Original Article | Thyroid

eISSN 2005-8330
https://doi.org/10.3348/kjr.2021.0230
Korean J Radiol 2021;22(9):1579-1586

Korean Journal of Radiology
KJR

Diagnostic Performance of the Modified Korean Thyroid Imaging Reporting and Data System for Thyroid Malignancy: A Multicenter Validation Study

Sae Rom Chung[1], Hye Shin Ahn[2], Young Jun Choi[1], Ji Ye Lee[3], Roh-Eul Yoo[3], Yoo Jin Lee[4], Jee Young Kim[5], Jin Yong Sung[6], Ji-hoon Kim[3], Jung Hwan Baek[1]

[1]Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea; [2]Department of Radiology, Chung-Ang University Hospital, Chung-Ang University College of Medicine, Seoul, Korea; [3]Department of Radiology, Seoul National University Hospital, Seoul, Korea; [4]Department of Radiology, Busan Paik Hospital, Inje University College of Medicine, Busan, Korea; [5]Department of Radiology, Yeouido St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea; [6]Department of Radiology, Thyroid Center, Daerim Saint Mary's Hospital, Seoul, Korea

Objective: To evaluate the diagnostic performance of the modified Korean Thyroid Imaging Reporting and Data System (K-TIRADS), and compare it with the 2016 version of K-TIRADS using the Thyroid Imaging Network of Korea.

Materials and Methods: Between June and September 2015, 5708 thyroid nodules (≥ 1.0 cm) from 5081 consecutive patients who had undergone thyroid ultrasonography at 26 institutions were retrospectively evaluated. We used a biopsy size threshold of 2 cm for K-TIRADS 3 and 1 cm for K-TIRADS 4 (modified K-TIRADS 1) or 1.5 cm for K-TIRADS 4 (modified K-TIRADS 3). The modified K-TIRADS 2 subcategorized the K-TIRADS 4 into 4A and 4B, and the cutoff sizes for the biopsies were defined as 1 cm for K-TIRADS 4B and 1.5 cm for K-TIRADS 4A. The diagnostic performance and the rate of unnecessary biopsies of the modified K-TIRADS for detecting malignancy were compared with those of the 2016 K-TIRAD, which were stratified by nodule size (with a threshold of 2 cm).

Results: A total of 1111 malignant nodules and 4597 benign nodules were included. The sensitivity, specificity, and unnecessary biopsy rate of the benign nodules were 94.9%, 24.4%, and 60.9% for the 2016 K-TIRADS; 91.0%, 39.7%, and 48.6% for the modified K-TIRADS 1; 84.9%, 45.9%, and 43.5% for the modified K-TIRADS 2; and 76.1%, 50.2%, and 40.1% for the modified K-TIRADS 3. For small nodules (1–2 cm), the diagnostic sensitivity of the modified K-TIRADS decreased by 5.2–25.6% and the rate of unnecessary biopsies reduced by 19.2–32.8% compared with those of the 2016 K-TIRADS ($p < 0.001$). For large nodules (> 2 cm), the modified K-TIRADSs maintained a very high sensitivity for detecting malignancy (98%).

Conclusion: The modified K-TIRADSs significantly reduced the rate of unnecessary biopsies for small (1–2 cm) nodules while maintaining a very high sensitivity for malignancy for large (> 2 cm) nodules.

Keywords: *Thyroid nodule; Thyroid cancer; Diagnostic performance; Ultrasonography; Fine-needle aspiration*

INTRODUCTION

The rate of detection of thyroid nodules and carcinomas has increased with the widespread use of ultrasonography (US), which is the primary diagnostic tool for the assessment of thyroid nodules [1]. With an associated increase in the incidence of thyroid cancer, the use of thyroid fine-needle aspiration (FNA) has increased rapidly [2,3]. Although FNA

Received: March 24, 2021 **Revised:** April 20, 2021 **Accepted:** April 21, 2021
This research was supported by research fund from the Korean Society of Radiology through Radiology Imaging Network of Korea for Clinical Research (RINK-CR).
Corresponding author: Hye Shin Ahn, MD, Department of Radiology, Chung-Ang University Hospital, Chung-Ang University College of Medicine, 102 Heukseok-ro, Dongjak-gu, Seoul 06973, Korea.
• E-mail: ach0224@gmail.com

is a simple and safe diagnostic tool for thyroid nodules, the procedure causes discomfort to patients, and additional costs are associated with repeat biopsies or more expensive testing [4,5]. Therefore, identifying the optimal indications for FNA for US-detected thyroid nodules are important. Several international societies have developed guidelines using US-risk stratification systems to assist physicians in determining when it is appropriate to perform FNA or follow-up examinations of thyroid nodules in patients [6-10]. Recently, several studies have compared the diagnostic performance of various US- risk stratification systems for diagnosing thyroid malignancy and the rate of unnecessary biopsies [11-16]. A previous comparative study revealed that the Korean Thyroid Imaging Reporting and Data System (K-TIRADS) exhibited higher sensitivity and a higher rate of unnecessary biopsies than other US-risk stratification systems [11-16]. The lower size cutoff for biopsies is known to be the main cause of the high rate of unnecessary biopsies associated with K-TIRADS [11,16]. In light of this finding, it is necessary to modify the size cutoff for biopsies in the 2016 K-TIRADS to reduce the rate of unnecessary biopsies while maintaining an appropriate sensitivity for malignancy.

The nodule size should be considered when determining the optimal biopsy size cutoff, as the size of a thyroid cancer is an important prognostic factor [17]. A previous study revealed that a size threshold of 2 cm maximizes the prognostic discrimination of tumors. A size of > 2 cm is associated with a five times higher risk of recurrence than a size of ≤ 2 cm [18]. Therefore, it is necessary to evaluate the diagnostic performance and rate of unnecessary biopsies associated with the US-risk stratification system based on nodule size.

Therefore, this study aimed to evaluate the diagnostic performance and rate of unnecessary biopsies associated with the modified K-TIRADS, which has different cutoff sizes for biopsy, and to compare the results with those of the 2016 K-TIRADS. Furthermore, we performed a subgroup analysis to determine the optimal cutoff size based on nodule size with a threshold of 2 cm.

MATERIALS AND METHODS

Institutional Review Board approval was obtained from all participating institutions for this retrospective study. The requirement for informed consent was waived due to the retrospective nature of the analyses and the use of anonymized medical records. The study methods and data reporting were performed following the Standards for Reporting of Diagnostic Accuracy Studies [19].

Study Population

This retrospective analysis was based on patient data collected from 26 different hospitals in Korea (Thyroid Imaging Network of Korea registry). Consecutive patients who underwent thyroid US between June 2015 and September 2015 were enrolled in this study. Only patients who met the following criteria were included: 1) patients who had nodules of size ≥ 1 cm, and 2) patients who had undergone FNA, core-needle biopsy (CNB), or surgery for nodules. Patients were excluded from the study if 1) the thyroid nodule was smaller than 1 cm, 2) there was no reference standard test (biopsy or surgery), or 3) the image quality was suboptimal. Among 22775 consecutive patients who had undergone thyroid US from 26 institutions, 16679 patients were excluded due to the absence of a reference standard test (biopsy or surgery) (n = 4304), a thyroid nodule size less than 1 cm (n = 12130), or suboptimal image quality (n = 245). Among them, 1015 patients with 1102 nodules were further excluded because of inconclusive biopsy results (Fig. 1).

Finally, 5708 thyroid nodules (1111 malignant and 4597 benign nodules) in 5081 consecutive patients (4176 female and 905 male; mean age, 53.2 years; age range, 19–76 years) were included in the study. For malignant nodules (n = 1111), a final diagnosis was confirmed during surgery (n = 947, 85.2%) or FNA or CNB (n = 164, 14.8%). For

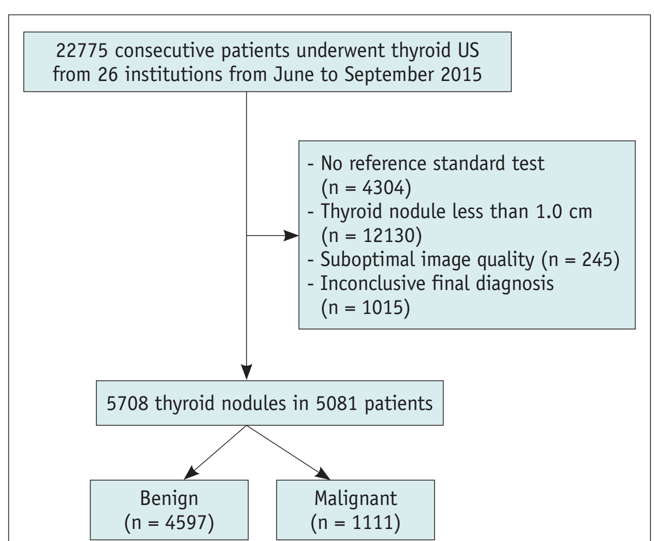

Fig. 1. Flowchart of the study. US = ultrasonography

https://doi.org/10.3348/kjr.2021.0230

benign nodules (n = 4597), a pathologic diagnosis was confirmed during surgery (n = 394, 8.6%), single FNA or CNB (n = 3602, 78.3%), and repeated FNA or CNB on at least two occasions (n = 601, 13.1%).

US Examination and Image Analysis

All US examinations were performed using high-resolution US systems and 5–14 MHz linear probes. US images in the Digital Imaging and Communications in Medicine format were reviewed by 17 experienced radiologists using an online program (AIM AiCRO: https://study.aim-aircro.com). Initially, two meetings were held to establish a baseline consensus in the lexicon for the US criteria. The US features of the nodules were strictly assessed using the definitions in the US lexicon of the K-TIRADS to minimize the misclassification of nodules [8]. The reviewers evaluated the images of 15 biopsy-proven thyroid nodules that were not included in the study and were asked to assess the US features including internal content, echogenicity of the solid portion, shape, orientation, margin, calcifications, presence of a spongiform appearance, and comet-tail artifacts. In the modified K-TIRADS, nonparallel orientation was defined as a longer anteroposterior diameter than the transverse diameter of a nodule in the transverse plane. The benign category included spongiform or partially cystic nodules with intracystic echogenic foci and comet-tail artifacts, regardless of suspicious US features. The echogenicity of the solid portion was classified as hyperechoic, isoechoic, mild hypoechoic, and marked hypoechoic. A nodule with marked hypoechogenicity is hypoechoic or has similar echogenicity relative to the anterior neck muscle [20]. All reviewers, who were blinded to the biopsy results and final diagnoses, assessed the US features of the thyroid nodules.

Biopsy Size Thresholds for the Modified K-TIRADS

Three different biopsy size cutoffs (modified K-TIRADS 1–3) were simulated and used to compare the diagnostic performance and rate of unnecessary biopsies. The modified K-TIRADS 1 simulated biopsy size cutoffs for the low suspicion category (K-TIRADS 3) from 1.5 cm to 2 cm. The modified K-TIRADS 2 subcategorized the intermediate suspicion category (K-TIRADS 4) into 4A and 4 B, based on the malignancy risk of the US features (Supplementary Table 1) [20]. The size cutoff for the biopsies was subdivided into 1 cm for K-TIRADS 4B, 1.5 cm for K-TIRADS 4A, and 2 cm for low suspicion (K-TIRADS 3) nodules. The modified K-TIRADS 3 simulated a biopsy size cutoff of 2 cm for the low suspicion category (K-TIRADS 3) and 1.5 cm for the intermediate suspicion category (K-TIRADS 4) (Table 1).

Statistical Analyses

The thyroid nodules were dichotomized into two groups, biopsy-indicated and biopsy-not-indicated (test positivity and test negativity, respectively), based on the criteria for biopsy of each K-TIRADS. The diagnostic performance for the detection of thyroid cancer and sonographic recommendations for biopsy were evaluated based on sensitivity, specificity, positive predictive value, negative predictive value, and accuracy for each guideline (at 95% confidence intervals). The unnecessary biopsy rate for the diagnosis of thyroid cancer was calculated as a percentage of the benign nodules among the fine-needle aspiration biopsy-required nodules. The results were compared to those for the 2016 K-TIRADS using the generalized estimating equation method. Subgroup analysis was conducted for the biopsies with a cutoff size of 2 cm. Statistical analyses were performed using SPSS for Windows (version 23.0; IBM Corp.) and SAS for Windows (version 9.2; SAS Institute). The differences were considered significant at $p < 0.05$.

RESULTS

The sizes of the thyroid nodules ranged from 1.0 cm to 10.0 cm (mean, 2.1 cm). The maximum diameter of the nodules was 1–2 cm (small) in 3576 nodules (62.6%) and > 2 cm (large) in 2132 nodules (37.4%). The malignancy rate was higher for the small thyroid nodules (1–2 cm) than for the large nodules (> 2 cm) (22.7% [810 of 3576] and 14.1% [301 of 2132]) (Table 2). The diagnoses of malignancy based on histological examination included papillary carcinomas (n = 1011), follicular carcinomas (n = 62), medullary carcinomas (n = 12), anaplastic carcinomas (n = 6), poorly differentiated carcinomas (n = 7),

Table 1. Simulated Size Criteria for Biopsy in Modified K-TIRADS

Category	2016 K-TIRADS (cm)	Modified K-TIRADS 1 (cm)	Modified K-TIRADS 2 (cm)	Modified K-TIRADS 3 (cm)
5 High suspicion	≥ 1.0	≥ 1.0	≥ 1.0	≥ 1.0
4 Intermediate suspicion	≥ 1.0	≥ 1.0	4A ≥ 1.5 4B ≥ 1.0	≥ 1.5
3 Low suspicion	≥ 1.5	≥ 2.0	≥ 2.0	≥ 2.0
2 Benign	N/A	N/A	N/A	N/A

K-TIRADS = Korean Thyroid Imaging Reporting and Data System, N/A = not applicable

 https://doi.org/10.3348/kjr.2021.0230

lymphomas (n = 3), squamous cell carcinomas (n = 1), metastasis (n = 5), and unspecified malignancy (n = 4). Supplementary Table 2 shows the frequencies and the risks of malignancy for the nodules for the 2016 K-TIRADS and modified K-TIRADS categories.

Diagnostic Performance and Rate of Unnecessary Biopsy according to Nodule Size Cutoff Simulations

The 2016 K-TIRADS showed high sensitivity (94.9%) and a high rate of unnecessary biopsies (60.9%). Table 3 compares the diagnostic performance in the prediction of thyroid malignancy and the rate of unnecessary biopsies of the modified K-TIRADS and the 2016 K-TIRADS for all nodules. The modified K-TIRADS 1, 2, and 3 showed significantly lower diagnostic sensitivity, higher specificity, and higher diagnostic accuracy than the 2016 K-TIRADS ($p < 0.001$). The modified K-TIRADS reduced the rate of unnecessary biopsies by 12.3-20.8% compared with those of the 2016 K-TIRADS ($p < 0.001$).

Subgroup Analysis in Small Nodules (1–2 cm)

Table 4 shows the diagnostic performance and rate of unnecessary biopsies for the modified K-TIRADS for small nodules (1–2 cm). The modified K-TIRADS 1, 2, and 3 showed significantly lower diagnostic sensitivities, higher specificities, and higher diagnostic accuracies than the 2016 K-TIRADS ($p < 0.001$). The diagnostic sensitivities of modified K-TIRADS 1, 2, and 3 were significantly decreased by 5.2%, 13.6%, and 25.6%, respectively, compared with those of the 2016 K-TIRADS ($p < 0.001$). The modified K-TIRADS, compared with the 2016 K-TIRADS, reduced the rate of unnecessary biopsies by 19.2–32.8% ($p < 0.001$).

Subgroup Analysis in Large Nodules (> 2 cm)

Table 5 shows the diagnostic performance and rate of unnecessary biopsies of the modified K-TIRADS for large

Table 2. Demographic Data of 5708 Nodules in 5081 Patients in This Study

	Total	Benign	Malignant
Patients			
Number	5081	4063	1018
Age, year*	53.2 ± 12.7	54.3 ± 12.2	48.9 ± 13.9
Sex			
Female	4176 (82.2)	3393 (83.5)	783 (76.9)
Male	905 (17.8)	670 (16.5)	235 (23.1)
Nodules			
Number	5708	4597	1111
Size of nodule, cm*	2.1 ± 1.1	2.1 ± 1.1	1.9 ± 1.1
≤ 2	3576 (62.6)	2766 (60.2)	810 (72.9)
> 2	2132 (37.4)	1831 (39.8)	301 (27.1)

Data are number of patients or nodules with percentage in parentheses, unless specified otherwise. *Data are mean ± standard deviation.

Table 3. Diagnostic Performance of Biopsy Criteria by the Modified K-TIRADS and 2016 K-TIRADS in All Nodules (n = 5708)

Category	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Diagnostic Accuracy (%)	Unnecessary Biopsy Rate (%)
K-TIRADS	94.9 (93.4–96.0)	24.4 (23.2–25.7)	23.3 (22.1–24.5)	95.2 (93.8–96.3)	38.1 (36.9–39.4)	60.9 (59.6–62.1)
Modified K-TIRADS 1	91.0 (89.2–92.5)	39.7 (38.3–41.1)	26.7 (25.3–28.2)	94.8 (93.7–95.7)	49.7 (48.4–51.0)	48.6 (47.3–49.9)
Modified K-TIRADS 2	84.9 (82.6–86.9)	45.9 (44.5–47.4)	27.5 (26.0–29.0)	92.6 (91.5–93.6)	53.5 (52.2–54.8)	43.5 (42.3–44.8)
Modified K-TIRADS 3	76.1 (73.6–78.6)	50.2 (48.7–51.6)	27.0 (25.5–28.6)	89.7 (88.5–90.8)	55.2 (53.9–56.5)	40.1 (38.9–41.4)

Data in parentheses are 95% confidence intervals. Modified K-TIRADS 1 simulate the size cutoff for low suspicion category as 2 cm. Modified K-TIRADS 2 subdivide intermediate suspicion category as 4A and 4B and simulate the size cutoff for low suspicion category as 2 cm, 4A as 1.5 cm, and 4B as 1 cm. Modified K-TIRADS 3 simulate the size cutoff for low suspicion category as 2 cm and intermediate suspicion as 1.5 cm. K-TIRADS = Korean Thyroid Imaging Reporting and Data System, NPV = negative predictive value, PPV = positive predictive value

https://doi.org/10.3348/kjr.2021.0230

Table 4. Diagnostic Performance of Biopsy Criteria by the Modified K-TIRADS and 2016 K-TIRADS in Smaller Nodules (≤ 2 cm) (n = 3576)

Category	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Diagnostic Accuracy (%)	Unnecessary Biopsy Rate (%)
K-TIRADS	93.6 (91.7–95.1)	34.9 (33.1–36.7)	29.6 (27.9–31.4)	94.9 (93.4–96.1)	48.2 (46.5–49.8)	50.4 (48.7–52.0)
Modified K-TIRADS 1	88.4 (86.0–90.4)	59.7 (57.9–61.5)	39.1 (36.9–41.4)	94.6 (93.5–95.6)	66.2 (64.7–67.8)	31.2 (29.7–32.7)
Modified K-TIRADS 2	80.0 (77.1–82.6)	70.1 (68.4–71.8)	44.0 (41.4–46.5)	92.3 (91.1–93.4)	72.4 (70.9–73.8)	23.1 (21.7–24.5)
Modified K-TIRADS 3	68.0 (64.7–71.1)	77.2 (75.6–78.7)	46.6 (43.8–49.5)	89.2 (87.9–90.4)	75.1 (73.7–76.5)	17.6 (16.4–18.9)

Data in parentheses are the raw data used to calculate the percentages, and data in brackets are 95% confidence intervals. Modified K-TIRADS 1 simulate the size cutoff for low suspicion category as 2 cm. Modified K-TIRADS 2 subdivide intermediate suspicion category as 4A and 4B and simulate the size cutoff for low suspicion category as 2 cm, 4A as 1.5 cm, and 4B as 1 cm. Modified K-TIRADS 3 simulate the size cutoff for low suspicion category as 2 cm and intermediate suspicion as 1.5 cm. K-TIRADS = Korean Thyroid Imaging Reporting and Data System, NPV = negative predictive value, PPV = positive predictive value

Table 5. Diagnostic Performance of Biopsy Criteria by the Modified K-TIRADS and 2016 K-TIRADS in Larger Nodules (> 2 cm) (n = 2132)

Category	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Diagnostic Accuracy (%)	Unnecessary Biopsy Rate (%)
K-TIRADS	98.3 (96.1–99.3)	8.6 (7.4–10.0)	15.0 (13.5–16.7)	96.9 (92.8–98.7)	21.3 (19.6–23.1)	78.5 (76.7–80.2)
Modified K-TIRADS 1	98.0 (95.6–99.1)	9.4 (8.1–10.8)	15.1 (13.6–16.8)	96.6 (92.7–98.5)	21.9 (20.2–23.7)	77.8 (76.0–79.5)
Modified K-TIRADS 2	98.0 (95.6–99.1)	9.4 (8.1–10.8)	15.1 (13.6–16.8)	96.6 (92.7–98.5)	21.9 (20.2–23.7)	77.8 (76.0–79.5)
Modified K-TIRADS 3	98.0 (95.6–99.1)	9.4 (8.1–10.8)	15.1 (13.6–16.8)	96.6 (92.7–98.5)	21.9 (20.2–23.7)	77.8 (76.0–79.5)

Data in parentheses are the raw data used to calculate the percentages, and data in brackets are 95% confidence intervals. Modified K-TIRADS 1 simulate the size cutoff for low suspicion category as 2 cm. Modified K-TIRADS 2 subdivide intermediate suspicion category as 4A and 4B and simulate the size cutoff for low suspicion category as 2 cm, 4A as 1.5 cm, and 4B as 1 cm. Modified K-TIRADS 3 simulate the size cutoff for low suspicion category as 2 cm and intermediate suspicion as 1.5 cm. K-TIRADS = Korean Thyroid Imaging Reporting and Data System, NPV = negative predictive value, PPV = positive predictive value

nodules (> 2 cm). The diagnostic sensitivity, specificity, and diagnostic accuracy of the modified K-TIRADS 1, 2, and 3 were 98.0%, 9.4%, and 21.9%, respectively, revealing no significant change compared with those of the 2016 K-TIRADS (p = 0.317). The rate of unnecessary biopsy decreased by 0.7% for the modified K-TIRADS compared with the 2016 K-TIRADS (p < 0.001).

DISCUSSION

Our study evaluated the diagnostic performance and rate of unnecessary biopsies for the modified K-TIRADS, which raised the size cutoffs for biopsies within the intermediate and low suspicion category and those of the 2016 K-TIRADS. For small nodules (1–2 cm), the modified K-TIRADSs significantly reduced the rate of unnecessary biopsy (19.2–32.8%) and the diagnostic sensitivity (5.2–25.6%), compared with the 2016 K-TIRADS. All the modified K-TIRADSs maintained a very high sensitivity of 98.0% for large thyroid nodules (> 2 cm).

There are several US-risk stratification systems for assessing thyroid nodules, and each system assigns different size thresholds to identify nodules that require FNA. Previous comparative studies revealed that K-TIRADS showed the highest sensitivity and highest rate of unnecessary biopsies compared with the other US-risk stratification systems [11-15]. A study simulating different size cutoffs for biopsies for each US-risk stratification system revealed that a high rate of unnecessary biopsies in the K-TIRADS resulted from the lower size cutoffs for biopsies [11]. US-risk stratification systems are designed to identify nodules with a low risk of malignancy and deferrable cytologic assessment [13,20]. In light of this finding, it is necessary to modify the size criteria for biopsies for the K-TIRADS to reduce the rate of unnecessary biopsies while maintaining an appropriate sensitivity for malignancy.

Nodules that were within the high suspicion category were associated with biopsies when they were larger than 1 cm, whereas nodules within the benign category were not associated with biopsies [6-10]. However, the size cutoffs for biopsies related to the low- and intermediate-suspicion categories vary among US-risk stratification systems. We used three different simulated biopsy thresholds for the low- and intermediate-suspicion categories and the data of a large population to evaluate the diagnostic performance and rates of unnecessary biopsies for the 2016 K-TIRADS. The modified K-TIRADS 1 reduced the rate of unnecessary biopsy by 12.3%, but it maintained a high sensitivity of 91.0%. The modified K-TIRAD 3 showed higher reductions in the rate of unnecessary biopsy by 20.8% and diagnostic sensitivity by 18.8%. The rate of unnecessary biopsies and the diagnostic sensitivity of the modified K-TIRADS 2 showed a moderate reduction compared with the modified K-TIRADS 1 and 3. The biopsy size threshold can be set within the range of 1–1.5 cm by considering the estimated malignancy risk and location of a nodule, clinical risk factors, and patient factors (age, comorbidity, and preference).

The sensitivity and specificity of US-risk stratification systems for detecting thyroid cancers remain controversial. The size of the nodule should be considered in relation to this issue. For small thyroid nodules (1–2 cm), a higher reduction in the rate of unnecessary biopsies may be appropriate, even though the sensitivity also slightly decreases. Most patients with small thyroid cancers have an excellent prognosis, and preventing overdiagnosis and the harm caused by overtreatment is an important consideration. At the same time, a higher sensitivity and a high rate of unnecessary biopsies may be appropriate for large nodules (> 2 cm), considering the higher likelihood of aggressive interventions for large malignant tumors [20,21]. For our study cohort, the subgroup analysis according to nodule size revealed that the modified K-TIRADS 3 reduced the rate of unnecessary biopsies by 32.8% for small thyroid nodules (≤ 2 cm), while maintaining a high sensitivity of 98.0% for nodules larger than 2 cm.

Our study had limitations. First, the evaluation of cases was retrospective, and included thyroid nodules that had undergone biopsy. Hence, there was an unavoidable selection bias. To minimize this limitation, we conducted a multicenter study involving a large sample. Second, the final diagnoses of non-surgical benign nodules (78.3%) were based on repeated biopsies. This may have resulted in false-negative results. Finally, the features were described by different radiologists, and this may have resulted in interobserver variability.

In conclusion, the modified K-TIRADS significantly reduced the rate of unnecessary biopsies for small thyroid nodules (1–2 cm), while maintaining a very high sensitivity for large nodules (> 2 cm). These results may be helpful for the adjustment of size cutoffs for biopsies in future guidelines.

Supplement

The Supplement is available with this article at https://doi.org/10.3348/kjr.2021.0230.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Acknowledgments

We would like to express our gratitude to all doctors from 26 different hospitals who provided ultrasound data on thyroid nodules for Thyroid Imaging Network of Korea registry
Ji Eun Shin[1], Younghee Yim[2], Nami Choi[3], Hwa Seon Shin[4], Seongjun Bak[4], Sung-Hye You[5], Beomsu Kim[6], Min Ji Hong[7], Hyun Jeong Kim[8], Dae Bong Kim[9], Eun Kyoung Lee[9], Dong Gyu Na[10], Wooyul Paik[10], Sun-Won Park[11], Jung Hyo Rhim[11], Bo Eun Lee[11], Chu a jung[11], Jong yoon Lee[11], Chang Yoon Lee[12], Hyun Kyung Lim[13], Ra Gyoung Yoon[14], Yeo Koon Kim[15], Joon Hyung Lee[16], So Lyung Jung[17], Tae Yoon Kim[18], Eun Ju Ha[19], Seul Kee Kim[20], Yun Young Lee[20], Young Jun Choi[21], Se Jin Cho[21], Donghyun Kim[21], Chong Hyun Suh[21], Tae Jin Yun[22], Inpyeong Hwang[22], Koung Mi Kang[22]
[1]Department of Radiology, Cha University Gangnam Medical Center
[2]Department of Radiology, Chung-Ang University Hospital, Chung-Ang University College of Medicine
[3]Department of Radiology, Konkuk University Medical Center, Konkuk University Scool of Medicine

[4]Department of Radiology, Gyeongsang National University Hospital
[5]Department of Radiology, Korea University Anam Hospital
[6]Department of Radiology, Kosin University College of Medicine
[7]Department of Radiology, Chung-Ang University Hospital, Chung-Ang University College of Medicine
[8]Department of Radiology, Daejeon St. Mary's Hospital, College of Medicine, The Catholic University of Korea
[9]Department of Radiology, Dongguk University Ilsan Hospital
[10]Department of Radiology, GangNeung Asan Hospital, University of Ulsan College of Medicine
[11]Department of Radiology, Seoul Metropolitan Government Seoul National University Boramae Medical Center
[12]Department of Radiology, National Cancer Center
[13]Department of Radiology, Soonchunhyang University Seoul Hospital
[14]Department of Radiology, Nowon Eulji Medical Center, Eulji University
[15]Department of Radiology, Seoul National University Bundang Hospital
[16]Department of Radiology, Inje university haeundae paik hospital
[17]Department of Radiology, Seoul St. Mary's Hospital College of Medicine, the Catholic University of Korea
[18]Department of Radiology, Hanyang University Guri Hospital, Hanyang University College of Medicine
[19]Department of Radiology, Ajou University School of Medicine
[20]Department of Radiology, Chonnam National University Hwasun Hospital
[21]Department of Radiology, University of Ulsan College of Medicine, Asan Medical Center
[22]Department of Radiology, Seoul National University Hospital

Author Contributions

Conceptualization: Sae Rom Chung, Hye Shin Ahn, Ji-hoon Kim, Jung Hwan Baek. Data curation: all authors. Formal analysis: all authors. Funding acquisition: Sae Rom Chung, Hye Shin Ahn, Jin Yong Sung, Ji-hoon Kim, Jung Hwan Baek. Investigation: all authors. Methodology: Sae Rom Chung, Hye Shin Ahn, Ji-hoon Kim, Jung Hwan Baek. Project administration: Jin Yong Sung, Ji-hoon Kim, Jung Hwan Baek. Resources: Hye Shin Ahn, Jin Yong Sung, Ji-hoon Kim, Jung Hwan Baek. Software: Sae Rom Chung, Hye Shin Ahn. Supervision: all authors. Validation: Sae Rom Chung, Ji-hoon Kim, Jung Hwan Baek. Visualization: Sae Rom Chung, Hye Shin Ahn. Writing—original draft: Sae Rom Chung, Hye Shin Ahn. Writing—review & editing: all authors.

ORCID iDs

Sae Rom Chung
https://orcid.org/0000-0003-4219-7166
Hye Shin Ahn
https://orcid.org/0000-0001-7260-7467
Young Jun Choi
https://orcid.org/0000-0001-7098-5042
Ji Ye Lee
https://orcid.org/0000-0002-3929-6254
Roh-Eul Yoo
https://orcid.org/0000-0002-5625-5921
Yoo Jin Lee
https://orcid.org/0000-0003-4701-7339
Jee Young Kim
https://orcid.org/0000-0002-2812-8159
Jin Yong Sung
https://orcid.org/0000-0002-8163-4624
Ji-hoon Kim
https://orcid.org/0000-0002-6349-6950
Jung Hwan Baek
https://orcid.org/0000-0003-0480-4754

REFERENCES

1. Lee JY, Baek JH, Ha EJ, Sung JY, Shin JH, Kim JH, et al. 2020 Imaging guidelines for thyroid nodules and differentiated thyroid cancer: Korean Society of Thyroid Radiology. *Korean J Radiol* 2021;22:840-860
2. Sosa JA, Hanna JW, Robinson KA, Lanman RB. Increases in thyroid nodule fine-needle aspirations, operations, and diagnoses of thyroid cancer in the United States. *Surgery* 2013;154:1420-1426; discussion 1426-1427
3. Zevallos JP, Hartman CM, Kramer JR, Sturgis EM, Chiao EY. Increased thyroid cancer incidence corresponds to increased use of thyroid ultrasound and fine-needle aspiration: a study of the Veterans Affairs health care system. *Cancer* 2015;121:741-746
4. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda system for reporting thyroid cytopathology: a meta-analysis. *Acta Cytol* 2012;56:333-339
5. Na DG, Kim JH, Sung JY, Baek JH, Jung KC, Lee H, et al. Core-needle biopsy is more useful than repeat fine-needle aspiration in thyroid nodules read as nondiagnostic or atypia

of undetermined significance by the Bethesda system for reporting thyroid cytopathology. *Thyroid* 2012;22:468-475
6. 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
7. Gharib H, Papini E, Garber JR, Duick DS, Harrell RM, Hegedüs 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
8. 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
9. 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
10. 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
11. 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
12. 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
13. 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
14. 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
15. 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
16. 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
17. Mazzaferri EL, Jhiang SM. Long-term impact of initial surgical and medical therapy on papillary and follicular thyroid cancer. *Am J Med* 1994;97:418-428
18. Tran B, Roshan D, Abraham E, Wang L, Garibotto N, Wykes J, et al. The prognostic impact of tumor size in papillary thyroid carcinoma is modified by age. *Thyroid* 2018;28:991-996
19. Choi YJ, Chung MS, Koo HJ, Park JE, Yoon HM, Park SH. Does the reporting quality of diagnostic test accuracy studies, as defined by STARD 2015, affect citation? *Korean J Radiol* 2016;17:706-714
20. 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* 2020 Dec [Epub]. https://doi.org/10.14366/usg.20148
21. Machens A, Holzhausen HJ, Dralle H. The prognostic value of primary tumor size in papillary and follicular thyroid carcinoma. *Cancer* 2005;103:2269-2273