ORIGINAL ARTICLE

Comparison among TIRADS (ACR TI-RADS and KWAK- TI-RADS) and 2015 ATA Guidelines in the diagnostic efficiency of thyroid nodules

Luying Gao[1] · **Xuehua Xi**[1] · **Yuxin Jiang**[1] · **Xiao Yang**[1] · **Ying Wang**[1] · **Shenling Zhu**[1] · **Xingjian Lai**[1] · **Xiaoyan Zhang**[1] · **Ruina Zhao**[1] · **Bo Zhang** [1]

Abstract
Purpose To compare the value of Thyroid Imaging Reporting and Data Systems proposed by Kwak (KWAK-TIRADS) and ACR (ACR TI-RADS) and 2015 American Thyroid Association (ATA) guidelines in the diagnosis of surgically resected thyroid nodules.
Methods From January 2015 to December 2015, 2544 thyroid nodules in 1758 patients who underwent thyroidectomy at our center were included. The KWAK-TIRADS category, ACR TI-RADS and ultrasound (US) pattern based on ATA guidelines were assigned to each thyroid nodule. Nodules were divided into groups according to their maximal diameter further.
Results Of all the nodules, 863 (33.9%) were benign, whereas 1681 (66.1%) were malignant. The malignancy percentage of ACR TI-RADS category 1, 2, 3, 4, and 5 were 0%, 1.3%, 9.1%, 52.5%, and 88.8%, respectively. KWAK-TIRADS and ATA guidelines showed a better diagnostic efficiency than ACR TI-RADS ($P < 0.01$). ACR TI-RADS demonstrated a higher specificity (79.7%, $P < 0.05$), whereas the ATA US pattern had a higher sensitivity (95.5%, $P < 0.01$). The TIRADS (KWAK-TIRADS and ACR TI-RADS) category and ATA guidelines performed better in differentiating nodules >1 cm. KWAK-TIRADS showed better diagnostic efficiency than the other methods in differentiating nodules >1 cm (AUC: 0.92, $P < 0.01$).
Conclusions KWAK-TIRADS and ATA guidelines provide a better diagnostic efficiency than ACR TI-RADS. The TIRADS (KWAK-TIRADS and ACR TI-RADS) category and ATA guidelines perform better in differentiating nodules >1 cm than nodules ≤1 cm. KWAK-TIRADS perform better in differentiating nodules >1 cm than other methods.

Keywords Thyroid nodules · Ultrasound · 2015 American Thyroid Association management guideline · Thyroid Imaging Reporting and Data System

Introduction

Thyroid nodules are a very common medical problem with a prevalence of 19–68% in the general population [1, 2]. Approximately 7–15% of thyroid nodules are thyroid

These authors contributed equally: Luying Gao, Xuehua Xi

✉ Bo Zhang
thyroidus@163.com

1   Department of Ultrasound, Chinese Academy of Medical Sciences & Peking Union Medical College Hospital, 100730 Beijing, China

cancer, and it has been estimated that 96% of all new endocrine organ cancers originate from the thyroid gland [3, 4]. Among them, palpable nodules account for only 4–7%, and most are incidentalomas in the general population that are detected by ultrasound (US) [3, 5]. US is useful for not only detection but also discriminating between benign and malignant lesions. This technique is used as guidance for fine-needle aspiration biopsy (FNAB) and further treatment and is also an important tool to assess the risk of recurrence. To date, there are many established guidelines for the interpretation of thyroid US [4, 6–8]. On the basis of a number of suspicious US features, the Thyroid Imaging Reporting and Data System classification proposed by Kwak (KWAK-TIRADS) was first established in 2011 [9], and 2015 American Thyroid Association (ATA) management guidelines have provided a risk stratification from

very low suspicion to high suspicion for malignancy [10]. Recently, ACR Thyroid Imaging Reporting and Data System (ACR TI-RADS) provided an up-to-date suggestion to stratify the nodules according to sonographic features [11]. These differences regarding the categorization of thyroid nodules may affect the diagnostic performances [12, 13]. We compared the diagnostic efficiency of KWAK-TIRADS pattern, ACR TI-RADS pattern, and ATA guidelines, and clarified the impact of nodule size on the performance of the three classification systems further.

Materials and methods

Patients

We retrospectively reviewed the medical records of all 1994 patients with 3004 thyroid nodules who underwent thyroidectomy at our center between January 2015 and December 2015. Among this initial cohort, only patients who met the following criteria were included: (1) total or nearly total thyroidectomy or lobectomy performed; (2) complete preoperative US of thyroid nodules; and (3) surgical pathology. Non-mass-forming lesions and nodules that failed to meet the criteria for any pattern of ATA guidelines were excluded. A total of 1758 patients with 2544 nodules were included finally. Thyroid nodules were divided into two groups according to the maximal diameter.

Thyroid US examination and retrospective evaluation

All US examinations were performed with Philips HDI 5000, IU 22, GE Logiq 9, or Logiq 7 devices equipped with either a 5–12 MHz or an 8–15 MHz linear-array transducer. US images were retrospectively reviewed by two radiologists who were experienced in thyroid US and blind to the patients' clinical data and pathological results (staff radiologists with 8 and 9 years of experience). Two experienced radiologists classified the degree of suspicion of thyroid nodule according to TI-RADS (proposed by Kwak and ACR) and ATA guideline independently. If there were differences, they discussed to get agreement.

According to the US classification of the 2015 ATA guidelines [12], thyroid nodules were assigned to one of the following degrees of suspicion: (1) high suspicion: solid hypoechoic nodule or solid hypoechoic component of a partially cystic nodule with one or more of the following features, including irregular margins (infiltrative, microlobulated), microcalcifications, taller-than-wide shape, disrupted rim calcifications with small extrusive soft tissue components, or evidence of extra-thyroidal extension; (2) intermediate suspicion: hypoechoic solid nodule with

smooth margins without microcalcifications, extra-thyroidal extension, or taller-than-wide shape; (3) low suspicion: isoechoic or hyperechoic solid nodule or partially cystic nodule with eccentric solid areas, without microcalcification, irregular margin or extra-thyroidal extension, or taller-than-wide shape; (4) very low suspicion: spongiform or partially cystic nodules without any of the sonographic features described in the low, intermediate or high suspicion patterns; and (5) benign: purely cystic nodules (no solid component).

Then all thyroid nodules were evaluated on the basis of the TIRADS patterns proposed by Kwak and ACR, respectively, [9, 11]. In Kwak version, suspicious US features included solid component, hypoechogenicity, marked hypoechogenicity, microlobulated or irregular margins, microcalcifications, and taller-than-wide shape. The nodules without any suspicious US features were classified as TIRADS category 3, and the other nodules were classified as TIRADS category 4a (with one suspicious US feature), 4b (with two suspicious US features), 4c (with three or four suspicious US features), or 5 (with five suspicious US features). TIRADS category 2 consisted of benign lesions (including simple cysts, spongiform nodules, isolated macrocalcifications, and typical subacute thyroiditis). In newly published TI-RADS patterns proposed by ACR [11], points are given for all the ultrasound features in a nodule, with more suspicious features being awarded additional points. The point total determines the nodule's ACR TI-RADS level, which ranges from TR1 (benign) to TR5 (high suspicion of malignancy). The ultrasound features in the ACR TI-RADS are categorized as benign (TR1, 0 point), not suspicious (TR2, 2 points), mildly suspicious (TR3, 3 points), moderately suspicious (TR4, 4–6 points), or highly suspicious (TR5, 7 points or more) for malignancy. Points are added from five categories to determine TI-RADS level. Composition: cystic or almost completely cystic, 0 points; spongiform, 0 points; mixed cystic and solid, 1 point; solid or almost completely solid, 2 points. Echogenicity: anechoic, 0 points; hyperechoic or isoechoic, 1 point; hypoechoic, 2 points; very hypoechoic, 3 points. Shape: wider-than-tall, 0 points; taller-than-wide, 3 points. Margin: smooth, 0 points; ill-defined, 0 points; lobulated or irregular, 2 points; extra-thyroidal extension, 3 points. Echogenic foci: none or large comet-tail artifacts, 0 points; macrocalcifications, 1 point; peripheral (rim) calcifications, 2 points; punctate echogenic foci, 3 points [Fig. 1].

Statistical analysis

Quantitative data are presented as the mean ± standard deviation (SD). Qualitative data are presented as frequencies. The Shapiro–Wilk test was used to determine the presence of a normal distribution. For nonparametric data,

Fig. 1 US scans illustrate ATA guideline pattern, KWAK-TIRADS categories, and ACR TI-RADS categories. **a** High suspicion for malignancy; KWAK-TIRADS Category 5; ACR TI-RADS categories 5. **b** Indeterminate suspicion for malignancy; KWAK-TIRADS Category 4b; ACR TI-RADS categories 4. **c** Low for malignancy; KWAK-TIRADS Category 4a; ACR TI-RADS categories 3. **d** Very low suspicion for malignancy; KWAK-TIRADS Category 3; ACR TI-RADS categories 2. **e** Very low suspicion for malignancy; KWAK-TIRADS Category 2; ACR TI-RADS categories 2. **f** Benign; KWAK-TIRADS Category 2; ACR TI-RADS categories 1
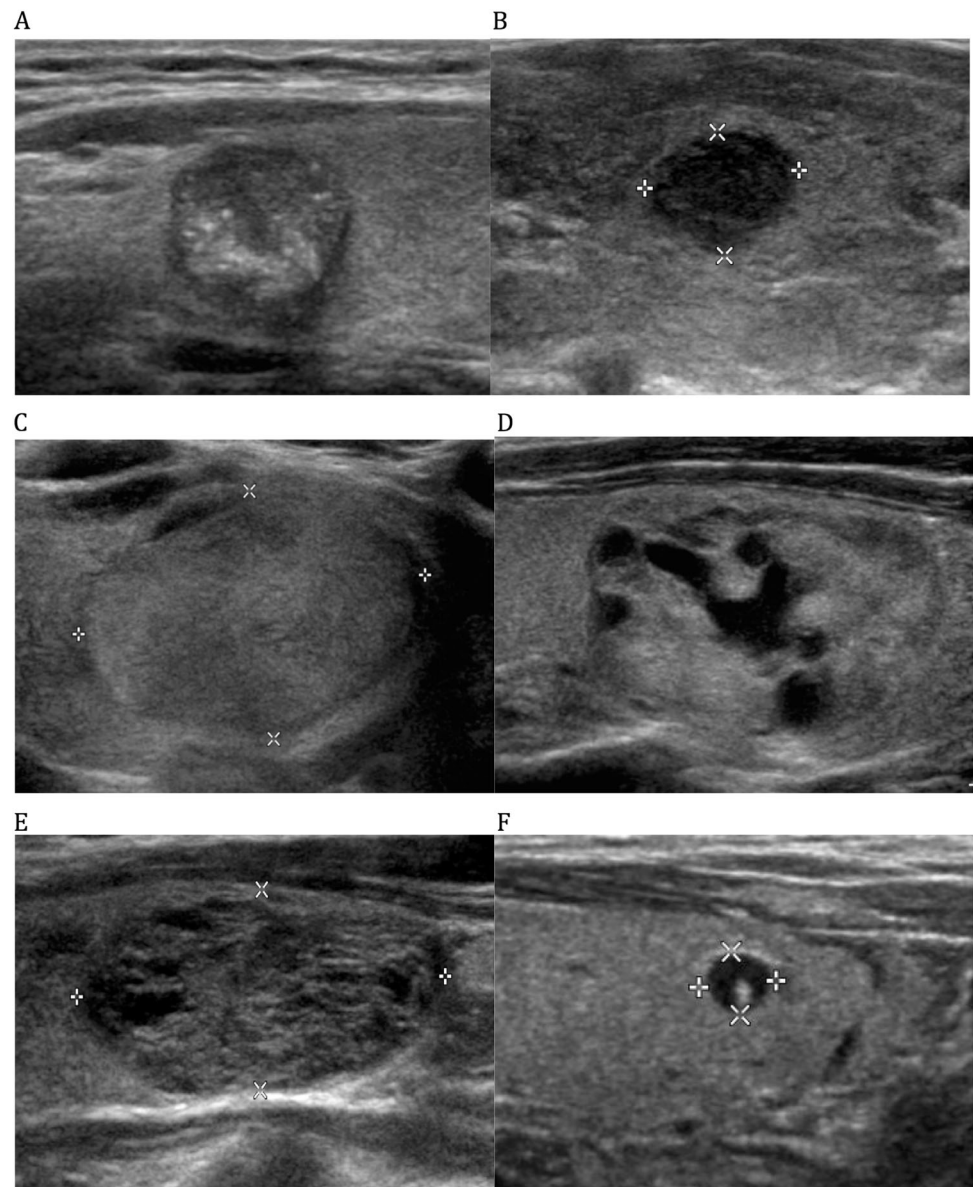
differences between groups were analyzed using a Mann–Whitney U test. For parametric data, an unpaired t-test was used to evaluate differences between two groups. The χ^2 test with Yates' correction and Fisher's exact test were used to compare categorical variables. The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were calculated through a comparison with the pathological findings. The Spearman rank test was used to assess the relationship between each category and the pathology findings. A receiver operating characteristic (ROC) curve analysis was used to compare KWAK-TIRADS, ACR-TIRADS, and ATA guidelines, and to calculate the optimal cutoff value. We calculated the value of kappa to assess the inter-observer variability. A value of $P < 0.05$ was considered statistically significant.

Statistical analyses were performed with SPSS software (Version 19.0, SPSS Chicago, IL, USA) and MedCalc 11.4.2.0 software (MedCalc Software, Ostend, Belgium).

Results

Demographic features of the patients

Of the 2544 thyroid nodules, 1681 (66.1%) were malignant, and 863 (33.9%) were benign. The distribution of demographic features of the patients are listed in Table 1 and Supplemental Table 1. The mean age of the patients with benign nodules were 48.5 ± 12.0 years, and those with malignant nodules were 43.2 ± 10.7 years. Age and sex

Table 1 Summary of demographic features

| | Pathological findings | | P-value |
	Benign	Malignant	
No. of nodules	863	1681	
No. of patients	531	1233	
Age (y)			
$\bar{x} \pm s$	48.5 ± 12.0	43.2 ± 10.7	<0.01
Sex			
Male	288 (33.4%)	468 (27.8%)	<0.01
Female	575 (66.6%)	1213 (72.2%)	
Size (cm)			<0.01
≤1	233 (27.0%)	884 (52.6%)	
[1.0–2.0]	218 (25.3%)	663 (39.4%)	
[2.0–3.0]	145 (16.8%)	77 (4.6%)	
[3.0–4.0]	113 (13.1%)	42 (2.5%)	
>4	154 (17.8%)	15 (0.9%)	

were significantly different between the two groups ($P < 0.01$). The size of malignant nodules was significantly smaller than that of benign (1.1 ± 0.7 vs. 2.3 ± 1.6 cm, $P < 0.01$).

Inter-observer agreement between the two observers was analyzed. Observer consistency of two radiologists was obtained for the assessment of Kwak-TIRADS category (kappa = 0.82; $P < 0.01$), ACR TI-RADS category (kappa = 0.84; $P < 0.01$), and ATA guidelines (kappa = 0.86; $P < 0.01$).

Correlations between the KWAK-TIRADS category and pathological findings

On the basis of the KWAK-TIRADS US categories, the percentages of malignancy in KWAK-TIRADS category 2, 3, 4a, 4b, 4c, and 5 were 0%, 1.9%, 10.9%, 55.2%, 88.8%, and 87.1%, respectively, the differences were statistically significant ($P < 0.01$). The correlation coefficient between the KWAK-TIRADS category and the malignancy was 0.65 [Table 2]. The ROC curves demonstrated that the best cutoff for the KWAK-TIRADS category was 4c. The sensitivity, specificity, PPV, NPV, accuracy, and area under the curve (AUC) were 89.4%, 77.4%, 88.5%, 78.9%, 85.3% and 0.86% (95% CI: 0.84–0.88), respectively [Table 3].

Correlations between the ACR TI-RADS category and pathological findings

On the basis of the ACR TI-RADS US categories, the percentages of malignancy in ACR TI-RADS category 1, 2, 3, 4, and 5 were 0%, 1.3%, 9.1%, 52.5%, and 88.8%,

respectively, the differences was statistically significant ($P < 0.01$). The correlation coefficient between the ACR TI-RADS category and the malignancy was 0.65. The ROC curves demonstrated that the best cutoff for the ACR-TIRADS category was TR5 [Table 2]. The sensitivity, specificity, PPV, NPV, accuracy and area under the curve (AUC) were 81.6%, 79.7%, 88.7%, 68.9%, 80.9%, and 0.81% (95% CI: 0.78–0.85), respectively [Table 3].

Correlations between the ATA category and pathological findings

On the basis of the ATA US categories, the percentages of malignancy in the nodules with benign, very low, low, intermediate, and high suspicion for malignancy were 0%, 0%, 5.6%, 33.9%, and 87.3%, respectively, and the differences was statistically significant ($P < 0.01$). The correlation coefficient between the ATA US category and the malignancy was 0.74 [Table 2]. The ROC curves demonstrated that the best cutoff of ATA pattern was high suspicion. Sensitivity, specificity, PPV, NPV, accuracy, and AUC were 95.5%, 73.0%, 87.3%, 89.4%, 87.8%, and 0.85% (95%CI, 0.83–0.87), respectively [Table 3].

Comparison of KWAK-TIRADS, ACR TI-RADS, and ATA guidelines in the diagnostic efficiency of thyroid nodules

Compared with the ACR TI-RADS, KWAK-TIRADS, and ATA guideline showed a higher AUC separately ($P < 0.01$). The ACR TI-RADS US pattern demonstrated a statistically higher specificity (79.7%, $P < 0.05$), whereas the ATA US pattern yielded a statistically higher sensitivity (95.5%, $P < 0.01$).

For the 1427 nodules with >1 cm, the ROC curves demonstrated that the best cutoff of the ATA, KWAK-TIRADS, and ACR TI-RADS categories were high suspicion, 4c and TR5, respectively. KWAK-TIRADS demonstrated a higher AUC (0.92, $P < 0.05$). The KWAK-TIRADS and ACR TI-RADS US pattern showed a significantly higher specificity than ATA guideline ($P < 0.01$), whereas the ATA US pattern yielded a significantly higher sensitivity (96.1%, $P < 0.01$).

For the 1117 nodules with a size ≤ 1 cm, the ROC curves demonstrated that the best cutoffs of the ATA, KWAK-TIRADS, and ACR TI-RADS categories were high suspicion, 4c and TR5, respectively. The ACR TI-RADS US pattern showed a statistically higher specificity (57.1%, $P < 0.05$), whereas the sensitivity of the ATA guidelines was higher than that of the TIRADS category (95.0%, $P < 0.01$). The AUC showed no statistically significant difference between the three patterns ($P > 0.05$) [Table 4].

Table 2 Malignant rates in the categories of KWAK-TIRADS, ACR TI-RADS and ATA guidelines

Category	Pathological diagnosis		Recommended malignancy risk (%)	Calculated malignancy rate (%)	R	P-value
	Benign	Malignant				
KWAK-TIRADS					0.65	<0.01
2	8	0	0	0		
3	403	8	1.7	1.9		
4a	131	16	3.3	10.9		
4b	126	155	9.2	55.2		
4c	159	1258	44–72.4	88.8		
5	36	244	87.5	87.1		
ATA					0.71	<0.01
Benign	6	0	<1	0		
Very low suspicion	178	0	<3	0		
Low suspicion	339	20	5–10	5.6		
Intermediate suspicion	107	55	10–20	33.9		
High suspicion	233	1606	>70–90	87.3		
ACR TI-RADS					0.65	<0.01
TR1	8	0		0		
TR2	144	2		1.3		
TR3	280	28		9.1		
TR4	252	279		52.5		
TR5	179	1372		88.8		

Table 3 Diagnostic efficiency of KWAK-TIRADS, ACR TI-RADS, and ATA guidelines

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	AUC
All						
KWAK-TIRADS	89.4	77.4	88.5	78.9	85.3	0.86
ATA	95.5	73.0	87.3	89.4	87.8	0.85
ACR TI-RADS	81.6	79.7	88.7	68.9	80.9	0.81
Nodules >1 cm						
KWAK-TIRADS	89.2	87.8	90.2	86.5	88.6	0.92
ATA	96.1	84.1	88.5	94.5	90.8	0.91
ACR TI-RADS	85.2	88.1	90.1	82.5	86.4	0.87
Nodules ≤1 cm						
KWAK-TIRADS	89.5	49.4	87.0	55.3	81.1	0.70
ATA	95.0	42.9	86.3	69.4	84.2	0.70
ACR TI-RADS	78.3	57.1	87.4	40.09	73.9	0.68

Discussion

Since Horvath first established the TIRADS classification, it has been widely applied to assess thyroid nodules. In our study, the malignancy rates of KWAK-TIRADS category 2, 3, 4a, 4c, 5 nodules were 0%, 1.9%, 10.9%, 88.7%, and 88.1%, respectively, which were comparable to the recommended rates. The malignancy rate of KWAK-TIRADS category 4b was 55.2%, which was much higher than the recommended rate but comparable

to those reported in previous studies [14]. These differences between studies may be partly due to the reference standards, inter-observer variability and the study population. In the present study, we calculated that the sensitivity and specificity of KWAK-TIRADS were 0.89 and 0.77, respectively, and the AUC was 0.86, thus indicating a high diagnostic accuracy. Our results are comparable to those from a meta-analysis reported recently, which has indicated a pooled sensitivity of 0.79 and a pooled specificity of 0.71 for the US reporting system in the

Table 4 Comparison of KWAK-TIRADS, ACR TI-RADS, and ATA guidelines in terms of diagnostic efficiency

	P-value		
	Sensitivity	Specificity	AUC
All			
KWAK-TIRADS vs. ATA	<0.01	<0.01	0.07
ATA vs. ACR TI-RADS	<0.01	<0.01	<0.01
K-TIRAD vs. ACR TI-RADS	<0.01	0.02	<0.01
Nodules >1 cm			
KWAK-TIRADS vs. ATA	<0.01	<0.01	0.01
ATA vs. ACR TI-RADS	<0.01	<0.01	<0.01
K-TIRAD vs. ACR TI-RADS	<0.01	0.86	0.01
Nodules ≤1 cm			
KWAK-TIRADS vs. ATA	<0.01	<0.01	0.67
ATA vs. ACR TI-RADS	<0.01	<0.01	0.37
K-TIRAD vs. ACR TI-RADS	<0.01	0.01	0.26

differential diagnosis of thyroid nodules [15]. Recently, committees convened by the ACR published white papers that presented an approach to incidental thyroid nodules for ultrasound reporting. In our study, the malignancy rates of ACR TI-RADS category 1, 2, 3, 4, and 5 nodules were 0%, 1.3%, 9.1%, 52.5%, and 88.8%, respectively. The malignancy rate of ACR TI-RADS category 4 was 52.5%, which was relatively high. In the present study, we calculated that the sensitivity and specificity of ACR-TIRADS were 0.82 and 0.80, respectively, and the AUC was 0.81. Recently, similar results indicated that ACR-TIRADS had a sensitivity of 0.80 and a specificity of 0.69 in the differential diagnosis of thyroid nodules [16]. Moreover, Grani's study showed that The ACR-TIRADS outperformed in its ability to reduce the number of unnecessary thyroid nodule FNAs than the other guidelines (such as ATA guidelines) [17], thus indicating a high diagnostic accuracy of ACR-TIRADS guidelines.

2015 ATA guidelines have suggested risk stratification on the basis of a constellation of sonographic features. In our study, the malignancy rates of benign, very low, low, and high suspicion nodules were 0%, 0%, 5.6%, and 87.3%, respectively, which were comparable to the recommended rates. The malignancy rate for the intermediate-suspicion pattern was 33.9%, which was higher than the recommended rate. This difference may be due to lack of consideration of its solid nature in the prediction of malignancy, although the solid nature of a nodule has been considered to be an independent risk factor for malignancy in a previous study [9, 18].

The TIRADS category and US pattern have previously been applied to 1293 thyroid nodules ($d > 1$ cm). The authors have found that TIRADS and ATA guidelines provide effective malignancy risk stratification for nodules. In particular, in that study, TIRADS showed a higher sensitivity, whereas the specificity, PPV, and accuracy were higher in ATA guidelines [19]. In our study, the TIRADS (KWAK-TIRADS and ACR TI-RADS) and ATA guidelines also performed well in differentiating thyroid nodules. Unlike Yoon's study, we found that TIRADS showed a higher specificity, whereas the ATA US pattern yielded a higher sensitivity. These findings were consistent with a recent study on 902 nodules of East Asians, which has confirmed that ACR TI-RADS guidelines were significantly less sensitive and had a higher specificity than ATA guidelines [16]. This difference between Yoon's study and our study may be partly due to the study population, additionally, in Yoon's study, some nodules were regarded as benign lesions on the basis of cytology. Recently, Xu's study has indicated that the ATA guidelines might yield a higher specificity than TIRADS for nodules larger than 2 cm [14]. We also found that TIRADS and ATA guidelines showed a better diagnostic efficiency in differentiating nodules >1 cm, whereas KWAK-TIRADS showed a better diagnostic efficiency than ACR TI-RADS and ATA guidelines. Similarly to our results, Cheng's study has reported that the TIRADS pattern is more reliable than ATA guidelines for larger thyroid nodules [18].

There are several limitations to our study. First, all analyses were based on the recorded static images and thus may have led to misdiagnosis by TIRADS and ATA guidelines. Second, all of the patients underwent thyroidectomy, which may have led to selection bias resulting in the underestimation of NPV and the overestimation of PPV for KWAK-TIRADS, ACR TI-RADS patterns, and ATA guidelines.

We found that ACR TI-RADS had a higher specificity, whereas the ATA guideline yielded a higher sensitivity. Moreover, nodule sizes may impact the diagnostic efficiency of the three patterns, and both the TIRADS and ATA guidelines perform better in differentiating nodules >1 cm. For nodules >1 cm, KWAK-TIRADS demonstrated better diagnostic efficiency than ACR TI-RADS and ATA guidelines. For nodules with a size ≤1 cm, there was no difference of diagnostic efficiency among the three guidelines.

Author contributions L.G. and X.X. participated in the study design, performed the statistical analysis, and drafted the manuscript. Y.J. conceived of the study, participated in its design. X.Y., Y.W., S.Z., R. Z., X.L., and X.Z. carried out the selection and collection of samples. B.Z. participated in the study design, performed the statistical analysis, and review the manuscript. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards, and the requirement for informed consent was waived for this retrospective study.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. G.H. Tan, H. Gharib, Thyroid incidentalomas: management approaches to nonpalpable nodules discovered incidentally on thyroid imaging. Ann. Intern. Med. **126**, 226–231 (1997)

2. S. Guth, U. Theune, J. Aberle, A. Galach, C.M. Bamberger, Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. Eur. J. Clin. Invest. **39**, 699–706 (2009)

3. L. Hegedüs, The thyroid nodule. NEJM **351**, 1764–1771 (2004)

4. L. Leenhardt, M.F. Erdogan, L. Hegedus, S.J. Mandel, R. Paschke et al. European thyroid association guidelines for cervical ultrasound scan and ultrasound-guided techniques in the postoperative management of patients with thyroid cancer. Eur. Thyroid J. **2**, 147–59 (2013)

5. G.H. Tan, H. Gharib, Thyroid incidentalomas: management approaches to nonpalpable nodules discovered incidentally on thyroid imaging. Ann. Intern Med. **126**, 226–231 (1997)

6. P. Perros, K. Boelaert, S. Colley, C. Evans, R.M. Evans et al. Guidelines for the management of thyroid cancer. Clin. Endocrinol. **81**, 1–122 (2014)

7. J.H. Shin, J.H. Baek, J. Chung, E.J. Ha, J.H. Kim et al. Ultrasonography diagnosis and imaging-based management of thyroid nodules: revised Korean Society of Thyroid Radiology Consensus Statement and Recommendations. Korean J. Radiol. **17**, 370–395 (2016)

8. H. Gharib, E. Papini, R. Paschke, D.S. Duick, R. Valcavi 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. **22**, 622–639 (2016)

9. J.Y. Kwak, K.H. Han, J.H. Yoon, H.J. Moon, E.J. Son et al. Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk. Radiology **260**, 892–899 (2011)

10. B.R. Haugen, E.K. Alexander, K.C. Bible, G.M. Doherty, S.J. Mandel et al. 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 **26**, 1–133 (2015)

11. F.N. Tessler, W.D. Middleton, E.G. Grant, J.K. Hoang, L.L. Berland et al. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. J. Am. Coll. Radiol. **14**, 587–595 (2017)

12. W.D. Middleton, S.A. Teefey, C.C. Reading, J.E. Langer, M.D. Beland 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. **210**, 1–7 (2018)

13. J. Koh, S.Y. Kim, H.S. Lee, E.K. Kim, J.Y. Kwak et al. Diagnostic performances and interobserver agreement according to observer experience: a comparison study using three guidelines for management of thyroid nodules. Acta Radiol. **59**, 028418511774400 (2017)

14. T. Xu, J.Y. Gu, X.H. Ye, S.H. Xu, Y. Wu et al. Thyroid nodule sizes influence the diagnostic performance of TIRADS and ultrasound patterns of 2015 ATA guidelines: a multicenter retrospective study. Sci. Rep. **7**, 43183 (2017)

15. X. Wei, Y. Li, S. Zhang, M. Gao, Meta-analysis of thyroid imaging reporting and data system in the ultrasonographic diagnosis of 10,437 thyroid nodules. Head Neck **38**, 309–315 (2016)

16. E.J. Ha, D.G. Na, W.J. Moon, Y.H. Lee, N. Choi, 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 Congress of Radiology Guidelines. Thyroid **28**, 1532–1537 (2018)

17. G. Grani, L. Lamartina, V. Ascoli, D. Bosco, M. Biffoni et al. Reducing the number of unnecessary thyroid biopsies while improving diagnostic accuracy: toward the "right" TIRADS. J. Clin. Endocrinol. Metab. **104**, 95–102 (2019)

18. S.P. Cheng, J.J. Lee, J.L. Lin, S.M. Chuang, M.N. Chien et al. Characterization of thyroid nodules using the proposed thyroid imaging reporting and data system (TI-RADS). Head. Neck. **35**, 541–547 (2013)

19. J.H. Yoon, H.S. Lee, E.K. Kim, H.J. Moon, J.Y. Kwak, Malignancy risk stratification of thyroid nodules: comparison between the thyroid imaging reporting and data system and the 2014 American Thyroid Association Management guidelines. Radiology **278**, 917 (2016)