Original Article

J Korean Soc Radiol 2025;86(1):114-126
https://doi.org/10.3348/jksr.2024.0016
eISSN 2951-0805

JOURNAL of THE KOREAN SOCIETY of RADIOLOGY

Ultrasound Findings Suggestive of Malignancy in Thyroid Nodules Classified as Follicular Lesion of Undetermined Significance or Follicular Neoplasm based on the 2017 Bethesda System for Reporting Thyroid Cytopathology

2017 베데스다 시스템 갑상선 세포병리검사에서 불확정 여포성 병변 또는 여포성 종양으로 진단된 갑상선 결절에서 악성 종양을 시사하는 초음파 소견

Heui Jin Jung, MD[1], Na Lae Eun, MD[1], Eun Ju Son, MD[1], Jeong-Ah Kim, MD[1]*, Ji Hyun Youk, MD[1], Hye Sun Lee, PhD[2], Soyoung Jeon, PhD[2]

[1]Department of Radiology, Yonsei University College of Medicine, Gangnam Severance Hospital, Seoul, Korea
[2]Biostatistics Collaboration Unit, Yonsei University College of Medicine, Seoul, Korea

Received January 15, 2024
Revised May 1, 2024
Accepted August 13, 2024

***Corresponding author**
Jeong-Ah Kim, MD
Department of Radiology,
Yonsei University College of Medicine,
Gangnam Severance Hospital,
211 Eonju-ro, Gangnam-gu,
Seoul 06273, Korea.

Tel 82-2-2019-3510
Fax 82-2-3462-5472
E-mail jakim@yuhs.ac

Purpose To identify US findings suggestive of malignancy in thyroid nodules with follicular lesions of undetermined significance (FLUS) or follicular neoplasm (FN) on fine-needle aspiration cytology (FNAC) and evaluate the diagnostic performance.

Materials and Methods Seventy FLUS (*n* = 57) or FN (*n* = 13) nodules on FNAC that underwent surgical excision between February 2018 and November 2020 were selected. US findings were retrospectively reviewed. Orientation, margin, echogenicity, calcification, additional findings of the rim, echogenicity, heterogeneity of the solid portion, and the ratio of anterior posterior diameter to lateral diameter (criteria) were assessed. The diagnostic performances of US findings, criteria, and the Korean Society of Thyroid Radiology Thyroid Imaging Reporting and Data System (K-TIRADS) were evaluated using logistic regression analysis.

Results Microcalcification, homogeneous solid echotexture, and thickened rims were suggestive of malignancy. Our criteria showed a highest area under the ROC curve (AUC) value of 0.771, sensitivity of 97.14%, accuracy of 77.14%, positive predictive value of 93.33%, negative predictive value of 95.24%, and specificity of 97.14%. The criteria showed a significantly higher AUC value than K-TIRADS.
Conclusion US findings of homogenous solid portions, thick rims, and microcalcifications suggested malignancy in nodules with FLUS or FN on FNAC. These additional US findings could improve the diagnostic performance of K-TIRADS.

Index terms Thyroid Neoplasms; Thyroid Nodule; Ultrasonography; Biopsy, Fine-Needle

INTRODUCTION

US and fine-needle aspiration cytology (FNAC) are essential for the evaluation of thyroid nodules. The 2017 Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) categorizes thyroid FNAC samples into six groups, from benign (II) to malignant (VI). The categories of atypia of undetermined significance (AUS)/follicular lesions of undetermined significance (FLUS) (III) and follicular neoplasm (FN)/suspicious for a FN (IV) are indeterminate, with a risk of malignancy (ROM) of 10%–30% for category III and 25%–40% for category IV (1, 2).

US features linked to malignancy have previously been identified (3-8), with the Thyroid Imaging Reporting and Data System (TIRADS) and the Korean TIRADS (K-TIRADS) focusing on microcalcification, non-parallel orientation, and spiculated/microlobulated margins as malignancy markers (4-6).

Thyroid FN is an interpretative rather than a specific diagnosis. A diagnosis of thyroid FN is implied when the FNAC results show FN or FLUS. FN or FLUS is an indeterminate diagnosis in thyroid cytopathology and is typically applied to specimens with at least moderate cellularity and microfollicular architecture (9). FNAC has a limited role in the diagnosis of FN or FLUS because a definitive diagnosis is possible only after assessing the invasion of the entire tumor capsule, and surgical dissection is usually recommended for a definite diagnosis (1, 10).

US diagnosis of FN is challenging because of the overlapping features of benign and malignant pathologies (11-15). Despite attempts to differentiate between follicular adenoma (FA) and follicular carcinoma (FC) using US, the effectiveness of current classification systems such as the ATA, K-TIRADS, and ACR-TIRADS in stratifying the risk of Bethesda category IV nodules is limited (16). Therefore, this study aimed to identify additional US features suggestive of malignancy in nodules suggestive of thyroid FNs diagnosed as FN/FLUS on FNAC.

The purposes of this study were to identify US findings suggestive of malignancy in nodules classified as FLUS or FN via FNAC and to assess the diagnostic accuracy of these US features.

MATERIALS AND METHODS

STUDY POPULATIONS

This retrospective study was conducted at our thyroid cancer center between February

JOURNAL of THE KOREAN SOCIETY of RADIOLOGY

2018 and November 2020. The study was approved by the Institutional Review Board of the Yonsei University College of Medicine (IRB No. 3-2023-0299), and the requirement for informed consent or patient approval was waived. This retrospective analysis was conducted using data from the FNAC database. This study analyzed patients who underwent FNAC at our institution, were diagnosed with FLUS of TBSRTC categories III or IV (FN), and subsequently underwent surgery with pathological confirmation at our institution. Cases with only AUS of TBSRTC III at the first FNAC were excluded to evaluate thyroid FNs.

From the FNAC database, 3831 FNA procedures were performed during the study period. We extracted 270 FLUS of TBSRTC III and 24 TBSRTC IV FNAC results, with at least one FLUS of TBSRTC III or IV results. Surgical excision was performed if the FNA results were TBSRTC IV, more than two consecutive FNA results of FLUS, or one FLUS result with suspicious US findings (K-TIRADS 4 or 5). Surgical excision was performed in 99 patients with FLUS and 23 patients with FN. Among them, 70 patients with 70 nodules >10 mm were included in this study, which included TBSRTC IV (n = 13), more than two consecutive FNA results of TBSRTC III (FLUS, n = 48), or one FLUS result with suspicious US findings (n = 9), comprising 55 female patients and 15 male patients with a mean age of 45.5 years (range: 25–77).

US FINDINGS ANALYSIS

US findings were retrospectively reviewed without knowledge of the histopathological diagnosis by two radiologists with 25 and 3 years of experience.

The US equipment used included the following: ALXPLORER (Supersonic Imagine, Aix-en-Provence, France), ALXPLORER ULTIMATE (Supersonic Imagine), and EPIQ7 (Philips, Bothell, WA, USA). We retrospectively analyzed the US findings of the nodules, and our FNAC database was prepared by one of six board-certified radiologists with more than five years of experience. The US parameters of the database included size, composition_1 (cystic, mixed (solid portion <50%), predominantly solid [solid portion >50%]), echogenicity (hyperechoic, isoechoic, hypoechoic, markedly hypoechoic), margin (well, irregular microlobulated, irregular (infiltrative)), calcification (negative, macro, micro), orientation (parallel, non-parallel), and vascularity (none, central, peripheral, both).

Additional US findings were retrospectively analyzed by the two aforementioned radiologists, and a consensus was reached. We analyzed additional US findings, focusing on the homogeneity of the solid portion, shape, and rim of the nodules from our clinical experiences. Homogeneity was assessed in terms of the echotexture and echogenicity of the solid portion (Figs. 1, 2). These additional findings included homogeneity of solid portion in echotexture and echogenicity; (heterogeneous, homogeneous), echogenicity of solid portion (hyperechoic, isoechoic, hypoechoic, markedly hypoechoic), composition_2 (cystic >50%, 10% < cystic <50%, cystic <10%, no cyst), composition 3 (cystic, no cystic), anterior-posterior (AP, mm) diameter and lateral (LAT, mm) diameter on the transverse image, ratio 1 (AP/LAT on transverse image), shortest diameter (S, mm) and longest diameter (L, mm) of nodules on any plane, ratio 2 (S/L), rim margin (no rim, regular thickening, irregular thickening), rim thickness (maximal thickness of rim measured on a transverse plane, mm), and rim echogenicity compared with thyroid parenchyma (iso- to hyperechoic, hypoechoic, markedly hypoechoic).

Our criteria (benign, low suspicious, suspicious) were obtained by incorporating these addi-

Fig. 1. The US findings of FVPTC in a 29-year-old female patient.
A, B. A 5.3-cm sized, oval shape, solid mass showing relatively homogeneous hypoechogenicity and markedly hypoechoic rim (arrows) with 3-mm thickness is noted in the left thyroid gland (A: Transverse, B: Longitudinal image). US category was suspicious based on our criteria and K-TIRADS 4 (intermediate suspicious). Fine needle aspiration cytology indicated category IV (follicular neoplasm or suspicious for a follicular neoplasm). Left total thyroidectomy was performed and the pathological diagnosis was FVPTC with infiltrative type.
FVPTC = follicular variant of papillary thyroid carcinoma, K-TIRADS = Korean Society of Thyroid Radiology Thyroid Imaging Reporting and Data System

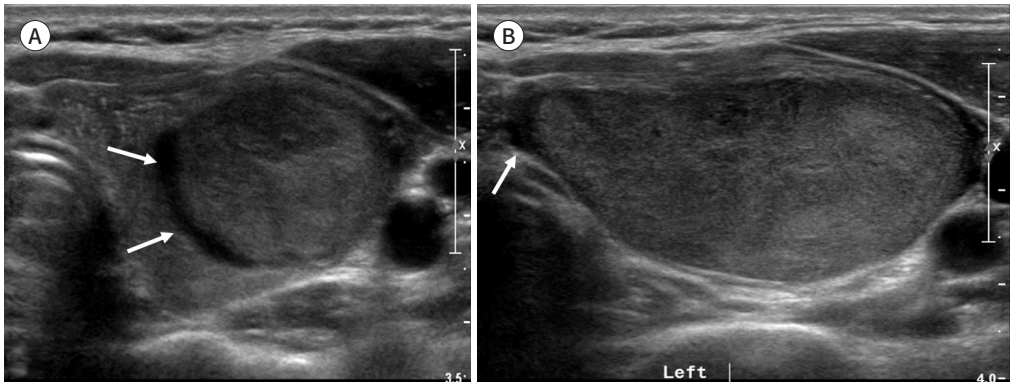
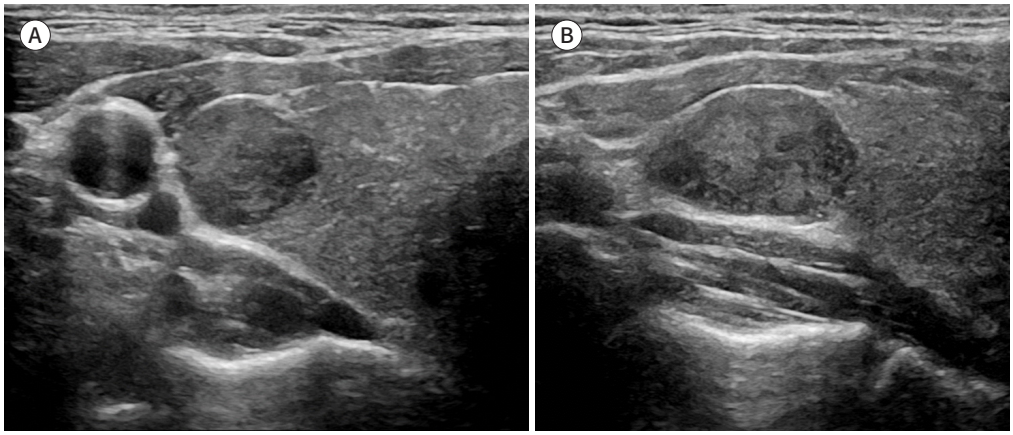


Fig. 2. The US findings of follicular adenoma in a 45-year-old male patient.
A, B. A 2-cm sized, oval shape, solid nodule showing heterogeneous echogenicity and a thin, regular, hypoechoic rim is noted in the right thyroid (A: Transverse, B: Longitudinal image). The US category was probably benign based on our criteria and K-TIRADS 4 (intermediate suspicious). Fine needle aspiration cytology indicated category III (follicular lesion of undetermined significance with focal nuclear atypia). Right thyroid lobectomy with frozen sectional diagnosis was performed and the pathological diagnosis was follicular adenoma.
K-TIRADS = Korean Society of Thyroid Radiology Thyroid Imaging Reporting and Data System



tional findings into the original US findings assessed by FNA performers. Benign nodules were those with iso/hyperechoic solid portions, or heterogeneous echoic nodules without any suspicious findings (irregular rims, microcalcifications, non-parallel orientation, or marked hypoechogenicity). The low-suspicious nodules were homogeneously hypoechoic, without any suspicious findings. Suspicious nodules were homogeneously hypoechoic with one of the suspicious findings. K-TIRADS was retrospectively assessed to compare the diagnostic performance of our criteria.

JOURNAL of
THE KOREAN SOCIETY of
RADIOLOGY

HISTOPATHOLOGIC FINDINGS ANALYSIS

The surgical histopathological results of the nodules that underwent FNA biopsy were retrospectively analyzed 3 months after US imaging for blinded analysis of the US images. The pathological diagnosis of the nodules that underwent FNAC was determined by comparing the locations and sizes of the nodules mentioned in the pathological reports. The final diagnosis was categorized as either benign or malignant based on surgical histopathological reports.

STATISTICAL ANALYSIS

US findings were compared with surgical pathology (benign vs. malignant) using an independent two-sample *t*-test or Chi-square test (or Fisher's exact test). All analyses were performed using SAS version 9.2 (SAS Institute, Cary, NC, USA). Statistical significance was set at $p < 0.05$. The first step involved identifying US findings that predict malignancy through univariate analysis, including patient sex, age, US findings, and surgical pathology. Logistic regression was then used to identify predictors of malignancy and to calculate odds ratios (OR) for the relevant findings. Subsequently, multivariate analysis was conducted using significant variables from the univariate model, excluding sex and age. The diagnostic performances of each US finding, criterion, and K-TIRADS in predicting malignancy were evaluated using logistic regression analysis. The diagnostic performance of each criterion was compared based on the area under the receiver operator characteristic curve curve (AUC), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy.

RESULTS

HISTOPATHOLOGICAL DIAGNOSIS

The FNAC results for the 70 nodules were 57 nodules classified as category III (FLUS) and 13 nodules classified as category IV (FN). Among the FLUS nodules, surgical histopathological diagnosis revealed 29 benign nodules (50.9%), including 19 cases of adenomatous hyperplasia (27.0%), 5 cases of FA (7.1%), 2 cases of Hurthle cell adenoma (2.8%), 2 cases of lymphocytic thyroiditis (2.8%), and 1 case of fibrocystic nodule (1.4%). Additionally, 28 FLUS nodules were malignant (49.1%), consisting of 18 cases of papillary thyroid carcinoma (PTC) (25.7%), 7 cases of follicular variant of papillary thyroid carcinoma (FVPTC) (10%), 1 case of medullary carcinoma, 1 case of noninvasive follicular thyroid tumor with papillary-like nuclear features (NIFTP) (1.4%), and 1 case of follicular carcinoma (1.4%).

Among the category IV nodules, 6 nodules were benign (46.2%), including three cases of adenomatous hyperplasia (4.3%), two cases of FA (2.8%), and one case of Hurthle cell adenoma (1.4%). In contrast, seven nodules were malignant (53.8%), comprising five cases of PTC (7.1%) and two cases of FVPTC (2.8%).

US FINDINGS SUGGESTING MALIGNANCY

Table 1 presents the US findings in relation to the histopathological diagnosis (benign vs. malignant). Table 1 presents the US findings that exhibited a significant correlation with the histopathological diagnosis ($p < 0.05$). These findings included echogenicity, margin, calcification, homogeneity of the solid portion, echogenicity of the solid portion, composition_2,

Table 1. US Findings according to Histopathological Diagnosis

Variables	Overall (*n* = 70)	Benign (*n* = 35)	Malignant (*n* = 35)	*p*-Value
Sex				0.771
Female	55 (78.57)	27 (77.14)	28 (80.00)	
Male	15 (21.43)	8 (22.86)	7 (20.00)	
Age, yrs	45.486 ± 11.589	47.143 ± 11.340	43.829 ± 11.761	0.234
Echogenicity				0.002
Hyperechoic	None			
Isoechoic	26 (37.14)	18 (51.43)	8 (22.86)	
Hypoechoic	37 (52.86)	17 (48.57)	20 (57.14)	
Markedly hypoechoic	7 (10.00)	0 (0.00)	7 (20.00)	
Margin				0.005
Well	52 (74.29)	31 (88.57)	21 (60.00)	
Microlobulated	11 (15.71)	4 (11.43)	7 (20.00)	
Irregular	7 (10.00)	0 (0.00)	7 (20.00)	
Calcification				<0.001
Negative	50 (71.43)	32 (91.43)	18 (51.43)	
Macrocalcification	11 (15.71)	3 (8.57)	8 (22.86)	
Microcalcification	9 (12.86)	0 (0.00)	9 (25.71)	
Homogeneity of solid portion				0.001
Heterogeneous	47 (67.14)	30 (85.71)	17 (48.57)	
Homogeneous	23 (32.86)	5 (14.29)	18 (51.43)	
Echogenicity of solid portion				0.002
Hyperechoic	None			
Isoechoic	22 (31.43)	17 (48.57)	5 (14.29)	
Hypoechoic	30 (42.86)	14 (40.00)	16 (45.71)	
Markedly hypoechoic	18 (25.71)	4 (11.43)	14 (40.00)	
Composition 2				0.046
Cystic >50%	1 (1.43)	0 (0.00)	1 (2.86)	
10< cystic <50%	11 (15.71)	8 (22.86)	3 (8.57)	
Cystic <10%	11 (15.71)	8 (22.86)	3 (8.57)	
No cyst	47 (67.14)	19 (54.29)	28 (80.00)	
Composition 3				0.032
Cystic	22 (31.88)	15 (44.12)	7 (20.00)	
No cystic	47 (68.12)	19 (55.88)	28 (80.00)	
Ratio 1 (AP/LAT)	0.677 ± 0.198	0.630 ± 0.148	0.724 ± 0.231	0.047
Ratio 2 (shortest/longest)	0.570 ± 0.248	0.503 ± 0.138	0.638 ± 0.309	0.022
Rim margin				<0.001
No	13 (18.57)	7 (20.00)	6 (17.14)	
Regular	28 (40.00)	22 (62.86)	6 (17.14)	
Irregular	29 (41.43)	6 (17.14)	23 (65.71)	
Rim Thickness, mm	1.187 ± 1.235	0.457 ± 0.412	1.917 ± 1.352	<0.001
Rim echogenicity				<0.001
Hyperechoic	None			
Isoechoic	3 (5.26)	2 (7.14)	1 (3.45)	
Hypoechoic	22 (38.60)	19 (67.86)	3 (10.34)	
Markedly hypoechoic	32 (56.14)	7 (25.00)	25 (86.21)	

Data are presented as mean ± SD or *n* (%) values.
AP = anterior-posterior diameter, LAT = lateral diameter, SD = standard deviation

JOURNAL of THE KOREAN SOCIETY of RADIOLOGY

ratios 1 and 2, rim margin, rim thickness, and rim echogenicity (Table 1).

Univariate analysis using logistic regression was conducted to identify US findings indicating malignant pathology (Table 2). The results (Table 2) revealed several US findings that significantly correlated with malignancy ($p < 0.05$), including marked hypoechogenicity, microcalcification, homogenous solid portion, marked hypoechogenicity of the solid portion, ratio 2, irregular rim, and rim thickness.

Table 2. Urinvariable Analysis of US Findings for Predicting Malignant Pathology

Variables	OR (95% CI)	*p*-Value	Variables	OR (95% CI)	*p*-Value
Sex			Echogenicity of solid portion 1		
Female	Ref		Iso	Ref	
Male	0.844 (0.269–2.649)	0.771	Hypo	3.886 (1.138–13.271)	0.030
Age	0.975 (0.935–1.017)	0.233	Markedly hypo	11.900 (2.674–52.959)	0.001
FNA size	1.007 (0.971–1.045)	0.690	Echogenicity of solid portion 2		
Composition 1			Iso	Ref	
Mixed	Ref		Hypo & markedly hypo	5.667 (1.784–17.999)	0.003
Predominantly solid	4.387 (0.465–41.402)	0.197	Composition 2		
Echogenicity			Cystic >50%	Ref	
Iso	Ref		10< cystic <50%	0.137 (0.001–15.234)	0.408
Hypo	2.550 (0.892–7.287)	0.081	Cystic <10%	0.137 (0.001–15.234)	0.408
Markedly hypo	32.653 (1.371–777.812)	0.031	No cyst	0.487 (0.005–46.741)	0.757
Margin			AP diameter, mm	1.038 (0.964–1.117)	0.322
Well	Ref		LAT diameter, mm	1.012 (0.969–1.058)	0.581
Microlobulated	2.442 (0.639–9.328)	0.192	Ratio 1 (AP/LA)	14.080 (0.950–208.705)	0.055
Irregular	21.981 (0.980–492.968)	0.052	Shortest diameter, mm	1.047 (0.978–1.121)	0.185
Calcification			Longest diameter, mm	1.011 (0.975–1.048)	0.555
Negative	Ref		Ratio 2 (shortest/longest)	95.126 (2.191–4130.513)	0.018
Macro	4.266 (1.029–17.690)	0.046	Rim margin 1		
Micro	33.368 (1.577–706.105)	0.024	No	Ref	
Orientation			Regular	0.318 (0.077–1.311)	0.113
Parallel	Ref		Irregular	4.472 (1.089–18.371)	0.038
Nonparallel	3.414 (0.639–18.248)	0.151	Rim margin 2		
Vascularity			No & regular	Ref	
No	Ref		Irregular	9.264 (3.015–28.461)	<0.001
Central	6.431 (0.117–353.936)	0.363	Rim thickness, mm	5.201 (2.265–11.939)	<0.001
Peripheral	1.800 (0.340–9.522)	0.489	Rim echogenicity		
Both	0.890 (0.171–4.640)	0.890	Iso	Ref	
Underlying echo			Hypo	0.316 (0.021–4.660)	0.401
Homogeneous	Ref		Markedly hypo	7.143 (0.562–90.802)	0.130
Heterogeneous	1.168 (0.391–3.488)	0.781			
Homogeneity of solid portion					
Heterogeneous	Ref				
Homogeneous	6.353 (2.000–20.179)	0.002			

AP = anterior-posterior diameter, CI = confidence interval, FNA = fine-needle aspiration, LA = LAT lateral, LAT = lateral, OR = odds ratio, Ref = reference

Multivariate analysis was conducted using significant US variables ($p < 0.05$) identified in the univariate analysis (Table 3). Regarding the echogenicity of the solid portion, hypoechogenicity and marked hypoechogenicity were combined into one group, whereas no rims and regular rims were grouped for statistical analysis. Microcalcification, homogeneous echogenicity of the solid portion, and rim thickness were identified as findings suggestive of malignancy (Fig. 1).

According to the criteria presented in this study, 21 cases were classified as benign, 34 as low suspicion, and 15 as suspicious.

DIAGNOSTIC PERFORMANCES OF US FINDINGS TO SUGGEST MALIGNANCY

To evaluate the diagnostic performance of each US finding in predicting malignancy, the AUC, sensitivity, specificity, accuracy, PPV, and NPV were calculated (Table 4). As rim thickness is a continuous variable, a cut-off point was determined using the Youden index, and a cut-off value of 0.7 mm was selected for rim thickness. Among the findings, rim thickness (≥0.7 mm) demonstrated the highest AUC (0.829), accuracy (82.86%), and NPV (79.49%). Microcalcification exhibited the highest specificity (100%) and PPV (100%).

To assess the diagnostic performance of each criterion for predicting malignancy, the AUC, sensitivity, specificity, accuracy, PPV, and NPV were determined (Table 5). The criteria were categorized into two groups for evaluation (Table 5: benign, low suspicious vs. suspicious or

Table 3. Multivariable Analysis of US Findings to Predict Malignant Pathology

US Findings	OR (95% CI)	*p*-Value
Echogenicity		
Iso	Ref	
Hypoechoic	17.794 (0.128–2468.395)	0.253
Markedly hypoechoic	138.489 (0.252–76187.66)	0.126
Calcification		
Negative	Ref	
Macrocalcificaiton	6.923 (0.435–110.258)	0.171
Microcalcification	83.739 (1.905–3681.903)	0.022
Homogeneity of solid portion		
Heterogeneous	Ref	
Homogeneous	9.340 (1.496–58.312)	0.017
Echogenicity of solid portion		
Isoechoic	Ref	
Hypo/markedly hypoechoic	0.373 (0.004–36.319)	0.673
Ratio 2 (shortest/longest)	2.220 (0.099–49.691)	0.615
Rim margin		
No & regular	Ref	
Irregular	0.460 (0.017–12.762)	0.647
Rim thickness, mm	10.127 (1.137–90.174)	0.038

CI = confidence interval, OR = odds ratio, Ref = reference

Table 4. Diagnostic Performances of Each US Finding in Predicting Malignancy

Variables	Value	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)	PPV (95% CI)	NPV (95% CI)
Calcification	Micro	0.629 (0.555–0.702)	25.71 (11.23–40.19)	100 (100–100)	62.86 (51.54–74.18)	100 (100–100)	57.38 (44.97–69.79)
	Macro & micro	0.700 (0.604–0.796)	48.57 (32.01–65.13)	91.43 (82.16–100)	70.00 (59.26–80.74)	85.00 (69.35–100)	64.00 (50.70–77.30)
Homogeniety of solid portion	Homogeneous	0.686 (0.583–0.788)	51.43 (34.87–67.99)	85.71 (74.12–97.30)	68.57 (57.69–79.45)	78.26 (61.40–95.12)	63.83 (50.09–77.57)
Rim_thickness	≥0.7 (cut-off)	0.829 (0.740–0.917)	77.14 (63.23–91.05)	88.57 (78.03–99.11)	82.86 (74.03–91.69)	87.10 (75.30–98.90)	79.49 (66.82–92.16)

Since thickness is a continuous variable, the cut-off point was obtained and analyzed using the Youden index.
AUC = area under the receiver operator characteristic curve, CI = confidence interval, NPV = negative predictive value, PPV = positive predictive value

Table 5. Diagnostic Performances of Each Criterion in Predicting Malignancy

Variables	Value	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)	PPV (95% CI)	NPV (95% CI)
K-TIRADS (1)	Low suspicious & Suspicious	0.629 (0.521–0.736)	80.00 (66.75–93.25)	45.71 (29.21–62.21)	62.86 (51.54–74.18)	59.57 (45.54–73.60)	69.57 (50.77–88.37)
Criteria (2)	Suspicious	0.686 (0.599–0.773)	40.00 (23.77–56.23)	97.14 (91.62–100)	68.57 (57.69–79.45)	93.33 (80.70–100)	61.82 (48.98–74.66)
Criteria (3)	Low suspicious & Suspicious	0.771 (0.684–0.859)	97.14 (91.62–100)	57.14 (40.74–73.54)	77.14 (67.30–86.98)	69.39 (56.49–82.29)	95.24 (86.13–100)
(1) vs. (2)		0.341	<0.001	<0.001	0.478	<0.001	0.309
(1) vs. (3)		0.015	0.023	0.195	0.014	0.039	0.013
(2) vs. (3)		0.151	<0.001	<0.001	0.300	0.002	<0.001

AUC = area under the receiver operator characteristic curve, CI = confidence interval, K-TIRADS = Korean Society of Thyroid Radiology Thyroid Imaging Reporting and Data System, NPV = negative predictive value, PPV = positive predictive value

benign vs. low suspicious, suspicious). The K-TIRADS scores were analyzed by dividing the categories into 2, 3 and 4, 5.

Focusing on the US assessment results, the criteria (benign vs. low suspicious, suspicious) exhibited a highest AUC value of 0.771, sensitivity of 97.14%, accuracy of 77.14%, and NPV of 95.24%. The highest specificity was observed for criteria (97.14%, benign, low suspicious, vs. suspicious). In terms of predicting malignancy, criteria displayed the highest PPV (93.33%; benign, low suspicious, vs. suspicious).

Notably, criteria exhibited a significantly higher AUC value compared with K-TIRADS.

DISCUSSION

Based on histological findings, thyroid FN represents a wide range of potential diagnoses, ranging from hyperplastic nodules to FC (10).

FNA is the initial diagnostic modality for thyroid nodules; however, it has limitations; notably, the potential for inconclusive outcomes in Bethesda categories III and IV and non-diagnostic results. Surgical resection or core needle biopsy is typically advised for nodules cate-

gorized as FLUS or FN through FNAC because of the limitations of FNAC in evaluating crucial features such as capsular or vascular invasion (1, 2).

Several studies have investigated US features that can help differentiate between FA and FC and reported iso/hypoechoic echogenicity, predominantly solid or mixed echotexture, microcalcifications, rim calcifications, and absent or irregular thick halos as US findings more commonly associated with FCs (11-14). However, US findings are often similar between benign and malignant cases. FNs typically appear as well-defined, iso/hypoechoic solid nodules that do not exhibit suspicious US findings, such as irregular margins, taller-than-wide shape, or microcalcifications (17-20). Therefore, additional US findings are necessary for the evaluation of nodules with indeterminate FNAC results of FLUS or FNs. We conducted this study focusing on the US findings of homogeneity of solid portion, orientation, and rim of the nodules from our clinical experiences.

We conducted a meticulous analysis of the US findings and examined additional US features, including homogeneity of the solid portion, ratios of nodular diameters, margin, thickness, and echogenicity of nodular rims, to identify the predictors of malignancy in FNs. Our findings revealed that homogenous hypoechogenicity of the solid portion, increased ratio 2 (S/L diameter), and irregular and thick rims were suggestive of malignancy, in addition to the classic suspicious findings of marked hypoechogenicity and microcalcifications (Fig. 1). Microcalcifications, homogeneous echogenicity of the solid portion, and a thick rim were identified as significant predictors of malignancy through multivariate analysis.

In this study, the overall malignancy rate based on surgical pathology was 50% (49.1% for FLUS and 53.8% for FN). The relatively high malignancy rate was attributed to a potential selection bias, as we only included nodules diagnosed by surgical histopathology. The malignant diagnoses included 23 cases of PTCs (65.7%), 9 cases of FVPTCs (25.7%), 1 case of medullary carcinoma (2.8%), 1 case of NIFTP (2.8%), and 1 case of FC (2.8%).

FVPTC is the second most common subtype of PTC, accounting for approximately 9%–22% of all PTC cases (17). US findings of FVPTCs show a relatively benign appearance compared with classic PTC, with a smooth margin, isoechogenicity, and an oval/round shape, which are similar to other follicular cell tumors (18). Recently, non-invasive, encapsulated FVPTCs were reclassified as noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP), which can be considered a low-risk, "in situ" or "dysplastic" process that can be adequately treated with lobectomy (2, 10, 17). NIFTP lacks malignant features on US and exhibits more benign characteristics, such as hyper- or isoechogenicity, circumscribed margins, and the absence of calcification, compared with non-NIFTP lesions (19, 20).

Distinguishing between benign and malignant FNs using US alone can be challenging, as FNs often lack the suspicious US findings commonly observed in classic PTCs. In our study, we identified additional US features–homogeneous solid echogenicity and a thick or irregular nodular rim–which may serve as indicators for predicting malignancy in nodules diagnosed as FLUS or FNs on FNAC. Our detailed US criteria could improve diagnostic performance in predicting malignancy.

We observed that the diagnostic performance of our criteria, which categorized nodules as low suspicious or higher based on additional US features, was superior to that of the K-TI-RADS (4 or higher) system (Table 5). Our criteria demonstrated higher AUC, sensitivity, accu-

JOURNAL of THE KOREAN SOCIETY of RADIOLOGY

racy, PPV, and NPV values than the K-TIRADS system.

A recent study reported that TIRADS has limitations in the categorization of other malignancies, such as FVPTC, FN, medullary thyroid carcinoma, and lymphoma, as category 5. This challenge prompted the development of the follicular TIRADS (F-TIRADS) to refine these shortcomings (21).

This study has several limitations. First, it was a retrospective study with a relatively small number of cases. Second, the inclusion criteria were limited to nodules diagnosed as only FLUS or FN using FNAC and confirmed by surgical histopathology, which introduced a selection bias and a high rate of malignancy. However, confirming the diagnosis of FNs is possible with surgical excision; therefore, there could be an inevitable selection bias. Moreover, only nodules with FLUS cytology were included among those with AUS/FLUS cytology, which could affect the generalizability of the findings. Third, classic PTCs were the most common malignancy, and US findings may be influenced by this population. FNAC diagnosis of AUS/FLUS or FN is inconclusive, and there are overlapping features and interobserver variability among cytopathologists (1, 2). This could result in a high prevalence of classic PTCs. Future prospective multicenter studies are required to validate our observations. Fourth, one case of NIFTP was classified as malignant in the statistical analysis. This classification might have influenced the overall results. Therefore, a larger prospective study is required to validate our findings. Despite these limitations, our detailed US evaluation provides valuable insights for predicting malignancy and guiding the management of thyroid nodules with indeterminate FLUS or FN diagnoses using FNAC.

In conclusion, US findings of homogeneous solid echogenicity, thick rim, and microcalcifications were suggestive of malignancy in thyroid nodules diagnosed using FNAC as FLUS or FN. US criteria incorporating these additional features have the potential to improve the diagnostic performance of K-TIRADS for thyroid FNs.

Author Contributions

Conceptualization, K.J.; data curation, J.H.J.; formal analysis, L.H.S., J.S.; supervision, K.J.; writing—original draft, J.H.J., K.J.; and writing—review & editing, J.H.J., E.N.L., S.E.J., K.J., Y.J.H.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

ORCID iDs

Heui Jin Jung https://orcid.org/0009-0002-8446-1228
Na Lae Eun https://orcid.org/0000-0002-7299-3051
Eun Ju Son https://orcid.org/0000-0002-7895-0335
Jeong-Ah Kim https://orcid.org/0000-0003-4949-4913
Ji Hyun Youk https://orcid.org/0000-0002-7787-780X
Hye Sun Lee https://orcid.org/0000-0001-6328-6948
Soyoung Jeon https://orcid.org/0000-0002-9916-1917

Funding

None

REFERENCES

1. Cibas ES, Ali SZ. The 2017 Bethesda system for reporting thyroid cytopathology. *Thyroid* 2017;27:1341-

1346
2. Cibas ES, Ali SZ. The 2017 Bethesda system for reporting thyroid cytopathology. *J Am Soc Cytopathol* 2017;6:217-222
3. Kim EK, Park CS, Chung WY, Oh KK, Kim DI, Lee JT, et al. New sonographic criteria for recommending fine-needle aspiration biopsy of nonpalpable solid nodules of the thyroid. *AJR Am J Roentgenol* 2002;178:687-691
4. 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
5. 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
6. Ha EJ, Chung SR, Na DG, Ahn HS, Chung J, Lee JY, et al. 2021 Korean thyroid imaging reporting and data system and imaging-based management of thyroid nodules: Korean Society of Thyroid Radiology consensus statement and recommendations. *Korean J Radiol* 2021;22:2094-2123
7. Gweon HM, Son EJ, Youk JH, Kim JA. Thyroid nodules with Bethesda system III cytology: can ultrasonography guide the next step? *Ann Surg Oncol* 2013;20:3083-3088
8. Hong HS, Lee JY. Diagnostic performance of ultrasound patterns by K-TIRADS and 2015 ATA guidelines in risk stratification of thyroid nodules and follicular lesions of undetermined significance. *AJR Am J Roentgenol* 2019;213:444-450
9. Wu S, DeMay RM, Papas P, Yan B, Reeves W. Follicular lesions of the thyroid: a retrospective study of 1,348 fine needle aspiration biopsies. *Diagn Cytopathol* 2012;40(Suppl 1):E8-E12
10. Ohori NP, Nishino M. Follicular neoplasm of thyroid revisited: current differential diagnosis and the impact of molecular testing. *Adv Anat Pathol* 2023;30:11-23
11. 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
12. Zhang JZ, Hu B. Sonographic features of thyroid follicular carcinoma in comparison with thyroid follicular adenoma. *J Ultrasound Med* 2014;33:221-227
13. Sillery JC, Reading CC, Charboneau JW, Henrichsen TL, Hay ID, Mandrekar JN. Thyroid follicular carcinoma: sonographic features of 50 cases. *AJR Am J Roentgenol* 2010;194:44-54
14. Williams MD, Suliburk JW, Staerkel GA, Busaidy NL, Clayman GL, Evans DB, et al. Clinical significance of distinguishing between follicular lesion and follicular neoplasm in thyroid fine-needle aspiration biopsy. *Ann Surg Oncol* 2009;16:3146-3153
15. Najafian A, Olson MT, Schneider EB, Zeiger MA. Clinical presentation of patients with a thyroid follicular neoplasm: are there preoperative predictors of malignancy? *Ann Surg Oncol* 2015;22:3007-3013
16. Yang W, Fananapazir G, LaRoy J, Wilson M, Campbell MJ. Can the American Thyroid Association, K-TIRADS, and ACR-TIRADS ultrasound classification systems be used to predict malignancy in Bethesda category IV nodules? *Endocr Pract* 2020;26:945-952
17. Passler C, Prager G, Scheuba C, Niederle BE, Kaserer K, Zettinig G, et al. Follicular variant of papillary thyroid carcinoma: a long-term follow-up. *Arch Surg* 2003;138:1362-1366
18. Yoon JH, Kwon HJ, Kim EK, Moon HJ, Kwak JY. The follicular variant of papillary thyroid carcinoma: characteristics of preoperative ultrasonography and cytology. *Ultrasonography* 2016;35:47-54
19. Hahn SY, Shin JH, Lim HK, Jung SL, Oh YL, Choi IH, et al. Preoperative differentiation between noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) and non-NIFTP. *Clin Endocrinol* (Oxf) 2017;86:444-450
20. Yang GCH, Fried KO, Scognamiglio T. Sonographic and cytologic differences of NIFTP from infiltrative or invasive encapsulated follicular variant of papillary thyroid carcinoma: a review of 179 cases. *Diagn Cytopathol* 2017;45:533-541
21. Li J, Li C, Zhou X, Huang J, Yang P, Cang Y, et al. US risk stratification system for follicular thyroid neoplasms. *Radiology* 2023;309:e230949

JOURNAL of
THE KOREAN SOCIETY of
RADIOLOGY

2017 베데스다 시스템 갑상선 세포병리검사에서 불확정 여포성 병변 또는 여포성 종양으로 진단된 갑상선 결절에서 악성 종양을 시사하는 초음파 소견

정희진[1] · 은나래[1] · 손은주[1] · 김정아[1]* · 육지현[1] · 이혜선[2] · 전소영[2]

목적 세침 흡인 검사에서 불확정 여포성 병변 또는 여포성 종양으로 진단된 갑상선 결절에서 악성 가능성을 예측할 수 있는 초음파 소견을 식별하고, 그러한 소견의 진단 성능을 평가한다.

대상과 방법 2018년 2월부터 2020년 11월까지 시행된 갑상선세침검사에서, 불확정 여포성 병변(n = 57)과 여포성 종양(n = 13)으로 진단되고, 수술로 조직학적 진단이 확인된 총 70개의 갑상선 결절을 대상으로 분석하였다. 결절의 방향, 경계, 에코, 석회화 소견과 추가적으로 테두리, 고형 성분의 에코와 이질성, 전후좌우 직경의 비율을 후향적으로 분석하였다. 각 초음파 소견과 본 연구에서 제시한 진단기준 및 K-TIRADS의 악성 예측의 진단 성능을 로지스틱 회귀 분석을 통해 평가하였다.

결과 균질한 고형 에코, 두꺼운 테두리, 그리고 미세 석회화는 악성을 예측하는 초음파 소견으로 확인되었다. 진단기준은 AUC 값이 0.771로 가장 높았으며, 민감도 97.14%, 정확도 77.14%, 양성 예측값 93.33%, 음성 예측값 95.24%, 특이도 97.14%를 나타냈다. 본 연구의 진단기준은 K-TIRADS 보다 높은 AUC 값을 보였다.

결론 세침흡인검사에서 불확정 여포성 병변 또는 여포성 종양으로 진단된 갑상선 결절에서 균질한 고형 에코, 두꺼운 테두리, 미세 석회화는 악성을 예측하는 초음파 소견으로 판명되었다. 추가적인 초음파 소견의 활용은 K-TIRADS와 비교하여 더 높은 진단 성능을 제공할 수 있다.

[1]강남세브란스병원 영상의학과,
[2]연세대학교 통계지원팀