Real-time Ultrasound Elastography in the Differential Diagnosis of Benign and Malignant Thyroid Nodules

Yurong Hong, MS, Xueming Liu, MD, Zhiyu Li, MD, Xiufang Zhang, BS, Meifeng Chen, BS, Zhiyan Luo, MS

Objective. The purpose of this study was to evaluate the diagnostic utility of real-time ultrasound elastography in differentiating benign from malignant thyroid nodules. Methods. A total of 90 consecutive patients with thyroid nodules who were referred for surgical treatment were examined in this prospective study. One hundred forty-five nodules in these patients were examined by B-mode ultrasound, color Doppler ultrasound, and ultrasound elastography. The final diagnosis was obtained from histologic findings. Tissue stiffness on ultrasound elastography was scored from 1 (low stiffness over the entire nodule) to 6 (high stiffness over the entire nodule and surrounding tissue). Results. On real-time ultrasound elastography, 86 of 96 benign nodules (90%) had a score of 1 to 3, whereas 43 of 49 malignant nodules (88%) had a score of 4 to 6 ($P < .001$), with sensitivity of 88%, specificity of 90%, a positive predictive value of 81%, and a negative predictive value of 93%. The predictivity of ultrasound elastographic measurement was independent of the nodule size. High sensitivity (88%) and specificity (93%) were also observed in 68 nodules that had a greatest diameter of 1 cm or less. Conclusions. Real-time ultrasound elastography is a promising imaging technique that is useful in the differential diagnosis of thyroid cancer. Key words: elastography; real-time; thyroid nodules; ultrasound.

Thyroid nodules are very common and are found in 4% to 8% of adults by palpation, 41% by ultrasound, and 50% by pathologic examination at autopsy. Most nodules are benign, with less than 5% of them being malignant. The challenge of managing thyroid nodules is to reassure the majority of patients who have benign disease and to diagnose the minority of patients who will prove to have malignant disease. Conventional ultrasound provides information regarding characteristics correlated with the risk of cancer, such as hypoechogenicity, blurred or spiculated margins, microcalcifications, an anteroposterior/transverse diameter (A/T) of 1 cm or greater, and intranodular vascularity. However, the sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV) for these criteria are extremely variable from study to study. No ultrasound feature has both high sensitivity and high specificity.
Ultrasound Elastography of Thyroid Nodules

Palpation is a basic and important method in the assessment of thyroid nodules. Malignant nodules tend to be much harder than benign ones. Conventional ultrasound does not provide direct information corresponding to the hardness of a nodule. Ultrasound elastography was developed to obtain information on tissue stiffness noninvasively. This technique can evaluate the degree of distortion of a tissue under application of an external force and is based on the principle that the softer parts of tissues deform easier than the harder parts under compression, thus allowing an objective determination of tissue stiffness. Ultrasound elastography has been successfully applied to breast lesions, the prostate, the pancreas, and lymph nodes.\(^9\)–\(^12\) The thyroid gland is well positioned for elastographic examination. It can be easily assessed and efficiently compressed against underlying anatomic structures with an ultrasound probe. Several studies have used ultrasound elastography for thyroid evaluation.\(^13\)–\(^19\) Tanaka et al\(^15\) showed sensitivity of 89.1% and specificity of 59.4%. However, as discussed below, these findings may be misleading and may have been due to a sample selection bias. Rago et al\(^14\) reported sensitivity of 97% and specificity of 100% for this technique.

The objective of this study was to evaluate the diagnostic utility of real-time ultrasound elastography in differentiating benign from malignant thyroid nodules.

Materials and Methods

Patients

Ninety consecutive patients who were referred for surgical treatment were examined in this prospective study. The study was conducted during a 34-month period from January 2006 to October 2008. All patients underwent surgery, and the final diagnosis was based on the results of histopathologic examination of resected thyroid gland tissue. Ethical approval for this study was granted by the Medical Research Ethics Committee of our university, and informed consent was obtained from all patients.

Conventional Thyroid Ultrasound Imaging and Real-time Ultrasound Elastography

Both conventional ultrasound imaging and real-time ultrasound elastography were performed in a single session with an EUB-8500 ultrasound system (Hitachi Medical Corporation, Tokyo, Japan) and a 6- to 13-MHz linear array transducer (EUP-L54M).

For all patients, the ultrasound examination started with B-mode imaging. A careful evaluation of the following ultrasound parameters was performed on all thyroid nodules: \(A/T\) \((\leq 1\) or \(>1\) cm); echogenicity (hyperechoic, isoechoic, or hypoechoic with respect to normal thyroid parenchyma); calcifications (no calcifications, microcalcifications [presence of hyperechoic spots \(<2\) mm with or without acoustic shadowing], coarse dense calcifications, or peripheral rimlike calcifications); and margins (well defined and smooth, blurred, or spiculated). The presence and pattern of blood flow on color Doppler imaging was also evaluated and classified as follows:\(^6\): type 1, absence of blood flow; type 2, perinodular and absent intranodular blood flow; and type 3, intranodular blood flow.

Ultrasound elastographic measurement was performed during the ultrasound examination with the same real-time instrument and the same probe. The probe was placed on the neck with light pressure, and a box was highlighted by the operator, which included the nodule and sufficient surrounding thyroid tissue to be evaluated. The principle of ultrasound elastography is to acquire 2 ultrasound images (before and after tissue compression by the probe) and track tissue displacement by assessing the propagation of the imaging beam. Dedicated software (Combined Autocorrelation Method; Hitachi Medical Corporation) able to provide an accurate measurement of tissue distortion was used. The ultrasound elastogram was displayed over the B-mode image in a color scale that ranged from red for components with the greatest elastic strain (ie, softest components) to blue for those with no strain (ie, hardest components). To minimize the interobserver and intraobserver variability, the freehand compression applied on the neck region was standardized by real-time measurement displayed on a numeric scale (graded 1–5) to maintain an intermediate level optimal for ultrasound elastographic evaluation (2–3).

It was important that the level of pressure was maintained constant throughout the examination. All examinations were performed by the
same operator (Y.H.), who had 8 years of experience in thyroid ultrasound. Static and moving images were also recorded and reviewed subsequently by a second radiologist (X.L.), who had 30 years of experience in thyroid ultrasound. The elastograms thus obtained were classified into 6 patterns according to the elasticity (Table 1).

**Histopathologic Diagnosis**

The histologic diagnosis served as the reference standard for comparison of conventional ultrasound imaging and ultrasound elastography. After scanning and elastography, all 145 nodules were examined histologically. Routine hematoxylin-eosin staining and immunohistologic examination were performed.

**Statistical Analysis**

The SPSS for Windows version 13.0 software package (SPSS Inc, Chicago, IL) was used for statistical data analysis. Data were expressed as mean ± SD. The χ² test was performed to determine whether the elasticity scores were different between the two groups. To assess the diagnostic value of conventional ultrasound imaging and ultrasound elastography compared with the histologic results, cross-table tests were performed. The diagnostic sensitivity, specificity, PPV, and NPV were calculated. The ability of ultrasound imaging and ultrasound elastography to independently differentiate benign from malignant nodules was analyzed by a binary logistic regression technique with the final histopathologic diagnosis as the reference standard. P < .05 was considered statistically significant.

**Results**

One hundred forty-five thyroid nodules were examined in this prospective study. Twenty-seven patients had 2 or more nodules. The mean age of the examined patients (16 men and 74 women) was 46 ± 13 years (range, 18–72 years). Histologic examination revealed 49 malignant nodules (44 papillary carcinomas, 2 medullary carcinomas, 1 follicular carcinoma, and 2 poorly differentiated metastatic adenocarcinomas) and 96 benign nodules (83 hyperplastic nodules, 8 follicular adenomas, 3 Hashimoto thyroiditis nodules, and 2 subacute thyroiditis nodules).

**Conventional Ultrasound Imaging**

Nodule hypoechogenicity, microcalcifications, blurred or spiculated margins, and an A/T of 1 cm or greater were the ultrasound patterns most predictive of malignancy. The intranodular blood flow pattern was not predictive of malignancy (Table 2).

**Ultrasound Elastography**

Scoring by the two examiners was coincident in 138 of 145 thyroid nodules. In 7 nodules, the final scores agreed after conjoint reexamination of the recorded movies.

**Table 1. Elasticity Scores**

<table>
<thead>
<tr>
<th>Score</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Low stiffness over the entire nodule; the entire nodule is evenly shaded green, as is the surrounding thyroid tissue</td>
</tr>
<tr>
<td>2</td>
<td>Low stiffness over most of the nodule; the nodule is almost completely green but with some blue spots</td>
</tr>
<tr>
<td>3</td>
<td>Low stiffness at the periphery; high stiffness in the center of the nodule; the central part of the nodule is blue; the peripheral part is green</td>
</tr>
<tr>
<td>4</td>
<td>High stiffness over most of the nodule; the nodule is almost completely blue but with some green spots</td>
</tr>
<tr>
<td>5</td>
<td>High stiffness over the entire nodule; the entire nodule is evenly shaded blue</td>
</tr>
<tr>
<td>6</td>
<td>High stiffness over the entire nodule and surrounding tissue; both the nodule and surrounding area are blue</td>
</tr>
</tbody>
</table>

**Table 2. Diagnostic Value of Different Ultrasound Features for Identification of Malignant Thyroid Nodules**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Benign, n (n = 96)</th>
<th>Malignant, n (n = 49)</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>PPV, %</th>
<th>NPV, %</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypoechogenicity</td>
<td>28</td>
<td>42</td>
<td>86</td>
<td>71</td>
<td>60</td>
<td>91</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Spot microcalcifications</td>
<td>6</td>
<td>27</td>
<td>55</td>
<td>94</td>
<td>82</td>
<td>80</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Blurred or spiculated margins</td>
<td>15</td>
<td>45</td>
<td>92</td>
<td>84</td>
<td>75</td>
<td>95</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>A/T ≥1 cm</td>
<td>4</td>
<td>12</td>
<td>24</td>
<td>96</td>
<td>75</td>
<td>71</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Intranodular blood flow</td>
<td>40</td>
<td>17</td>
<td>35</td>
<td>58</td>
<td>30</td>
<td>64</td>
<td>.416</td>
</tr>
</tbody>
</table>
Ultrasound Elastography of Thyroid Nodules

On ultrasound elastography, score 1 was found in 30 nodules, all benign lesions; score 2 in 55 nodules, 4 carcinomas and 51 benign; score 3 in 7 nodules, 2 carcinomas and 5 benign; score 4 in 21 nodules, 13 carcinomas and 8 benign; score 5 in 17 nodules, 15 carcinomas and 2 benign; and score 6 in 15 nodules, all carcinomas (Figure 1). Thus, 86 of 96 benign nodules had a score of 1 to 3, whereas 43 of 49 malignant nodules had a score of 4 to 6 ($P \lt 0.001$), with sensitivity of 88%, specificity of 90%, a PPV of 81%, and an NPV of 93% (Table 3).

Among benign nodules, the score was 5 in 2 hyperplastic nodules, both with peripheral rimlike calcifications. Six nodules had a score of 4 (2 subacute thyroiditis, 3 hyperplastic nodules with coarse dense calcifications, and 1 with microcalcifications). Another 2 hyperplastic nodules without calcifications also had a score of 4 (greatest diameter, 29 and 32 mm, respectively; both protruded from thyroid posterior capsule). Three hyperplastic nodules with coarse dense calcifications and 2 with microcalcifications had a score of 3. Among malignant nodules, the score was 2 in 2 poorly diff-

Figure 1. Elastograms (left) and conventional ultrasound images (right) of thyroid nodules in 6 patients. A, Follicular adenoma, score 1, in a 58-year-old woman. B, Follicular adenoma, score 2, in a 31-year-old woman. C, Follicular carcinoma, score 3, in a 20-year-old woman. D, Papillary carcinoma, score 4, in a 35-year-old woman. E, Papillary carcinoma, score 5, in a 34-year-old woman. F, Papillary carcinoma, score 6, in a 48-year-old woman.
differentiated metastatic adenocarcinomas and 2 papillary carcinomas (greatest diameter, 4 and 3 mm). One follicular carcinoma (greatest diameter, 20 mm) and 2 papillary carcinomas (greatest diameter, 7 and 13 mm; both protruded from thyroid anterior capsule) had a score of 3.

The predictivity of ultrasound elastographic measurement was independent of the nodule size, with high sensitivity and specificity also observed in 68 nodules that had a greatest diameter of 1 cm or less (Table 3).

A receiver operating characteristic curve for distinguishing malignant from benign nodules according to the elasticity score is shown in Figure 2. The area under the receiver operating characteristic curve for diagnosing malignant thyroid nodules was 0.94 (95% confidence interval, 0.9–0.98). As a critical value, an elasticity score of 4 corresponded to sensitivity of 88%, specificity of 90%, a PPV of 81%, and an NPV of 93%.

**Discussion**

Elastography is a technique that uses ultrasound to analyze the stiffness of a nodule by measuring the amount of distortion that occurs when the nodule is subjected to external pressure. The elasticity scoring system initially proposed by Ueno and Itoh was useful for comparing breast ultrasound elastographic results. In our study, tissue stiffness was scored from 1 to 6 on the basis of subjective analysis of the elastograms. Scores of 4 to 6 showed an elevated predictive value for malignancy, with sensitivity of 88%, specificity of 90%, a PPV of 81%, and an NPV of 93%.

Elastography examines the mechanical and elastic properties of soft tissue, which rely on the composition and structural organization of the macromolecules. Some pathologic conditions induce considerable changes in the soft tissue structure, modifying the elastic properties and leading to increased firmness and reduced mobility of the involved tissue. Malignant thyroid nodules tend to be much harder than benign ones, such as papillary thyroid carcinoma, the most common histologic type of thyroid malignancy. Papillary thyroid carcinoma has complex papillae with a central fibrovascular stalk. Psammoma bodies and fibrosis are often found in them. In our study, 41 of 44 papillary thyroid carcinomas (93%) had a score of 4 to 6. However, not all malignant tumor tissue has increased stiffness. In our study, 2 poorly differentiated metastatic adenocarcinomas (100%) had a score of 2. Another follicular carcinoma had a score of 3. The gross anatomy and cellular patterns of follicular carcinoma overlap with those of benign follicular adenoma, and this

<table>
<thead>
<tr>
<th>Size, cm</th>
<th>Score</th>
<th>Benign, n</th>
<th>Malignant, n</th>
<th>P</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>PPV, %</th>
<th>NPV, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤1</td>
<td>1–3</td>
<td>39</td>
<td>3</td>
<td>&lt;.001</td>
<td>88</td>
<td>93</td>
<td>88</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>4–6</td>
<td>3</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;1</td>
<td>1–3</td>
<td>47</td>
<td>3</td>
<td>&lt;.001</td>
<td>87</td>
<td>87</td>
<td>74</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>4–6</td>
<td>7</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>1–3</td>
<td>86</td>
<td>6</td>
<td>&lt;.001</td>
<td>88</td>
<td>90</td>
<td>81</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>4–6</td>
<td>10</td>
<td>43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Receiver operating characteristic curve for distinguishing between benign and malignant thyroid nodules. The area under the curve for diagnosing malignant thyroid nodules was 0.94 (95% confidence interval, 0.9–0.98).
kind of thyroid malignancy can be differentiated from benign follicular adenoma only when capsular or vascular invasion are discovered at histologic examination. Some benign nodules may have increased stiffness. In our study, 2 nodules (100%) with the histologic diagnosis of subacute thyroiditis (disappearance of the follicular epithelium, replaced by a rim of histiocytes and giant cells, interstitial fibrosis, infiltration of lymphocytes, and plasma cells) had a score of 4. Calcifications within nodules may increase nodular stiffness. In our study, 2 hyperplastic nodules with peripheral rimlike calcifications had a score of 5; 3 with coarse dense calcifications and 1 with microcalcifications had a score of 4; and 3 with coarse dense calcifications and 2 with microcalcifications had a score of 3. 

Rago et al reported that the predictivity of ultrasound elastographic measurement was independent of the nodular size, with sensitivity of 100% and specificity of 100% being observed in 9 nodules that had a greatest diameter of 0.8 to 1 cm. In our study, ultrasound elastograms predicted malignancy with 88% sensitivity and 93% specificity in nodules that had a greatest diameter of 1 cm or less. Two papillary thyroid microcarcinomas (greatest diameters, 4 and 3 mm) had a score of 2, and 1 (greatest diameter, 4 mm) had a score of 3. This may have been due to nonaxial and out-of-plane motion during freehand external compression. Another issue pertains to the relatively large section thickness (~5 mm) of the ultrasound probe used in this study. A thicker section and sample volume would mean that information from smaller lesions would be averaged out with that from other tissues. Further studies will be necessary to understand whether ultrasound elastographic measurement can give reliable results in nodules that have a greatest diameter of 1 cm or less.

During elastography, by comparing echoes made with compression, one can obtain information about how hard or soft the tissues are relative to their surroundings. In our study, 2 papillary thyroid carcinomas (greatest diameter, 7 and 13 mm) protruded from the thyroid anterior capsule, and most of the surrounding tissues were not normal thyroid tissue but cervical muscular tissue or other fibrous connective tissues. They both had a score of 2. Two other hyperplastic nodules (greatest diameter, 29 and 32 mm) protruded from the thyroid posterior capsule and extended to the posterior sternum. They both had a score of 4.

There were some limitations in this study that need to be addressed. First, fine-needle aspiration was not used to evaluate the thyroid nodules. It may have increased the percentage of benign nodules in the series. Second, only 2 poorly differentiated metastatic adenocarcinomas and 1 follicular thyroid carcinoma were included in the study group. Whether elastography has limited capabilities in the differentiation of follicular or metastatic cancers is still uncertain. Third, the impact of background Hashimoto thyroiditis or other abnormalities of the thyroid parenchyma was not adequately assessed. Because of the design of this study, we were unable to assess a given observer’s ability to diagnose thyroid gland cancer on the basis of elastographic findings. Further studies are required to evaluate the interobserver and intraobserver variability and reliability of ultrasound elastography for thyroid cancer diagnosis.

In conclusion, ultrasound elastography is a promising imaging technique that is useful in the differential diagnosis of thyroid cancer. Larger prospective studies are needed to confirm our results and establish the diagnostic accuracy of this technique.

References


