Original Article
The diagnosis value of acoustic radiation force impulse (ARFI) elastography for thyroid malignancy without highly suspicious features on conventional ultrasound

Bo-Ji Liu¹², Feng Lu¹², Hui-Xiong Xu¹²³, Le-Hang Guo¹², Dan-Dan Li¹², Xiao-Wan Bo¹², Xiao-Long Li¹², Yi-Feng Zhang¹², Jun-Mei Xu¹², Xiao-Hong Xu³, Shen Qu²⁴
¹Department of Medical Ultrasound, Shanghai Tenth People’s Hospital, Ultrasound Research and Education Institute, Tongji University School of Medicine, Shanghai 200072, China; ²Thyroid Institute, Tongji University School of Medicine, Shanghai 200072, China; ³Department of Ultrasound, Guangdong Medical College Affiliated Hospital, Zhanjiang 524001, China; ⁴Department of Endocrinology and Metabolism, Shanghai Tenth People’s Hospital, Tongji University School of Medicine, Shanghai 200072, China. *Equal contributors.

Received July 13, 2015; Accepted September 1, 2015; Epub September 15, 2015; Published September 30, 2015

Abstract: Objective: The aim of this study was to evaluate the potential diagnostic performance of acoustic radiation force impulse (ARFI) elastography in identifying malignancy in nodules that do not appear highly suspicious on conventional ultrasound (US). Methods: 330 pathologically confirmed thyroid nodules (40 malignant and 290 benign; mean size, 22.0±11.6 mm) not suspicious of malignancy on conventional US in 330 patients (mean age 52.8±11.7 years) underwent ARFI elastography before surgery. ARFI elastography included qualitative ARFI-induced strain elastography (SE) and quantitative point shear wave elastography (p-SWE). ARFI-induced SE image was assessed by SE score, while p-SWE was denoted with shear wave velocity (SWV, m/s). The diagnostic performance of four criteria sets was evaluated: criteria set 1 (ARFI-induced SE), criteria set 2 (p-SWE), criteria set 3 (either set 1 or 2), criteria set 4 (both set 1 and 2). Receiver operating characteristic curve (ROC) analyses were performed to assess the diagnostic performance. Results: SE score ≥4 was more frequently found in malignant nodules (32/40) than in benign nodules (30/290, P<0.001). The mean SWV of malignant nodules (3.64±2.23 m/s) was significantly higher than that of benign nodules (2.02±0.69 m/s) (P<0.001). ARFI-induced SE (set 1) had a sensitivity of 80.0% (32/40) and a specificity of 89.7% (260/290) with a cut-off point of SE score ≥4; p-SWE (set 2) had a sensitivity of 80.0% (32/40) and a specificity of 57.9% (168/290) with a cut-off point of SWV ≥2.15 m/s. When ARFI-induced SE and p-SWE were combined, set 3 had the highest sensitivity (92.5%, 37/40) while set 4 had the highest specificity (95.2%, 276/290). Conclusion: ARFI elastography can be used for differential diagnosis of malignant thyroid nodules without highly suspicious features on US. The combination of ARFI-induced SE and p-SWE leads to improved sensitivity and specificity.

Keywords: Ultrasound, thyroid, elastography, acoustic radiation force impulse, thyroid malignancy

Introduction
According to the recommendations of American Thyroid Association (ATA) guideline about thyroid nodules, thyroid ultrasound (US) examination should be performed in all patients with thyroid nodules [1]. Approximately 68% of the general population have thyroid nodules by high-resolution US and 5%-15% of them are malignant [2]. The differential diagnosis between benign and malignant thyroid nodules on US depends on the prevalence of suspicious characteristics including solid component, hypoechoegenicity or marked hypoechoegenicity, microcalcification, taller-than-wide shape and microlobulated or irregular margin [3-5]. However, up to 55% of benign nodules are also hypoechoic compared with thyroid parenchyma, making the feature of nodule hypoechoegenicity less specific [4]. In addition, papillary thyroid carcinomas (PTCs) without highly suspicious features on conventional US account for approximately 15% of all PTCs [6]. Some malignant nodules may be misdiagnosed due to the absence of typical suspicious features on US. Therefore, the sensitivity of conventional US in
ARFI elastography for malignancy without highly suspicious features

diagnosing malignant thyroid nodules varies from 38.8% to 90.9% [7-10], which causes a dilemma in management of those thyroid malignancies without highly suspicious features on US.

Fine needle aspiration cytology (FNAC) is the most accurate method for diagnosis of thyroid nodule and for selection of nodules for surgery, with a sensitivity range of 83%-100% and a specificity range of 60%-96% [11-13]. FNAC is recommended for solid thyroid nodule more than 1 cm in diameter according to ATA guidelines [1]. However, since it is an invasive and somewhat expensive procedure, many patients are reluctant to undergo FNAC. In addition, 3% to 22% FNAC results are non-diagnostic, so approximately 10%-30% of patients undergo diagnostic thyroid lobectomy for a benign diagnosis [14-16]. Finally, in consideration of a large amount of thyroid nodules detected in daily clinical practice, it is impossible to perform FNAC for all of them. Acoustic radiation force impulse (ARFI) elastography is a new technique used in parenchymal organs in recent years [17-20]. It is a novel US-based elastography method enabling qualitative and quantitative evaluation of tissue stiffness. Transverse shear wave propagation is generated when the target tissue within the sampling box is excited by an acoustic push pulse transmitted from the transducer. The strain change caused by the push pulse is named as virtual touch tissue imaging (VTI) (Siemens Medical Solutions, Mountain View, Calif, USA), which is a type of ARFI-induced strain elastography (SE) and is displayed as a grayscale image; while quantitative evaluation of the tissue stiffness is virtual touch tissue quantification (VTQ) (Siemens Medical Solutions, Mountain View, Calif, USA), which is point shear wave elastography (p-SWE). The diagnostic value of ARFI elastography for thyroid nodules in general population has been confirmed by several studies that showed a sensitivity range of 74.2%-87.0% for ARFI-induced SE and a sensitivity range of 57.0-90.9% for p-SWE [21-26]. However, to the best of our knowledge, no reports have been published to explore the diagnostic performance of ARFI elastography for thyroid nodules without highly suspicious features on US.

The aims of this study were to retrospectively evaluate the potential diagnostic performance of ARFI elastography in identifying malignancy in nodules that do not appear highly suspicious on conventional US and to explore the potential improvement in this performance by combining ARFI-induced SE with p-SWE.

Methods and materials

Patients

This retrospective study was approved by the Ethical Committee of the university hospital and informed consent was waived. Patients were included if they met the following inclusion criteria: (A) nodules with diameter ≥5 mm showed on conventional US; (B) the patients received conventional US and ARFI elastography examinations before surgery; (C) pathological confirmation was obtained after the examinations. Patients were excluded when (A) one or more accepted typical suspicious features showed on conventional US; (B) cystic portion was more than 50%; (C) the nodule area without cyst and calcification was smaller than the 6x5 mm sampling region of interest (ROI); (D) complete data were not available on US and ARFI elastography. The accepted highly suspicious US features included marked hypoechoegenicity (echogenicity lower than that of the surrounding strap muscle), infiltrative border, taller-than-wide shape (greater in the antero-posterior dimension than the transverse dimension) and intranodular microcalcifications (calcifications equal to or less than 1 mm in diameter and visualized as tiny punctate hyperechoic foci) [3, 6, 7, 27]. The nodules included in this study were nodules without any of the above-mentioned accepted highly suspicious US features.

One thyroid nodule was selected for analysis per patient. If nodules with and without highly suspicious features were coexisting, nodule without highly suspicious features was selected; When multiple nodules that were not highly suspicious of malignancy existed, the largest one was selected for evaluation. The flowchart of nodule selection was presented in Figure 1. Finally, 330 nodules (mean size, 22.0±11.6 mm; range, 5-62 mm) from 330 patients (80 men and 250 women; mean age, 52.8±11.7 years old; range, 21-78 years old) were included from April 2011 to September 2013. All of the nodules in this study underwent surgery. The indications for thyroidectomy were as follows: (1) confirmed or suspicious of malignancy by
Patients received US and ARFI elastography examination before surgery (n=752)

- i. incomplete data (n=14)
- ii. malignant-looking nodules on US (n=349)
- iii. cystic part ≥ 50% (n=59)

330 nodules in 330 patients were included (n=330)

Malignant nodules (n=40)
Benign nodules (n=290)

Figure 1. The diagram of nodule selection for the included patients. Malignant-looking nodule was defined as nodule without marked hypoechogenicity, infiltrative border, taller-than-wide shape or microcalcification on ultrasound image. US, ultrasound; ARFI, acoustic radiation force impulse; n, number of patients.

FNAC; (2) non-diagnostic, indeterminate results or follicular lesions detected by FNAC; (3) patients' anxiety and refusal of FNAC; (4) compressive symptoms caused by large nodules or associated large nodules.

Conventional ultrasound (US) examination

Conventional US and ARFI elastography were performed using the same S2000 US scanner (Siemens Medical Solutions, Mountain View, Calif, USA) containing a 4-9 MHz linear transducer by one of two investigators with ten years of experience in US and three years of experience in US elastography. The patients were lying in the supine position with dorsal flexion of the head to make the neck skin fully exposed. The coupling medium was smeared on the US transducer to reduce interference of air with skin and probe. For each thyroid nodule, the maximal diameter was measured and US images were recorded.

ARFI elastography

ARFI elastography was then performed by the same investigator after conventional US as described in the previous literatures [9, 10, 21]. A sampling box was placed on the target nodule while performing real-time B-mode US imaging with some surrounding thyroid tissue. To avoid the influence of pre-compression on the elastic properties of the tissue, the investigator was asked to place the transducer gently on the skin surface with minimum pressure. The patient was required to hold the breath and the transducer was motionless during the acquisition of the elastic images [28]. Tissue within the sampling box was mechanically excited using acoustic pulses to generate localized tissue displacements, which resulted in ARFI-induced SE [21]. ARFI-induced SE which qualitatively evaluated the tissue stiffness was displayed as a gray scale image and was assessed by SE score. Afterwards, quantitative measurement of p-SWE was started and a region of interest (ROI) was placed on the target nodule with the cystic or calcified areas avoided. The placement of p-SWE ROI was guided by the ARFI-induced SE image. Generally, the ROI was placed at the center of small nodules. As for relatively large nodules, ROI was placed at the solid part, usually the inner peripheral portion of the nodule, with an aim to avoid potential liquefaction or necrosis in the central region. When the nodule was heterogeneous, the ROI was placed at the stiffest location indicated by the ARFI-induced SE image and repeated measurement was then carried out. The quantitative measurement was performed by calculating the shear wave velocity (SWV, m/s). In p-SWE, SWVs of the nodule were measured at least 7 times. The ARFI elastography images were recorded. The whole process of US and ARFI elastography examination lasted approximately 10 minutes.

Image interpretation

Two investigators who were blind to the pathological results independently reviewed the
ARFI elastography for malignancy without highly suspicious features

Table 1. Clinical and conventional US characteristics of the patients with thyroid nodules

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Overall</th>
<th>Benign</th>
<th>Malignant</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>330</td>
<td>290 (87.9%)</td>
<td>40 (12.1%)</td>
<td>NA</td>
</tr>
<tr>
<td>Number of nodules</td>
<td>330</td>
<td>290 (87.9%)</td>
<td>40 (12.1%)</td>
<td>NA</td>
</tr>
<tr>
<td>Gender (men/women)</td>
<td>80/250</td>
<td>75/215</td>
<td>5/35</td>
<td>0.076</td>
</tr>
<tr>
<td>Mean age (years)</td>
<td>52.8±11.7</td>
<td>52.8±12.0</td>
<td>51.5±11.8</td>
<td>0.504</td>
</tr>
<tr>
<td>Range</td>
<td>21-78</td>
<td>21-77</td>
<td>23-78</td>
<td></td>
</tr>
<tr>
<td>Nodule size (mm)</td>
<td>22.0±11.6</td>
<td>22.8±11.5</td>
<td>16.2±11.5</td>
<td>0.001*</td>
</tr>
<tr>
<td>Range</td>
<td>5-62</td>
<td>6-62</td>
<td>5-52</td>
<td></td>
</tr>
<tr>
<td>Nodule distribution (single/multiple)</td>
<td>94/236</td>
<td>81/209</td>
<td>13/27</td>
<td>0.577</td>
</tr>
<tr>
<td>Background of Hashimoto’s thyroiditis</td>
<td>59</td>
<td>44 (15.2%)</td>
<td>15 (37.5%)</td>
<td>0.002*</td>
</tr>
</tbody>
</table>

Ultrasound features

<table>
<thead>
<tr>
<th>Internal component</th>
<th></th>
<th></th>
<th></th>
<th>0.034*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete solid</td>
<td>301</td>
<td>261 (90.0%)</td>
<td>40 (100%)</td>
<td></td>
</tr>
<tr>
<td>Predominantly solid</td>
<td>29</td>
<td>29 (10.0%)</td>
<td>0 (0%)</td>
<td></td>
</tr>
</tbody>
</table>

Echogenicity

| Hyperechoic                            | 4       | 4 (1.4%) | 0 (0%)  | <0.001* |
| Isoechoic                              | 92      | 91 (31.4%) | 1 (2.5) |         |
| Hypoechoic                             | 114     | 80 (27.6)   | 34 (85.0%) |         |
| Mixed echoic                           | 120     | 115 (39.7%) | 5 (12.5%) |         |
| Halo sign (present/absent)             | 208/122 | 200/90 | 8/32     | <0.001* |

Note: *indicates statistically significant difference. NA, not applicable.

Table 2. ARFI elastography features of benign and malignant thyroid nodules without highly suspicious features on US

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Benign nodules (n=290)</th>
<th>Malignant nodules (n=40)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE score</td>
<td>&lt;0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score 1</td>
<td>1 (0.3%)</td>
<td>0 (0%)</td>
<td></td>
</tr>
<tr>
<td>Score 2</td>
<td>65 (22.4%)</td>
<td>1 (2.5%)</td>
<td></td>
</tr>
<tr>
<td>Score 3</td>
<td>194 (66.9%)</td>
<td>7 (17.5%)</td>
<td></td>
</tr>
<tr>
<td>Score 4</td>
<td>29 (10.0%)</td>
<td>22 (55.0%)</td>
<td></td>
</tr>
<tr>
<td>Score 5</td>
<td>1 (0.3%)</td>
<td>10 (25.0%)</td>
<td></td>
</tr>
<tr>
<td>Score 6</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td></td>
</tr>
<tr>
<td>p-SWE</td>
<td>&lt;0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean SWV (m/s)</td>
<td>2.02±0.69</td>
<td>3.64±2.23</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>0.50-4.40</td>
<td>1.12-8.40</td>
<td></td>
</tr>
</tbody>
</table>

Note: ARFI, acoustic radiation force impulse; SE, strain elastography; p-SWE, point shear wave elastography; SWV, shear wave velocity. *indicates statistically significant difference.

Table 1. Clinical and conventional US characteristics of the patients with thyroid nodules

The gray scales of ARFI-induced SE image were classified into 2 categories (white and black) through the comparison between the nodule and the surrounding thyroid tissue. ARFI-induced SE score was categorized according to Xu’s VTI scoring system [10, 24] as follows: SE score-1, nearly all white; SE score-2, predominantly white with a few black portions; SE score-3, equal white and black portions; SE score-4, predominantly black with a few white portions; SE score-5, almost completely black; and SE score-6, completely black. Higher SE scores indicated stiffer tissue and higher risk of malignancy. For p-SWE, SWV results were expressed in m/s (range, 0.5-8.4 m/s). The maximal and the minimal SWVs were eliminated and the mean SWV of the nodule were calculated by the remaining 5 measurements. SWVs displayed as “x. xx m/s” were regarded as invalid measurements.

Statistical analysis

To explore the diagnostic performance in differentiating thyroid malignant nodules without
highly suspicious features on US from benign nodules, the following four sets of criteria were specified: criteria set 1, ARFI-induced SE alone (SE score ≥4); criteria set 2, p-SWE alone (SWV ≥ cut-off point); criteria set 3, either ARFI-induced SE or p-SWE (in compliance with either criteria set 1 or 2); criteria set 4, combination of ARFI-induced SE and p-SWE (in compliance with both criteria set 1 and 2).

Continuous variables were expressed as mean ± standard deviation (SD) and range. Two sets of measurement data and enumeration data were compared by t test or χ² test, while categorical data by using Mann-Whitney rank sum test of non-parametric tests. The sensitivities, specificities, positive predictive values (PPVs), negative predictive values (NPVs) and accuracies of four criteria sets were calculated. Differences in sensitivity, specificity and accuracy were compared using the Mc Nemar test, while PPV and NPV using χ² test. The diagnostic performance of the criteria sets were evaluated by receiver operating characteristic curve (ROC) analysis which was expressed as the area under ROC (AUC) and its 95% confidence interval (CI). Comparison of AUCs was performed by using the method described by Hanley [29]. The cut-off point was obtained when Youden index (sensitivity + specificity-1) reached the maximum value. Odds ratio (OR) value and 95% CI were used to evaluate the association of four criteria sets with malignancy. A two-tailed P-value <0.05 indicated statistically significant
ARFI elastography for malignancy without highly suspicious features

Statistical analysis was performed using SPSS software (version 14.0, Chicago, IL, USA).

Results

Pathological results

All the nodules were confirmed by pathological results after surgery. Of the 330 nodules, 40 were malignant and the remaining 290 were benign. Among the 40 malignant nodules, there were 36 PTCs (90.0%), 2 follicular carcinomas (5.0%), 1 medullary carcinoma (2.5%) and 1 diffuse large B-cell lymphoma (2.5%). The 290 benign nodules included 235 nodular goiters (81.0%), 35 adenomas (12.1%) and 20 Hashimoto's nodules (6.9%).

Conventional US and ARFI elastography features

The differences in conventional US features between malignant and benign nodules without highly suspicious features on US were presented in Table 1. Statistically significant differences were found between the two groups in background of Hashimoto’s thyroiditis, internal component (P=0.034), echogenicity (P<0.001) and halo sign (P<0.001). Background of Hashimoto’s thyroiditis, complete solid component, hypoechoogenicity and presence of halo sign were more frequently found in malignant nodules.

The ARFI elastography characteristics were presented in Table 2. In ARFI-induced SE, no

Figure 3. Images in a 45-year-old woman with papillary thyroid carcinoma without highly suspicious features on US. At conventional ultrasound, a 13 mm nodule in the left lobe of the thyroid is solid, isoechoic and hypoechoic, absence of halo sign (arrows) (A). At ARFI elastography, SE score-5 (arrows) (B), and SWV of 2.30 m/s at p-SWE are assigned (C). The nodule is confirmed to be papillary thyroid carcinoma pathologically (D) (H&E staining, magnification, 10×20). SE, strain elastography; p-SWE, point shear wave elastography; SWV, shear wave velocity.
ARFI elastography for malignancy without highly suspicious features

Table 3. The diagnostic performance of ARFI elastography for 4 criteria sets

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Set 1 (ARFI-induced SE)</th>
<th>Set 2 (P-SWE)</th>
<th>Set 3 (ARFI-induced SE or p-SWE)</th>
<th>Set 4 (ARFI-induced SE + P-SWE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cut-off point</td>
<td>SE score-4</td>
<td>2.15 m/s</td>
<td>score-4 or 2.15 m/s</td>
<td>score-4 and 2.15 m/s</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>80.0 (32/40)</td>
<td>80.0 (32/40)</td>
<td>92.5 (37/40)</td>
<td>67.5 (27/40)</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>89.7* (260/290)</td>
<td>57.9 (168/290)</td>
<td>52.4 (152/290)</td>
<td>95.2* (276/290)</td>
</tr>
<tr>
<td>Positive predictive value (%)</td>
<td>51.6* (32/62)</td>
<td>20.8 (32/154)</td>
<td>21.1 (37/175)</td>
<td>65.9 (27/41)</td>
</tr>
<tr>
<td>Negative predictive value (%)</td>
<td>97.0 (260/268)</td>
<td>95.5 (168/176)</td>
<td>98.1 (152/155)</td>
<td>95.5 (276/289)</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>88.5* (292/330)</td>
<td>60.6 (200/330)</td>
<td>57.3 (189/330)</td>
<td>91.8* (303/330)</td>
</tr>
<tr>
<td>Area under ROC</td>
<td>0.87</td>
<td>0.78</td>
<td>0.73</td>
<td>0.81</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>0.80-0.94</td>
<td>0.70-0.87</td>
<td>0.66-0.80</td>
<td>0.72-0.90</td>
</tr>
<tr>
<td>Odds ratio</td>
<td>34.7</td>
<td>5.5</td>
<td>13.6</td>
<td>40.9</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>14.6-82.1</td>
<td>2.5-12.4</td>
<td>4.1-45.1</td>
<td>17.5-96.0</td>
</tr>
</tbody>
</table>

Note: ARFI, acoustic radiation force impulse; SE, strain elastography; SWE, shear wave elastography; ROC, receiver operating characteristic curve. 4 criteria sets: Criteria set 1, SE score ≥4; Criteria set 2, SWV value ≥2.15 m/s; Criteria set 3, in compliance with either criteria set 1 or criteria set 2; Criteria set 4, in compliance with both criteria set 1 and criteria set 2. *P<0.001, in comparison with those in set 2. #P=0.018, in comparison with that in set 1. ^P<0.001, in comparison with that in set 2. ΔP<0.001, in comparison with those in set 2 and set 3.

The diagnostic performance of 4 criteria sets

The sensitivities, specificities, PPVs, NPVs, accuracies, AUCs and their 95% CI, ORs and their 95% CI of four criteria sets were shown in Table 3 and Figure 4. The sensitivity and NPV of set 1 was similar to set 2 (both P>0.05), while the specificity, PPV and accuracy of set 1 were statistically higher than that of benign nodules (both P<0.05) and lower than that of malignant nodules (both P<0.001).

The sensitivities, specificities, PPVs, NPVs, accuracies, AUCs and their 95% CI, ORs and their 95% CI of four criteria sets were shown in Table 3 and Figure 4. The sensitivity and NPV of set 1 was similar to set 2 (both P>0.05), while the specificity, PPV and accuracy of set 1 were statistically higher than that of benign nodules (both P<0.05). The sensitivity was improved by set 3 compared with that of set 1 (92.5% vs. 80.0%, P=0.193) and set 2 (92.5% vs. 80.0%, P=0.193). The specificity of set 4 (95.2%) was statistically higher than that of benign nodules (both P<0.05) and lower than that of malignant nodules (both P<0.001).
P<0.001) and set 3 (P<0.001). No significant differences were found among the criteria sets 1, 2 and 4 in terms of AUCs (all P>0.05), and AUC of criteria set 3 (0.725) was statistically lower than that of criteria set 1 (0.871, P=0.003).

The mean size of false positive (FP) nodules was significantly larger than that of true positive (TP) ones for both ARFI-induced SE (20.5±9.8 mm vs. 13.2±6.4 mm, P=0.001) and p-SWE (21.1±10.9 mm vs. 15.3±10.9 mm, P=0.009). The FP nodules for criteria set 1 (ARFI-induced SE) and set 2 (p-SWE) were 30 (including 26 nodular goiters and 4 Hashimoto’s nodules) and 122 (including 107 nodular goiters, 9 Hashimoto’s nodules and 6 adenomas) respectively; 16 (53.3%) and 108 (88.5%) of them were correctly diagnosed by criteria set 4. The false negative (FN) nodules for criteria set 1 and set 2 were 8 (6 PTCs, 1 follicular carcinoma and 1 lymphoma) and 8 (7 PTCs and 1 follicular carcinoma) respectively; 5 (62.5%) and 5 (62.5%) of them were correctly diagnosed by criteria set 3.

Discussion
Thyroid nodules detected by palpation or radiography method are recommended to be evaluated by US to assess possible malignancy [1]. Nevertheless, it is difficult for conventional US to distinguish malignant thyroid nodules without highly suspicious features on US from authentic benign nodules.

The mean size of malignant nodules was statistically smaller than that of benign nodules in accordance with previous studies [3, 10, 30]. With regard to the US features, there was statistically significant difference in background of Hashimoto’s thyroiditis, which is consistent with previous studies that higher prevalence of thyroid cancer was found coexisting with Hashimoto’s thyroiditis than in the overall population [31, 32]. Complete solid component and hypoechogenicity were more commonly found in malignant thyroid nodules in this series. However, their specificities were hampered since up to 55% of benign nodules were hypoechogenic and 47.7-75.9% of benign nodules were solid [3, 4, 10]. Absence of halo sign was more frequently found in malignant thyroid nodules, which is in agreement with previous studies [9, 10, 33].

In the current study, the diagnostic performance of ARFI elastography for differentiation between malignant nodules without highly suspicious features on US and benign nodules was firstly evaluated. The sensitivity (80.0%) and specificity (89.7%) of ARFI-induced SE score in the current study were similar to those reported by Xu et al [10] and the cut-off point was also SE score-4. Thus, ARFI-induced SE could also be used for differential diagnosis in malignant thyroid nodules without highly suspicious features on US that would not cause any loss in diagnostic performance. For p-SWE, it had a higher sensitivity (80.0%) and lower specificity (57.9%) compared with several previous studies [9, 10, 21]. The mean SWV of benign nodules (2.02 m/s) was nearly the same as that in some previous studies (range: 2.10-2.34 m/s); while it was interesting that the mean SWV of malignant nodules without highly suspicious features on US (3.64 m/s) was markedly lower than that in previous studies (range: 4.82-6.34 m/s) [9, 10, 34]. As a consequence, the specificity of p-SWE decreased in comparison with general population. The finding indicates that the stiffness of thyroid malignant nodules without highly suspicious features on US is not as obvious as that of typical thyroid malignancies, which is possibly related to the uniform pathological changes in thyroid malignancy without highly suspicious features on US. However, future study is needed to confirm this hypothesis. Related to this, the cut-off point of SWV for p-SWE (2.15 m/s) was observably lower than that in the published reports (range: 2.42-2.87 m/s) [9, 10, 21, 25, 34, 35]. The result is relevant in clinical practice, which reminds us that a lower SWV value does not necessarily indicate a benign nodule. The thyroid malignancy without highly suspicious features on US should be excluded when making a diagnosis of benign nodule.

Distinguished from simple method of combination in previous studies [36, 37], 2 combination approaches were firstly evaluated in the current study. Criteria set 3 was either ARFI-induced SE or p-SWE, while criteria set 4 was the combination of ARFI-induced SE and p-SWE. Although no significant improvement were found for the combination criteria sets in terms of AUCs, the sensitivity reached the highest (92.5%) for criteria set 3 whereas the specificity was the highest (95.2%) for criteria set 4.
In addition, the accuracy was the highest (91.8%) for criteria set 4. Therefore, ARFI-induced SE or p-SWE should not be used independently. The two combination approaches should be considered together with an aim to achieve both high sensitivity and high specificity. As can be seen in the results section, the false positive and false negative nodules reduced significantly after applying the two combination approaches.

There were some limitations in the current study. Firstly, only nodules ≥5 mm were included because of the limit of 6×5 mm sampling ROI when measuring SWV of p-SWE. Therefore, malignant nodules of <5 mm might be omitted. However, thyroid malignancy <5 mm generally does not need aggressive management thus the limitation will not influence the clinical practice. Secondly, selection bias might be present for included patients who were referred for surgery, and it couldn’t represent the situation of common population completely. Additionally, it was a single-center and retrospective study. Multi-center and prospective studies with large case series were needed to verify the usefulness of ARFI elastography in diagnosing thyroid malignancy without highly suspicious features on US.

Conclusions

ARFI elastography is helpful for differentiation between malignant thyroid nodules without highly suspicious features on US and real benign nodules. The combination of ARFI-induced SE and p-SWE leads to improved sensitivity and specificity. However, different combination methods should be applied. Finally, the stiffness in thyroid malignant nodules without highly suspicious features on US might be lower than that in typical thyroid malignancies; therefore, attention should be paid to make a benign diagnosis solely based on the results of SWV measurement on p-SWE.

Acknowledgements

This work was supported in part by Grant SHDC12014229 from Shanghai Hospital Development Center, Grant 14441900900 from Science and Technology Commission of Shanghai Municipality, Grant 2012045 from Shanghai Municipal Human Resources and Social Security Bureau, and Grant 81401417 from the National Natural Science Foundation of China.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Hui-Xiong Xu, Department of Medical Ultrasound, Shanghai Tenth People’s Hospital, Ultrasound Research and Education Institute, Tongji University School of Medicine, No. 301 Yanchangzhong Road, Shanghai 200072, China. E-mail: xuhuixiong@126.com

References

ARFI elastography for malignancy without highly suspicious features


