Ultrasonographic United Stiffness Score System: Resolving the Discrepancy between Acoustic Radiation Force Impulse and Real-time Elastography in Evaluating the Thyroid Nodules Stiffness

Jia Zhan1,2, Xuehong Diao1*, Jiamei Jin1, Lin Chen1 and Yue Chen1

1Department of Ultrasound, Fudan University, Shanghai, China
2Shanghai Key Laboratory of Clinical Geriatric Medicine

Abstract

Objective: The aim of this study was to assess the ultrasonographic united stiffness score system (UUSSS) in diagnosing thyroid nodules, and retrospectively analyze the discrepancy between acoustic radiation force impulse (ARFI) and real-time elastography (RTE).

Methods: 170 conventional ultrasound (US) proven thyroid nodules in 70 patients were included and all were examined by RTE and ARFI. RTE and ARFI were first analyzed respectively, comparing with pathological findings. Then nodules have discrepancy between ARFI and RTE were retrospectively analyzed by UUSSS.

Results: The AUC (area under curve) of ARFI combined RTE in 170 nodules was 0.87 (sensitivity=79.4% (54/68), specificity=94.3% (80/102), PPV=77.1% (54/70), NPV=86.0% (86/100), accuracy for ARFI were 82.4% (140/170), relatively to RTE was 0.83, 80.0%, 65.7%, 61.1%, 83.8%, 71.8%.The AUC of UUSSS was 0.96 (sensitivity=93.8% (57/62), specificity=87.3% (89/102), PPV=81.4% (57/70), NPV=89.0% (89/100), accuracy for UUSSS was 85.9% (146/170). The difference between area under UUSSS and RTE ROC curves was statistically significant (p<0.05), but was insignificant between UUSSS and ARFI (p=0.2245). There were 47 thyroid nodules had discrepancy in assessing the stiffness between ARFI and RTE. In these nodules, the accuracy of UUSSS was 83.0% (39/47), there were significant difference compare to RTE (20/47) (p=0.000), and ARFI (27/47) (p=0.012).

Conclusions: UUSSS is useful to increase the diagnostic accuracy compared to independent RTE and ARFI, especially for those had discrepancy between ARFI and RTE. UUSSS is easy and convenient to re-evaluate stiffness on RTE and ARFI basis, and it place no additional cost burden.

Keywords: Real-time elastography; Acoustic radiation force impulse; Thyroid nodule; Differential diagnosis; Ultrasonographic united stiffness score system (UUSSS)

Introduction

Thyroid nodules are a common clinical problem, and the incidence of thyroid cancer has increased significantly in recent decades [1]. Thyroid lesions are found about 3-7% by palpation [2] and 19-76% of randomly selected individuals by high-resolution ultrasound (US) [3-4]. Most thyroid nodules are benign but less than 5% are malignant [5]. The presence of calcification, hypoechogenicity, irregular margins, absence of a halo, a deeper than wide shape and intranodular vascularity in the sonographic image are the key features associated with an increased risk of malignancy [6]. However, no sonographic feature has both high sensitivity and a high positive predictive value (PPV) for thyroid cancer [7].

Tissue stiffness is another feature that may reflect the nature of the thyroid nodule, neoplastic diseases can alter tissue composition and structure, that may reflect the nature of the thyroid nodule, with the malignant nodule tends to be hard and the benign nodule be soft [8]. Elastography is a new sonographic technology that reflects tissue stiffness. It depicts the deformation and displacement of pressurized tissues to reflect the elastic characteristics of tissues [9]. There were two kinds of ultrasonographic methods for evaluating the stiffness of thyroid nodules: RTE as qualitative one and ARFI as quantitative one.

ARFI imaging involves two parts in the past literature: VTI (virtual touch imaging) and VTQ (virtual touch tissue quantification). In VTI mode, similar to the RTE, the sampling box encircled the nodule with adequate surrounding thyroid tissue. The VTI images and the grayscale images were displayed simultaneously in a split-screen mode. VTI elastography images were divided into six grades, as a qualitative evaluation. In VTQ mode, tissue within an ARFI “measuring box” (6 x 5 mm) is mechanically stimulated using short-duration (262 µs) acoustic pulses [10]. In present study, we regard ARFI as VTQ-a quantitative measurement.

In the comparative study with a small sample (30 patients with 58 nodules), ARFI as the quantitative assessment for thyroid nodules stiffness was more reliable than RTE [11]. However, the majority of the previous studies only retrospectively analyzed the stiffness data of thyroid ARFI, giving the cut-off Shear wave velocity (SWV) value as well as the associated diagnostic value after thyroidectomy. In clinical reality, conventional US is generally performed first, and then elastography, including RTE and ARFI, is carried out. Such a study design does not conform to clinical reality; without knowing the cut-off, the real impact of ARFI elastography before the operation is questionable; in addition, discrepancy between ARFI and RTE was also not mentioned [12]. The aim of the present study was to assess the ultrasonographic...
united stiffness score system (UUSSS) in diagnosing thyroid nodules, and retrospectively analyze the discrepancy between ARFI and RTE.

Patients and Methods

Patients

Informed consent was obtained from all patients and the study was performed in accordance with the ethical guidelines of the Helsinki Declaration and approved by ethics committee of Huadong hospital. From December 2013 to August 2015, 170 thyroid nodules in 70 patients were enrolled in the study, and all the nodules were confirmed by histopathology after thyroidectomy (Ten patients had one nodule, 30 patients had two nodules, 20 patients had three nodules, and ten patients had four nodules). The mean age of the examined patients-21 men and 49 women-was 49.7 ± 11.6 years (age range, 23-70 years).

Median size of the nodules was 2.5 × 1.8 × 2.2 cm (range, 3.2 × 2.5 × 2.9 cm to 0.6 × 0.6 × 0.5 cm).

The inclusion criteria for the patients were as follows: 1. Nodule size larger than 0.6 cm in diameter, and smaller than 3.5 cm. 2. Solid or almost solid (<25% cystic) nodules on conventional US. 3. Enough thyroid tissue surrounding the nodule at the same depth. 4. At least one nodule was suspicious malignant. 5. The patients underwent thyroidectomy.

Exclusion criteria were: 1. The size of thyroid nodules ≥ 3.5 cm or <0.6 cm in diameter. 2. The anatomy surrounding the thyroid nodules consists of movable structures like the jugular vein. 3. The shear wave velocity showed “Vs=X.XX cm/s”. The flowchart of the patient selection is shown in Figure 1.

Figure 1: The flowchart for the patient selection.
Conventional ultrasound (B-mode and Doppler)

All patients received an ultrasound examination of the thyroid gland including Color Doppler Ultrasound using a 9 MHz transducer (Acuson S2000, Siemens, Germany). The patients were positioned in a supine position with dorsal flexion of the head. The ultrasound examination was performed by experienced examiners. Thyroid nodules were evaluated for size, volume, echogenicity, echotexture, presence or absence of halo sign, presence or absence of microcalcification. Then B-mode ultrasound, color Doppler and Duplex imaging were performed.

RTE

After conventional US, RTE was performed for each nodule. A 9 MHz linear transducer (Acuson S2000, Siemens) was placed lightly on the patient's neck, with US gel placed to create a stand-off pad. No external compression was applied for elastography because the carotid arterial pulsation was used as the compression source. The ROI used for obtaining elasticity image was set to include sufficient surrounding normal thyroid tissues. RTE and grey-scale ultrasound images were displayed simultaneously in dual mode.

Video clips and single images were stored. According to the Acuson S2000 RTE color coding characteristics and the RTE scored experiences by other authors, RTE was classified into the following five different patterns (Table 1).

The entire examination was carried out in approximately 5-10 min per patient. The ES was discussed and made by two lead sonographers with more than 3 years of experience in conventional ultrasound examination and more than 200 RTE measurements performed. Tissue within a “measuring box” of experience in conventional ultrasound examination and more than 200 ARFI measurements performed. Tissue within a “measuring box” (6 × 5 mm) is mechanically stimulated using short-duration (262 µs) acoustic pulses [8]. The stiffer (nonelastic) the tissue through which the wave passes, the greater the shear velocity is. The shear wave velocity is called “Vs” for short in this study, and it's expressed in m/s with a measurement range of 0-9 m/s [13].

Five successful measurements per patients were performed with the ROI placed in the each thyroid nodule.

Ultrasoundographic united stiffness score system (UUSSS)

Details are given in Table 1.

Statistical analysis

The ES was used for statistical analysis. The sensitivity and specificity of the method for the diagnosis of malignancy was assessed by calculating PPV and NPV, results, using X² test.

For ARFI-imaging the mean of all 5 measurements per nodule or healthy thyroid gland was calculated and used for further analysis. Values of ARFI-imaging were expressed as mean ± SD, median and range. T test was used to compare ARFI-measurements in benign thyroid nodules and malignant thyroid nodules.

ROC analysis was performed and performances of RTE, ARFI and UUSSS were assessed from the areas under ROC curves. The areas under ROC curve was compared with Z test.

All statistical tests were performed by using commercially available software (Stata, version 10.0; StataCorp, College Station, TX). For all tests, a p-value <0.05 was considered to indicate a statistically significant difference.

Results

Pathological findings

On thyroidectomy, sixty-five of the 170 nodules (38%) were diagnosed as benign nodular goiters, twenty (12%) as adenomas, five (3%) as follicular neoplasms of oxyphilic cell type (Hürthle cell adenoma), twelve (7%) as Hashimoto’s thyroiditis; sixty-five (38%) as papillary thyroid carcinoma, three (2%) as medullary thyroid carcinoma.

Basic characteristics of US features of thyroid nodules

The characteristics of the US features of the thyroid nodules are presented in Table 2.

RTE

The distribution of ES values for thyroid nodules according to their pathological findings is reported: Six of 170 (4%) nodules had ES value of 1, twenty-four (14%) had ES value of 2, fifty (29%) had ES value of 3, 46 (27%) had ES value of 4 and 44 (26%) had ES value of 5.

For purposes of analysis of the sensitivity and specificity of RTE, a diagnosis of malignant was assigned to nodules with an ES of 4 or 5 and a diagnosis of benign was assigned to nodules with an ES of 1, 2 or 3. The ROC curve analysis was performed in order to confirm the choice of the cutoff point for malignancy between an ES of 3 and 4. ROC curve analyses showed that the AUC for the RTE was 0.83 (95% CI: 0.768, 0.885). Using this criterion, the sensitivity and specificity of US elastography for the diagnosis of malignancy were 80.9% (55/68) and 65.7% (67/102), respectively. The positive predictive value (PPV), negative predictive value (NPV) and accuracy were 61.1% (55/90), 83.8% (67/80) and 71.8% (122/170), respectively.

ARFI

The results obtained indicate that the shear wave velocity measured in benign and malignant thyroid nodules ranged from 1.15-3.56 m/s, and 1.68-6.25 m/s. The median Vs of ARFI-imaging in benign thyroid nodules was 2.25 m/s (mean: 2.29 ± 0.52 m/s). The median Vs of ARFI-imaging in the malignant thyroid nodules was 3.25 m/s (mean: 3.39 ± 0.92 m/s). Significant difference in mean Vs was found between benign and malignant thyroid nodules (p<0.001). Vs was significantly higher in malignant nodules than in benign nodules. The cutoff level of Vs for malignancy was estimated as 2.75 m/s. With this cutoff value, from the ROC curves the characteristics of Vs (AUC: 0.87, 95% CI: 0.815, 0.926) to predict malignancy were: sensitivity=79.4% (54/68), specificity=84.3% (86/102), PPV=77.1% (54/70), NPV=86.0% (86/100), accuracy for Vs were 82.4% (140/170).

Ultrasoundographic united stiffness score system (UUSSS)

In ultrasoundographic united stiffness score system, the corresponding values from 0 to 4 were 11 (6.4%), 37 (21.8%), 52 (30.6%), 42 (24.7%), and 28 (16.5%) respectively. The sensitivity and specificity of UUSSS for the diagnosis of malignancy were 83.8% (57/68) and 87.3% (89/102), respectively. The positive predictive value (PPV), negative predictive value (NPV), and accuracy were 61.1% (55/90), 83.8% (67/80) and 71.8% (122/170), respectively.
value (NPV) and accuracy were 81.4% (57/70), 89.0% (89/100) and 85.9% (146/170) when the 0-2 were grouped as negative and 3-4 were grouped as positive. ROC curve analyses showed that the AUC for the UUSSS was 0.89 (95% CI: 0.841, 0.939). The difference between area under UUSSS and RTE ROC curves was statistically significant (p<0.05), but was insignificant between UUSSS and ARFI (p=0.2245) (Figure 2).

![Figure 2](image_url)

Figure 2: The comparison between RTE (blue line), ARFI (red line) and UUSSS (green line) in differential diagnosis of thyroid nodules. The area under the ROC curve of UUSSS (0.890) was larger than RTE (0.827) (P<0.05) and ARFI (0.871) (P>0.05).
Retrospectively analyze the special nodules which have discrepancy stiffness evaluation between ARFI and RTE

There were 47 thyroid nodules had discrepancy in assessing the stiffness between ARFI and RTE: nineteen nodules were ES<4 but Vs ≥ 2.75 m/s (Figure 3), twenty-eight nodules were ES ≥ 4 but Vs<2.75 m/s (Figure 4). The accuracy were 42.6% (20/47) with the cutoff value ES ≥ 4, or 57.4% (27/47) with the cutoff value Vs ≥ 2.75 m/s as positive findings, independently. Correspondingly, the accuracy were 83.0% (39/47) with the cutoff value UUSSS ≥ 3, there were significant difference compare to RTE (p=0.000), and ARFI (p=0.012) (Table 3).

Figure 3: Images of a 60-y-old woman with benign nodular goiters. (A) Conventional US showed hypo-echogenicity, regular margins. (B) Color Doppler ultrasound showed intra-nodular blood flow. After reading the conventional US and CDFI images, the reader classified the lesion as TI-RADS: 4A. (C) RTE: Predominantly green but some areas of yellow are present. The margins are not clearly identified from surrounding tissue, ES=2. (D) The Vs of the nodule was 3.04 m/s. This nodule had discrepancy between ARFI and RTE: ES<4, but Vs ≥ 2.75 m/s. The stiffness were re-evaluated by UUSSS: 0 (ES=2)+2(Vs>3 m/s)=2 (indeterminate). In this case, UUSSS corrected ARFI.

Figure 4: Images of a 56-y-old woman with benign nodular goiters. (A) Conventional US showed hypo-echogenicity, regular margins and calcification of the nodule. (B) Color Doppler ultrasound showed absence of Doppler signals. After reading the conventional US and CDFI images, the reader classified the lesion as TI-RADS: 4A. (C) RTE: Predominantly yellow but a few small areas of red are present, ES=4. (D) The Vs of the nodule was 1.73 m/s. This nodule had discrepancy between ARFI and RTE: ES ≥ 4, but Vs<2.75 m/s. The stiffness were re-evaluated by UUSSS: 1(ES=4)+0 (Vs<2 m/s)=1 (probably soft). In this case, UUSSS corrected ARFI.
Discussion

Stiffness is an important tissue property for clinical differentiation between benign and malignant tumors [14]. Thus, we concerned about which technology was more accurate to reflect the stiffness of thyroid nodules and to predict the malignance-RTE or ARFI? In our impression, ARFI as the quantitative assessment for thyroid nodules stiffness was more reliable than RTE.

In fact, related data also support this conclusion: the sensitivity and specificity of RTE for the diagnosis of malignant thyroid nodules were not steady: 79% and 77% [15], respectively to 92% and 90% [16]-both results originate recent meta-analysis, while the sensitivity and specificity of ARFI were 0.86 and 0.90 [17] respectively to 0.80 and 0.85 [18]. Considering RTE was performed by different ultrasound device in different color-coded maps with different ES scales, this may be some of the reasons why there is considerable variation in the reported sensitivity and specificity of RTE. In comparison, ARFI was played with the same equipment.

ARFI is a novel technology for evaluating tissue stiffness, and gives an objective numerical evaluation of the tissue stiffness by calculating the Vs. It was indicated that ARFI was valuable in differentiating malignant from benign thyroid lesions. However, cutoff from previous study had to be consult when ARFI would be used to prospective predict the malignance [19]. Commonly Vs<2.50 m/s were regarded as benign; V>3.0 m/s were suspected to be malignant. And the stiffness overlaps between the benign and malignant thyroid nodules exits in ranged from 2.5 to 3 m/s. In recent meta-analysis, a total of 11 in 16 studies [18] reported the SWV cutoffs were ranged in 2.5-3 m/s which called “Grey Zone” of ARFI measurement. Therefore it is full of challenges to differential diagnosis the thyroid nodules in “Grey Zone” before the operation.

In UUSSS, similar to qualitative RTE, we could prospective predict thyroid nodules scored 3 (probably hard) and 4 (definitely hard) as cancerous instead of consulting cutoff from previous study. UUSSS is also believed to be useful to solve the discrepancy between ARFI and RTE. The accuracy of 47 nodules had discrepancy were significant improved (39/47) after re-evaluation the stiffness by UUSSS (20/47 by RTE, 27/47 by ARFI).

The range of VS is 0-9 m/s. Value beyond these range is displayed as “X.XX m/s” , thus t test could be used to compare ARFI-measurements in benign thyroid nodules and malignant thyroid nodules.

In fact, the thyroid nodules displayed as “X.XX m/s” could be tried to be included in UUSSS: cystic ones could be classified as 0 (the softest), and solid ones could be classified as 4 (the hardest) (Figure 5). Scores by UUSSS were rank variation-unlike Vs by ARFI were continuous variations, are not limited by normal distribution. Besides these, limited by its measuring box (6 × 5 mm), ARFI was indeterminate in small thyroid nodules which were exactly suitable for assessment by RTE [21]. Individual cases showed the assessment by UUSSS are promising in small thyroid nodules diameter in 4 mm (Figure 6). Furthermore, UUSSS is easy and convenient, just take a few minutes to re-evaluate stiffness on RTE and ARFI basis, and that's more, this procedure cost no additional burden.

The present study had some limitations: First, only 170 nodules were included in this retrospective study, which might be relatively small for sensitivity comparison, thus, larger prospective studies are needed to confirm the role of this new system in the diagnosis of thyroid nodules. Secondly, in 170 thyroid nodules the sensitivity, specificity and accuracy of UUSSS was a little higher but not significant compared to ARFI. While in 47 thyroid nodules had discrepancy, accuracy of UUSSS was significant higher than ARFI which means UUSSS was recommended in thyroid nodules had discrepancy in assessing the stiffness between ARFI and RTE, but not in all thyroid nodules. In consideration of the just passable performance of ARFI in “Grey Zone”, we strongly suggest re-evaluating those thyroid nodules whose Vs ranged from 2.5 to 3 m/s by UUSSS. Thirdly, to facilitate statistical calculation, cases displayed as “X.XX m/s” were excluded, which might lead to a selection bias. As a result of this, the population might not reflect a normal distribution of disease.

In conclusion, UUSSS is useful to increase the diagnostic accuracy compare to independent RTE and ARFI and to rule out the possibility of thyroid malignancy, especially for those had discrepancy between ARFI and RTE. Lesions in “Grey Zone” whose stiffness overlaps between the benign and malignant are strongly recommend being re-evaluated by UUSSS. The nodules displayed as “X.XX m/s” could be properly disposed in this system and thyroid nodules slightly smaller ARFI measuring box may be included in UUSSS. Further and larger prospective studies are needed to confirm this result.

<table>
<thead>
<tr>
<th>Stiffness feature</th>
<th>Pathologic findings</th>
<th>ES&lt;4 but Vs ≥ 2.75 m/s</th>
<th>ES ≥ 4 but Vs&lt;2.75 m/s</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign ones</td>
<td>UUSSS 0-2</td>
<td>9</td>
<td>18</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>UUSSS 3-4</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Malignant ones</td>
<td>UUSSS 0-2</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>UUSSS 3-4</td>
<td>7</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>19</td>
<td>28</td>
<td>47</td>
</tr>
</tbody>
</table>

UUSSS: Re-evaluating nodules stiffness twice by two independent elastographic methods with a new stiffness scoring system “0” (ES=1.2 or Vs≥2m/s), “1” (ES=3.4 or 2m/s<Vs<3m/s), “2” (ES=5 or 3m/s<Vs). Add up two scores to arrive at a five-point scale of confidence level, that is definitely soft (0+0=0), probably soft (0+1=1 or 1+0=1), indeterminate (0+2=2, 1+1=2 or 2+0=2), probably hard (1+2=3 or 2+1=3), and definitely hard (2+2=4).

Table 3: 47 nodules have discrepancy between ARFI and RTE.
Acknowledgments

This work was supported by the Medical Guide Project of the Shanghai Science and Technology Commission (14411970400) and Shanghai Municipal Commission of Health and Family Planning, Key developing disciplines (2015ZB0501).

References


