RESEARCH

Assessment of perinodular stiffness in differentiating malignant from benign thyroid nodules

Lei Hu[ID]1,*, **Xiao Liu1,***, **Chong Pei2,***, **Li Xie1,*** and **Nianan He1**

1Department of Ultrasound, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui, China
2Department of Respiratory and Critical Care Medicine, The First People's Hospital of Hefei City, The Third Affiliated Hospital of Anhui Medical University, Hefei, Anhui, China

Correspondence should be addressed to Nianan He: henianan@ustc.edu.cn

*(L Hu, X Liu, C Pei and L Xie contributed equally to this work)

Abstract

Objective: We evaluated the diagnostic accuracy of perinodular stiffness, four risk stratification systems (RSSs) (KWAK-TIRADS, ACR-TIRADS, EU-TIRADS, and C-TIRADS), and the combination of perinodular stiffness and the four RSSs in differentiating malignant from benign thyroid nodules (TNs).

Methods: A total of 788 TNs in 726 patients were examined with conventional ultrasound (US) examination and sound touch elastography (STE). All TNs were classified by each of the four RSSs. The stiffness inside (E) the TNs was measured by STE. The stiffness of the 2.0-mm perinodular region (Eshell) was measured with the Shell measurement function of STE. The diagnostic performances of four RSSs, the E values, and the Eshell values were evaluated. All TNs were further divided into subgroups based on size ($\leq$ 10 mm group and > 10 mm group).

Results: Ninety-six TNs were classified as benign and 692 as malignant. Among the single-method approaches, ACR-TIRADS showed the highest AUC (0.77) for differentiating malignant from benign TNs for all TNs included. Eshell showed the highest AUC (0.75) in differentiating malignant from benign TNs for TNs with sizes $\leq$ 10 mm, and there were no significant differences in AUC among all single methods for diagnosis of TNs with sizes > 10 mm (P > 0.05). The combination of C-TIRADS and Eshell/E yielded the highest AUC for all TNs (0.83) and for TNs with size $\leq$ 10 mm (0.85) compared with other combinations.

Conclusions: Eshell/E combined with conventional US improves the diagnostic accuracy in TNs and may reduce unnecessary fine-needle aspiration.

Introduction

Thyroid nodules (TNs) are a common endocrine disease in China (1, 2, 3). It is estimated that 7–15% of TNs are thyroid cancer. Ultrasound (US) is useful not only for detection but also for discrimination between benign and malignant TNs. A series of TN risk stratification systems (RSSs) have been proposed for US radiologists to define TNs objectively (2, 3, 4, 5). For example, the thyroid imaging reporting and data system classification proposed by Kwak (KWAK-TIRADS) was established in 2011 (2). In 2017, the ACR-TIRADS was issued by the

American College of Radiology, and European Thyroid Association management provided EU-TIRADS (3, 4). Recently, Chinese researchers provided Chinese-TIRADS (C-TIRADS) to stratify TNs depending on their US features and total points (5). These RSSs allow standardization of the diagnosis of benign and malignant TNs by radiologists; however, the differences regarding the categorization of TNs may affect the diagnostic performance and confuse clinical treatment (6, 7, 8, 9, 10, 11).

Ultrasound elastography (UE) has been shown to improve differentiation between benign and malignant TNs (12, 13, 14, 15, 16, 17, 18, 19). Sound touch elastography (STE) and the Shell measurement function enable the quantitative assessment of interior and perinodular stiffness of a TN. Earlier studies have shown that the surrounding tissue stiffness measured by the Shell measurement function improved the differentiation of benign and malignant TNs and breast lesions (20, 21, 22, 23, 24, 25). In particular, the 2 mm-perinodular stiffness of TNs measured by the Shell measurement function may be an accurate diagnostic index for differentiating malignant from benign TNs (20, 21).

There have been many studies comparing different RSS versions and tissue stiffness in differentiating malignant from benign TNs (26, 27, 28, 29, 30). However, few studies have reported the diagnostic performance of a combination of TIRADS categories and perinodular stiffness to discriminate between benign and malignant TNs. Hence, based on previous studies, we compared the diagnostic efficiency of KWAK-TIRADS, ACR-TIRADS, EU-TIRADS, C-TIRADS, and these RSSs combined with perinodular stiffness of TNs in differentiating malignant from benign TNs.

Materials and methods

Patient selection

This prospective study was approved by the institutional ethics committee of the First Affiliated Hospital of the University of Science and Technology of China (USTC), and informed consent was obtained from all patients to include their data for this study. From February 2020 to December 2020, 756 consecutive patients with 1005 TNs were detected by conventional US, STE, and Shell measurement function at our center. The inclusion criteria were as follows: (i) solid or nearly solid (<20% cystic) TN; (ii) sufficient thyroid tissue surrounding the TN, so that 2-mm perinodular tissue could be measured and (iii) verified benign or malignant diagnosis based on pathohistology or fine-needle aspiration (FNA). In patients with multiple TNs in the same thyroid lobe, the most suspicious for malignancy was first selected, otherwise, the largest TN was included. Eventually, 726 patients (aged 16 to 77 years) with 788 TNs were enrolled in the study. The flowchart for TN selection is shown in Fig. 1. All TNs were divided into two subgroups depending on the maximal diameter (≤ 10 mm group and > 10 mm group).

US instrument

All conventional US and STE examinations of TNs were performed with a Resona 7 US diagnostic system (Mindray Medical Solutions, Shenzhen, China) and a 11L3 transducer, with the STE and the Shell measurement software. STE is a two-dimensional (2D) real-time shear wave elastography (SWE) technology with Shell measurement function, which automatically enables the quantitative assessment
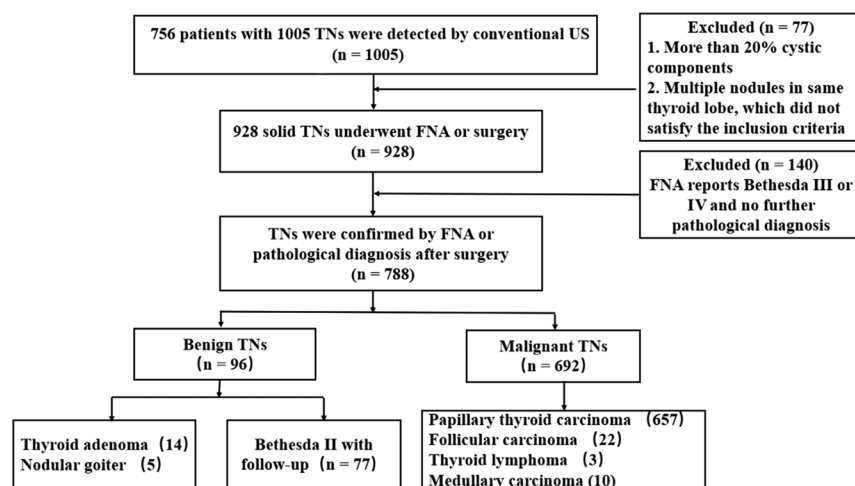

Figure 1
Flowchart of the selection of patients with 788 solid thyroid nodules (TNs).

of perinodular stiffness (the width of 0.5–9 mm) in 0.5-mm increments from the outline of the TN (Fig. 2).

Conventional US examination and retrospective evaluation

Conventional US images of all TNs were obtained by two radiologists (X L and L X) with 10 and 12 years of experience in TN US. The TNs were classified according to four RSS patterns (KWAK-TIRADS, ACR-TIRADS, EU-TIRADS, and C-TIRADS) (2, 3, 4, 5). If there were different opinions, the radiologists discussed them to reach an agreement. Both radiologists were blinded to the patients' clinical data and pathological results.

UE image acquisition

The region of interest (ROI) set for STE examination was adjusted as follows. Both the TN and at least 5 mm of the surrounding tissue were included in the ROI, and the TN was placed in the center of the ROI on the longitudinal section of the thyroid lobe. We first used the tracing method to outline the margin of the TN for measurement of the interior stiffness (Young's modulus max value, recorded as E). Subsequently, the operator activated the Shell measurement function key, selected 2.0 mm surrounding the TN, and the software automatically measured the 2 mm-perinodular stiffness of the TNs (Young's modulus max value, recorded as Eshell) (Figs 3 and 4). The patients were instructed to hold their breath during STE and Shell measurement. UE was examined by the same radiologist (L H) with 10 years of experience in thyroid UE. These examinations were repeated by the same operator with a 1-day interval, and the mean values were determined.
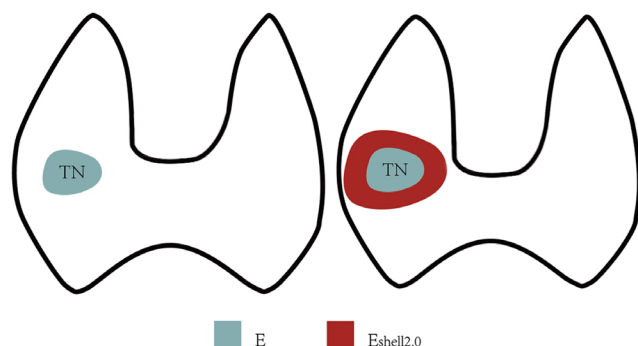

Figure 2
Shell measurement diagram. (A) The E value was the interior stiffness of the TN. (B) The Eshell value refers to 2 mm-perinodular stiffness of the TN.

Four RSSs of TNs combined with the ratio of Eshell to E values

Based on the ratio of Eshell to E values, all TNs were reclassified to KWAK-TIRADS + Eshell/E, ACR-TIRADS + Eshell/E, EU-TIRADS + Eshell/E, and C-TIRADS + Eshell/E. The grade of TN was upgraded one level if the ratio of Eshell to E values was higher than 1. The grade of TN was declined one level if the ratio of Eshell to E values was below 1. The grade of TN remained unchanged when the ratio of Eshell to E values equaled 1. When the TNs were already at the lowest grade or highest grade in our study, the grade of TNs remained unchanged even if the ratio of Eshell to E values was below 1 or higher than 1.

Cytological and pathological diagnosis of TNs

US-guided FNA was performed after US and UE examination by the same radiologist (L H). The cytological reports were classified based on the Bethesda System for Reporting Thyroid Cytopathology (31) by one of three thyroid cytopathologists with more than 5 years of experience. All TNs enrolled in the study had a verified benign or malignant diagnosis based on pathohistology or a definitive FNA report (Bethesda category II, V, and VI). When an FNA report was a Bethesda category I, the FNA was repeated after 1 week. When an FNA report showed a Bethesda category III or IV, the TN was included in this study if there was further surgical pathology or definitive FNA report. The FNA-benign TNs were followed up at 6 months. If they showed less than 20% increase in TN maximum diameter or not more than a 50% increase in volume on the conventional US, they were considered benign (31). All cytological indications of malignant TNs have been further verified by pathohistological diagnosis.

Statistical analysis

Statistical analysis was performed using SPSS software, version 20.0 (IBM corporation). Quantitative data are shown as the mean ± s.d. Qualitative data are shown as frequencies. The Shapiro–Wilk test was used to determine the presence of normal distribution. We compared normally distributed data using Student's unpaired t-test and non-normally distributed data using the Mann–Whitney U-test. The χ^2 test and Fisher's exact probability test were used to compare categorical variables. Correlations between the E and Eshell values were assessed using Spearman's correlation coefficient.
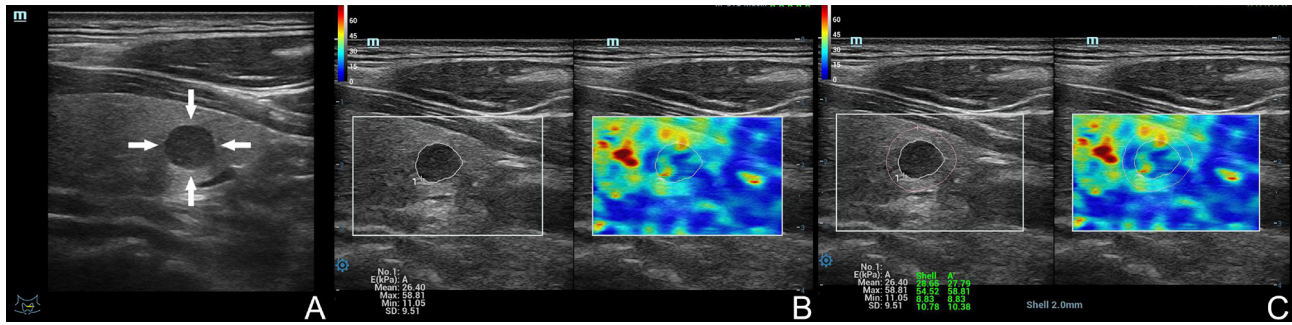
Figure 3

Images showing a solid TN in a 46-year-old woman. The FNA diagnosis was a benign TN. (A) Conventional US image of the TN; the arrows refer to the TN. KWAK-TIRADS: Category 4a; ACR-TIRADS: TR4; EU-TIRADS: 4, indeterminate risk; C-TIRADS: C-TR4B. (B) The E value of the TN was 58.81 kPa. (C) The Eshell value of the TN was 54.52 kPa.

Receiver operating characteristic (ROC) analyses were performed to assess the diagnostic performance of the different RSSs of TNs, as well as RSSs combined with Eshell/E. ROC was also used to determine the optimal cut-off values, and to calculate the corresponding sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the ROC curve (AUC). A P-value < 0.05 was considered statistically significant.

Results

Demographic and conventional US features

Of the 788 TNs, 692 (66.1%) were malignant and 96 (33.9%) were benign. Age range of the included patients was between 16 and 77 years (mean age, 35.1 ± 13.3 years), and 73% of patients were female. The average maximum diameter of all TNs as shown on grayscale US was 17.8 ± 2.5 mm (range, 6.5–25.6 mm). There were 427 (42 benign vs 385 malignant) TNs with sizes ≤ 10 mm

and 361 (54 benign vs 307 malignant) TNs with sizes > 10 mm. There were no significant differences in age and sex between the benign and malignant TNs (both $P > 0.05$). The TN size was significantly smaller in malignant TNs than in benign TNs (12 ± 0.7 vs 17 ± 1.6 mm, $P < 0.01$). Age, sex, and the numbers of benign and malignant TNs were not significantly different between the subgroups based on TNs sizes (all $P > 0.05$). All TNs classified by the four RSSs are shown in Table 1.

Cytological and pathological diagnoses

Cytological diagnosis suggested malignancy in 692 TNs, including 223 TNs of Bethesda category V and 469 TNs of Bethesda category VI, and 77 TNs with Bethesda category II were shown to be benign at follow-up with 6 months. All cytological diagnoses suggesting malignant TNs were finally diagnosed by postoperative pathology (Fig. 1). Nineteen benign TNs were removed by thyroidectomy because they were causing compressive symptoms (Fig. 1). All the surgical pathology diagnoses were made by one of
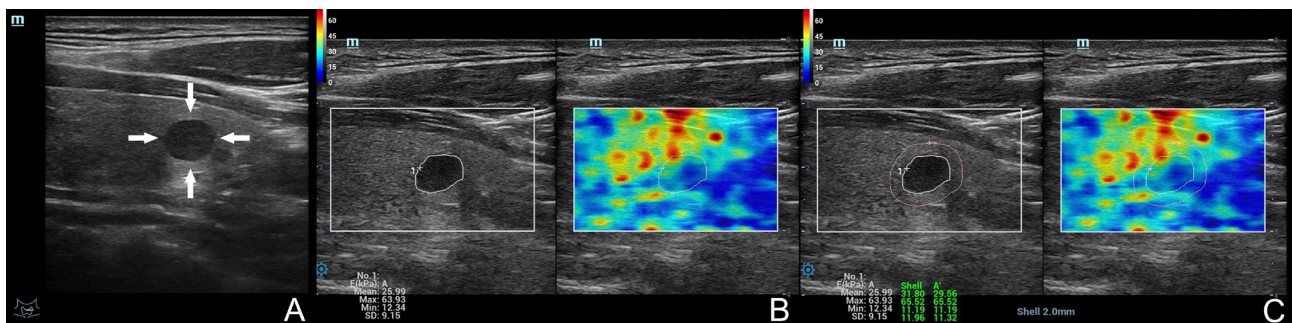

Figure 4

Images showing a solid TN in a 24-year-old woman. The pathological diagnosis after surgery was papillary thyroid carcinoma. (A) Conventional US image of the TN; the arrows refer to the TN. KWAK-TIRADS: Category 4a; ACR-TIRADS: TR4; EU-TIRADS: 4, indeterminate risk; C-TIRADS: C-TR4B. (B) The E value of the TN was 63.93 kPa. (C) The Eshell value of the TN was 65.52 kPa.

Table 1 Thyroid nodules (TNs) classified according to the four risk stratification systems (RSSs) and four RSSs combined with Eshell/E values.

	ALL TNs		TN ≤ 10 mm		TN > 10 mm			ALL TNs		TN ≤ 10 mm		TN > 10 mm	
RSS	B (n = 96)	M (n = 692)	B (n = 42)	M (n = 385)	B (n = 54)	M (n = 307)	**RSS + Eshell/E**	B (n = 96)	M (n = 692)	B (n = 42)	M (n = 385)	B (n = 54)	M (n = 307)
KWAK-TIRADS													
3	19	2	9	1	10	1	3	30	2	13	1	17	1
4a	55	230	15	110	40	120	4a	44	218	22	119	22	99
4b	18	290	15	135	3	155	4b	19	297	5	144	14	153
4c	3	153	2	125	1	28	4c	2	153	1	111	1	42
5	1	17	1	14	0	3	5	1	22	1	10	0	12
ACR-TIRADS													
TR3	19	4	7	3	12	1	TR3	36	19	16	12	20	7
TR4	68	476	29	263	39	213	TR4	52	377	21	233	31	144
TR5	9	212	6	119	3	93	TR5	8	296	5	140	3	156
EU-TIRADS													
3	19	4	8	3	11	1	3	35	22	23	14	12	8
4	66	562	29	286	37	276	4	53	392	15	216	38	176
5	11	126	5	96	6	30	5	8	278	4	155	4	123
C-TIRADS													
C-TR3	19	2	9	1	10	1	C-TR3	51	2	22	1	29	1
C-TR4A	49	298	16	142	33	156	C-TR4A	30	192	17	85	13	139
C-TR4B	24	222	14	114	10	108	C-TR4B	12	279	2	137	10	114
C-TR4C	3	156	2	120	1	36	C-TR4C	2	187	1	145	1	42
C-TR5	1	14	1	8	0	6	C-TR5	1	32	0	17	1	11

three pathologists with more than 8 years of experience in thyroid pathology.

E and Eshell values

The average E and Eshell values were significantly higher in malignant TNs than in benign TNs (all *P* < 0.05). The Eshell values were significantly higher than the E values in malignant TNs (*P* < 0.05) but not significantly different in benign TNs (*P* = 0.66) (Table 2). Average E and Eshell values were significantly higher in TNs with sizes ≤ 10 mm than in TNs > 10 mm (*P* < 0.05). Average Eshell values were significantly higher than the E values in TNs with sizes ≤ 10 mm (*P* < 0.05) but not significantly different in TNs with sizes > 10 mm (*P* = 0.56) (Table 2).

Diagnostic performance of the four RSSs, E values, and Eshell values

Based on ROC analysis, Table 3 shows the best cut-off values of the KWAK-TIRADS, ACR-TIRADS, EU-TIRADS, C-TIRADS, the E values, and Eshell values in all TNs and the two subgroups of TNs sizes.

The ACR-TIRADS showed the highest AUC for differentiating benign from malignant TNs compared with other single methods in all TNs (AUC = 0.77, *P* < 0.05). The ACR-TIRADS and the EU-TIRADS pattern demonstrated significantly higher sensitivity (84.4 and 83.4%, respectively, *P* < 0.05), and the C-TIRADS pattern yielded significantly higher specificity (67.8%, *P* < 0.05) compared with other methods in differentiating malignant from benign TNs (Table 3).

In TNs with sizes ≤ 10 mm, the Eshell values demonstrated a higher AUC (0.75, *P* < 0.05), sensitivity (89.7%, *P* < 0.05), and specificity (67.9%, *P* < 0.01) compared with other methods in differentiating malignant from benign TNs (Table 3).

In TNs with sizes > 10 mm, the KWAK-TIRADS pattern showed significantly higher sensitivity (89.3%, *P* < 0.05),

and the ACR-TIRADS pattern showed higher specificity (87.3%, *P* < 0.05) compared with other methods in differentiating malignant from benign TNs. The AUCs did not differ significantly between the four RSSs (*P* > 0.05) but they were higher than those for the E and Eshell values (all P < 0.05) (Table 3).

Diagnostic performance of the four RSSs combined with the ratio of Eshell to E values

All of the TNs were reclassified by the four RSSs combined with the Eshell/E as shown in Table 1. The sensitivity, specificity, cut-off values and AUC of the four RSS combined with Eshell/E in differentiating malignant from benign TNs are shown in Table 4.

The sensitivity, specificity and AUC of the four RSSs combined with Eshell/E were improved in differentiating malignant from benign TNs compared with any original single method for all TNs and TNs with sizes ≤ 10 mm (all *P* < 0.05), but there were no significant differences between the original single methods and the combinations for TNs with sizes > 10 mm (all *P* > 0.05) (Tables 3 and 4).

In all TNs, compared with other combinations, C-TIRADS + Eshell/E showed the highest AUC (AUC = 0.83, *P* < 0.05) and significantly higher sensitivity (90.4%, *P* < 0.05), while the KWAK-TIRADS + Eshell/E yielded significantly higher specificity (79.2%, *P* < 0.05) in differentiating malignant from benign TNs (Table 4).

In TNs with sizes ≤ 10 mm, the C-TIRADS + Eshell/E showed the highest AUC (AUC = 0.85, *P* < 0.05) and significantly higher sensitivity (89.5%, *P* < 0.05), while KWAK-TIRADS + Eshell/E yielded a significantly higher specificity (71.1%, *P* < 0.05) in differentiating malignant from benign TNs compared with other combinations (Table 4).

Discussion

Many previous studies have pointed out that RSSs can improve the accuracy of the diagnosis of benign and malignant TNs, but the results are not completely identical when different RSSs are compared (26, 27, 28, 29, 30). This is because their diagnostic accuracy is influenced both by multiple versions and the radiologists' diagnostic experience. In fact, each RSS has its own advantages and disadvantages. In this study, we showed that the ACR-TIRADS had the highest AUC (0.77) and significantly highest sensitivity (84.4%) for distinguishing between benign and malignant TNs compared with other

Table 2 Average E and E$_{shell}$ values of benign and malignant and different sizes of thyroid nodules (TNs).

Group	**E** (kPa)	**Eshell** (kPa)	***P***
Benign	65.92 ± 9.76	64.28 ± 8.75	0.66
Malignant	76.49 ± 11.04	84.27 ± 10.15	< 0.01
P	< 0.01	< 0.01	
TNs with sizes > 10 mm	64.35 ± 8.86	63.96 ± 7.95	0.56
TNs with sizes ≤ 10 mm	78.49 ± 15.04	89.49 ± 10.04	< 0.01
P	< 0.01	< 0.01	

Table 3 Diagnostic efficiency of four risk stratification systems, E values, and Eshell values.

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Cut-off value	Accuracy (%)	AUC
All							
KWAK-TIRADS	71.6	67.2	80.6	77.2	4b	81.3	0.72
ACR TIRADS	84.4	54.0	92.4	72.0	TR4	86.5	0.77
EU-TIRADS	83.4	58.8	88.4	70.8	4	84.5	0.75
C-TIRADS	78.7	67.8	85.7	74.5	C-TR4B	83.5	0.73
E	76.6	64.4	85.6	75.4	68.46 kPa	81.5	0.72
Eshell	77.4	60.5	85.4	72.6	69.76 kPa	82.2	0.74
Nodules ≤ 10 mm							
KWAK-TIRADS	69.5	44.5	77.9	59.4	4b	77.8	0.68
ACR TIRADS	70.4	45.7	77.8	55.5	TR4	73.6	0.67
EU-TIRADS	71.6	43.9	76.4	54.4	4	75.4	0.69
C-TIRADS	74.4	46.5	75.7	65.7	C-TR4B	79.5	0.70
E	87.6	65.7	82.6	70.3	72.75 kPa	82.0	0.73
Eshell	89.7	67.9	85.9	71.5	78.89 kPa	85.4	0.75
Nodules > 10 mm							
KWAK-TIRADS	89.3	76.7	88.0	76.6	4b	85.2	0.77
ACR TIRADS	72.6	87.3	79.4	89.4	TR4	86.4	0.78
EU-TIRADS	73.8	84.1	77.1	86.5	4	86.3	0.78
C-TIRADS	85.3	70.6	86.7	77.7	C-TR4B	85.6	0.77
E	74.3	81.3	84.5	83.6	64.28 kPa	82.6	0.72
Eshell	75.3	82.9	85.6	84.5	68.45 kPa	82.7	0.73

single methods in all TNs (Table 3), which is consistent with previous findings (26, 27). The perinodular stiffness combined with the four RSSs in differentiating malignant from benign TNs was further analyzed in this study. We found that the AUC of the combination of C-TIRADS and Eshell/E in discriminating between malignant and benign TNs for all TNs (0.83) and TNs with sizes ≤ 10 mm (0.85) was the highest compared with other combinations (Table 4).

The E values of malignant TNs were higher than the E values of benign TNs; this was true for the Eshell values as well (Table 2). The average E and Eshell values of TNs with sizes ≤ 10 mm were significantly higher compared

with those of TNs > 10 mm, which may stem from the fact that malignant TNs were significantly smaller than benign TNs in our study (Table 2). Moreover, the Eshell values of malignant TNs were significantly higher than their E values because the malignant nature not only increased the internal stiffness of the TN (E) but also stimulated fibroblasts resulting in increased perinodular tissue stiffness (Eshell) (21, 22, 23, 24, 25). The AUC of the E values (0.73) and Eshell values (0.75) in differentiating benign from malignant TNs with sizes ≤ 10 mm were higher compared with the four RSSs, and Eshell values showed the highest AUC compared with other single methods (Table 3). Although the smaller-sized malignant

Table 4 Diagnostic efficiency of Eshell/E combined with each of the four risk stratification systems.

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Cut-off value	Accuracy (%)	AUC
All							
KWAK-TIRADS +Eshell/E	84.6	79.2	80.6	77.2	4b	83.5	0.75
ACR-TIRADS+Eshell/E	86.7	74.0	92.4	72.0	TR4	85.3	0.78
EU-TIRADS+Eshell/E	87.6	77.8	88.4	70.8	4	83.2	0.77
C-TIRADS+Eshell/E	90.4	71.5	85.7	74.5	C-TR4B	86.8	0.83
Nodules ≤ 10 mm							
KWAK-TIRADS +Eshell/E	82.6	71.1	80.6	77.2	4b	82.3	0.73
ACR-TIRADS+Eshell/E	86.7	67.7	92.4	72.0	TR4	82.5	0.72
EU-TIRADS+Eshell/E	86.6	67.5	88.4	70.8	4	81.5	0.72
C-TIRADS+Eshell/E	89.5	68.7	85.7	74.5	C-TR4B	89.5	0.85
Nodules > 10 mm							
KWAK-TIRADS +Eshell/E	89.4	75.7	88.3	75.6	4b	85.7	0.78
ACR-TIRADS+Eshell/E	73.6	86.3	79.5	85.4	TR4	85.4	0.77
EU-TIRADS+Eshell/E	73.9	83.1	76.1	86.5	4	85.6	0.78
C-TIRADS+Eshell/E	85.5	70.7	86.7	75.7	C-TR4B	86.3	0.77

TNs showed no signs of malignancy on 2D US images, they stimulated fibroblasts to increase both their interior and perinodular stiffness. These results are consistent with our previous studies that reported that 2-mm perinodular stiffness of TNs is a better diagnostic index for TN differentiation than the stiffness inside the TN (20, 21).

In this study, an Eshell/E higher than 1 indicated a higher risk of malignancy, considering that perinodular stiffness increased by cancer-induced fibroblasts. Conversely, an Eshell/E below or equal to 1 indicated a low risk of malignancy. We believe that the perinodular stiffness of TNs (Eshell) compared with the stiffness inside the TNs (E) is not affected by the heterogeneous echo inside the TNs. Moreover, compared with Eshell or E alone, the Eshell/E is a more objective indicator that avoids the interobserver inconformity and lack of recognized cut-off values between benign and malignant TNs (20, 21). Hence, in this study, the AUCs of the Eshell/E combined with any RSS were all higher compared with each single method in diagnosing all TNs and in TNs with sizes ≤ 10 mm (Table 4).

The combination of C-TIRADS and Eshell/E not only yielded the highest AUC but also showed significantly higher sensitivity in differentiating malignant from benign TNs for all TNs (0.83; 90.4%) and TNs with sizes ≤ 10 mm (0.85; 89.5%), compared with other combinations (Table 4). This may be because the C-TIRADS not only adopted the scoring method but also adopted three subgroups for the category 4 TNs, implementing various advantages of KWAK-TIRADS, ACR-TIRADS, and EU-TIRADS. Hence, C-TIRADS is more suitable for the combining with Eshell/E to improve the diagnosis of benign and malignant TNs compared with other combinations.

RSSs are crucial to select those patients with TNs, in order to select those who should have an FNA performed. The main disadvantage of all RSSs is that they focus on recommending FNA for TNs (1, 2, 5, 21), which leads to unnecessary FNA. Ultrasound elastography has been reported to be helpful in the diagnosis of benign and malignant TNs (12, 13, 32, 33, 34, 35, 36, 37). When all TNs were reclassified according to perinodular stiffness combined with the RSSs, 11 KWAK-4a TNs were redefined as grade 3 TNs, 17 ACR TR4 TNs were redefined as TR3 TNs, 16 EU-TIRADS intermediate suspicion TNs were redefined as low suspicion, and 32 C-TR4A TNs were redefined as C-TR3 TNs in our study (Table 1). Thus, the combination of perinodular stiffness with the RSSs would reduce unnecessary FNA, while single traditional RSS methods may not.

Although the combinations of RSS with different US technologies have been applied to improve the accuracy of diagnosis of benign and malignant TNs, they only focused on a single RSS version (13, 14, 15, 16, 17, 18, 19, 22, 38, 39). Thus, we combined multiple RSSs versions with the perinodular stiffness index, which would be easier to generalize due to its broad applicability.

This study had some limitations. First, the number of benign TNs was small because FNA and surgery were not required for most of the TIRADS grade 3 TNs. Secondly, the Shell measurement function was not applied in all TNs, and the inapplicable TNs were excluded in patients selection (e.g. when the TNs were close to the capsule or the trachea).

Finally, we conclude that Eshell/E combined with RSSs improve the diagnostic accuracy in TNs, especially for TNs with sizes ≤ 10 mm, and may reduce unnecessary FNA procedures in intermediate suspicious TNs of RSS.

References

1 Zhu CL, Li SX, Gao X, Zhu GC, Song ML & Gao F. Retrospective analysis of thyroid nodules: thyroid cancer risk factors in Suzhou, China. *Clinical Laboratory* 2018 **64** 333–338. (https://doi.org/10.7754/Clin.Lab.2017.170829)

2 Kwak JY, Han KH, Yoon JH, Moon HJ, Son EJ, Park SH, Jung HK, Choi JS, Kim BM & Kim EK. 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. (https://doi.org/10.1148/radiol.11110206)

3 Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, Gronan JJ, Beland MD, Desser TS, Frates MC, *et al*. ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS Committee. *Journal of the American College of Radiology* 2017 **14** 587–595. (https://doi.org/10.1016/j.jacr.2017.01.046)

4 Russ G, Bonnema SJ, Erdogan MF, Durante C, Ngu R & Leenhardt L. European Thyroid Association guidelines for ultrasound

malignancy risk stratification of thyroid nodules in adults: the EU-TIRADS. *European Thyroid Journal* 2017 **6** 225–237. (https://doi.org/10.1159/000478927)

5 Zhou JQ, Yin LX, Wei X, Zhang S, Song YY, Luo BM, Li JC, Qian LX, Cui LG, Chen W, *et al.* 2020 Chinese guidelines for ultrasound malignancy risk stratification of thyroid nodules: the C-TIRADS. *Endocrine* 2020 **70** 256–279. (https://doi.org/10.1007/s12020-020-02441-y)

6 Yang GC, Liebeskind D & Messina AV. Ultrasound-guided fifine-needle aspiration of the thyroid assessed by ultrafast Papanicolaou stain: data from 1135 biopsies with a two- to six-year follow-up. *Thyroid* 2001 **11** 581–589. (https://doi.org/10.1089/105072501750302895)

7 Bakhos R, Selvaggi SM, DeJong S, Gordon DL, Pitale SU, Herrmann M & Wojcik EM. Fine-needle aspiration of the thyroid: rate and causes of cytohistopathologic discordance. *Diagnostic Cytopathology* 2000 **23** 233–237. (https://doi.org/10.1002/1097-0339(200010)23:4<233::aid-dc3>3.0.co;2-l)

8 Gharib H, Papini E, Valcavi R, Baskin HJ, Crescenzi A, Dottorini ME, Duick DS, Guglielmi R, Hamilton CR, Zeiger MA, *et al.* American Association of Clinical Endocrinologists and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules. *Endocrine Practice* 2006 **12** 63–102. (https://doi.org/10.4158/EP.12.1.63)

9 Gharib H & Goellner JR. Fine-needle aspiration biopsy of the thyroid: an appraisal. *Annals of Internal Medicine* 1993 **118** 282–289. (https://doi.org/10.7326/0003-4819-118-4-199302150-00007)

10 Zhang YF, Xu JM, Xu HX, Liu C, Bo XW, Li XL, Guo LH, Liu BJ, Liu LN & Xu XH. Acoustic radiation force impulse elastography: a useful tool for differential diagnosis of thyroid nodules and recommending fine-needle aspiration: a Diagnostic Accuracy Study. *Medicine* 2015 **94** e1834. (https://doi.org/10.1097/MD.0000000000001834)

11 Chen L, Shi YX, Liu YC, Zhan J, Diao XH, Chen Y & Zhan WW. The values of shear wave elastography in avoiding repeat fine-needle aspiration for thyroid nodules with nondiagnostic and undetermined cytology. *Clinical Endocrinology* 2019 **91** 201–208. (https://doi.org/10.1111/cen.13992)

12 Cosgrove D, Barr R, Bojunga J, Cantisani V, Chammas MC, Dighe M, Vinayak S, Xu JM & Dietrich CF. WFUMB guidelines and recommendation on the clinical use of ultrasound elastography: part 4. Thyroid. *Ultrasound in Medicine and Biology* 2017 **43** 4–26. (https://doi.org/10.1016/j.ultrasmedbio.2016.06.022)

13 Mao F, Xu HX, Zhou H, Bo XW, Li XL, Li DD, Liu BJ, Zhang YF, Xu JM & Qu S. Assessment of virtual touch tissue imaging quantification and the ultrasound thyroid imaging reporting and data system in patients with thyroid nodules referred for biopsy. *Journal of Ultrasound in Medicine* 2018 **37** 725–736. (https://doi.org/10.1002/jum.14413)

14 Zhao CK, Chen SG, Alizad A, He YP, Wang Q, Wang D, Yue WW, Zhang K, Qu S, Wei Q, *et al.* Three-dimensional shear wave elastography or differentiating benign from malignant thyroid nodules. *Journal of Ultrasound in Medicine* 2018 **37** 1777–1788. (https://doi.org/10.1002/jum.14531)

15 Sun CY, Lei KR, Liu BJ, Bo XW, Li XL, He YP, Wang D, Ren WW, Zhao CK & Xu HX. Virtual touch tissue imaging and quantification (VTIQ) in the evaluation of thyroid nodules: the associated factors leading to misdiagnosis. *Scientific Reports* 2017 **7** 41958. (https://doi.org/10.1038/srep41958)

16 Zhang FJ, Zhao XM, Han RL, Du M, Li P & Ji XH. Comparison of acoustic radiation force impulse imaging and strain elastography in differentiating malignant from benign thyroid nodules. *Journal of Ultrasound in Medicine* 2017 **36** 2533–2543. (https://doi.org/10.1002/jum.14302)

17 Hahn S, Lee YH, Lee SH & Suh JS. Value of the strain ration ultrasonic elastography for differentiation of benign and malignant

soft tissue tumors. *Journal of Ultrasound in Medicine* 2017 **36** 121–127. (https://doi.org/10.7863/ultra.16.01054)

18 Dong FJ, Li M, Jiao Y, Xu JF, Xiong Y, Zhang L, Luo H & Ding ZM. Acoustic radiation force impulse imaging for detecting thyroid nodules: a systematic review and pooled metaanalysis. *Medical Ultrasonography* 2015 **17** 192–199. (https://doi.org/10.11152/mu.2013.2066.172.hyr)

19 Zhang YF, He Y, Xu HX, Xu XH, Liu C, Guo LH, Liu LN & Xu JM. Virtual touch tissue imaging on acoustic radiation force impulse elastography: a new technique for differential diagnosis between benign and malignant thyroid nodules. *Journal of Ultrasound in Medicine* 2014 **33** 585–595. (https://doi.org/10.7863/ultra.33.4.585)

20 Hu L, He NA, Ye L, Zhou HC, Zhong W & Zhang XS. Evaluation of the stiffness of tissues surrounding thyroid nodules with shear wave elastography. *Journal of Ultrasound in Medicine* 2018 **37** 2251–2261. (https://doi.org/10.1002/jum.14578)

21 Hu L, He NA, Xie L, Ye XJ, Liu X, Pei C, Zhou HC & Zhong W. Evaluation of the perinodular stiffness potentially predicts the malignancy of the thyroid nodules. *Journal of Ultrasound in Medicine* 2020 **39** 2183–2193. (https://doi.org/10.1002/jum.15329)

22 Zhang L, Ding ZM, Dong F, Wu HY, Liang WY, Tian HT, Ye XQ, Luo H & Xu JF. Diagnostic performance of multiple sound touch elastography for differentiating benign and malignant thyroid nodules. *Frontiers in Pharmacology* 2018 **9** 1359. (https://doi.org/10.3389/fphar.2018.01359)

23 Zhang L, Xu JF, Wu DY, Liang WY, Ye XQ, Tian HT & Dong FJ. Screening breast lesions uing shear modulus and its 1mm shell in sound touch elastography. *Ultrasound in Medicine and Biology* 2018 **55** 1–10. (https://doi.org/10.1016/j.ultrasmedbio.2018.11.013)

24 Zhou JQ, Zhan WW, Dong YJ, Yang ZF & Zhou C. Stiffness of the surrounding tissue of breast lesions evaluated by ultrasound elastography. *European Radiology* 2014 **24** 1659–1667. (https://doi.org/10.1007/s00330-014-3152-7)

25 Zhou JQ, Zhan WW, Chang C, Zhang XX, Jia Y, Dong YJ, Zhou C, Sun J & Grant EG. Breast lesion: evaluation with shear wear ealstography, with special emphasis on the 'stiff rim' sign. *Radiology* 2014 **272** 63–72. (https://doi.org/10.1148/radiol.14130818)

26 Gao L, Xi X, Jiang Y, Yang X, Wang Y, Zhu S, Lai X, Zhang X, Zhao R & Zhang B. Comparison among TIRADS (ACR TI-RADS and KWAK-TI-RADS) and 2015 ATA Guidelines in the diagnostic efficiency of thyroid nodules. *Endocrine* 2019 **64** 90–96. (https://doi.org/10.1007/s12020-019-01843-x)

27 Li JM, Li HR, Yang Y, Zhang XL & Qian LX. The KWAK TI-RADS and 2015 ATA guidelines for medullary thyroid carcinoma: combined with cell block-assisted ultrasound-guided thyroid fine-needle aspiration. *Clinical Endocrinology* 2020 **92** 450–460. (https://doi.org/10.1111/CEN.14121)

28 Zhang WB, Xu HX, Zhang YF, Guo LH, Xu SH, Zhao CK & Liu BJ. Comparisons of ACR TI-RADS, ATA guidelines, Kwak TI-RADS, and KTA/KSThR guidelines in malignancy risk stratification of thyroid nodules. *Clinical Hemorheology and Microcirculation* 2020 **75** 219–232. (https://doi.org/10.3233/CH-190778)

29 Xiao J, Xiao Q, Cong W, Li T, Ding SL, Shao CC, Zhang Y, Liu JN, Wu M & Jia HY. Discriminating malignancy in thyroid nodules: the nomogram versus the Kwak and ACR TI-RADS. *Otolaryngology: Head and Neck Surgery* 2020 **163** 1156–1165. (https://doi.org/10.1177/0194599820939071)

30 Wang D, Du LY, Sun JW, Hou XJ, Wang H, Wu JQ & Zhou XL. Evaluation of thyroid nodules with coexistent Hashimoto's thyroiditis according to various ultrasound-based risk stratification systems: a retrospective research. *European Journal of Radiology* 2020 **131** 109059. (https://doi.org/10.1016/j.ejrad.2020.109059)

31 Cibas ES & Ali SZ. The Bethesda system for reporting thyroid cytopathology. *Thyroid* 2009 **19** 1159–1165. (https://doi.org/10.1089/thy.2009.0274)

32 Moon WJ, Baek JH, Jung SL, Kim DW, Kim EK, Kim JY, Kwak JY, Lee JH, Lee JH, Lee YH, *et al*. Ultrasonography and the ultrasound based management of thyroid nodules: consensus statement and recommendations. *Korean Journal of Radiology* 2011 **12** 1–14. (https://doi.org/10.3348/kjr.2011.12.1.1)

33 Horvath E, Majlis S, Rossi R, Franco C, Niedmann JP, Castro A & Dominguez M. An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. *Journal of Clinical Endocrinology and Metabolism* 2009 **94** 1748–1751. (https://doi.org/10.1210/jc.2008-1724)

34 Kim HJ, Kwak MK, Choi IH, Jin SY, Park HK, Byun DW, Suh K & Yoo MH. Utility of shear wave elastography to detect papillary thyroid carcinoma in thyroid nodules: efficacy of the standard deviation elasticity. *Korean Journal of Internal Medicine* 2019 **34** 850–857. (https://doi.org/10.3904/kjim.2016.326)

35 Yang BR, Kim EK, Moon HJ, Yoon JH, Park VY & Kwak JY. Qualitative and semiquantitative elastography for the diagnosis of intermediate suspicious thyroid nodules based on the 2015 American Thyroid Association Guidelines. *Journal of Ultrasound in Medicine* 2018 **37** 1007–1014. (https://doi.org/10.1002/jum.14449)

36 Wang F, Chang C, Gao Y, Chen YL, Chen M & Feng LQ. Does shear wave elastography provide additional value in the evaluation of thyroid nodules that are suspicious for malignancy? *Journal of Ultrasound in Medicine* 2016 **35** 2397–2404. (https://doi.org/10.7863/ultra.15.09009)

37 Chong Y, Shin JH, Ko ES & Han BK. Ultrasonographic elastography of thyroid nodules: is adding strain ratio to colour mapping better? *Clinical Radiology* 2013 **68** 1241–1246. (https://doi.org/10.1016/j.crad.2013.06.023)

38 Dobruch-Sobczak K, Adamczewski Z, Szczepanek-Parulska E, Migda B, Wolinski K, Krauze A, Prostko P, Ruchala M, Lewinski A, Jakubowski W, *et al*. Histopathological verification of the diagnostic performance of the EU-TIRADS classification of thyroid nodules-results of a multicenter study performed in a previously iodine-deficient region. *Journal of Clinical Medicine* 2019 **8** 1781. (https://doi.org/10.3390/jcm8111781)

39 Szczepanek-Parulska E, Wolinski K, Dobruch-Sobczak K, Antosik P, Ostalowska A, Krauze A, Migda B, Zylka A, Lange-Ratajczak M, Banassiewicz T, *et al*. S-detect software vs. EU-TIRADS classification: a dual-center validation of diagnostic performance in differentiation of thyroid nodules. *Journal of Clinical Medicine* 2020 **9** 2495. (https://doi.org/10.3390/jcm9082495)