Original Article | Thyroid

http://dx.doi.org/10.3348/kjr.2016.17.5.811
pISSN 1229-6929 · eISSN 2005-8330
Korean J Radiol 2016;17(5):811-821

Korean Journal of Radiology
KJR

A Multicenter Prospective Validation Study for the Korean Thyroid Imaging Reporting and Data System in Patients with Thyroid Nodules

Eun Ju Ha, MD[1], Won-Jin Moon, MD, PhD[2], Dong Gyu Na, MD[3], Young Hen Lee, MD[4], Nami Choi, MD[2], Soo Jin Kim, MD[3, 5], Jae Kyun Kim, MD[6]

[1]Department of Radiology, Ajou University School of Medicine, Suwon 16499, Korea; [2]Department of Radiology, Konkuk University Medical Center, Konkuk University School of Medicine, Seoul 05030, Korea; [3]Department of Radiology, Human Medical Imaging and Intervention Center, Seoul 06524, Korea; [4]Department of Radiology, Ansan Hospital, Korea University School of Medicine, Ansan 15355, Korea; [5]Department of Radiology, New Korea Hospital, Gimpo 10086, Korea; [6]Department of Radiology, Chung-Ang University Hospital, Seoul 06973, Korea

Objective: To validate a new risk stratification system for thyroid nodules, the Korean Thyroid Imaging Reporting and Data System (K-TIRADS), using a prospective design.

Materials and Methods: From June 2013 to May 2015, 902 thyroid nodules were enrolled from four institutions. The type and predictive value of ultrasonography (US) predictors were analyzed according to the combination of the solidity and echogenicity of nodules; in addition, we determined malignancy risk and diagnostic performance for each category of K-TIRADS, and compared the efficacy of fine-needle aspiration (FNA) with a three-tier risk categorization system published in 2011.

Results: The malignancy risk was significantly higher in solid hypoechoic nodules, as compared to partially cystic or isohyperechoic nodules (each $p < 0.001$). The presence of any suspicious US features had a significantly higher malignancy risk (73.4%) in solid hypoechoic nodules than in partially cystic or isohyperechoic nodules (4.3–38.5%; $p < 0.001$). The calculated malignancy risk in K-TIRADS categories 5, 4, 3, and 2 nodules were 73.4, 19.0, 3.5, and 0.0%, respectively; and the sensitivity, specificity, positive predictive value, negative predictive value, and accuracy for malignancy were 95.5, 58.6, 44.5, 96.9, and 69.5%, respectively, in K-TIRADS categories 4 and 5. The efficacy of FNA for detecting malignancy based on K-TIRADS was increased from 18.6% (101/544) to 22.5% (101/449), as compared with the three-tier risk categorization system ($p < 0.001$).

Conclusion: The proposed new risk stratification system based on solidity and echogenicity was useful for risk stratification of thyroid nodules and the decision for FNA. The malignancy risk of K-TIRADS was in agreement with the findings of a previous retrospective study.

Keywords: *Thyroid nodules; Thyroid cancer; Thyroid imaging reporting and data system; Ultrasonography; Fine needle aspiration; Risk stratification; Thyroid neoplasm; Malignancy risk; Core needle biopsy*

Received May 6, 2016; accepted after revision June 15, 2016.
Corresponding author: Won-Jin Moon, MD, PhD, Department of Radiology, Konkuk University Medical Center, Konkuk University School of Medicine, 120-1 Neungdong-ro, Gwangjin-gu, Seoul 05030, Korea.
• Tel: (822) 2030-5544 • Fax: (822) 2030-5549
• E-mail: mdmoonwj@naver.com

INTRODUCTION

Thyroid nodules are a common clinical problem (1-6). Ultrasonography (US) is used as a primary diagnostic tool for assessing malignancy risk, fine-needle aspiration (FNA) decisions, and management decisions after FNA in patients with thyroid nodules. Thus, the Korean Society of Thyroid Radiology (KSThR) published a consensus recommendation for US-based management of thyroid nodules in 2011, based on a three-tier risk categorization system (5).

Thyroid nodules are divided into three categories: suspicious malignant, indeterminate, and probably benign. The nodules with at least one of the suspicious US findings including taller than wide shape, spiculated margin, marked hypoechogenicity, microcalcifications, and macrocalcifications were defined as suspicious malignant. In contrast, a simple cyst, a predominantly cystic or cystic nodule with reverberating artifacts and a nodule with a spongiform appearance were defined as probably benign. Indeterminate nodules included nodules with neither malignant nor benign features on US (5).

Although the KSThR recommended three-tier system for thyroid nodules is useful in clinical practice, personalized and optimal management of thyroid nodules requires a revised risk stratification system that can increase the efficacy of FNA, avoid unnecessary procedures, and provide supplementary information on thyroid nodules after FNA (7-14). The combined risk stratification with US and cytologic results after FNA could achieve more timely detection of thyroid cancer and provide an optimal management decision in cases with thyroid nodules. Thus, the KSThR recently suggested a new risk stratification system for thyroid nodules i.e., Korean Thyroid Imaging Reporting and Data System (K-TIRADS), which was developed from a retrospective study (15, 16). However, there is no prospective validation study of K-TIRADS for the risk stratification of thyroid nodules. The aim of this multicenter study was to validate the new risk stratification system for thyroid nodules using a prospective design.

MATERIALS AND METHODS

The prospective multicenter study was approved by the Institutional Review Boards of the four participating centers.

Study Population

Patient data collected from four different hospitals (Stratification Of MAlignancy Risk by Thyroid UltraSonography registry, SOMARTUS) were analyzed prospectively. From June 2013 to May 2015, in total, 1109 thyroid nodules (> 5 mm) in 928 consecutive patients who had undergone thyroid US were enrolled in the study. Among the 1109 nodules, 198 nodules in 169 patients were excluded because a final diagnosis was not obtained (32 nondiagnostic and 166 indeterminate cytologic results) and the US characteristics could not be analyzed in 9 entirely calcified nodules. Thus, in total, 902 nodules in 750 patients were finally included in the study (594 women and 156 men; mean age, 49.2 years; age range, 9–81 years) (Fig. 1).

Final diagnoses were determined from the surgical pathology or cytopathological results based on the Bethesda system (17). Cytological results of non-diagnostic lesions and lesions of indeterminate significance (atypia of undetermined significance, and suspicions of follicular neoplasm and malignancy) without surgical confirmation were excluded from the study.

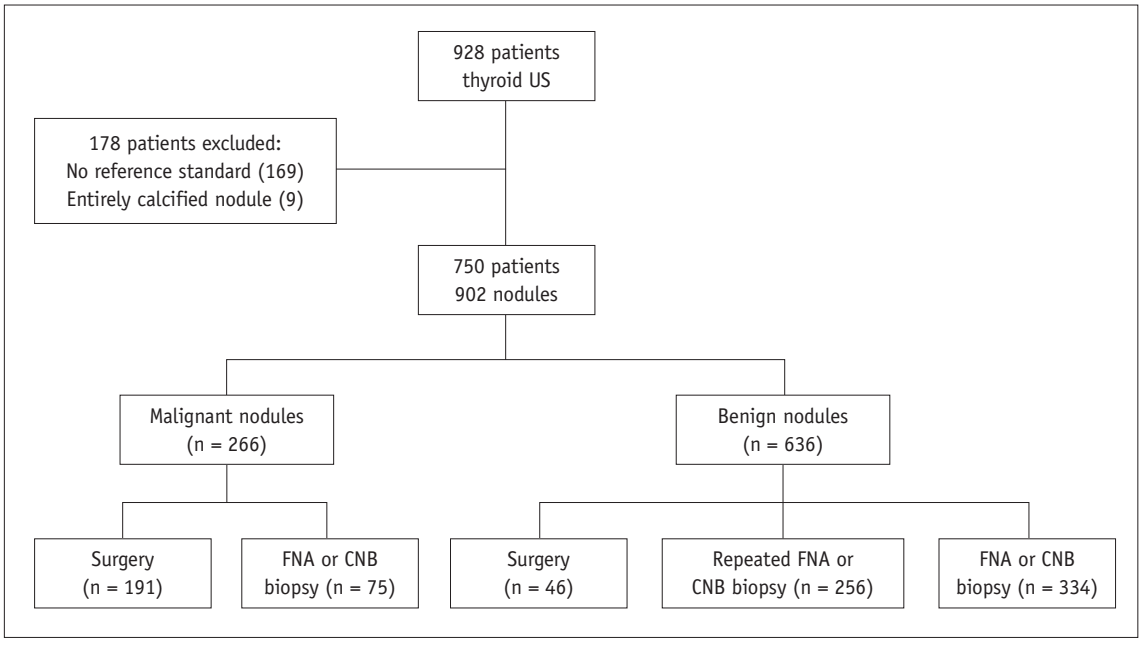

Fig. 1. Flow chart of study group. CNB = core needle biopsy, FNA = fine needle aspiration, US = ultrasonography

Korean Journal of Radiology
KJR

US Examination

All US examinations were performed with a 10–16 MHz linear probe and a real-time US system (EUB-7500, Hitachi Medical Systems, Tokyo, Japan; iU22 and HDI-5000, Philips Healthcare, Bothell, WA, USA; Aplio SSA-770A, Toshiba Medical Systems Corporation, Otawara, Japan; and Accuvix XG, Samsung Medison, Seoul, Korea). The scanning protocol in all cases included both transverse and longitudinal real-time imaging of the thyroid nodules. Five board-certified radiologists, in four different hospitals specializing in thyroid imaging, performed the examinations.

Image Analysis

Before starting the multicenter study, three training sessions were held to establish a baseline consensus in the lexicon for the US criteria. Five experienced radiologists evaluated images of 20 biopsy-proven masses not included in the study and were asked to assess the lexicon for the US criteria during a consensus meeting. The criteria included internal content, echogenicity of the solid portion, shape, orientation, margin, calcifications, presence of a spongiform appearance, and comet-tail artifacts (16).

Ultrasonography features of the nodules were prospectively assessed in each patient during US examination; subsequently, the nodules were classified according to the K-TIRADS followed by the three-tier risk categorization system. The internal content of a nodule was categorized according to the ratio of the cystic portion in the entire nodule: solid (< 10% cystic), predominantly solid (< 50% cystic), predominantly cystic (> 50% cystic), and cystic (> 90% cystic). The nodule echogenicity was categorized as hypoechogenicity (marked or mild), isoechogenicity, and hyperechogenicity by the predominant echogenicity with the reference of normal thyroid gland and anterior neck muscle. The shape of the nodule was categorized as ovoid-to-round or irregular, and the orientation was categorized as parallel (when the anteroposterior diameter of a nodule is equal to or less than its transverse or longitudinal diameter) and non-parallel (when the anteroposterior diameter of a nodule is longer than its transverse or longitudinal diameter on a transverse or longitudinal plane) (16). The margin of a nodule was categorized as smooth, spiculated/microlobulated, or ill-defined. Calcification was classified as microcalcification (tiny, punctate echogenic foci of ≤ 1 mm either with or without posterior shadowing), macrocalcification (echogenic foci of > 1 mm in size), or rim calcification (nodule with peripheral curvilinear or egg-shell calcification) (16). When measuring nodule size, the calipers were positioned at the outer margin of the halo of the nodule (18).

US-Guided FNA and CNB Procedures

Ultrasonography-guided FNAs or core needle biopsies (CNBs) were performed by the same radiologists who performed the thyroid US. US-guided FNAs were performed with 23-gauge needles and a combination of capillary and aspiration FNA techniques. CNB was performed using a disposable 18-gauge, single- or double-action spring-activated needle (TSK Acecut or Stericut; Create Medic, Yokohama, Japan). FNA was usually performed for thyroid nodules > 1 cm, with the exception of pure cystic nodules, partially cystic nodules with comet-tail artifacts, and spongiform nodules that usually underwent FNA for therapeutic cyst aspiration, ethanol or radiofrequency ablation therapy, or nodule size of > 2 cm in case of spongiform nodule. FNA was performed for thyroid nodules < 1 cm in case of suspicious US features, or for decisions on surgical planning. The interpretation of FNA was based on the Bethesda system for reporting thyroid cytopathology and CNB results were diagnosed with a six-tier pathology reporting system (17, 19).

Statistical Analyses

The type and predictive value of US predictors for malignancy were evaluated in all and subgroups according to solidity and echogenicity. The χ^2 or Fisher's exact test was used to evaluate US features associated with malignancy and compare the frequency and malignancy risk of US features in all and subgroups according to solidity and echogenicity. A multivariable logistic regression analysis was used to determine independent US predictors for malignancy.

The malignancy risk and the diagnostic performances in each category of K-TIRADS were calculated as percentages and further stratified by size. The area under the receiver operating characteristics curve (Az) with 95% confidence intervals (CIs) were calculated for both K-TIRADS and the three-tier risk categorization system. The efficacy of FNA for detecting malignancy was compared between the two risk stratification systems, with the reduction in the number of FNAs. The efficacy of FNA was calculated as the ratio of the total number of detecting malignancy among the total number of FNAs, which were indicated according to each risk stratification system.

Statistical analyses were performed with SPSS for Windows (ver. 23.0; IBM, Armonk, NY, USA) and MedCalc for Windows software (ver. 15.0; MedCalc, Ostend, Belgium). A significant difference was defined as a p value < 0.05.

RESULTS

Nodule Characteristics

The mean size of nodules was 1.5 ± 1.1 cm (range: 0.5–10.0 cm). Final diagnoses in 902 nodules were 636 (70.5%) benign nodules and 266 (29.5%) malignant nodules. Final diagnoses were determined by surgical resections in 191 of 266 (71.8%) malignant nodules, which included 186 papillary thyroid carcinomas (PTCs) including 24 follicular variant PTCs and 5 follicular carcinomas. Seventy-five malignant nodules diagnosed by FNA or CNB were PTCs. The 46 surgically confirmed benign nodules included 41 nodular hyperplasias, 4 follicular adenomas, and 1 thyroiditis. The others were diagnosed based on the repetitive benign FNA results or benign FNA results with US follow-up studies.

US Features for Predicting Malignant Thyroid Nodules

The mean size of the benign nodules was 1.6 ± 1.1 cm, which was significantly larger than that of malignant nodules (1.1 ± 0.9 cm; p < 0.001); however, it was not an independent predictor for malignancy in the multivariate analysis (p = 0.488). The suspicious US features of microcalcifications, non-parallel orientation, and spiculated/microlobulated margin were all significant and independent predictors for malignancy on univariate and multivariate analysis (each p < 0.001). Macrocalcification was significant on univariate analysis (p = 0.032); however, it was not an independent predictor of malignancy in the multivariate analysis (p = 0.798). There was no malignant tumor in 37 partially cystic nodules with comet-tail artifacts or in 24 spongiform nodules (p < 0.001 and p = 0.001, respectively).

Frequency and Risk of Malignant Thyroid Nodules According to Solidity, Echogenicity, and the Combination of Solidity and Echogenicity

Table 1 showed the frequency and risk of malignancy for thyroid nodules according to solidity, echogenicity, and the combination of solidity and echogenicity. According to solidity, the frequency and risk of malignancy were significant higher in solid nodules than in partially cystic (predominantly solid/cystic) nodules (95.1% vs. 4.9%; 35.4% vs. 7.3%, respectively; p < 0.001). According to echogenicity, the frequency and risk of malignancy

Table 1. Frequency and Risk of Malignant Thyroid Nodules According to Solidity, Echogenicity, and Combination of Solidity and Echogenicity

Pattern	No. of Nodules (n = 902)	No. of Malignant Nodules (n = 266)	Malignancy Risk (%)
Solidity			
Solid	715 (79.3)	253 (95.1)	35.4
Predominantly solid	134 (14.9)	11 (4.1)	8.2
Predominantly cystic	45 (5.0)	2 (0.8)	4.4
Cystic	8 (0.8)	0 (0.0)	0.0
Partially cystic (predominantly solid/cystic)	179 (19.8)	13 (4.9)	7.3
Echogenicity			
Markedly hypoechoic	76 (8.4)	53 (19.9)	69.7
Mildly hypoechoic	398 (44.1)	197 (74.1)	49.5
Isoechoic	423 (46.9)	15 (5.7)	3.5
Hyperechoic	5 (0.6)	1 (0.4)	20.0
Hypoechoic (marked or mild)	474 (52.5)	250 (94.0)	52.7
Isohyperechoic	428 (47.5)	16 (6.0)	3.7
Combination of solidity and echogenicity			
Solid hypoechoic	436 (48.3)	242 (91.0)	55.5
Solid isohyperechoic	279 (30.9)	11 (4.1)	3.9
Partially cystic hypoechoic	38 (4.2)	8 (3.0)	21.1
Partially cystic isohyperechoic	141 (15.6)	5 (1.9)	3.5
Partially cystic or isohyperechoic	466 (51.7)	24 (9.0)	5.2

Numbers in parentheses are percentages.

Korean Journal of Radiology
KJR

were significantly higher in hypoechoic nodules than in isohyperechoic nodules (94.0% vs. 6.0%; 52.7% vs. 3.7%, respectively; $p < 0.001$). When combining solidity with echogenicity, the risk of malignancy was the highest in solid hypoechoic nodules (55.5%) followed by partially cystic hypoechoic (21.1%), solid isohyperechoic (3.9%), and partially cystic isohyperechoic nodules (3.5%), in decreasing order. The overall frequency and risk of malignancy were significantly higher in solid hypoechoic nodules than in partially cystic or isohyperechoic nodules (91.0% vs. 9.0%; 55.5% vs. 5.2%, respectively; $p < 0.001$).

US Features Predicting for Malignant Thyroid Nodules According to the Combination of Solidity and Echogenicity

Table 2 listed US features for predicting malignant thyroid nodules according to the combination of solidity and echogenicity. Three suspicious US features including microcalcifications, non-parallel orientation, and spiculated/microlobulated margin were all independent predictors for malignancy in solid hypoechoic nodules (each $p < 0.001$), while microcalcifications ($p = 0.003$) and spiculated/microlobulated margins ($p < 0.001$) were independent predictors in partially cystic or isohyperechoic nodules. Non-parallel orientation was not an independent predictor in partially cystic or isohyperechoic nodules ($p = 0.889$).

Table 3 showed the frequency and malignancy risk of suspicious US features according to the combination of solidity and echogenicity. The malignancy risks of microcalcifications and non-parallel orientation were significantly higher in solid hypoechoic nodules (74.1 and 77.0%, respectively) than those in partially cystic or isohyperechoic nodules (21.7 and 10.5%, respectively; $p < 0.001$ and $p < 0.001$, respectively). The malignancy risk of a spiculated/microlobulated margin was higher in solid hypoechoic nodules (84.1%) than in partially cystic or isohyperechoic nodules (70.0%); however, it was not statistically significant ($p = 0.248$). Suspicious US feature had a high malignancy risk (73.4%) in solid hypoechoic nodules, and a low-to-intermediate malignancy risk in solid isohyperechoic, partially cystic hypoechoic, partially cystic isohyperechoic, and partially cystic or isohyperechoic

Table 2. US Features Predicting Malignant Thyroid Nodules According to Combination of Solidity and Echogenicity

Solidity and Echogenicity	US Features	No. of Benign Nodules	No. of Malignant Nodules	*P*	Multivariable Analysis	
					Odds Ratio (95% CI)	*P*
All (n = 902)	Overall	636 (70.5)	266 (29.5)			
	Solid	462 (64.6)	253 (35.4)	< 0.001	1.92 (0.93–3.97)	0.077
	Hypoechoic	224 (47.3)	250 (52.7)	< 0.001	10.35 (5.79–18.51)	< 0.001
	Irregular shape	24 (3.8)	25 (9.4)	0.001	0.35 (0.16–0.78)	0.009
	Microcalcification	73 (38.6)	116 (61.4)	< 0.001	3.58 (2.29–5.61)	< 0.001
	Non-parallel orientation	75 (34.7)	141 (65.3)	< 0.001	3.54 (2.23–5.60)	< 0.001
	Spiculated/microlobulated margin	27 (16.8)	134 (83.2)	< 0.001	6.29 (3.73–10.62)	< 0.001
	Comet tail artifact	37 (5.8)	0 (0.0)	< 0.001	N/A	0.998
	Spongiform	24 (3.8)	0 (0.0)	0.001	N/A	0.998
Solid hypoechoic (n = 436)	Overall	194 (44.5)	242 (55.5)			
	Irregular shape	22 (46.8)	25 (53.2)	0.735		
	Microcalcification	37 (25.9)	106 (74.1)	< 0.001	3.46 (2.11–5.67)	< 0.001
	Non-parallel orientation	41 (23.0)	137 (77.0)	< 0.001	3.55 (2.19–6.76)	< 0.001
	Spiculated/microlobulated margin	24 (15.9)	127 (84.1)	< 0.001	4.99 (2.93–8.50)	< 0.001
	Comet tail artifact	2 (100.0)	0 (0.0)	0.197		
	Spongiform	0 (0.0)	0 (0.0)	N/A		
Partially cystic or isohyperechoic (n = 466)	Overall	442 (94.8)	24 (5.2)			
	Irregular shape	2 (100.0)	0 (0.0)	1.000		
	Microcalcification	36 (78.3)	10 (21.7)	< 0.001	0.21 (0.07–0.58)	0.003
	Non-parallel orientation	34 (89.5)	4 (10.5)	0.118	1.13 (0.21–5.98)	0.889
	Spiculated/microlobulated margin	3 (30.0)	7 (70.0)	< 0.001	0.03 (0.01–0.14)	< 0.001
	Comet tail artifact	35 (100.0)	0 (0.0)	0.242		
	Spongiform	24 (100.0)	0 (0.0)	0.626		

Numbers in parentheses are percentages. Malignant US features were indicated by spiculated/microlobulated margin, non-parallel orientation, and microcalcifications. CI = confidence interval, US = ultrasonography

Table 3. Frequency and Malignant Risk of Suspicious Ultrasonography (US) Features According to Solidity, Echogenicity, and Combination of Solidity and Echogenicity

Group	Microcalcifications (n = 189)		Non-Parallel Orientation (n = 216)		Spiculated/Microlobulated Margin (n = 161)		Any Three Suspicious US Features (n = 368)		None of Three Suspicious US Features (n = 534)	
	Malignancy Risk (%)	Frequency (%)	Malignancy Risk (%)	Frequency (%)	Malignancy Risk (%)	Frequency (%)	Malignancy Risk (%)	Frequency (%)	Malignancy Risk (%)	Frequency (%)
All (n = 902)	61.4 (116/189)	21.0 (189/902)	65.3 (141/216)	23.9 (216/902)	83.2 (134/161)	17.8 (161/902)	60.3 (222/368)	40.8 (368/902)	8.2 (44/534)	59.2 (534/902)
Solidity										
Solid (n = 715)	64.5 (111/172)	24.1 (172/715)	70.4 (138/196)	27.4 (196/715)	84.0 (131/156)	21.8 (156/715)	65.3 (216/331)	46.3 (331/715)	9.6 (37/384)	53.7 (384/715)
Partially cystic (n = 179)	31.3 (5/16)	8.9 (16/179)	15.0 (3/20)	11.2 (20/179)	60.0 (3/5)	2.8 (5/179)	16.7 (6/36)	20.1 (36/179)	4.9 (7/143)	79.9 (143/179)
Echogenicity										
Hypoechoic (n = 474)	74.0 (111/150)	31.6 (150/474)	75.5 (139/184)	38.8 (184/474)	83.3 (130/156)	32.9 (156/474)	71.9 (215/299)	63.1 (299/474)	20.0 (35/175)	36.9 (175/474)
Isohyperechoic (n = 428)	12.8 (5/39)	9.1 (39/428)	6.3 (2/32)	7.5 (32/428)	80.0 (4/5)	1.2 (5/428)	10.1 (7/69)	16.1 (69/428)	2.5 (9/359)	83.9 (359/428)
Combination										
Solid hypoechoic (n = 436)	74.1 (106/143)	32.8 (143/436)	77.0 (137/178)	40.8 (178/436)	84.1 (127/151)	34.6 (151/436)	73.4 (210/286)	65.6 (286/436)	21.3 (32/150)	34.4 (15/436)
Solid isohyperechoic (n = 279)	17.2 (5/29)	10.4 (29/279)	5.6 (1/18)	6.5 (18/279)	80.0 (4/5)	1.8 (5/279)	13.3 (6/45)	16.1 (45/279)	2.1 (5/234)	83.9 (234/279)
Partially cystic hypoechoic (n = 38)	71.4 (5/7)	18.4 (7/38)	33.3 (2/6)	15.8 (6/38)	60.0 (3/5)	13.2 (5/38)	38.5 (5/13)	34.2 (13/38)	12.0 (3/25)	65.8 (25/38)
Partially cystic isohyperechoic (n = 141)	0 (0/9)	6.4 (9/141)	7.1 (1/14)	9.9 (14/141)	0 (0/0)	0 (0/141)	4.3 (1/23)	16.3 (23/141)	3.4 (4/118)	83.7 (118/141)
Partially cystic or isohyperechoic (n = 466)	21.7 (10/46)	9.9 (46/466)	10.5 (4/38)	8.2 (38/466)	70.0 (7/10)	2.1 (10/466)	14.6 (12/82)	17.6 (82/466)	3.1 (12/384)	82.4 (384/466)

Korean Journal of Radiology
KJR

nodules (13.3, 38.5, 4.3, and 14.6%, respectively). Solid hypoechoic nodules with or without any suspicious US feature had a significantly higher malignancy risk than the partially cystic or isohyperechoic nodules ($p < 0.001$ and $p < 0.001$, respectively).

Malignancy Risk According to Category in K-TIRADS

In this study, the malignancy risk of thyroid nodules was stratified into the 5 categories of K-TIRADS according to US patterns by combining solidity, echogenicity, and suspicious US features (Table 4). The overall malignancy risk rates in K-TIRADS categories (nodules) 5, 4, 3, and 2 were 73.4, 19.9, 3.5, and 0.0%, respectively, and all were estimated within the range of the suggested malignancy risk in K-TIRADS.

The malignancy risk of thyroid nodules in K-TIRADS was

Table 4. Malignancy Risk According to Category in Korean Thyroid Imaging Reporting and Data System

Category	US Features	Malignant Risk (%)	Calculated Malignant Risk (%)	Frequency (%)
5 High suspicion	Solid hypoechoic nodule with any of three suspicious US features*	> 60	73.4 (210/286)	31.7 (286/902)
4 Intermediate suspicion	1) Solid hypoechoic nodule without any of three suspicious US features* 2) Partially cystic or isohyperechoic nodule with any of three suspicious US features*	15–50	19.0 (44/231)	25.6 (231/902)
3 Low suspicion	Partially cystic or isohyperechoic nodule without any of three suspicious US features*	3–15	3.5 (12/340)	37.7 (340/902)
2 Benign	1) Spongiform 2) Pure cystic or partially cystic nodule with comet-tail artifact	< 3 < 1	0.0 (0 /45)	5.0 (45/902)
1 Normal	-	-	-	

*Malignant ultrasonography (US) features were spiculated/microlobulated margins, non-parallel orientation, and microcalcifications.

Table 5. Diagnostic Performance for Prediction of Malignant Thyroid Nodules in Korean Thyroid Imaging Reporting and Data System

Category	Overall		Sensitivity	Specificity	PPV	NPV	Accuracy
	Benign	Malignancy					
Overall	636	266					
Category 5	76	210	78.9% (210/266)	88.1% (560/636)	73.4% (210/286)	90.9% (560/616)	85.4% (770/902)
Category 4 + 5	263	254	95.5% (254/266)	58.6% (373/636)	44.5% (254/517)	96.9% (373/385)	69.5% (627/902)
Category 3 + 4 + 5	591	266	100.0% (266/266)	7.1% (45/636)	31.0% (266/857)	100.0% (45/45)	34.5% (311/902)
Macro-nodules (≥ 10 mm)	485	101					
Category 5	28	67	66.3% (67/101)	94.2% (457/485)	70.5% (67/95)	93.1% (457/491)	89.4% (524/586)
Category 4 + 5	156	94	93.1% (94/101)	67.8% (329/485)	37.6% (94/250)	97.9% (329/336)	72.2% (423/586)
Category 3* + 4 + 5	347	101	100.0% (101/101)	28.5% (138/485)	22.5% (101/448)	100.0% (138/138)	40.8% (239/586)
Micro-nodules (< 10 mm)	151	165					
Category 5	48	143	86.7% (143/165)	68.2% (103/151)	74.9% (143/191)	82.4% (103/125)	77.8% (246/316)

*Category 3 nodules were included when nodule size was ≥ 15 mm. NPV = negative predictive value, PPV = positive predictive value

further stratified by size, and the nodules were divided into two groups: macronodules ≥ 10 mm and micronodules < 10 mm. In total, 544 nodules were allocated in macronodules (101 malignant nodules, 443 benign nodules) and 313 nodules were in micronodules (165 malignant nodules 148 benign nodules). The malignancy risks of macronodules in K-TIRADS categories 5, 4, 3, and 2 were 70.5, 17.3, 2.4, and 0.0%, respectively, and those of micronodules were 74.9, 22.7, 10.6, and 0.0%, respectively. The risks of malignancy in both groups were all estimated within the range of the suggested malignancy risk except for macronodules in K-TIRADS category 3, which showed a slightly lower risk of malignancy (2.4%) than the suggested malignancy risk in K-TIRADS (3–15%). The difference in malignancy risk between the macronodules and micronodules was not statistically significant in K-TIRADS categories 5, 4, or 2 (p = 0.434, 0.331, and p > 0.999, respectively); however, it was significant in K-TIRADS category 3 (p = 0.004).

Diagnostic Performance for Prediction of Malignant Thyroid Nodules in K-TIRADS

Table 5 showed the diagnostic performance for the prediction of thyroid malignancy in K-TIRADS. According to the criteria for K-TIRADS category 5, the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy for malignancy were 78.9, 88.1, 73.4, 90.9, and 85.4%, respectively. Application of the FNA criteria of K-TIRADS based on the size and category of nodules (category 4 and 5 nodules, ≥ 1.0 cm; and category 3 nodules, ≥ 1.5 cm) showed sensitivity, specificity, PPV, NPV, and accuracy for malignancy of 100.0, 28.5, 22.5, 100.0, and 40.8%, respectively.

The predicted probability of malignancy tended to rise along with the K-TIRADS categories (p < 0.001) with the Az value of 0.878 (95% CI = 0.855–0.899). The predictive power (Az = 0.878) of the K-TIRADS was significantly superior to that (Az = 0.805) of the three-tier risk categorization system (p < 0.001).

Considering the efficacy of FNA for nodules ≥ 1 cm in K-TIRADS, 7 nodules with a suspicious malignant category and 88 nodules with an indeterminate category, based on the three-tier risk categorization system, were not required to be aspirated based on the FNA criteria of K-TIRADS (95/586, 16.2%); however, all had a benign FNA result without malignancy. Thus, the efficacy of FNA based on K-TIRADS was increased from 18.6% (101/544) to 22.5% (101/449).

DISCUSSION

The current study confirmed that the three suspicious US features of microcalcifications, non-parallel orientation, and spiculated/microlobulated margin were independent predictors for malignancy; however, the malignancy risk and predictive value of suspicious US features depended on the solidity and echogenicity of the thyroid nodules. The presence of any suspicious US features showed a high malignancy risk (73.4%) in solid hypoechoic nodules, while it was a low-to-intermediate risk (4.3–38.5%) in partially cystic or isohyperechoic nodules. The results of malignancy risk of K-TIRADS confirmed the findings of a former retrospective study and increased the efficacy of FNA by avoiding unnecessary procedures.

The use of high-resolution US for thyroid disease has markedly increased the detection of thyroid nodules with an increase in the number of FNAs (1, 3-6, 20). Although many guidelines and studies have suggested several suspicious US features for decision–making in FNA, the terminology used for thyroid nodules and suspicious US features for malignancy has not been consistent among the studies and standardized report form for the results of thyroid US is lacking (1-6). Thus, a standardized and simplified report form for thyroid US has been suggested for effective communication between referring physicians and cytopathologists, as well as to increase the efficacy of FNA, avoid unnecessary procedures, and provide supplementary information for thyroid nodules after FNA (7-14).

The K-TIRADS was established recently by the KSThR as a practical and convenient risk stratification system for thyroid nodules (16). Previously reported TIRADS classifications have limitations in clinical applicability and none are adopted widely in Korea (7-10). The proposed new risk stratification system, K-TIRADS, is based on the previous KSThR guidelines, published in 2011, with some modifications in terminology and suspicious US features for simple and easy application in clinical practice (5). The malignant risk stratification of a nodule was assessed by combining solidity, echogenicity, and suspicious US features using a five-point malignancy rating scale based on retrospective study results (15, 21); however, prospective validation study has not been conducted yet. In this prospective study, we validated suspicious US features of microcalcifications, non-parallel orientation, and spiculated/microlobulated margin as independent predictors for malignancy. Thyroid nodules with isolated

Korean Journal of Radiology
KJR

macrocalcification were not predictive of malignancy (22). In addition, our results indicated that the frequency and risk of malignancy were significantly different according to solidity and echogenicity. Thyroid nodules with minimal cystic changes or isohyperechogenicity had a low risk of malignancy (23); thus, K-TIRADS, combining solidity, echogenicity, and suspicious US features, is a useful diagnostic approach for the risk stratification of thyroid nodules.

The malignancy risk and predictive value of suspicious US features were also dependent on the solidity and echogenicity of the thyroid nodules. The malignancy risks of microcalcifications and non-parallel orientation were significantly higher in solid hypoechoic nodules than those in partially cystic or isohyperechoic nodules (15, 24-26). The lower malignancy risk of partially cystic or isohyperechoic nodules may be due to several factors (15). First, we regarded punctate echogenic foci of ≤ 1 mm with or without posterior shadowing as microcalcifications; however, the origins of punctate echogenic foci could be variable, other than psammomatous microcalcification. Thus, they may emanate from the back walls of tiny unresolved cysts, which can be seen more prominently in partially cystic nodules; in addition, colloid materials could also manifest as echogenic foci within solid portion of a nodule. Second, non-parallel orientation is indicative of centrifugal growth across the normal tissue plane and decreased compressibility for thyroid cancer; however, partially cystic nodules could change orientation more readily in the confined, narrow space. Although a spiculated/microlobulated margin did not show a different malignant risk between the two groups, it has low clinical impact because it is rarely detected (2.1%, 10/466) in partially cystic or isohyperechoic nodules. As a result, the presence of any suspicious US features showed a low-to-intermediate malignant risk (4.3–38.5%) in partially cystic or isohyperechoic nodules, unlike a high malignant risk (73.4%) in solid hypoechoic nodules. Thus, partially cystic or isohyperechoic nodules should be separated from solid hypoechoic nodules in a risk categorization system.

The malignancy risk of K-TIRADS in our study corroborated the findings of a former retrospective study. The malignancy risks of each category corresponded well to the ranges of suggested malignancy risk in K-TIRADS, regardless of size. Thus, it could be applicable to thyroid nodules < 10 mm, as well as those ≥ 10 mm. However, the malignancy risk of thyroid nodules ≥ 10 mm with K-TIRADS category 3 was significant lower than that of thyroid nodules < 10 mm, and there was no case of malignancy up to 17 mm with K-TIRADS category 3. The small proportion of follicular carcinomas among malignant tumors, compared with the study by Na et al. (15) (1.5% vs. 10.6%), may be a causative factor for the lower malignant risk of thyroid nodules ≥ 10 mm with K-TIRADS category 3.

Regarding diagnostic performance, the overall sensitivity of K-TIRADS categories 4 and 5 for malignancy was 95.5%, similar to that reported by Russ et al. (10) (95.7%) and higher than those of Na et al. (15) and Horvath et al. (7) (78.8 and 88.0%, respectively). The higher sensitivity in our study may be explained by the high proportion of PTC (up to 97.4%) among the malignant tumors. Because other malignant tumors such as follicular carcinomas have no suspicious US features in many cases compared with PTC, the smaller proportion of these tumors may increase the sensitivity (27-29). The PPV of K-TIRADS categories 4 and 5 (44.5%) was similar to the PPVs of Na et al. (15) and Horvath et al. (7) (44.6 and 49%, respectively), and higher than the PPV (9%) of Russ et al. (10).

The TIRADS has a clinical role in increasing the efficacy of FNA by avoiding unnecessary procedures (7-10). Because thyroid cancers are slow-growing and less aggressive than other malignancies, recent guidelines recommend a more conservative approach for FNA of nodules < 10 mm (1). The suggested malignancy risk of thyroid nodules of K-TIRADS category 3 is low, hence, FNA is not recommended in K-TIRADS category 3 nodules < 1.5 cm, and macrocalcifications were also excluded from the suspicious US features compared with the three-tier risk categorization system (5, 16). In our study, the predictive value of K-TIRADS for malignancy was superior to that of the three-tier risk categorization system and the efficacy of FNA was increased, by avoiding 16.2% of unnecessary FNA among the indications of FNA based on the three-tier risk categorization system. Thus, K-TIRADS is a more effective tool for selecting patients than the three-tier system, allowing for better selection of nodules for FNA, avoiding unnecessary procedures, and ultimately improving patient management.

There were several limitations to this study. First, we included only thyroid nodules that had undergone US-guided FNA, which was usually performed on a thyroid nodule with suspicious US features or on the largest nodule > 10 mm if no suspicious US feature was detected. Therefore, selection bias could have led to the lower malignancy risk of thyroid nodules ≥ 10 mm with K-TIRADS category 3

in this study. Second, the final diagnoses were based on the cytopathological results, as well as surgical histology, which may cause false-negative and false-positive results. However, the probability of false diagnosis in Bethesda 2 and 6 categories is low at < 3 and < 1%, respectively, as compared to histopathology (17).

In conclusion, the proposed new risk stratification system, K-TIRADS, based on a combination of suspicious US features, solidity, and echogenicity, was useful for the risk stratification of thyroid nodules and management decisions for FNA. The malignancy risk of K-TIRADS confirmed findings of a previous retrospective study.

REFERENCES

1. 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
2. Moon WJ, Jung SL, Lee JH, Na DG, Baek JH, Lee YH, et al. Benign and malignant thyroid nodules: US differentiation--multicenter retrospective study. *Radiology* 2008;247:762-770
3. Gharib H, Papini E, Paschke R, Duick DS, Valcavi R, Hegedüs L, et al. American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and European Thyroid Association Medical guidelines for clinical practice for the diagnosis and management of thyroid nodules: executive summary of recommendations. *Endocr Pract* 2010;16:468-475
4. Perros P, Boelaert K, Colley S, Evans C, Evans RM, Gerrard Ba G, et al. Guidelines for the management of thyroid cancer. *Clin Endocrinol (Oxf)* 2014;81 Suppl 1:1-122
5. Moon WJ, Baek JH, Jung SL, Kim DW, Kim EK, Kim JY, et al. Ultrasonography and the ultrasound-based management of thyroid nodules: consensus statement and recommendations. *Korean J Radiol* 2011;12:1-14
6. Wémeau JL, Sadoul JL, d'Herbomez M, Monpeyssen H, Tramalloni J, Leteurtre E, et al. Guidelines of the French society of endocrinology for the management of thyroid nodules. *Ann Endocrinol (Paris)* 2011;72:251-281
7. Horvath E, Majlis S, Rossi R, Franco C, Niedmann JP, Castro A, et al. An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. *J Clin Endocrinol Metab* 2009;94:1748-1751
8. Park JY, Lee HJ, Jang HW, Kim HK, Yi JH, Lee W, et al. A proposal for a thyroid imaging reporting and data system for ultrasound features of thyroid carcinoma. *Thyroid* 2009;19:1257-1264
9. 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;260:892-899
10. Russ G, Royer B, Bigorgne C, Rouxel A, Bienvenu-Perrard M, Leenhardt L. Prospective evaluation of thyroid imaging reporting and data system on 4550 nodules with and without elastography. *Eur J Endocrinol* 2013;168:649-655
11. Cheng SP, Lee JJ, Lin JL, Chuang SM, Chien MN, Liu CL. Characterization of thyroid nodules using the proposed thyroid imaging reporting and data system (TI-RADS). *Head Neck* 2013;35:541-547
12. Kwak JY, Jung I, Baek JH, Baek SM, Choi N, Choi YJ, et al. Image reporting and characterization system for ultrasound features of thyroid nodules: multicentric Korean retrospective study. *Korean J Radiol* 2013;14:110-117
13. Russ G. Risk stratification of thyroid nodules on ultrasonography with the French TI-RADS: description and reflections. *Ultrasonography* 2016;35:25-38
14. Seo H, Na DG, Kim JH, Kim KW, Yoon JW. Ultrasound-based risk stratification for malignancy in thyroid nodules: a four-tier categorization system. *Eur Radiol* 2015;25:2153-2162
15. Na DG, Baek JH, Sung JY, Kim JH, Kim JK, Choi YJ, et al. Thyroid imaging reporting and data system risk stratification of thyroid nodules: categorization based on solidity and echogenicity. *Thyroid* 2016;26:562-572
16. 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
17. Cibas ES, Ali SZ. The Bethesda system for reporting thyroid cytopathology. *Thyroid* 2009;19:1159-1165
18. Choi YJ, Baek JH, Hong MJ, Lee JH. Inter-observer variation in ultrasound measurement of the volume and diameter of thyroid nodules. *Korean J Radiol* 2015;16:560-565
19. Jung CK, Min HS, Park HJ, Song DE, Kim JH, Park SY, et al. Pathology reporting of thyroid core needle biopsy: a proposal of the Korean endocrine pathology thyroid core needle biopsy study group. *J Pathol Transl Med* 2015;49:288-299
20. Lee YH, Baek JH, Jung SL, Kwak JY, Kim JH, Shin JH, et al. Ultrasound-guided fine needle aspiration of thyroid nodules: a consensus statement by the korean society of thyroid radiology. *Korean J Radiol* 2015;16:391-401
21. Brito JP, Gionfriddo MR, Al Nofal A, Boehmer KR, Leppin AL, Reading C, et al. The accuracy of thyroid nodule ultrasound to predict thyroid cancer: systematic review and meta-analysis. *J Clin Endocrinol Metab* 2014;99:1253-1263
22. 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
23. Na DG, Kim JH, Kim DS, Kim SJ. Thyroid nodules with minimal cystic changes have a low risk of malignancy. *Ultrasonography* 2016;35:153-158
24. Kim DW, Lee EJ, In HS, Kim SJ. Sonographic differentiation of partially cystic thyroid nodules: a prospective study. *AJNR Am J Neuroradiol* 2010;31:1961-1966

25. Lee MJ, Kim EK, Kwak JY, Kim MJ. Partially cystic thyroid nodules on ultrasound: probability of malignancy and sonographic differentiation. *Thyroid* 2009;19:341-346
26. Park JM, Choi Y, Kwag HJ. Partially cystic thyroid nodules: ultrasound findings of malignancy. *Korean J Radiol* 2012;13:530-535
27. Seo HS, Lee DH, Park SH, Min HS, Na DG. Thyroid follicular neoplasms: can sonography distinguish between adenomas and carcinomas? *J Clin Ultrasound* 2009;37:493-500
28. Ozdemir D, Ersoy R, Cuhaci N, Arpaci D, Ersoy EP, Korukluoglu B, et al. Classical and follicular variant papillary thyroid carcinoma: comparison of clinical, ultrasonographical, cytological, and histopathological features in 444 patients. *Endocr Pathol* 2011;22:58-65
29. Rhee SJ, Hahn SY, Ko ES, Ryu JW, Ko EY, Shin JH. Follicular variant of papillary thyroid carcinoma: distinct biologic behavior based on ultrasonographic features. *Thyroid* 2014;24:683-688