FNAB) [1,2]. Therefore, many international societies have proposed US-based risk stratification systems (RSS) for the clinical management of thyroid nodules [1–5]. The Korean Thyroid Imaging Reporting and Data System (K-TIRADS) is a US-based RSS proposed by the Korean Society of Thyroid Radiology (KSThR) in 2016 [1]. Since its publication, many studies have compared the diagnostic performance of the K-TIRADS for thyroid malignancy to other RSS, and have reported its clinical usefulness [6–12]. As it allows intuitive and easy assessment of the US pattern of nodules, the K-TIRADS has been applied widely in clinical practice to classify nodules during real-time US examinations.

However, recent comparative studies of RSS reported that the US-based FNAB criteria of the K-TIRADS for nodules ≥1 cm had the highest sensitivity (91.7%–100%) and lowest specificity (15.4%–28.7%) for thyroid cancers, and the highest unnecessary FNAB rate for benign nodules (71.3%–84.6%) [6–9,12–15]. The diagnostic performance was similar among the different RSS at the same biopsy size threshold for the classification categories in simulation studies [8,15]. The need to modify the biopsy criteria of the K-TIRADS to reduce the unnecessary FNAB rate while maintaining high sensitivity for detecting thyroid cancer has been suggested [8,9,12,15].

Therefore, this study was performed to compare the diagnostic performance of the US-based FNAB criteria of the modified K-TIRADS for thyroid malignancy with the RSS of three international societies: the American College of Radiology (ACR)-TIRADS, European (EU)-TIRADS, and American Association of Clinical Endocrinologists/American College of Endocrinology/Associazione Medici Endocrinologi (AACE/ACE/AME).

Materials and Methods

Compliance with Ethical Standards
The institutional review boards of the 26 participating centers approved this study (AJIRB-MED-MDB-21-062). The requirement for informed consent was waived for this retrospective review.

Study Population
Patient data collected by 26 different hospitals in Korea (Thyroid Imaging Network of Korea registry, THINK) were analyzed. From June to September 2015, 22,775 consecutive patients underwent thyroid US at the 26 participating institutions. Among them, 16,679 were excluded due to the lack of a reference standard test (biopsy or surgery) (n=4,304), thyroid nodules <1.0 cm (n=12,130), or

suboptimal image quality (n=245). A further 1,015 patients with 1,102 nodules were excluded due to inconclusive biopsy results. A total of 5,708 thyroid nodules (≥1 cm) in 5,081 consecutive patients were finally included in this study (4,176 women, 905 men; mean age, 53.2 years; age range, 19 to 93 years) (Fig. 1).

Malignant nodules were diagnosed based on the histopathological results after surgery (n=947) or malignant fine-needle aspiration (FNA) or core-needle biopsy (CNB) results (n=164). Benign nodules were diagnosed based on the histopathological results after surgery (n=394), at least two benign FNA or CNB results (n=601), and one benign FNA or CNB result (n=3,602) [16–18].

US Examination and Image Analysis
All US examinations were performed with a 10–12 or 5–14 MHz linear probe. The US images were retrospectively reviewed by 1 of 17 experienced radiologists with 8–22 years of experience in performing thyroid US using an online program (AIM AiCRO, https://study.aim-aicro.com).

Prior to the commencement of this multicenter study, training sessions were held to establish a baseline consensus regarding the US criteria. The 17 radiologists evaluated images of 15 biopsy-proven masses not included in the study, and were asked to assess the US criteria during a consensus meeting, including composition, echogenicity, margin, calcification, orientation (taller-than-wide), spongiform appearance, and intracystic echogenic foci with a comet-tail artifact. In the modified K-TIRADS, nonparallel orientation was defined as the anteroposterior diameter of a nodule being longer than its transverse diameter in the transverse plane. The benign category included spongiform or partially cystic nodules with intracystic echogenic foci with comet-tail artifact regardless of suspicious US features. All of the reviewers, who were blind to the FNAB results and final diagnoses, then assessed the US features of the thyroid nodules [3–5]. However, extrathyroidal extension status was not evaluated in this study due to a lack of standardized US criteria. An isolated macrocalcification was defined as an entirely calcified nodule with posterior acoustic shadowing, in which no soft tissue component was identified due to dark shadows on the US image [19–21]. Such nodules were classified as "intermediate suspicion" based on the modified K-TIRADS, "moderate suspicion" (4 points) based on the ACR-TIRADS, and "unclassified" based on the other RSS [1,3–5].

Biopsy Size Thresholds for the Modified K-TIRADS and Other RSS
The size threshold for biopsy was increased from 1.5 cm to 2.0 cm for low-suspicion (K-TIRADS 3) nodules in the modified K-TIRADS. The size threshold for biopsy was subdivided into 1.0 and 1.5 cm
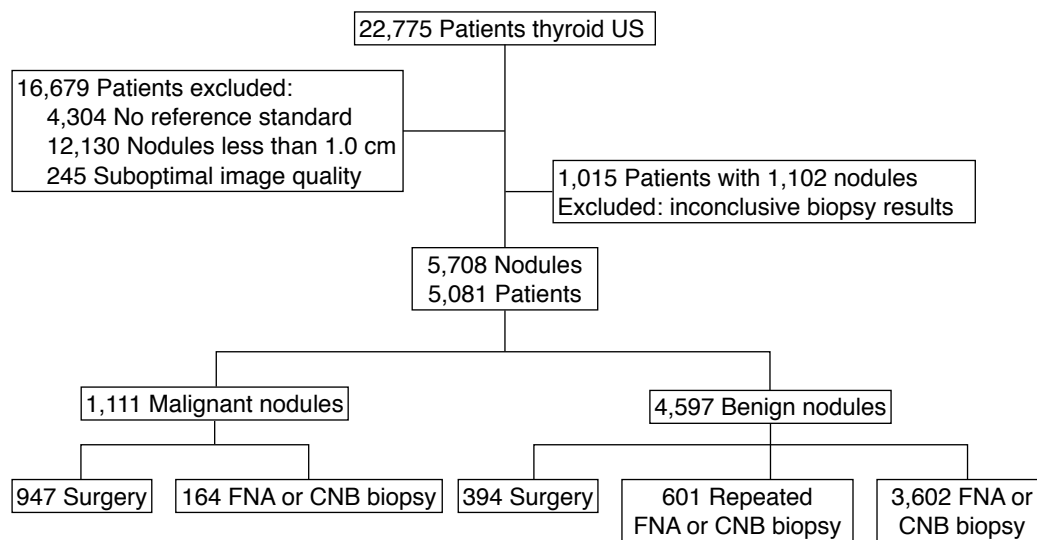
Fig. 1. **Flowchart showing the study population.** US, ultrasonography; FNA, fine-needle aspiration; CNB, core-needle biopsy.

for K-TIRADS 4 (K-TIRADS$_{1.0cm}$ and K-TIRADS$_{1.5cm}$, respectively). The biopsy size thresholds in the RSS are shown in the Supplementary Table 1. Nodules classified as low-risk based on the AACE/ACE/AME guidelines, not suspicious (TR2) or benign (TR1) based on the ACR-TIRADS, or benign based on the EU-TIRADS and modified K-TIRADS were not indicated for biopsy in this study.

Statistical Analysis

The nodules were classified according to the categories defined by the different RSS, by one radiologist (D.G.N.) with 22 years of experience in performing thyroid US and interventional procedures (Supplementary Table 1). The malignancy risk for each RSS category was calculated as a percentage. The associations between all categories and the final diagnoses were evaluated using the linear-by-linear association test.

Thyroid nodules were classified into two groups: nodules in which US-guided FNAB was not indicated and nodules in which US-guided FNAB was indicated based on the criteria of each RSS (i.e., US features and nodule size). The performance of the US-based FNAB criteria of each RSS for diagnosing thyroid cancer was evaluated based on the sensitivity, specificity, positive predictive value, negative predictive value, and accuracy, and compared using the generalized estimating equation (GEE) method.

The potentially unnecessary FNAB rate was calculated according to the number of benign nodules among FNAB-required nodules in the total of thyroid nodules, and the number of benign nodules among FNAB-required nodules in the total of benign thyroid nodules. These values were calculated as percentages, and the results were compared among the guidelines using the GEE method. Statistical analyses were performed using SPSS for Windows version 23.0) (IBM Corp., Armonk, NY, USA) and SAS for Windows version 9.2

(SAS Institute Inc., Cary, NC, USA). In all analyses, P<0.05 was taken to indicate statistical significance.

Results

Of the 5,708 thyroid nodules, 4,597 (80.5%) were diagnosed as benign and 1,111 (19.5%) as malignant. The 1,111 malignant nodules included 1,011 papillary thyroid carcinomas (91.0%) and 62 follicular carcinomas (5.6%). The mean size of the nodules was 20.7±10.8 mm (range, 10 to 100 mm). There were 3,576 (62.6%) small nodules (1–2 cm) and 2,132 (37.4%) large (>2 cm) nodules. The malignancy rate was significantly higher in small tumors (1–2 cm) than in large tumors (>2 cm) (22.7% [810/3,576] vs. 14.1% [301/2,132], respectively; P<0.001).

Malignancy Risk by Category in the Different RSS

Table 1 lists the malignancy risk for each category in the different RSS. The malignancy probability was higher in the higher risk categories of all guidelines (P<0.001). Overall, 1.0% (59 of 5,708) of nodules were unclassified according to the AACE/ACE/AME and EU-TIRADS; however, 33.9% (20 of 59) of those were actually malignant.

Diagnostic Performance of the Modified K-TIRADS and Other RSS

Table 2 shows the diagnostic performance of the different RSS for thyroid cancer. The sensitivity and negative predictive value were highest for the modified K-TIRADS$_{1.0cm}$ (91.0% and 94.8%, respectively), followed by the EU-TIRADS (84.6% and 91.4%, respectively), AACE/ACE/AME (80.5% and 91.1%, respectively), ACR-TIRADS (76.1% and 91.4%, respectively), and K-TIRADS$_{1.5cm}$

Table 1. Frequency and malignancy risk of tumors according to the different guidelines

Guideline	Benign nodules (n=4,597)	Malignant nodules (n=1,111)	Total, n (%) (n=5,708)	Estimated risk of malignancy (%)	Actual rate of malignancy (%)	P-value
ACR-TIRADS 2017						<0.001
Highly suspicious	388	609	997 (17.5)	>20	61.1	
Moderately suspicious	1,433	350	1,783 (31.2)	5−20	19.6	
Mildly suspicious	1,756	114	1,870 (32.8)	5	6.1	
Not suspicious	795	31	826 (14.5)	<2	3.8	
Benign	225	7	232 (4.1)	<2	3.0	
AACE/ACE/AME 2015						<0.001
High-risk	850	735	1,585 (27.8)	50−90	46.4	
Intermediate-risk	3,439	349	3,788 (66.4)	5−15	9.2	
Low-risk	269	7	276 (4.8)	1	2.5	
Unclassified	39	20	59 (1.0)	−	33.9	
EU-TIRADS 2017						<0.001
High-risk	858	736	1,594 (27.9)	26−87	46.2	
Intermediate-risk	1,335	231	1,466 (25.7)	6−27	14.8	
Low-risk	2,279	122	2,401 (42.1)	2−4	5.1	
Benign	86	2	88 (1.5)	0	2.3	
Unclassified	39	20	59 (1.0)	−	33.9	
Modified K-TIRADS						<0.001
High suspicion	244	536	780 (13.7)	>60	68.7	
Intermediate suspicion	1,180	362	1,542 (27.0)	15−50	23.5	
Low suspicion	2,864	204	3,068 (53.7)	3−15	6.6	
Benign	309	9	318 (5.6)	<3	2.8	

ACR, American College of Radiology; TIRADS, Thyroid Imaging Reporting and Data System; AAACE/ACE/AME, American Association of Clinical Endocrinologists/American College of Endocrinology/Associazione Medici Endocrinologi; EU, European; K-TIRADS, Korean Thyroid Imaging Reporting and Data System.

(76.1% and 89.7%, respectively).

The modified K-TIRADS$_{1.0cm}$ had significantly higher diagnostic sensitivity than the AACE/ACE/AME, ACR-TIRADS, and EU-TIRADS (all P<0.001), while it showed lower specificity than the AACE/ACE/AME and ACR-TIRADS (both P<0.001) and similar specificity to the EU-TIRADS (P=0.492). The modified K-TIRADS$_{1.5cm}$ had lower diagnostic sensitivity than the AACE/ACE/AME and EU-TIRADS (both P<0.001), and comparable sensitivity to the ACR-TIRADS (P=0.923). It also showed higher specificity than the AACE/ACE/AME and EU-TIRADS (both P<0.001).

When applying the FNAB criteria to large nodules (>2.0 cm), the sensitivity and negative predictive values increased to 89.7%–99.7% and 96.0%–97.8%, respectively. Both the modified K-TIRADS$_{1.0cm}$ and K-TIRADS$_{1.5cm}$ had high diagnostic sensitivity (98.0%) and negative predictive values (96.6%), respectively. The modified K-TIRADS$_{1.5cm}$ had higher diagnostic sensitivity but lower specificity than the ACR-TIRADS (98.3% and 89.7%, respectively, P<0.001; and 9.4% and 40.4%, respectively, P<0.001).

Unnecessary FNAB Rates for the Modified K-TIRADS and Other RSS

Among the RSS, the unnecessary FNAB rates for all nodules and benign nodules were lowest with the ACR-TIRADS (30.8% and 38.2%, respectively) followed by the modified K-TIRADS$_{1.5cm}$ (40.1% and 49.8%, respectively), AACE/ACE/AME (41.5% and 51.5%, respectively), modified K-TIRADS$_{1.0cm}$ (48.6% and 60.3%, respectively), and EU-TIRADS (48.9% and 60.7%, respectively) (Table 3). The modified K-TIRADS$_{1.5cm}$ had significantly lower unnecessary FNAB rates for all nodules than the AACE/ACE/AME, EU-TIRADS, and modified K-TRADS$_{1.0cm}$ (all P<0.001), but higher unnecessary FNAB rates than the ACR-TIRADS (P<0.001).

When applying the FNAB criteria to small nodules (≤2.0 cm), the unnecessary FNAB rate was lowest with the modified K-TIRADS$_{1.5cm}$ (17.6% and 22.8%, respectively), followed by the ACR-TIRADS (18.6% and 24.0%, respectively), AACE/ACE/AME (19.3% and 24.9%, respectively), EU-TIRADS (28.1% and 36.3%, respectively), and modified K-TIRADS$_{1.0cm}$ (31.2% and 40.3%, respectively). The

Table 2. Diagnostic performance of the various guidelines for detecting malignant thyroid nodules

Guideline	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Accuracy (%)
All nodules (n=5,708)					
ACR-TIRADS	76.1 (845/1,111) [73.5−78.5]	61.8 (2,841/4,597) [60.4−63.2]	32.5 (845/2,601) [30.7−34.3]	91.4 (2,841/3,107) [90.4−92.4]	64.6 (3,686/5,708) [63.3−65.8]
AACE/ACE/AME	80.5 (894/1,111) [78.0−82.7]	48.5 (2,229/4,597) [47.0−49.9]	27.4 (894/3,262) [25.9−29.0]	91.1 (2,229/2,446) [89.9−92.2]	54.7 (3,123/5,708) [53.4−56.0]
EU-TIRADS	84.6 (940/1,111) [82.4−86.6]	39.3 (1,807/4,597) [37.9−40.7]	25.2 (940/3,730) [23.8−26.6]	91.4 (1,807/1,978) [90.0−92.5]	48.1 (2,747/5,708) [46.8−49.4]
Modified K-TIRADS$_{1.0cm}$	91.0 (1,011/1,111) [89.2−92.5]	39.7 (1,824/4,597) [38.3−41.1]	26.7 (1,011/3,784) [25.3−28.2]	94.8 (1,824/1,924) [93.7−95.7]	49.7 (2,835/5,708) [48.4−51.0]
Modified K-TIRADS$_{1.5cm}$	76.1 (846/1,111) [73.6−78.6]	50.2 (2,307/4,597) [48.7−51.6]	27.0 (846/3,136) [25.5−28.6]	89.7 (2,307/2,572) [88.5−90.8]	55.2 (3,153/5,708) [53.9−56.5]
Nodules >2.0 cm (n=2,132)					
ACR-TIRADS	89.7 (270/301) [85.7−92.7]	40.4 (739/1,831) [38.1−42.6]	19.8 (270/1,362) [17.8−22.0]	96.0 (739/770) [94.3−97.2]	47.3 (1,009/2,132) [45.2−49.4]
AACE/ACE/AME	98.3 (296/301) [96.1−99.3]	8.3 (152/1,831) [7.1−9.7]	15.0 (296/1,975) [13.5−16.6]	96.8 (152/157) [92.6−98.7]	21.0 (448/2,132) [19.3−22.8]
EU-TIRADS	99.7 (300/301) [97.7−100.0]	2.5 (45/1,831) [1.8−3.3]	14.4 (300/2,086) [12.9−16.0]	97.8 (45/46) [86.1−99.7]	16.2 (345/2,132) [14.7−17.8]
Modified K-TIRADS$_{1.0cm}$	98.0 (295/301) [95.6−99.1]	9.4 (172/1,831) [8.1−10.8]	15.1 (296/1,969) [13.6−16.8]	96.6 (158/163) [92.7−98.5]	21.9 (467/2,132) [20.2−23.7]
Modified K-TIRADS$_{1.5cm}$	98.0 (295/301) [95.6−99.1]	9.4 (172/1,831) [8.1−10.8]	15.1 (295/1,954) [13.6−16.8]	96.6 (172/178) [92.7−98.5]	21.9 (467/2,132) [20.2−23.7]
Nodules ≤2.0 cm (n=3,576)					
ACR-TIRADS	71.0 (575/810) [67.8−74.0]	76.0 (2,102/2,766) [74.4−77.5]	46.4 (575/1,239) [43.6−49.2]	89.9 (2,102/2,337) [88.7−91.1]	74.9 (2,677/3,576) [73.4−76.3]
AACE/ACE/AME	73.8 (598/810) [70.7−76.7]	75.1 (2,077/2,766) [73.4−76.7]	46.5 (598/1,287) [43.8−49.2]	90.7 (2,077/2,289) [89.5−91.9]	74.8 (2,675/3,576) [73.4−76.2]
EU-TIRADS	79.0 (640/810) [76.1−81.7]	63.7 (1,762/2,766) [61.9−65.5]	38.9 (640/1,644) [36.6−41.3]	91.2 (1,762/1,932) [89.9−92.4]	67.2 (2,402/3,576) [65.5−68.7]
Modified K-TIRADS$_{1.0cm}$	88.4 (716/810) [86.0−90.4]	59.7 (1,652/2,766) [57.9−61.5]	39.1 (716/1,830) [36.9−41.4]	94.6 (1,652/1,746) [93.5−95.6]	66.2 (2,368/3,576) [64.7−67.8]
Modified K-TIRADS$_{1.5cm}$	68.0 (551/810) [64.7−71.1]	77.2 (2,135/2,766) [75.6−78.7]	46.6 (551/1,182) [43.8−49.5]	89.2 (2,135/2,394) [87.9−90.4]	75.1 (2,686/3,576) [73.7−76.5]

Values in parentheses are the raw data used to calculate the percentages, and values in brackets are 95% confidence intervals.
ACR, American College of Radiology; TIRADS, Thyroid Imaging Reporting and Data System; AAACE/ACE/AME, American Association of Clinical Endocrinologists/American College of Endocrinology/Associazione Medici Endocrinologi; EU, European; K-TIRADS, Korean Thyroid Imaging Reporting and Data System.

modified K-TIRADS$_{1.5cm}$ had significantly lower unnecessary FNAB rates than all the other RSS (all P<0.05).

Discussion

Our study demonstrated the clinical utility of the modified K-TIRADS. The modified K-TIRADS$_{1.5cm}$ significantly reduced the unnecessary FNAB rates compared to the AACE/ACE/AME and EU-TIRADS in all nodules, while maintaining high sensitivity for detecting malignancy in >2.0 cm nodules (up to 98.0%). Compared to the ACR-TIRADS, it had a significantly lower unnecessary FNAB rate for small nodules (≤2.0 cm), but higher diagnostic sensitivity for detecting malignancy in nodules >2.0 cm. Although the modified K-TIRADS$_{1.0cm}$ increased

the diagnostic sensitivity for detecting small malignancies, it should be used only after careful consideration of the risks and benefits to the individual patient.

US-based RSS has an essential role in predicting the malignancy risk of thyroid nodules and the decision to perform FNAB [1,2]. However, the sensitivity and specificity of US-based RSS for detecting thyroid cancers remain controversial. As most patients with low-risk small thyroid cancers have an excellent prognosis and low likelihood of mortality, it is also important to reduce overdiagnosis and the harm caused by overtreatment [22]. US-based RSS is expected to reduce the unnecessary FNAB rate while also maintaining appropriate sensitivity for detecting thyroid cancer. Many studies have suggested that increasing sensitivity should be the focus,

Table 3. Unnecessary FNA rates for the diagnosis of thyroid cancer according to the various guidelines

Guideline	No. of nodules indicated for FNAB	No. of malignant nodules among all FNAB nodules	No. of benign nodules among all FNAB nodules (A)	Unnecessary FNAB rate (A/all nodules)	Unnecessary FNAB rate (A/benign nodules)
All nodules (n=5,708)					
ACR-TIRADS	2,601	845 (32.5)	1,756 (67.5)	30.8 [29.6−32.0]	38.2 [36.8−39.6]
AACE/ACE/AME	3,262	894 (27.4)	2,368 (72.6)	41.5 [40.2−42.8]	51.5 [50.1−53.0]
EU-TIRADS	3,730	940 (25.2)	2,790 (74.8)	48.9 [47.6−50.2]	60.7 [59.3−62.1]
Modified K-TIRADS$_{1.0cm}$	3,784	1,011 (26.7)	2,773 (73.3)	48.6 [47.3−49.9]	60.3 [58.9−61.7]
Modified K-TIRADS$_{1.5cm}$	3,136	846 (27.0)	2,290 (73.0)	40.1 [38.9−41.4]	49.8 [48.4−51.3]
Nodules >2.0 cm (n=2,132)					
ACR-TIRADS	1,362	270 (19.8)	1,092 (80.2)	51.2 [49.1−53.3]	59.6 [57.4−61.9]
AACE/ACE/AME	1,975	296 (15.0)	1,679 (85.0)	78.8 [77.0−80.4]	91.7 [90.3−92.9]
EU-TIRADS	2,086	300 (14.4)	1,786 (85.6)	83.8 [82.1−85.3]	97.5 [96.7−98.2]
Modified K-TIRADS$_{1.0cm}$	1,954	295 (15.1)	1,659 (84.9)	77.8 [76.0−79.5]	90.6 [89.2−91.9]
Modified K-TIRADS$_{1.5cm}$	1,954	295 (15.1)	1,659 (84.9)	77.8 [76.0−79.5]	90.6 [89.2−91.9]
Nodules ≤2.0 cm (n=3,576)					
ACR-TIRADS	1,239	575 (46.4)	664 (53.6)	18.6 [17.3−19.9]	24.0 [22.5−25.6]
AACE/ACE/AME	1,287	598 (46.5)	689 (53.5)	19.3 [18.0−20.6]	24.9 [23.3−26.6]
EU-TIRADS	1,644	640 (38.9)	1,004 (61.1)	28.1 [26.6−29.6]	36.3 [34.5−38.1]
Modified K-TIRADS$_{1.0cm}$	1,830	716 (39.1)	1,114 (60.9)	31.2 [29.7−32.7]	40.3 [38.5−42.1]
Modified K-TIRADS$_{1.5cm}$	1,182	551 (46.6)	631 (53.4)	17.6 [16.4−18.9]	22.8 [21.3−24.4]

Values are presented as number (%) and values in brackets are 95% confidence intervals.
FNA, fine-needle aspiration; FNAB, fine-needle aspiration biopsy; ACR, American College of Radiology; TIRADS, Thyroid Imaging Reporting and Data System; AAACE/ACE/AME, American Association of Clinical Endocrinologists/American College of Endocrinology/Associazione Medici Endocrinologi; EU, European; K-TIRADS, Korean Thyroid Imaging Reporting and Data System.

regardless of the rate of unnecessary FNAB of large nodules (>2 cm), considering the higher risk of aggressive behavior in large cancers. On the other hand, reducing unnecessary FNAB for small nodules (1−2 cm) is important, given the slow growth rate of small cancers [9]. In simulation studies, when the biopsy size threshold for K-TIRADS 3 increased from 1.5 to 2 cm, there was a decrease (14.1%−16.2%) in the unnecessary FNAB rate of benign nodules, and a minimal decrease (1%−1.9%) in the sensitivity (91.6%− 94.7%) for thyroid cancer, compared to the original K-TIRADS (sensitivity, 93.5%−95.7%) [8,15]. When the biopsy size threshold was increased from 1.0 to 1.5 cm for K-TIRADS 4, there was a substantial decrease (27%−28.9%) in the unnecessary FNAB rate of benign nodules and a further significant decrease (15%−27.1%) in the sensitivity (66.4%−80.7%) for thyroid cancer [8,15]. Based on these results, it was planned to modify the K-TIRADS biopsy size threshold to 2.0 cm for K-TIRADS 3 and 1.0 or 1.5 cm for K-TIRADS 4 in this study.

The modified K-TIRADS$_{1.5cm}$ reduced the unnecessary FNAB rate to 17.6% in small nodules (≤2.0 cm). This was significantly lower than the rate for the AACE/ACE/AME, EU-TIRADS, and ACR-TIRADS (18.6%−28.1%). Moreover, it maintained high sensitivity (up to

98.0%) for detecting large cancers (>2.0 cm). Compared to the ACR-TIRADS, the modified K-TIRADS$_{1.5cm}$ had a higher unnecessary FNAB rate (40.1% and 30.8%, respectively) for all nodules, a lower unnecessary FNAB rate (17.6% and 18.6%, respectively) for small nodules (≤2.0 cm), and significantly higher sensitivity for detecting large cancers (98.0% and 89.7%, respectively). The modified K-TIRADS$_{1.0cm}$ had a diagnostic sensitivity up to 88.4% for small cancers, but the unnecessary FNAB rate was significantly higher compared to the ACR-TIRADS and AACD/ACE/AME guidelines. Therefore, it is suggested that the modified K-TIRADS$_{1.0cm}$ be applied only after careful consideration of clinical factors, family history, and US features.

This study had several limitations. First, it only included thyroid nodules that had undergone US-guided FNAB, which is usually performed in thyroid nodules with suspicious US features, or in the largest nodule if no suspicious US features are detected; this may have led to selection bias. Second, the final diagnoses of benign nodules were based on the cytopathological results and surgical histological findings, which may have resulted in false-negative or false-positive results. Third, this study included a large proportion of papillary thyroid carcinomas. If the proportion of follicular cancers

increase, the diagnostic performance may be changed.

In conclusion, the modified K-TIRADS$_{1.5cm}$ reduced the unnecessary FNAB rates in all nodules compared to the AACE/ACE/AME and EU-TIRADS, while maintaining high sensitivity for detecting malignancies >2.0 cm. In particular, it showed higher sensitivity for detecting malignancies >2.0 cm, but a lower unnecessary FNAB rate for small nodules (1.0–2.0 cm) than the ACR-TIRADS.

ORCID: Eun Ju Ha: https://orcid.org/0000-0002-1234-2919; Jung Hee Shin: https://orcid.org/0000-0001-6435-7357; Dong Gyu Na: https://orcid.org/0000-0001-6422-1652; So Lyung Jung: https://orcid.org/0000-0002-3267-8399; Young Hen Lee: https://orcid.org/0000-0001-7739-4173; Wooyul Paik: https://orcid.org/0000-0001-9617-6227; Min Ji Hong: https://orcid.org/0000-0001-7818-4346; Yeo Koon Kim: https://orcid.org/0000-0002-8729-9484; Chang Yoon Lee: https://orcid.org/0000-0001-8645-3569

*Author affiliations
[1]Department of Radiology, Ajou University School of Medicine, Suwon; [2]Department of Radiology and Center for Imaging Science, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul; [3]Department of Radiology, GangNeung Asan Hospital, University of Ulsan College of Medicine, Gangneung; [4]Department of Radiology, Yeouido St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul; [5]Department of Radiology, Korea University Ansan Hospital, Korea University College of Medicine, Ansan; [6]Department of Radiology, Chung-Ang University Hospital, Chung-Ang University College of Medicine, Seoul; [7]Department of Radiology, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam; [8]Department of Radiology, Research Institute and Hospital, National Cancer Center, Goyang, Korea

Supplementary Material
Supplementary Table 1. Summary of US features and US-guided FNAB recommendations (https://doi.org/10.14366/usg.21056).

References

1. Shin JH, Baek JH, Chung J, Ha EJ, Kim JH, Lee YH, et al. Ultrasonography diagnosis and imaging-based management of thyroid nodules: revised Korean Society of Thyroid Radiology consensus statement and recommendations. Korean J Radiol 2016;17:370-395.

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 differentiated thyroid cancer: the American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid 2016;26:1-133.

3. Gharib H, Papini E, Garber JR, Duick DS, Harrell RM, Hegedus L, et al. American Association of Clinical Endocrinologists, American College of Endocrinology, and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules: 2016 update. Endocr Pract 2016;22:622-639.

4. 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;6:225-237.

5. 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;14:587-595.

6. Ha EJ, Na DG, Baek JH, Sung JY, Kim JH, Kang SY. US Fine-needle aspiration biopsy for thyroid malignancy: diagnostic performance of seven society guidelines applied to 2000 thyroid nodules. Radiology 2018;287:893-900.

7. Ha EJ, Na DG, Moon WJ, Lee YH, Choi N. Diagnostic performance of ultrasound-based risk-stratification systems for thyroid nodules: comparison of the 2015 American Thyroid Association Guidelines with the 2016 Korean Thyroid Association/Korean Society of Thyroid Radiology and 2017 American College of Radiology guidelines. Thyroid 2018;28:1532-1537.

8. Ha SM, Baek JH, Na DG, Suh CH, Chung SR, Choi YJ, et al. Diagnostic performance of practice guidelines for thyroid nodules: thyroid nodule size versus biopsy rates. Radiology 2019;291:92-99.

9. Na DG, Paik W, Cha J, Gwon HY, Kim SY, Yoo RE. Diagnostic performance of the modified Korean Thyroid Imaging Reporting and Data System for thyroid malignancy according to nodule size: a comparison with five society guidelines. Ultrasonography 2021;40:474-485.

10. Kim DH, Chung SR, Choi SH, Kim KW. Accuracy of thyroid imaging reporting and data system category 4 or 5 for diagnosing malignancy: a systematic review and meta-analysis. Eur Radiol 2020;30:5611-5624.

11. Kim PH, Suh CH, Baek JH, Chung SR, Choi YJ, Lee JH. Diagnostic performance of four ultrasound risk stratification systems: a systematic review and meta-analysis. Thyroid 2020;30:1159-1168.

12. Yim Y, Na DG, Ha EJ, Baek JH, Sung JY, Kim JH, et al. Concordance of three international guidelines for thyroid nodules classified by ultrasonography and diagnostic performance of biopsy criteria. Korean J Radiol 2020;21:108-116.

13. Middleton WD, Teefey SA, Reading CC, Langer JE, Beland MD, Szabunio MM, et al. Comparison of performance characteristics of American College of Radiology TI-RADS, Korean Society of Thyroid Radiology TIRADS, and American Thyroid Association guidelines. AJR Am J Roentgenol 2018;210:1148-1154.

14. Grani G, Lamartina L, Ascoli V, Bosco D, Biffoni M, Giacomelli L, et al. Reducing the number of unnecessary thyroid biopsies while improving diagnostic accuracy: toward the "Right" TIRADS. J Clin Endocrinol Metab 2019;104:95-102.

15. Yoon SJ, Na DG, Gwon HY, Paik W, Kim WJ, Song JS, et al. Similarities and differences between Thyroid Imaging Reporting and Data Systems. AJR Am J Roentgenol 2019;213:W76-W84.

16. Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. Thyroid 2017;27:1341-1346.

17. Jung CK, Baek JH, Na DG, Oh YL, Yi KH, Kang HC. 2019 Practice guidelines for thyroid core needle biopsy: a report of the Clinical Practice Guidelines Development Committee of the Korean Thyroid Association. J Pathol Transl Med 2020;54:64-86.

18. Na DG, Baek JH, Jung SL, Kim JH, Sung JY, Kim KS, et al. Core needle biopsy of the thyroid: 2016 consensus statement and recommendations from Korean Society of Thyroid Radiology. Korean J Radiol 2017;18:217-237.

19. Gwon HY, Na DG, Noh BJ, Paik W, Yoon SJ, Choi SJ, et al. Thyroid nodules with isolated macrocalcifications: malignancy risk of isolated macrocalcifications and postoperative risk stratification of malignant tumors manifesting as isolated macrocalcifications. Korean J Radiol 2020;21:605-613.

20. Na DG, Kim DS, Kim SJ, Ryoo JW, Jung SL. Thyroid nodules with isolated macrocalcification: malignancy risk and diagnostic efficacy of fine-needle aspiration and core needle biopsy. Ultrasonography 2016;35:212-219.

21. Paik W, Na DG, Gwon HY, Kim J. CT features of thyroid nodules with isolated macrocalcifications detected by ultrasonography. Ultrasonography 2020;39:130-136.

22. Ito Y, Miyauchi A, Oda H. Low-risk papillary microcarcinoma of the thyroid: a review of active surveillance trials. Eur J Surg Oncol 2018;44:307-315.