The application of qualitative shear wave elastography in differential diagnosis of thyroid nodules

Qian-Qi Liu¹, Zhuo-Wen Yang², Jie Tian¹, Ping Xing¹, Hai-Ling Tang¹, Qian-Yi Qiu¹, Chang-Jun Wu¹

Departments of ¹Ultrasound, ²Endocrinology, First Affiliated Hospital of Harbin Medical University, Harbin, China

Received October 8, 2017; Accepted January 11, 2019; Epub April 15, 2019; Published April 30, 2019

Abstract: Objective: To investigate the diagnostic performance of qualitative shear wave elastography (SWE) in the characterization of thyroid nodules, compared with quantitative SWE and grayscale ultrasound (US). Methods: A total of 174 thyroid nodules from 174 patients were preoperatively evaluated by grayscale US and SWE. The quantitative SWE metrics were measured. For the qualitative SWE analysis, a four-pattern classification was established according to the color distribution features of the elastic maps and assigned for each enrolled nodule, as 1 for no meaningful findings, 2 for a capsule-related color, 3 for marginal color, and 4 for interior color. Results: Intra-observer and inter-observer agreement of the qualitative classification proposed in this study were substantial (κ=0.793 and 0.756, respectively). The area under the receiver operating characteristic curve (AUC) of the proposed classification was similar to that of grayscale US (P=0.73) and quantitative SWE metrics (P=0.90). Compared with the quantitative SWE, the qualitative classification yielded a higher sensitivity (P<0.05) and a similar specificity (P=0.75). Adding SWE features to grayscale US, either qualitatively or quantitatively, improved the overall specificity (both P<0.001). Among all data sets, the optimal diagnostic performance came from the combined set of grayscale US and the qualitative classification (AUC, 0.890), with 84.4% sensitivity and 94.6% specificity. Conclusion: The qualitative four-pattern classification proposed in this study was feasible and highly reproducible in the SWE evaluation of thyroid nodules. SWE imaging features, especially the qualitative classification, have the potential to facilitate ultrasound diagnosis for thyroid nodules.

Keywords: Ultrasound, shear wave elastography, thyroid nodule, diagnosis

Introduction

Ultrasound (US) is regarded as a first-line imaging method for the detection and identification of thyroid nodules. US elastography, a novel modality capable of revealing tissue biomechanical properties, has been recommended as an additional tool for conventional US for the diagnosis and management of thyroid nodules [1-3]. There are two main types of ultrasound elastography technologies that are frequently used, strain elastography and shear wave elastography [4]. Strain elastography (SE), the most widely used type of elastography, has been shown to be both highly sensitive and specific in the evaluation of thyroid nodules [5]. However, there are some intrinsic defects of SE, including high operator dependence and substantial interobserver variability [6, 7]. The diagnostic efficacy of SE has been challenged [8, 9].

Recently, shear wave elastography (SWE), a highly reproducible technique, has received increasing attention [10]. This system enables the generation of focused acoustic impulses to induce transverse shear waves, whose propagation velocities can be measured and used to deduce tissue stiffness quantitatively, expressed as either shear wave speed (m/s) or Young’s modulus (kPa) [4, 11]. A growing number of studies shows promising applications of quantitative SWE in the differential diagnosis of thyroid nodules [10, 12-14]. However, the published thresholds for identifying thyroid cancers are fairly inconsistent, ranging from approximately 40 kPa to 90 kPa [13, 14]. There would be controversy with respect to introducing these results into clinical practice.
Quantitative assessment is not the only way to evaluate stiffness with the SWE technique. As confirmed by the literature, qualitative information provided by the color elastic map also helps in stiffness analysis [15-18]. Nonetheless, up until now, the value of qualitative SWE in the diagnosis of thyroid nodules has not been well established. One published study referring to this topic has proved the feasibility of qualitative SWE [19]. However, in their study, SWE images were assessed using Ueno and Rago scales, two widely utilized scoring systems for SE [20, 21]. Theoretically, there are many differences between SE and SWE, including the excitation methods and the measured physical quantities [4]. Moreover, SE images are generally analyzed according to the prevalence of 3 colors (e.g., red, blue, and green) [6, 22], whereas there are 6 main color levels in SWE images (i.e., black, dark blue, light blue, green, orange, red) [11]. These distinctions may lead to inaccuracy and confusion when we interpret SWE images using scoring systems for SE.

Therefore, the purpose of this study was to investigate the diagnostic performance of qualitative SWE in the differential diagnosis of thyroid nodules, compared with quantitative SWE and grayscale US.

Materials and methods

Patients

This retrospective study was approved by our Institutional Review Board, and written informed consent was waived. The methods in this study were performed in accordance with the approved study plan. From November 2016 to January 2017, 190 consecutive patients with thyroid nodules scheduled to undergo surgery were imaged by grayscale US and SWE. Among these included patients, sixteen were later excluded for the following reasons: six patients had previous invasive procedures for thyroid nodules, seven patients had cystic or almost cystic nodules, and three patients lacked histopathologic confirmation. One nodule per patient was included. For patients with multiple nodules, the most suspicious nodule on US or the largest one was selected for analysis. Ultimately, the study cohort consisted of 174 thyroid nodules in 174 patients. Surgical histopathologic results were used as the reference standard.

Image acquisition

Grayscale US and SWE imaging were performed using the same Aixplorer US instrument (SuperSonic Imagine, Aix-en-Provence, France) equipped with a 4-15 MHz linear-array transducer. All the examinations were performed by one of the two radiologists with more than 3 years of experience in US elastography. The patients were positioned in a supine position with their necks fully exposed.

The conventional US scanning was carried out according to the relevant practice guidelines [23]. Radiologists who performed the examination recorded nodule location and size (the maximal diameter measured on the US image). All grayscale US images were digitally stored for further analysis.

SWE images were acquired on a longitudinal plane following grayscale US by the same radiologist. When the SWE mode was activated, a rectangular field of view (FOV) was generated on the grayscale image, in which information of the tissue stiffness was obtained pixel by pixel and coded as a translucent color map ranging gradually from dark blue (soft) to red (stiff). During SWE data acquisition, the probe was kept stationary with minimal pressure. To reduce extra pre-compression, the FOV was first placed to include the anterior strap muscles of neck, and the patients were instructed to relax their necks, till the muscle layers were depicted without red/orange areas. At that point, the FOV was moved to include the target nodule. When the image was stable for several seconds, the operator froze and saved it for subsequent analysis. All SWE images were analyzed at the default color scale of 0-100 kPa.

Image analysis

The quantitative SWE was measured by the radiologist who performed the SWE imaging. The build-in circular region of interest (ROI) (Q-Box™; SuperSonic Imagine) was placed on the stiffest portion inside the nodule, taking care to avoid blood vessels, calcifications and cystic portions. The Q-box™ diameter was fixed at 2-mm in all cases. The SWE metrics of $E_{\text{max}}$ (the maximum elastic modulus), $E_{\text{mean}}$ (mean elastic modulus), $E_{\text{min}}$ (minimum elastic modulus), and $E_{\text{SD}}$ (elastic modulus standard deviation) were automatically calculated.
The grayscale US and qualitative SWE features were reviewed by another blinded radiologist with 6 years of experience in thyroid US. To reduce bias, the images were reviewed in a random order.

For each nodule, grayscale sonographic features including composition, echogenicity, margins, presence and type of calcifications, and shape were evaluated. Based on previous criteria, grayscale US features predictive of malignancy were defined as marked hypoechochogenicity (more hypoechoic compared with the adjacent strap muscle), poorly defined margin (a microlobulated or irregular margin), microcalcifications (tiny, punctate, hyperechoic foci without acoustic shadowing), and taller-than-wide shape [8, 24]. The nodules were classified as suspicious for malignancy when they had one or more of the malignant sonographic characteristics. Otherwise, they were considered to be unsuspicious.

Qualitative SWE were evaluated visually, according to the color distribution features of the elastic maps. Here, the nodules were classified into four patterns (Figure 1). Pattern 1 (no meaningful findings) was assigned when no meaningful high stiffness color signal was revealed, including homogeneous blue, vertical strip artifacts [11, 17] and irregular color patches (Figure 1A). Pattern 2 (capsule-related color) was assigned to cases with areas coded in high stiffness color related to the thyroid capsules, which may extend to the interior of the nodule or the superficial cervical fascia (Figure 1B). Pattern 3 (marginal color) was assigned to cases with localized color signal at the nodule margin adjacent to thyroid parenchyma (Figure 1C). Pattern 4 (interior color) was assigned when localized colors were observed interior of the lesion heterogeneously (Figure 1D). The localized color described in pattern 3 and 4, in particular, refers to increased stiffness that neither continues as vertical bands beyond the nodule nor belongs to color patches that distribute diffusely. Additionally, if a nodule indicated visual feature satisfying more than one patterns, the higher one was chosen (Figure 2). This four-pattern classification was modified from the criteria proposed by Tozaki et al. [17] for breast cancer (hereafter, the Tozaki criteria).

To evaluate the potential effect of adding SWE to grayscale US in predicting malignancy, combinations of grayscale US and SWE features were established, in which the thyroid nodules were divided into two categories. A nodule was classified as suspicious for malignancy when suspicious results were found by both grayscale US and SWE (i.e., greater than the cut-off value); otherwise, it would be classified as unsuspicious.

Observer agreement of the four-pattern classification

In a subset of 40 nodules, the observer agreement of the four-pattern classification was assessed. Three months after the first evaluation, the same radiologist reviewed the SWE images for intra-observer agreement calculation. Meanwhile, another two radiologists with 4 and 5 years of experience in thyroid US assessed all the elastograms separately, and the results were compared with each other for inter-observer agreement calculation.

Statistical analysis

The relevant findings on grayscale US and SWE were compared between the benign and malignant lesions. Categorical variables were compared using a chi-squared test and Fisher's exact test, while continuous variables were assessed using Student's t test. In addition, a Mann-Whitney U test was applied to compare continuous variables with a non-normal distribution. Receivers operating characteristic (ROC) curves were constructed for grayscale US, SWE features and the combined data sets. Areas under the ROC curve (AUC) were calculated and compared using a z test. For the SWE features, the optimal cut-off points to predict malignancy were determined using the Youden index (maximum of sensitivity+specificity-1). Sensitivity, specificity, accuracy, positive predictive values (PPVs) and negative predictive values (NPVs) were calculated for these cut-off points. Univariate and multivariate (logistic regression analysis) and odds ratios (OR) were calculated to determine independent predictors for thyroid malignancy. Inter-observer and intra-observer agreements were evaluated by cohen's κ statistic, and classified as slight (κ < 0.20), fair (κ=0.20-0.40), moderate (κ=0.40-0.60), substantial (κ=0.60-0.80), and almost perfect (κ=0.80-1.00) [25]. The statistical analysis was performed using the SAS software package (version 9.2, SAS Institute, Cary, NC). A two-tailed P < 0.05 was considered statistically significant.
Qualitative SWE in differentiation of thyroid nodules

Figure 1. Representative images of the qualitative four-pattern classification. A. Scans of a nodular goiter (arrows). Shear wave elastography (SWE) image (top) is coded mostly blue with irregular cyan patches (no meaningful findings; pattern 1). B. Scans of a follicular adenoma (arrows). The SWE image (top) shows high stiffness areas related to capsules, extending to the intrinsic nodule tissues and the superficial cervical fascia (capsule-related color; pattern 2). C. Scans of a papillary thyroid carcinoma (arrows). SWE image (top) shows colored areas at the nodule margin adjacent to parenchyma (marginal color; pattern 3). D. Scans of a papillary thyroid carcinoma (arrows). Increased stiffness areas are present inside the nodule on SWE image (top) (interior color; pattern 4).

Figure 2. Papillary thyroid carcinoma (arrows) in an 18-year-old female patient. Grayscale image (bottom) shows a mass with a poorly defined margin and microcalcifications that were classified as suspicious for malignancy. The shear wave elastography (SWE) image (top) shows a qualitative feature satisfying pattern 2 (capsule-related color) and pattern 4 (interior color) simultaneously and the higher one, pattern 4 was chosen. The elasticity metric of $E_{\text{max}}$ is 85.9 kPa.

Results

Demographic information and clinic features

A total of 174 patients (mean age, 47.4±10.5 years; range, 18-77 years) with 174 thyroid nodules (64 malignant, 110 benign) were evaluated in this study. These nodules underwent surgical excision mainly owing to the following reasons: biopsy-proven malignancy, compressive symptom, rapid growth, suspicion of malignancy, or patient anxiety. There were 140 women (mean age, 48.0±10.6 years; range, 18-77 years) and 34 men (mean age, 44.5±9.9 years; range, 20-58 years). Of the 174 thyroid nodules, 110 (63.2%) were benign and 64 (36.8%) were malignant. Amongst the demographic characteristics, age was associated with malignancy ($P=0.001$), but male sex was not ($P=0.342$). The histopathologic diagnoses are summarized in Table 1. The maximal diameter of the malignant nodules (14.4±12.7 mm; range, 3.7-78.0...
Table 1. Histopathologic results of the 174 thyroid lesions

<table>
<thead>
<tr>
<th>Histopathologic results</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Nodular goiter</td>
<td>95</td>
</tr>
<tr>
<td>Follicular adenoma</td>
<td>9</td>
</tr>
<tr>
<td>Focal thyroiditis</td>
<td>4</td>
</tr>
<tr>
<td>Others</td>
<td>2</td>
</tr>
<tr>
<td>Malignant</td>
<td></td>
</tr>
<tr>
<td>Papillary carcinoma</td>
<td>61</td>
</tr>
<tr>
<td>Follicular carcinoma</td>
<td>1</td>
</tr>
<tr>
<td>Medullar carcinoma</td>
<td>1</td>
</tr>
<tr>
<td>Poorly differentiated carcinoma</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>174</td>
</tr>
</tbody>
</table>

mm) was similar to the maximal diameter of the benign nodules (17.1±10.5 mm; range, 4.0-52.9 mm) (P=0.13).

Grayscale US features between benign and malignant nodules

Suspicious US features were more frequently seen in malignant nodules than benign nodules (all P < 0.001) (Table 2). The combined suspicious grayscale US features yielded an AUC of 0.813 (95% CI: 0.762, 0.864), and the corresponding sensitivity and specificity were 95.3%, 67.3%, respectively (Table 3).

Diagnostic performance of quantitative SWE

$E_{\text{mean}}, E_{\text{max}}, E_{\text{min}}$, and $E_{\text{SD}}$ were significantly higher in the malignant nodules than in benign nodules ($P<0.0001$ for all) (Table 2). The diagnostic performances of the quantitative elasticity metrics are summarized in Table 3. Among all metrics, $E_{\text{max}}$ yielded the highest AUC of 0.801 (95% CI: 0.732, 0.870), which showed no significant difference with that of grayscale US ($P=0.79$). At the optimal cut-off value of 52.1 kPa, a sensitivity of 73.4% and specificity of 76.4% were achieved. $E_{\text{max}}$ was selected for the combined diagnosis.

Diagnostic performance of qualitative SWE classification

A significant difference was shown in the malignancy rate of the four-pattern classification ($P<0.001$) (Table 2). The diagnostic performance of the qualitative classification is summarized in Table 3. The AUC value for the classification was 0.798 (95% CI: 0.731, 0.864), which was comparable to that of the grayscale US ($P=0.73$) and $E_{\text{max}}$ ($P=0.90$) (Figure 3A). Using the optimal cut-off point between pattern 2 and 3, sensitivity and specificity was 89.1% and 74.6%, respectively. No significant difference was shown between the specificity of the qualitative classification and $E_{\text{max}}$ ($P=0.75$), but the sensitivity of the classification was significantly higher ($P=0.02$) (Figure 4). Table 4 shows the details of the 7 carcinomas misdiagnosed by both the qualitative and quantitative methods (Figure 5).

Observer agreement of the four-pattern classification

The intra-observer and inter-observer agreement of the classification were both substantial, with $\kappa$ values of 0.793 and 0.756, respectively.

Diagnostic performance combined grayscale US and SWE for thyroid nodules

The corresponding diagnostic performances are shown in Table 3. Among all the data sets, the AUC of combination of US and the qualitative classification (US+pattern) was the highest, at 0.890 (95% CI: 0.840, 0.940), which was significantly higher than grayscale US alone ($P=0.008$) and US+$E_{\text{max}}$ ($P=0.024$) (Figure 3B). Also, the difference in AUC between US and US+$E_{\text{max}}$ was not significant ($P=0.489$). Adding SWE features to grayscale US, either qualitatively or quantitatively, improved the overall specificity ($P<0.0001$ for both). There was no significant difference of sensitivity between grayscale US and US+pattern ($P=0.079$), but the overall sensitivity for US+$E_{\text{max}}$ decreased significantly ($P=0.0003$).

Univariate and multivariate regression analysis

At univariate logistic regression analysis, the presence of solitary nodule (OR: 2.219; 95% CI: 1.184, 4.156; $P=0.013$), $E_{\text{max}}$ (OR: 8.932; 95% CI: 4.400, 18.131; $P<0.001$) and qualitative classification (OR: 23.847; 95% CI: 9.748, 58.335; $P<0.001$) were statistically significant predictors for thyroid malignancy, but gender ($P=0.288$) and nodule size ($P=0.080$) were not. On the other hand, older age (> 45 years; OR: 0.329; 95% CI: 0.171, 0.637; $P=0.001$) was predictive of thyroid benignity. At multivariate logistic regression analysis, the presence of a solitary nodule (OR: 2.692; 95% CI: 1.167,
Qualitative SWE in differentiation of thyroid nodules

Table 2. Grayscale US, qualitative SWE and quantitative SWE of benign and malignant thyroid nodules

<table>
<thead>
<tr>
<th>Pathology</th>
<th>Malignancy (n=64)</th>
<th>Benign (n=110)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grayscale US*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marked hypoehogenicity</td>
<td>31 (48.4%)</td>
<td>8 (7.3%)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>poorly defined margin</td>
<td>36 (56.3%)</td>
<td>23 (20.9%)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Microcalcification</td>
<td>41 (64.1%)</td>
<td>17 (15.5%)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Taller-than-wide shape</td>
<td>20 (31.3%)</td>
<td>5 (4.6%)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Qualitative classification*</td>
<td></td>
<td></td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>1- nomeaningful findings</td>
<td>6 (9.1%)</td>
<td>60 (90.9%)</td>
<td></td>
</tr>
<tr>
<td>2- capsule-related color</td>
<td>1 (4.4%)</td>
<td>22 (95.7%)</td>
<td></td>
</tr>
<tr>
<td>3- marginal color</td>
<td>24 (75.0%)</td>
<td>8 (25.0%)</td>
<td></td>
</tr>
<tr>
<td>4- interior color</td>
<td>33 (62.3%)</td>
<td>20 (37.7%)</td>
<td></td>
</tr>
<tr>
<td>Quantitative metrics (kPa)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_{\text{max}}$</td>
<td>75.20 (50.60, 113.95)</td>
<td>38.50 (29.10, 51.70)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>$E_{\text{mean}}$</td>
<td>64.70 (39.25, 90.65)</td>
<td>32.05 (23.30, 43.40)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>$E_{\text{min}}$</td>
<td>47.85 (23.65, 63.80)</td>
<td>23.75 (15.60, 32.30)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>$E_{\text{SD}}$</td>
<td>7.10 (4.75, 13.50)</td>
<td>3.70 (2.60, 5.90)</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

Note: *Data in parenthesis are medians with interquartile ranges, otherwise percentages.

Table 3. Diagnostic performance of grayscale US, SWE, and US Combined with SWE in differentiating thyroid nodules

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Accuracy (%)</th>
<th>AUC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grayscale US</td>
<td>NA</td>
<td>95.3</td>
<td>67.3</td>
<td>62.9</td>
<td>96.1</td>
<td>0.813 (0.762-0.864)</td>
</tr>
<tr>
<td>Quantitative metrics (kPa)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_{\text{max}}$</td>
<td>52.1</td>
<td>73.4</td>
<td>76.4</td>
<td>64.4</td>
<td>83.2</td>
<td>75.3</td>
</tr>
<tr>
<td>$E_{\text{mean}}$</td>
<td>44.1</td>
<td>68.8</td>
<td>77.3</td>
<td>63.8</td>
<td>81.0</td>
<td>74.1</td>
</tr>
<tr>
<td>$E_{\text{min}}$</td>
<td>35.3</td>
<td>64.1</td>
<td>80.0</td>
<td>65.1</td>
<td>79.3</td>
<td>74.1</td>
</tr>
<tr>
<td>$E_{\text{SD}}$</td>
<td>4.7</td>
<td>76.6</td>
<td>63.6</td>
<td>55.1</td>
<td>82.4</td>
<td>68.4</td>
</tr>
<tr>
<td>Qualitative classification</td>
<td>&gt; 2</td>
<td>89.1*</td>
<td>74.6</td>
<td>67.1</td>
<td>92.1</td>
<td>79.9</td>
</tr>
<tr>
<td>US+$E_{\text{max}}$</td>
<td>NA</td>
<td>71.9**</td>
<td>95.5**</td>
<td>90.2</td>
<td>85.4</td>
<td>86.8</td>
</tr>
<tr>
<td>US+pattern</td>
<td>NA</td>
<td>84.4</td>
<td>93.6**</td>
<td>88.5</td>
<td>91.2</td>
<td>90.2</td>
</tr>
</tbody>
</table>

Note: AUC: area under the receiver operating characteristic curve; PPV: positive predict value; NPV: negative predict value; US: ultrasound; SWE: shear wave elastography; NA: not available; CI: confidence interval. *P < 0.05, in comparison with Qualitative metrics. **P < 0.05, in comparison with US only. #P < 0.05, in comparison with US+$E_{\text{max}}$.

6.211; P=0.020) and a qualitative classification (OR: 26.542; 95% CI: 10.151, 69.397; P < 0.001) were identified as independent predictors of thyroid malignancy, but $E_{\text{max}}$ was excluded by the stepwise variable selection procedure (P=0.086). Likewise, older age (> 45 years; OR: 0.316; 95% CI: 0.133, 0.752; P=0.009) was identified to be an independent predictor for thyroid benignity.

Discussion

Quantitative SWE metrics have been used to evaluate thyroid nodules. Nevertheless, the optimal cut-off values in prior studies varied considerably. Liu et al. [13] proposed using an $E_{\text{mean}}$ of 39.3 kPa, Kim et al. [24] proposed using $E_{\text{mean}}$ at the higher level of 62 kPa, and Park et al. [14] proved that the optimal parameters for predicting malignancy were $E_{\text{mean}}$ of 85.2 kPa and $E_{\text{max}}$ of 94.0 kPa. This wide disparity could be ascribed to the different study designs, the inclusion and exclusion criteria, the scale of patient cohort, and the diverse proportion of malignancy among their study populations. Furthermore, the pre-compression level was not adequately standardized in
Qualitative SWE in differentiation of thyroid nodules

Figure 3. Receiver operating characteristic (roc) curves for all qualitative and quantitative SWE features (A) and US alone and US combined with SWE features (B).

Figure 4. Papillary thyroid microcarcinoma (arrows) in a 47-year-old woman. The grayscale image (bottom) shows a nodule considered to be suspicious for malignancy. The shear wave elastography (SWE) image (top) shows increased stiffness signals at the marginal region considered to be pattern3, but with $E_{\text{max}}$ lower than the cut-off value (52.1 kPa). This nodule was misdiagnosed by quantitative SWE.

Further studies with a larger population are necessary to determine the most appropriate quantitative metric in discriminating thyroid malignancy.

In previous studies, quantitative SWE features of thyroid nodules were obtained with the transducer orientated in different planes. Recently, Ganga-dhar et al. [27] reported that the diagnostic efficacy of thyroid nodule characterization was similar in both transverse and longitudinal orientations. Investigators have also adopted longitudinal planes to avoid potential artifacts generated by carotid, trachea and uneven acoustic forces distribution [10, 28]. Conversely, Samir et al. [29] indicated that the transverse plane was the optimal measurement plane to evaluate follicular-patterned thyroid lesions. In our clinical practice, however, we found it difficult to summarize certain qualitative patterns using SWE images obtained in transverse planes. Compared with longitudinal planes, anatomic structures in transverse planes are more complicated, and vary obviously as the transducer moves, which
Table 4. False-negative cases according to both qualitative and quantitative SWE

<table>
<thead>
<tr>
<th>Case</th>
<th>Sex</th>
<th>Age</th>
<th>Suspicious grayscale features*</th>
<th>$E_{\text{max}}$ value (kPa)</th>
<th>Qualitative classification</th>
<th>Maximal diameter (mm)</th>
<th>Pathology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>45</td>
<td>1</td>
<td>34.3</td>
<td>1</td>
<td>6.8</td>
<td>PTMC</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>46</td>
<td>3</td>
<td>24.0</td>
<td>1</td>
<td>3.2</td>
<td>PTMC</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>38</td>
<td>4</td>
<td>30.9</td>
<td>1</td>
<td>4.2</td>
<td>PTMC</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>35</td>
<td>2</td>
<td>25.4</td>
<td>1</td>
<td>8.7</td>
<td>PTMC</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>44</td>
<td>3</td>
<td>29.9</td>
<td>2</td>
<td>7.1</td>
<td>PTMC</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>45</td>
<td>3</td>
<td>28.7</td>
<td>1</td>
<td>6.5</td>
<td>PTMC</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>27</td>
<td>2</td>
<td>32.3</td>
<td>1</td>
<td>9.3</td>
<td>PTMC</td>
</tr>
</tbody>
</table>

Note: PTMC: papillary thyroid microcarcinoma; FTC: follicular thyroid carcinoma. *Number of suspicious grayscale features.

Figure 5. Scans of Case 2 in Table 4, a surgery-proven papillary thyroid microcarcinoma (arrows). It is a nodule considered to be suspicious on grayscale image (bottom), but negative on both quantitative and qualitative shear wave elastography (SWE) (top). The SWE image shows the nodule and the surrounding tissue coded homogeneously blue (pattern 1), with an $E_{\text{max}}$ of 24.0 kPa.

Multivariate logistic regression indicated that the proposed qualitative classification was an independent predictor for thyroid malignancy. As compared with quantitative metrics, the proposed classification showed a similar diagnostic performance ($P=0.90$) and specificity ($P=0.75$), while the sensitivity was significantly higher ($P=0.02$). Notably, with the cut-off point set between pattern 2 and 3, eight of fifteen carcinomas omitted by $E_{\text{max}}$ were detected by the proposed classification (Figure 4). These were mostly small nodules measured to be “soft” (Young’s modulus below the cut-off $E_{\text{max}}$).
of 52.1 kPa), but with patterns indicating malignancy. The qualitative SWE features, moreover, can be assessed independent of the numerical cut-off values, which were inconsistent as discussed earlier. Additionally, our results indicated that the four-pattern classification was highly reproducible with substantial intra-observer and inter-observer agreement. Therefore, in thyroid SWE, the proposed classification could serve as a time-saving and practical evaluation method for differential diagnosis.

In this study, all malignant nodules misclassified by both quantitative and qualitative SWE were papillary thyroid carcinomas (PTC) with maximal diameters ≤ 10 mm, namely, papillary thyroid microcarcinomas (PTMCs) (Table 4). This implied a suboptimal power of SWE in detecting PTMC, a finding compatible with previous results [13, 32]. The cause of this phenomenon is unclear. One possibility is that PTMC is soft due to its genetic and biological properties [13]. The soft appearance of these PTMCs, however, might also be spurious and partially related to technical limitations, such as the slightly poorer spatial resolution and/or an insufficient generation of regular shear waves in small nodules [33, 34]. Further study is needed to elucidate this topic. On the other hand, all the PTMCs misdiagnosed by the SWE method were categorized as suspicious on grayscale US. A comprehensive interpretation of B-mode findings and elastographic features is necessary for small nodules.

Previous studies have confirmed that the combination of conventional US and elastography could achieve better performance in detecting thyroid malignancy [12, 35, 36]. Nevertheless several studies have also reported limited additional value of quantitative SWE parameters as compared to conventional US alone [32, 37]. Similarly, in the present study, diagnostic performance of US were not significantly improved by adding $E_{\text{max}}$ ($P=0.489$). The combination of the qualitative classification and grayscale US, however, yielded the highest AUC among all data sets, which was significantly higher than that of conventional US ($P=0.008$) or the combined data of US and $E_{\text{max}}$ ($P=0.024$). Our results also revealed that adding SWE features to grayscale US, either qualitatively or quantitatively, yielded higher specificity than US alone ($P < 0.001$ for both).

There were some limitations to our study. First, all the SWE images were analyzed at the default color scale. However, color contrast will vary with the adjustment of the scale, by which further qualitative information about stiffness might be obtained [31]. Second, assessment based on a visual method is subjective, but in our study, a substantial observer agreement of the proposed classification was found. In the future, by using computer-aided-diagnosis systems, more meaningful results might be achieved [26]. Third, this is a retrospective study, and the patients enrolled were scheduled for thyroid surgery, so our results may be influenced by sampling bias. Fourth, the sample size of this preliminary study is relatively small, so a subgroup analysis of the pathological type was precluded. In addition, potential influencing factors for SWE measurements and classifications, such as the presence of calcifications, lesion position, and cystic change were not considered in our study [13, 37, 38].

**Conclusions**

Our initial study revealed that qualitative SWE assessed by the proposed four-pattern classification is feasible in thyroid nodule characterization. Compared with quantitative SWE, qualitative SWE offers a comparable diagnostic performance with a higher sensitivity, so it may represent SWE in clinical practice. SWE features, especially the qualitative classification, have the potential to be valuable adjuncts to conventional US. However, further studies are warranted to validate our results.

**Acknowledgements**

This work was supported by grants 81271648 and 81671697 from the National Natural Science Foundation of China.

**Disclosure of conflict of interest**

None.

**Address correspondence to:** Chang-Jun Wu, Department of Ultrasound, First Affiliated Hospital of Harbin Medical University, 23 Youzheng Road, Harbin, Heilonjiang, China. Tel: (+86-451)855559-05; E-mail: changjunwu1964@hotmail.com

---

**Qualitative SWE in differentiation of thyroid nodules**

Qualitative SWE in differentiation of thyroid nodules

References


Qualitative SWE in differentiation of thyroid nodules

nostic value of conventional ultrasonography and shear wave elastography in the prediction of thyroid lesions malignancy. PLoS One 2013; 8: e81532.


