Ultrasound elastography in thyroid disease

Sorin M. Dudea, Carolina Botar-Jid

Abstract
Thyroid elastography has become, lately, one of the main focuses in thyroid imaging. With more than one hundred papers published on this subject, the core of accumulated knowledge justifies the need for a comprehensive review on the topic. The paper presents the various elastographic techniques used for thyroid assessment. Both strain and shear wave elastography, with all their variants, are discussed. Thereafter the paper proceeds to a detailed description of the technical peculiarities, diagnostic value, limitations and pitfalls for each technique. Although the main accent is placed on thyroid nodules, the applications of elastography in other thyroid disease, diffuse or focal, are also presented. The results of the available metaanalyses are reviewed and the proven value of the technique is highlighted. Still unanswered questions and directions for future research are emphasized, whenever appropriate.

Keywords: elastography, strain elastography, shear wave elastography, ARFI, thyroid

Introduction
Thyroid disease is of common occurrence, especially in iodine deficient areas. Thyroid nodules have been reported in as many as 50% of the population, at autopsy [1]. They may be discovered in up to 41% of the patients at ultrasonography (US) [2]. However, the cancer rate is low. The incidence of malignancy in patients who undergo fine needle aspiration biopsy (FNAB) is in the range of 9-14%, independent of the number of nodules present. In multinodular thyroids the cancer risk per nodule is decreasing, approximately proportional to the increase of the number of nodules, therefore the cancer rate per patient remains the same as in single nodular thyroids [2].

Ultrasonography is widely used to diagnose and characterize thyroid nodules and to assess diffuse thyroid disease. The main US features indicative of nodular malignancy are solid consistence, hypoechogenicity, “taller-than-wide” shape, irregular margins or no halo, microcalcifications and intranodular vascularization at Doppler [2,3]. However, none of these signs displays sensibility and/or specificity at a level that would allow either positive diagnosis or exclusion of carcinoma with a high degree of confidence [2,4-6].

Fine-needle aspiration biopsy (FNAB) is the next accepted diagnostic step. Fine needle aspirates are reported as nondiagnostic, nonneoplastic, atypia/follicular lesion, suspicious of malignancy and malignant [7]. Up to 20% of aspirates are nondiagnostic and require repeat puncture [2]. Cancer rate varies from 20% in the atypia group, to 60% in the suspicious group [7]. Therefore, FNAB is far from being a golden standard, especially due to limitations in its ability to exclude malignancy.

Ultrasound elastography (USE) has become, in the last decade, the most important addition to the armamentarium of sonographic techniques. With its various approaches to analyzing stiffness and with more than 100 papers published, many of them reporting conflicting results, USE has been the main focus of thyroid US imaging research.

The aim of this paper is to present a comprehensive review of the current status of the knowledge concerning the application of USE in thyroid gland disease.

Elastographic techniques
As various manufacturers have different approaches to the assessment of tissue stiffness and new approaches
are still emerging, there is some confusion and, at times, even misunderstanding on what each technique means. The recently published guidelines of EFSUMB [8] provide an orderly approach to the understanding of various elastographic techniques. Some of the reviews on thyroid USE provided own classifications [9,10] while others complied with the guidelines [11].

Basically, two main approaches for elastography of the thyroid are used in clinical practice: strain and shear-wave elastography [8-11] (fig 1). Most of the times, offline analysis of the information is performed in experimental, work-in-progress models or on prototypes while machines for clinical applications use online signal processing.

1. Strain elastography (SE) assesses the elastic properties of tissues by analyzing tissue strain, that is tissue deformation parallel to the direction of the exploratory force. Deformation may be induced by a pure mechanic force or by ultrasound.

1.1. SE using mechanic force (also named quasistatic elastography) was the first commercially available elastographic imaging technique. A kinetic impulse is applied to the tissue. This impulse induces tissue deformation. By repeating the impulse, a to-and-fro displacement movement of the image speckles is generated. Elasticity information is derived from RF echo lines. Stiff, rigid tissues exhibit less displacement compared to elastic, soft ones. The mechanic force may be applied externally (actively induced displacement of tissue surface) or it may be internal (physiologic pulsations).

1.1.1. SE with external force. The transducer is slightly pressed, repeatedly, on the skin surface. In a peculiar technical approach (Virtual Touch Tissue Imaging – Siemens), it is the transducer itself that sends exploratory impulses. The resulting information is qualitative but it may be analyzed semiquantitatively, as well.

a. Qualitative SE with external force:
   i. the result may be displayed as a single image, representing relative tissue strain over a period defined by the examiner (Q-elastography, Toshiba) or
   ii. the result may be displayed as elastographic information superimposed on each gray-scale image, at a refresh rate equaling that of grey-scale imaging – real-time elastography (RTE) – originally developed by Hitachi and now available with many manufacturers. RTE analyzes the signal within congruent, closely spaced strain estimator windows [8]. The combined autocorrelation method is used to compensate for out-of-plane movement of stiff lesions [12].

b. Semiquantitative SE with external force – either on the reference image (Q-elastography) or on a representative image selected by the operator from the video sequence, a strain ratio or strain rate (SR) is computed. This is, usually, the ratio between the strain of normal parenchyma and the strain of the nodule or area under analysis.

1.1.2. SE with internal (physiologic) force. In this approach, no external pressure is required. Pulsations from neighboring structures – carotid artery, in the case of thyroid USE – induce the displacement necessary to assess tissue elasticity. The information may be:
1. **Elastographic imaging of the thyroid nodule.**

**I. Real-time strain elastography (RTE) of thyroid nodules**

**Technical considerations**

*Where should the ROI be placed?* It is of importance to choose the shortest path for the exploratory US beam to the nodule. This will allow the placement of the elastographic ROI as close to the transducer as possible, thus avoiding strain decay with distance.

*What is the size of the ROI and what should it encompass?* The elastographic ROI should cover the whole nodule. Some authors recommend it should also include “sufficient surrounding parenchyma” without further elaborating on this [16]. Others recommend that the lateral margins of the ROI should be set at 5 mm or more of the apparent limits of the nodule [17,18]. In strain elastography it is important to encompass in the image as much “normal” reference parenchyma as possible, to maximize the nodule to parenchyma contrast. It is also important to exclude from the ROI, as much as possible, the vessels (mainly carotid), the esotracheal structures, bones and the muscles [17].

*What are the basic technical requirements?* Breath hold and no swallowing are of paramount importance [17,19]. A generous amount of gel and slight skin contact are needed. Probe stabilizer is not suited for this application due to the anatomic peculiarities of the neck. The image focus is placed at or below the level of the nodule [17]. The displayed elasticity scale should be adapted according to the stiffness of the studied tissue [11]. In some technical approaches, the transducer is held still (no need for manual vibration) [17,20]. In most of the instances, the examiner applies slight and regular manual axial vibration to the transducer.

*How should compression quality indicators be used?* For the sake of reproducibility, most manufacturers provide a strain quality indicator, either numeric or graphic. On Hitachi machines, the compression quality scale should indicate levels 2-3 [19,21,22] or 2-4 [23,24]. However, we prefer to work only with quality factors 3-4, in agreement with other authors [16,25]. Only images obtained at these quality levels are fit for assessment.

On Siemens machines, a quality factor above 50 for 3-4, in agreement with other authors [16,25]. Only images obtained at these quality levels are fit for assessment. On Siemens machines, a quality factor above 50 for 3-4 successive frames has been recommended [17].

*How is strain information displayed on the image?* For most manufacturers, by default, stiff areas (little or no strain) are displayed in blue, elastic areas (high strain) are colored in red and intermediate strain areas appear green. Other manufacturers chose a scale ranging from red = hard to purple = soft. However, nowadays the color is no longer an issue as most manufacturers offer a gamut of color scales for the same information. However, care must be taken about the settings of the machine: optimization of the elastogram for an elastic material will produce improper display of the elasticity variation encountered in stiff tissues (cancer, fibrosis) [11].

*What is really displayed on the RTE image?* It is always important to keep in mind that RTE displays the relative strain of the structures in the ROI. The absolute
value of the strain depends on the initial compression applied by the transducer (variable and nonquantifiable) and on the exploring repeated compression used to produce the images – again, variable. Strain values do not represent the elasticity modulus [13]. As strain changes with the applied compression, its absolute (numeric) value (although measured and displayed by some machines) is completely inappropriate to compare two lesions or two individuals [13,14].

What are the technical peculiarities in the neck?

Quite often there is interference between the pulsations induced by the hand or by the US beam and the ones coming from the carotid artery, which leads to image degradation [13]. On the other hand, lateral systolic expansion of the carotid compresses the thyroid against the trachea and induces anteroposterior expansion of the gland, which may be detected as strain [14]. At least one commercially available application, developed by Samsung Medison, uses carotid artery pulsations as the sole strain inductor for elastography.

What is the role of the experience of the examiner?

Ko and coworkers [26] found that, in thyroid elastography, experienced physicians have greater specificity as compared to less experienced operators but no other differences in diagnostic performance were noted.

Qualitative scoring of the RTE images.

The normal thyroid is homogenously green (of intermediate stiffness) or it presents a mixture of green, red and blue. At times, inhomogeneous thyroid parenchyma may represent the balance between hyperplasia and involution [11]. The RTE appearance of thyroid nodules is assessed subjectively by attributing a score based on the amount of apparent stiffness on the image (elasticity score – ES). The first scoring system for thyroid nodules was published in 2004 [27]. Currently, the proposed systems of scoring the RTE image of thyroid nodules vary from 6 to 2 patterns, each of them with variants and different reported diagnostic value (tables I-V).

It is noteworthy that the initial studies, until 2010-2011, reported very encouraging results and thereafter, as the number of studies, patients and nodules increased, a more tailored and cautious approach to qualitative thyroid elastography emerged. Most studies rely either on the five or four pattern scoring systems. None of the scoring systems appears to be superior. The two main scoring systems are illustrated in figures 2 and 3 and representative images are shown in figures 4 – 6.

Over time, lower sensibilities and positive predictive values were reported, as compared to original studies. However, high to very high negative predictive values appear to be common to all studies, no matter of the design or time of publication. To explain the somehow conflicting results, one needs to analyze observations associated to RTE.

What is the intra- and interobserver agreement?

The first study on this topic, published by Park and coworkers...
### Table II. The 5 pattern scoring systems for RTE of thyroid nodules.

<table>
<thead>
<tr>
<th>Author/Reference</th>
<th>Year</th>
<th>Score 1</th>
<th>Score 2</th>
<th>Score 3</th>
<th>Score 4</th>
<th>Score 5</th>
<th>Cutoff for malignancy</th>
<th>SE (%)</th>
<th>SP (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Acc (%)</th>
<th>AUROC</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rago [28]</td>
<td>2007</td>
<td>Entirely elastic</td>
<td>Mostly elastic</td>
<td>Elasticity only at the periphery</td>
<td>No elasticity in the nodule</td>
<td>No elasticity in the nodule and behind the nodule</td>
<td>≥ 4</td>
<td>97</td>
<td>100</td>
<td>100</td>
<td>98</td>
<td>-</td>
<td>Adapted Tsukuba (Ueno, Itoh) score</td>
<td></td>
</tr>
<tr>
<td>Ghervan [29]</td>
<td>2009</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>---</td>
<td>≥ 4</td>
<td>100</td>
<td>79</td>
<td>50</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Wang [30]</td>
<td>2010</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>---</td>
<td>≥ 4</td>
<td>90.63</td>
<td>89.47</td>
<td>93.5</td>
<td>85</td>
<td>90.2</td>
<td>-</td>
<td>For solitary nodule &lt; 10mm</td>
</tr>
<tr>
<td>Moon [31]</td>
<td>2012</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>---</td>
<td>≥ 4</td>
<td>15.7</td>
<td>95.3</td>
<td>28</td>
<td>90.6</td>
<td>70.7</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Stoian [32]</td>
<td>2012</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>---</td>
<td>≥ 4</td>
<td>23.2</td>
<td>97.5</td>
<td>52</td>
<td>91.5</td>
<td>80.8</td>
<td>-</td>
<td>&gt; 10 mm</td>
</tr>
<tr>
<td>Rivo-Vazquez [33]</td>
<td>2013</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>---</td>
<td>≥ 4</td>
<td>90.9</td>
<td>98.96</td>
<td>100</td>
<td>98.6</td>
<td>-</td>
<td>-</td>
<td>≤ 10 mm</td>
</tr>
<tr>
<td>Shweel [18]</td>
<td>2013</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>---</td>
<td>≥ 4</td>
<td>75</td>
<td>85.5</td>
<td>71.4</td>
<td>90.5</td>
<td>86.7</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Wang [34]</td>
<td>2013</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>---</td>
<td>≥ 4</td>
<td>84.62</td>
<td>78.45</td>
<td>-</td>
<td>-</td>
<td>80.36</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Liu [16]</td>
<td>2014</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>---</td>
<td>≥ 4</td>
<td>79</td>
<td>84.4</td>
<td>61.2</td>
<td>90.5</td>
<td>82.8</td>
<td>0.880</td>
<td>Similar to shear-wave results for the same nodules</td>
</tr>
<tr>
<td>Şahin [35]</td>
<td>2014</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>---</td>
<td>≥ 4</td>
<td>82.6</td>
<td>50.9</td>
<td>87.34</td>
<td>97.14</td>
<td>-</td>
<td>-</td>
<td>For nodules &gt; 1 cm with concomitant Hashimoto thyroiditis</td>
</tr>
<tr>
<td>Cepero Calvetere [17]</td>
<td>2013</td>
<td>&gt; 45% elastic</td>
<td>Elasticity &lt;45 % or only at periphery</td>
<td>--</td>
<td>No elasticity in the nodule or surrounding tissue</td>
<td>≥ 4</td>
<td>89</td>
<td>-</td>
<td>85</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yerli [36]</td>
<td>2013</td>
<td>Mostly elastic (few blue spots)</td>
<td>Mosaic green – blue, predominant green</td>
<td>Mosaic green – blue, predominant blue</td>
<td>Entirely blue, eventually blue perinodular rim</td>
<td>≥ 4</td>
<td>83</td>
<td>89</td>
<td>56</td>
<td>97</td>
<td>88</td>
<td>-</td>
<td>Image guided, pressure applied only in diastole</td>
<td></td>
</tr>
</tbody>
</table>

SE – sensitivity, SP – specificity, PPV – positive predictive value, NPV – negative predictive value, Acc – accuracy
Table III. The 4 pattern scoring systems for RTE of thyroid nodules

<table>
<thead>
<tr>
<th>Author/Reference</th>
<th>Year</th>
<th>Score 1</th>
<th>Score 2</th>
<th>Score 3</th>
<th>Score 4</th>
<th>Cutoff for malignancy</th>
<th>SE (%)</th>
<th>SP (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Acc (%)</th>
<th>AUROC</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanaka [27]</td>
<td>2004</td>
<td>Entirely green</td>
<td>Green center, blue periphery</td>
<td>Color mixture, predominant blue</td>
<td>Entirely blue</td>
<td>Suggested scores 3&amp;4 represent papillary carcinoma</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ding [37]</td>
<td>2012</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>78</td>
<td>80</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Wang [38]</td>
<td>2012</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>89.4</td>
<td>84.5</td>
<td>89.4</td>
<td>84.9</td>
<td>85.9</td>
<td>-</td>
<td>Including zero score for cystic lesions with red-green-blue pattern</td>
</tr>
<tr>
<td>Shuzen [4]</td>
<td>2012</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>83.25</td>
<td>90.94</td>
<td>-</td>
<td>-</td>
<td>87.89</td>
<td>0.908</td>
<td>-</td>
</tr>
<tr>
<td>Ding [37]</td>
<td>2012</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>83.25</td>
<td>90.94</td>
<td>-</td>
<td>-</td>
<td>87.89</td>
<td>0.908</td>
<td>-</td>
</tr>
<tr>
<td>Wang [38]</td>
<td>2012</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>83.25</td>
<td>90.94</td>
<td>-</td>
<td>-</td>
<td>87.89</td>
<td>0.908</td>
<td>-</td>
</tr>
<tr>
<td>Guazzaroni [20]</td>
<td>2012</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>83.25</td>
<td>90.94</td>
<td>-</td>
<td>-</td>
<td>87.89</td>
<td>0.908</td>
<td>-</td>
</tr>
<tr>
<td>Asińa [40]</td>
<td>2008</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>83.25</td>
<td>90.94</td>
<td>-</td>
<td>-</td>
<td>87.89</td>
<td>0.908</td>
<td>-</td>
</tr>
<tr>
<td>Lippolis [42]</td>
<td>2011</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>92.3</td>
<td>73</td>
<td>-</td>
<td>-</td>
<td>87.5</td>
<td>0.86</td>
<td>-</td>
</tr>
<tr>
<td>Friedrich-Rust [43]</td>
<td>2010</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>92.3</td>
<td>73</td>
<td>-</td>
<td>-</td>
<td>87.5</td>
<td>0.86</td>
<td>-</td>
</tr>
<tr>
<td>Kagoya [44]</td>
<td>2010</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>92.3</td>
<td>73</td>
<td>-</td>
<td>-</td>
<td>87.5</td>
<td>0.86</td>
<td>-</td>
</tr>
<tr>
<td>An [47]</td>
<td>2013</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>92.3</td>
<td>73</td>
<td>-</td>
<td>-</td>
<td>87.5</td>
<td>0.86</td>
<td>-</td>
</tr>
<tr>
<td>Friedrich-Rust [6]</td>
<td>2013</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>92.3</td>
<td>73</td>
<td>-</td>
<td>-</td>
<td>87.5</td>
<td>0.86</td>
<td>-</td>
</tr>
<tr>
<td>Azizi [48]</td>
<td>2013</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>92.3</td>
<td>73</td>
<td>-</td>
<td>-</td>
<td>87.5</td>
<td>0.86</td>
<td>-</td>
</tr>
<tr>
<td>Cakal [49]</td>
<td>2014</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>92.3</td>
<td>73</td>
<td>-</td>
<td>-</td>
<td>87.5</td>
<td>0.86</td>
<td>-</td>
</tr>
<tr>
<td>El-Hanthri [50]</td>
<td>2014</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>92.3</td>
<td>73</td>
<td>-</td>
<td>-</td>
<td>87.5</td>
<td>0.86</td>
<td>-</td>
</tr>
<tr>
<td>Tranquart [51]</td>
<td>2008</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>≥ 3</td>
<td>92.3</td>
<td>73</td>
<td>-</td>
<td>-</td>
<td>87.5</td>
<td>0.86</td>
<td>-</td>
</tr>
</tbody>
</table>
Table IV. The 3 pattern scoring systems for RTE of thyroid nodules

<table>
<thead>
<tr>
<th>Author/Reference</th>
<th>Year</th>
<th>Score 1</th>
<th>Score 2</th>
<th>Score 3</th>
<th>Cutoff for malignancy</th>
<th>SE (%)</th>
<th>SP (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Acc (%)</th>
<th>AUROC</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merino [55] 2011</td>
<td>&gt; 90% of the nodule white</td>
<td>Inclusion in score 1 or 3 impossible</td>
<td>&gt; 90% of the nodule is black, stiffness greater than the grey-scale image of the nodule</td>
<td>&gt;2</td>
<td>100</td>
<td>40.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>elastography on grey-scale</td>
<td></td>
</tr>
<tr>
<td>Rago [56] 2010</td>
<td>Score Ueno 1 + 2 (soft nodule)</td>
<td>Score Ueno 3 (intermediate)</td>
<td>Score Ueno 4 + 5 (stiff nodule)</td>
<td>≥ 2</td>
<td>96.8</td>
<td>91.8</td>
<td>76.9</td>
<td>99</td>
<td>92.9</td>
<td>-</td>
<td>Nodules with indeterminate cytology</td>
<td></td>
</tr>
<tr>
<td>Capelli [25] 2012</td>
<td>Intermediate (mixture green and blue type 1)</td>
<td>Hard, predominantly blue (type 2)</td>
<td>≥ type 1</td>
<td>90</td>
<td>79.6</td>
<td>24.3</td>
<td>99.1</td>
<td>80.3</td>
<td>-</td>
<td>-</td>
<td>Indeterminate + non-diagnostic cytology</td>
<td></td>
</tr>
</tbody>
</table>

SE – sensitivity, SP – specificity, PPV – positive predictive value, NPV – negative predictive value, Acc – accuracy
reported no or poor interobserver agreement when using the Ueno score. Moderate interobserver agreement was also noted using the reverse color Asteria score [60]. However, later studies using the four score scale reported little interobserver variability [46], with almost perfect interobserver agreement [17], better for RTE than for the TIRADS score [6]. Excellent interobserver agreement was also reported using a three score gray scale system [55]. Although most studies state that using the quality indicator while performing RTE aims to reduce variability, little is known about the actual figures for intraobserver variability of scoring, although at least one study reported acceptable intraobserver reproducibility [61]. One of the aims of simplifying the scoring systems is to reduce variability, but this has not been fully proven, yet. The influence of a learning curve has been proven by the existence of inconsistent results over time [62].

What is the influence of the nodule size on the diagnosis? Early studies reported no correlation between the ES and the diameter of the nodule [21,28,54]. Most of later studies confirmed that the elastic appearance does not depend on the size, surface or volume of the nodule [17,25,32], although at least two studies reported slightly worse performance for nodules less than 1 cm in diameter compared to those larger than 1 cm [19,31]. The two score scale (table V) produced better performance for nodules less than 2 cm in diameter [57]. Elastography is not suited for nodules with a diameter less than 5 mm [21,24] although the exact lower limit of the diameter for RTE usability is not known [9]. It is also unsuitable for nodules larger than 3 cm as the nodule cannot be encompassed by the image [9,38]. In degenerated benign nodules with the diameter less than 1 cm, Asteria scores 1 and 2 (table V) allowed differentiation from papillary
carcinoma of the same size with sensibility of 92.3% and specificity of 100%. However, when the diameter was greater than 1 cm, sensibility dropped to 16.7% [63].

What is the influence of the pathology of the nodule on the elastographic appearance? Even the early reports noted that follicular carcinoma may appear elastic on RTE [27]. Today it is well documented that RTE is not useful in diagnosing this type of carcinoma [4,21,42,43,55] as it may produce up to 44% false negative results [9]. Medullary, nondifferentiated and metastatic carcinoma may also appear elastic [4,21]. In a population with multinodular goiter and low risk of cancer, RTE is not useful in detecting cancer, sensibility and specificity being the same whether the nodule is elastic or hard [64]. When partial cystic nodules are considered, again, the diagnostic value of RTE is uncertain, as only the solid component should be assessed [16,24,28,56] and no data concerning the usefulness are available [9]. The presence of nonmalignant disease may also influence the diagnostic value, as shown later.

Some pathologic consequences were inferred from elastographic observations. A stiff nodule (score 4 and 5) was found to be an independent predictor for extrathyroid extension of thyroid microcarcinoma (<10 mm) [22,63], although increased stiffness was not associated with an increased likelihood for positive central or lateral neck lymphadenopathy [22].

How does RTE compare with other US techniques? The reports on this topic are conflicting. Some studies found that the diagnostic performance of elastography was inferior to that of grey-scale US alone, irrespective of the level of experience of the practicing physician [26]. Unlütürk et al. [5] also reported limited sensibility and positive predictive value for USE, with values that were not superior to conventional grey-scale US. Another study, however, reported positive predictive value of 63.1% for RTE, higher than that of any conventional US criterion and a negative predictive value of 97.2%, suggesting that USE may be the best predictor for malignancy [48].

The combination of elastographic and grey-scale information may improve the diagnostic performance [39]. Combining the elastographic, grey-scale and Doppler US information leads to an increase of sensibility from 85% to 97% and of the negative predictive value from 91% to 97%, with concomitant drop of specificity, positive predictive value and accuracy [52]. Another group [18] also reported that the association of grey-scale and USE information yields sensibility 95.45%, specificity 94.8%, positive predictive value 82.3% and negative predictive value 98.8%, greater that either grey-scale or RTE alone. However, another study [31] stated that “elastography alone, as well as the combination of elastography and gray-scale US, showed inferior performance in the differentiation of malignant and benign thyroid nodules compared with grey-scale US features; elastography was not a useful tool in recommending fine-needle aspiration biopsy”.

One study [16] compared RTE and SWE for the same nodules. The authors found no differences in the diagnostic performance between the two techniques, whether medium sized (1-3 cm) or small (<1 cm) nodules were studied, although SWE produced more false negative results in the latter group. The study also showed a significant correlation between the elastographic score and the SWE values [16].

Limitations and causes of false results for RTE.

Scanning plane and interference with carotid pulsations. Transverse scans through the thyroid are more susceptible to interferences from carotid pulsations [9,59] and, therefore, less suitable for RTE with external force. They are, however, preferred for SE with internal force. Longitudinal scans are less susceptible to carotid pulsations and they, also, offer more thyroid reference tissue [45,54].

Compression grading. It is impossible to measure and quantify exactly the initial compression as well as the compression that induces repeated tissue strain [59]. The degree of initial compression and the variations of the applied strain force and frequency influence the perceived score and induce interobserver variability [9,11,15]. Most manufacturers provide quality control indicators that address this issue, although strain is not accurately quantified.

Slice thickness may be a source of misleading results. If the section thickness is around 5 mm, a thick sample volume induces averaging of data from small lesions with that of neighboring tissues [21]. This may not be an issue with highly focused transducers.

Out-of-plane motion of the nodule during compression. Although addressed by he combined autocorrelation method, this technical factor may impede on the RTE appearance of a nodule [11,15].

ROI size and content. As RTE displays relative elasticity within a ROI, it is highly advisable to have a ROI as large as possible and include in it as much “normal” reference thyroid tissue as available. If the whole ROI content is abnormal, the results are extremely unreliable [61]. On the other hand, one should avoid the inclusion of vessels, bones and, if possible, nonthyroid tissue.

Nodule position. Anterior nodules, protruding to the capsule, may be mislabeled as soft because the reference tissue is represented by strap muscles, not by the thyroid parenchyma [21]. Isthmic nodules are also difficult to
assess, being compressed between to hard planes (transducer and thyroid) and lacking reference tissue [30].

Deep located nodules are subject to the stress decay phenomenon. Stress transmission is reduced as the distance from the transducer increases. Less tissue dislocation in deep portions of the thyroid will induce artefactual hardening [54]. The nodules located in front of the common carotid artery are the most susceptible to pulsation interferences [51]. Nodules within residual parenchyma lack reference tissue for comparison [61].

**Nodule size.** As discussed above, larger than 3cm or lobar size nodules cannot be assessed with RTE [51]. Coalescent nodules are, also, unsuitable for elastography [24,38,61].

**Intranodular gross changes.**

Calcifications inside a nodule are associated with increased stiffness, irrespective of underlying pathology. Therefore, the presence of intranodular calcifications, whether micronodular or rim type, is a factor of unreliability [9,28,30,56,66].

Fibrosis inside longstanding benign nodules or associated with subacute or Hashimoto thyroiditis may also induce stiff nodules [4,11]. To the date, there are no studies on grading the correlation between nodule stiffness and degenerative pathologic changes [11].

Necrosis, even without liquefaction, may induce soft areas.

Intranodular colloid cystic changes: the presence of fluid inside a nodule dramatically alters the RTE appearance and such nodules are not suitable for elastography [11,24,54]. The assertion that, when partial cystic nodules are considered only the solid component should be assessed, is highly debatable.

**Nodule pathology,** as shown above, largely influences the RTE appearance. Of all thyroid cancers, mostly the papillary type is expected to appear hard, providing only 7% false negative results [9]. Follicular cancer may appear elastic and so do other types of malignancy. On the other hand, some benign nodules may be hard.

**Other technical causes of unreliability** of RTE are: less than 50% green thyroid reference parenchyma in the ROI; score discordance between three different images of the same lesion; color loss inside the nodule [66]; manual, arbitrary selection of an image considered to be representative, out of a cine sequence; subjective allocation of a score; different scoring systems, with lack of standardization [11,15]; compression force > 3 and arterial pulsations [66]; the color scale; the examiner skill and experience [11,15].

The limitations and discrepancies may also be related to the imperfection of the golden standard (FNAB) used in most studies [33].

The main causes for false positive and false negative RTE diagnosis of thyroid cancer are presented in table VI.

A recently published paper [17] reported that the results of RTE are not influenced by the learning curve, mean distance of the nodule from transducer, nodule size, surface, stiffness area ratio and quality factor, if above 50. However, the study was conducted on a Siemens machine that has a peculiar technical approach to elasticity imaging.

2. Virtual Touch Tissue Imaging (VTI).

**Technique.** The transducer is in slight contact with the skin. The nodule is scanned along its maximal length. The ROI should encompass the whole nodule, in the center of the image, as well as sufficient neighboring reference thyroid tissue. Breath hold should be maintained during the acquisition [67].

**Image assessment.** Soft tissues appear as white or pale grey areas. The stiffer a tissue, the darker the image. The resulting images may be classified in four types: 1= white or honeycomb; 2= pale grey, similar to neighboring tissues; 3= dark grey, with or without white spots; 4= entirely black [67]. A six type score, similar to the one described by Hong (see table 1) was also used [68]. The diagnostic value of VTI is summarized in table VII. Very good inter –

<table>
<thead>
<tr>
<th>Table VI. Causes for false RTE diagnosis of thyroid cancer.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>False positive (stiff nodule, no cancer)</strong></td>
</tr>
<tr>
<td>Calcification</td>
</tr>
<tr>
<td>Fibrosis</td>
</tr>
<tr>
<td>Subacute / chronic /Hashimoto thyroiditis</td>
</tr>
<tr>
<td>Large size</td>
</tr>
<tr>
<td>Deep location</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
and intraobserver agreement was reported for this method [68]. However, there are few published studies employing the method and further confirmation is required.


**Area ratio (AR)** is one of the approaches for semiquantitative assessment of strain images. The area of the nodule is measured first on the VTI image (A1) and the on the simultaneous grey-scale image (A2). The ratio between the two is computed. Technically, to insure an accurate measurement, the recommendation is to obtain five consecutive images, retain and use the three images with the best delineation, compute AR three times and average the results for a final value [67]. For a cutoff of AR>1.08 as indicator of malignancy, sensitivity was 91.3% with specificity of 86.6%, positive predictive value 82.3%, negative predictive value 93.4 and accuracy 88.5% [67]. However, it is noteworthy that AR >1 implies that stiffness area is greater than nodule area, which corresponds to type 5 on Ueno scale. On the other hand, this technique was employed on a machine with a peculiar technical approach to elastography.

**Hard area ratio** is computed as the ratio between the hard area in the lesion and the area of the whole lesion [59, 69]. A cutoff value >0.6 for malignancy yielded sensitivity of 92.9%, specificity of 91.3% and accuracy 92% while cutoff >0.45 had sensitivity 98.2% and specificity 68.1% [69]. This indicator was shown to lack interobserver reproducibility [59], the main drawback being the fact that the rigid area is drawn manually by the examiner on a subjectively chosen image [69].

**Strain ratio (SR)** was first reported in one of the earliest publications on thyroid elastography [70].

The technique implies placement of two similar sized ROIs, at similar depth from the transducer (depth difference should be less than 10mm), one in the nodule (A) and the second in neighboring normal parenchyma (B). SR is then computed as B/A [19,38,45,70]. As strain ratio takes into consideration the whole range of strain values, it may be more accurate than elasticity imaging, which is optimized either for soft or for hard tissues [11]. Three different measurements should be undertaken and their average should be considered as the final value [38,45,50]. The computing of SR is illustrated in figure 7. One particular approach to SR is the Q-elasto technique developed by Toshiba, where strain is estimated retrospectively from the stored RF waves, without real time image for control.

The diagnostic value of SR is presented in table VIII.

Even in the early studies, average SR for benign tumors was reported at values 2.11±0.9, while malignant tumors had SR values of 9.82±6.33 [70]. Another study confirmed these values, finding in benign nodules SR of 2.59±2.12 and in malignant ones 9.10±7.02 [45]. SR was unable to differentiate the different types of malignancy, as the reported values were 5.02±2.07 for papillary carcinoma, 4.95±2.12 for follicular carcinoma and 6.54±0.55 for undifferentiated carcinoma [38]. While one study reported for SR better performance that the elasticity scores, for nodules greater than 1cm [19], at least one other study (conducted on a different machine) found that adding SR to the color scale classification does not improve the diagnosis, on contrary, the association of the two leading to lower diagnostic efficacy [58]. When using the Q-elasto technique, SR provided better results than the grey scale score and was more sensitive than contrast enhanced ultrasound [73], with substantial to almost perfect interobserver agreement [3,74].

Limitations and causes of false results for SR.

These are mostly superposable to the ones described for the elasticity scores.

Technical considerations need to be accounted for. Transverse scans through the thyroid lobe are more susceptible to interference with carotid pulsations, therefore the longitudinal scans are preferred [19,45]. The ROI should only include the nodule and thyroid parenchyma, avoiding the carotid or other neighboring tissues. The

<table>
<thead>
<tr>
<th>Author/Reference</th>
<th>Year</th>
<th>Cutoff for malignancy</th>
<th>SE (%)</th>
<th>SP (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Acc (%)</th>
<th>AUROC</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang [68]</td>
<td>2014</td>
<td>≥ 4</td>
<td>87</td>
<td>95.8</td>
<td>91.8</td>
<td>93.1</td>
<td>92.7</td>
<td>-</td>
<td>6 type scoring system</td>
</tr>
<tr>
<td>Zhang [67]</td>
<td>2014</td>
<td>≥ 3</td>
<td>95.7</td>
<td>89.6</td>
<td>86.27</td>
<td>96.77</td>
<td>92.04</td>
<td>-</td>
<td>4 type scoring system</td>
</tr>
</tbody>
</table>

SE – sensitivity, SP – specificity, PPV – positive predictive value, NPV – negative predictive value, Acc – accuracy
Table VIII. Diagnostic value of SR

<table>
<thead>
<tr>
<th>Author/ Reference</th>
<th>Year</th>
<th>Cutoff for malignancy</th>
<th>SE (%)</th>
<th>SP (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Acc (%)</th>
<th>AUROC</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lyshchik [70]</td>
<td>2005</td>
<td>&gt; 4</td>
<td>82</td>
<td>96</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Experimental device</td>
</tr>
<tr>
<td>Ding [69]</td>
<td>2011</td>
<td>&gt;2.73</td>
<td>89.3</td>
<td>79.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Xing [19]</td>
<td>2011</td>
<td>&gt;3.79</td>
<td>97.8</td>
<td>85.7</td>
<td>88</td>
<td>97.8</td>
<td>-</td>
<td>-</td>
<td>overall</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;4.21</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.89</td>
<td>-</td>
<td>Nodules &lt; 1cm</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;3.98</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.94</td>
<td>-</td>
<td>Nodules &gt; 1cm</td>
</tr>
<tr>
<td>Cantisani [71]</td>
<td>2012</td>
<td>&gt;2</td>
<td>97.3</td>
<td>91.7</td>
<td>87.8</td>
<td>98.2</td>
<td>93.8</td>
<td>0.98</td>
<td>Q-elasto</td>
</tr>
<tr>
<td>Cantisani [72]</td>
<td>2012</td>
<td>&gt;2.05</td>
<td>87.5</td>
<td>92</td>
<td>81.4</td>
<td>94.8</td>
<td>89.8</td>
<td>-</td>
<td>Q-elasto, nodules with indeterminate FNAB</td>
</tr>
<tr>
<td>Ding [37]</td>
<td>2012</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>87.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ning [45]</td>
<td>2012</td>
<td>&gt;4.225</td>
<td>81.8</td>
<td>82.9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.88</td>
<td>Size independent</td>
</tr>
<tr>
<td>Wang [38]</td>
<td>2012</td>
<td>&gt;2.9</td>
<td>87</td>
<td>92</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cantisani [73]</td>
<td>2013</td>
<td>&gt;2</td>
<td>95</td>
<td>88</td>
<td>97</td>
<td>91</td>
<td>-</td>
<td>-</td>
<td>Q-elasto</td>
</tr>
<tr>
<td>Chong [58]</td>
<td>2013</td>
<td>&gt;1.21</td>
<td>65.8</td>
<td>53.6</td>
<td>60</td>
<td>59.7</td>
<td>59.9</td>
<td>-</td>
<td>On a peculiar type of machine</td>
</tr>
<tr>
<td>Wang [34]</td>
<td>2013</td>
<td>&gt;3.855</td>
<td>80.77</td>
<td>91.38</td>
<td>-</td>
<td>-</td>
<td>88.1</td>
<td>0.907</td>
<td>-</td>
</tr>
<tr>
<td>Cantisani [74]</td>
<td>2014</td>
<td>&gt;2.02</td>
<td>93</td>
<td>92</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Experienced operator</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;1.86</td>
<td>84</td>
<td>79</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Inexperienced operator</td>
</tr>
<tr>
<td>El-Hariri [50]</td>
<td>2014</td>
<td>&gt; 3.5</td>
<td>88</td>
<td>86.4</td>
<td>73.3</td>
<td>94.4</td>
<td>86.9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Şahin [35]</td>
<td>2014</td>
<td>&gt;2.45</td>
<td>73.9</td>
<td>73</td>
<td>80.9</td>
<td>97</td>
<td>-</td>
<td>-</td>
<td>Patients with Hashimoto thyroiditis and nodules &gt; 1cm</td>
</tr>
<tr>
<td>Cakal [49]</td>
<td>2014</td>
<td>&gt;3.75</td>
<td>83</td>
<td>95</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>In population with nodular goiter</td>
</tr>
</tbody>
</table>

SE – sensitivity, SP – specificity, PPV – positive predictive value, NPV – negative predictive value, Acc – accuracy

Med Ultrason 2015; 17(1): 74-96

The technique is therefore difficult to apply in multinodular goiter [45]. Nodules located in the lower pole or in the isthmus are difficult to compress and to compare with normal parenchyma [19]. Sampling errors may also occur. The whole nodule should be encompassed in the measurement area, and this may prove to be difficult, due to irregular shape. In large nodules, nonuniform nodule compression, areas with alternating stiffness as well as the lack of comparable size reference tissue may lead to confusing results. The measurement should be avoided when only diffusely heterogeneous thyroid parenchyma is available, as the reference tissue is unsuitable [11,34]. With Q-elasto, the need of optimal compression angle, the lack of simultaneous real-time image and offline computing without the possibility to change the image chosen for assessment add to technical difficulties [45,71]. Slight, constant pressure over the whole examination period is also important, as high pressure may alter the results [75].

Microcalcifications in benign nodules, subacute thyroiditis, fibrosis or atypical adenoma may all induce false positive high SR [13,19], although one study claims that the size of the nodule and the concomitant presence of thyroiditis do not influence SR significantly [72]. Microcarcinomas show lower SR values, between 1.74 – 2.96 and may be easily misdiagnosed [45]. Limited or no information is available on the SR values for follicular, medullary carcinomas of lymphoma. Among the cited causes for false negative results are: follicular carcinoma, very well differentiated carcinoma, carcinoma with central necrosis and degeneration, small papillary carcinoma in lymphoma and microcarcinoma [13,70]. One study showed a very steep learning curve, with constant and reproducible results after only seven patients [62]. However, more studies on the intra- and interobserver variability of SR are needed.

Nodule to sternocleidomastoid muscle SR. In this approach, the muscle was used as a reference tissue, instead of the normal thyroid parenchyma. A value greater than 1.5 showed 90% sensibility and 50 % specificity for cancer [44]. In another study, a cutoff value > 2.31 provided sensibility 85.7%, specificity 82.1% and accuracy 82.4% [75].

Strain index was defined as the ratio of the nodule strain divided by the strain of the softest part of the surrounding normal tissue. A cutoff value >2.905 produced greater sensibility and negative predictive value than grey scale criteria [76].

Stiffness ratio. With this peculiar approach on a Philips machine, the stiffness value of the nodule was divided by the stiffness value of neighboring thyroid. A cutoff value greater than 3.16 had associated sensibility of 92.3%, specificity of 87.4%, positive and negative predictive values of 54.5% and 98.6% respectively. How-
ever, the authors found this index to provide no supplemental information, compared to the elasticity score [20].

**Strain value.** Some groups analyzed the diagnostic value of the absolute A value for strain, as displayed by the machine when computing strain ratio [47,77]. However, this approach is highly disputable, since, as discussed above, RTE machines only compute the relative strain in an area and not the absolute strain. Again, strain values do not represent the elasticity modulus [13]. As strain changes with the applied compression, its numeric value, although displayed by some machines, is completely inappropriate to compare two lesions or two individuals [13,14].

**Systolic thyroid strain index or thyroid stiffness index (TSI).** In this approach, the strain near the carotid artery (high strain area) is divided by the strain of the thyroid nodule [13,14]. A stiff nodule, with little strain will induce a high index value. A 2x2 mm ROI is used on transverse scans, where both the thyroid and the common carotid artery are seen. The highest strain near the carotid artery is compared to the lowest strain in the nodule [78]. The diagnostic value of this index is presented in table IX.

**Elasticity contrast index (ECI).** This approach obtains a strain oscillation map. By applying a co-occurrence matrix, an elasticity contrast index is computed. Malignant lesions will display high contrast while in benign lesions there will be no or low contrast [15]. The technique, experimental in the beginning, has been implemented on commercially available machines by Samsung – Medison.

**Technique.** Measuring the ECI is based on a peculiar approach to elastography, that is the Elastoscan method, which represents a quasistatic, steady state physiological excitation technique based on carotid pulsations used to induce the strain [15,79]. The transducer is held in contact with the skin, using only slight pressure. A scale indicator is used to ascertain the reaching of the steady state. Strain oscillation is displayed in colors. After the image is obtained, the contour of the nodule is traced either intranodular (containing only the nodule), on axial or longitudinal scans, or perinodular (encompassing the nodule and some tissue around it) [79]. The best tracing technique was proven to be axial perinodular [79], however without defining exactly what “some tissue around the nodule” means. Two measurements should be taken on the transverse plane containing the maximal nodule diameter [80]. Additional measurements should be obtained on images displaying peculiar characteristics of the nodule (calcification, cyst) and which are not located in the maximal diameter plane [80]. If different ECI values are obtained, one does not average them, but the highest value is retained, to maximize sensibility of detection [15,80]. Nodule assessment through ECI is shown in figure 8.

The diagnostic value of ECI is presented in table X. Values above 3.5-4 are characteristic for malignancy [15]. Inter-and intraobserver agreement was excellent, even in inexperienced operators (37), after a training stage of 30 cases (73). The technique was found more useful than grey-scale US, with the potential to reduce by 35% the number of FNABs in nodules with calcifications, due to the difference in elasticity between the benign and malignant noncalcified parenchyma [80].

**Limitations.** The exam plane should always be axial. The assessment of different areas of the same nodule is not standardized, yet. Variations in carotid pulsations in-

---

**Table IX. Diagnostic value of TSI**

<table>
<thead>
<tr>
<th>Author/Reference</th>
<th>Year</th>
<th>Cutoff for malignancy</th>
<th>SE (%)</th>
<th>SP (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Acc (%)</th>
<th>AUROC</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dighe [14]</td>
<td>2008</td>
<td>&gt;18</td>
<td>87.8</td>
<td>77.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.903</td>
<td>Experimental device</td>
</tr>
<tr>
<td>Dighe [78]</td>
<td>2010</td>
<td>&gt;10</td>
<td>100</td>
<td>79.5</td>
<td>60</td>
<td>100</td>
<td>-</td>
<td>0.942</td>
<td>-</td>
</tr>
</tbody>
</table>

SE – sensitivity, SP – specificity, PPV – positive predictive value, NPV – negative predictive value, Acc – accuracy

---

![Fig 8. Elasticity contrast index: a) benign nodule; b) malignant nodule.](image)
duce different strain and lead to different images, therefore the technique should not be employed in pregnancy, heart failure, severe pulmonary hypertension, atrial fibrillation or in isthmic nodules [15,79]. Also, the results are influenced by age, atherosclerosis and systemic hypertension and a certain interobserver variability persists [79].

**Hard area ratio by color analysis.** One study transferred the red-green-blue elastography information in the hue-saturation-value (HSV) color space [37]. The authors defined the hard area as hard connected pixels and computed the hard area ratio (hard area / lesion area) for different threshold values for hardness. For a threshold between 144 and 152 hue for hard, they found the hard area ratio accuracy to be 95.2%, the greatest reported for all elastography derived techniques [37].

3. Shear wave elastography.

**Acoustic Radiation Force Impulse (ARFI) quantification**

**Methodology of examination**

ARFI quantification estimates the elasticity of a region of interest, by measuring the speed of the US wave. This technique is integrated in a conventional ultrasonographic system using a 9-12 MHz linear transducer [83,84]. The examination is performed with the patient placed in the same manner as in conventional ultrasound. The probe is gently placed on the cervical skin surface with slight pressure on the thyroid. The patient is asked to hold the breath and the ARFI option is turned on. The ROI, which measures 5 mm, should be placed in thyroid tissue or within the solid component of a nodule, avoiding cystic areas or those with calcifications [83,85]. For the assessment of a nodule it is recommended not to include the adjacent thyroid parenchyma. After activation of ARFI the velocity is displayed on the screen, with the depth measurement. Some authors recommended performing 10 valid measurements for obtaining a reliable value [74,83]. Other authors showed that there is no significant difference between 5 or 10 measurements for thyroid stiffness [86]. A measurement is considered valid when a numeric value of the velocity, between 0 and 8.4 (or 9) m/s is displayed on the screen [83,87]. Increased stiffness of a tissue is correlated with increased velocity of the US wave. When the value exceeds the upper limit of the speed interval the display “x.xx m/s” appears on the screen, meaning not applicable (NA) [83,87]. The technique is illustrated in figure 9.

**Diagnostic value**

There are only a few studies published regarding the normal values of velocity measured by the ARFI technique in the thyroid parenchyma (table XI).

Several studies tried to define a cut-off value of the velocity for differentiating benign and malignant nodules, as increased values were found in the malignant type (table XII). Zhang et al showed that for differentiation between benign and malignant nodules, the diagnostic performance of measured US wave velocity is better for a nodule diameter greater than 20 mm [67]. In a recently published study, Han et al reported that the best cutoff value of velocity for differentiation benign vs. malignant is 2.75 m/s [89]. The diagnostic value of ARFI quantification as assessed by different studies is presented in table XIII.
Limitations and causes of false results.

The contact and the pressure of the investigator’s hand on the skin surface could alter the measurements. Experience is necessary in order to perform even quantitative ultrasound elastography. It is recommended to use a generous amount of ultrasound gel [71].

The size of the nodule is a limitation, because the ROI size is standard (5x6 mm or 2x2 cm) and cannot be changed. Thus, small nodules are not accurately measured using ARFI because the ROI will contain not only the nodule, but also the surrounding thyroid tissue and the values of velocity measured for that nodule will not be real. On the other hand, Fukuhara et al showed that the measurements of the velocity in nodules smaller than 20 mm in diameter are not stable [92].

Nodules with fluid areas or calcifications should be excluded due to the impossibility to place the ROI inside of the parenchyma of the nodule [84].

The ARFI penetration depth is limited to 5.5 cm, so large thyroids or very large and deep located nodules cannot be properly assessed using ARFI quantification [93].

Another technical limitation is represented by the impossibility to measure velocities higher than 9 m/s. Thus, very hard nodules or areas in the thyroid parenchyma will not be assessed properly [67,74,87,94,95].

Real-time shear wave elastography (Supersonic elastography, SSE)

Methodology of examination

SSE is an operator-independent, reproducible and quantitative elastography method [96,97]. The examination is performed with the patient placed in the same manner like for the conventional ultrasound. The linear probe (4-15 MHz) is gently placed on the cervical skin surface with slight pressure on the thyroid. The patient is asked to hold the breath and the quantitative evaluation is turned on. The displayed color-coded image shows soft tissue in blue and rigid tissue in red. The quantitative information is provided by the elasticity index (EI), expressed in kilo-Pascal [96] (fig 10).

SSE is displayed alongside with grey-scale US (split screen mode) for localization of the lesions. The box for SSE measurement should encompass the lesion or as much of it as displays elasticity color signal and a small amount of surrounding thyroid tissue. Areas with no signal on the elastogram represent areas where shear wave velocity could not be measured due to the properties of

Table XI. The values of US wave velocity by ARFI in normal thyroid tissue

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Velocity value (m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bojunga [83]</td>
<td>2012</td>
<td>1.76 (range: 0.89–3.33)</td>
</tr>
<tr>
<td>Friedrich-Rust [85]</td>
<td>2012</td>
<td>1.98 (range: 1.20–3.63)</td>
</tr>
<tr>
<td>Zhan [88]</td>
<td>2013</td>
<td>2.16 m/s (range: 1.45–3.00)</td>
</tr>
</tbody>
</table>

Table XII. The mean US wave velocity by ARFI in thyroid nodules

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Mean velocity (m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bojunga [83]</td>
<td>2012</td>
<td>1.90 2.60</td>
</tr>
<tr>
<td>Friedrich-Rust [85]</td>
<td>2012</td>
<td>2.02 4.3</td>
</tr>
<tr>
<td>Zhang [87]</td>
<td>2012</td>
<td>2.34 4.82</td>
</tr>
<tr>
<td>Zhan [88]</td>
<td>2013</td>
<td>2.36 3.45</td>
</tr>
<tr>
<td>Zhang [67]</td>
<td>2014</td>
<td>2.34 6.65</td>
</tr>
</tbody>
</table>

Table XIII. Diagnostic value of ARFI quantification in the assessment of thyroid nodules

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Cutoff value (m/s)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friedrich-Rust [85]</td>
<td>2012</td>
<td>3.1-3.3</td>
<td>-</td>
<td>91 – 95</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Zhang [87]</td>
<td>2012</td>
<td>2.87</td>
<td>75</td>
<td>82.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gu [90]</td>
<td>2012</td>
<td>2.555</td>
<td>86.36</td>
<td>93.42</td>
<td>79.17</td>
<td>95.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.45</td>
<td>100</td>
<td>63.6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Zhan [88]</td>
<td>2013</td>
<td>2.85</td>
<td>94.4</td>
<td>85.3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hou [91]</td>
<td>2013</td>
<td>3.1</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Zhang [67]</td>
<td>2014</td>
<td>2.9</td>
<td>91.3</td>
<td>85.10</td>
<td>80.77</td>
<td>93.44</td>
</tr>
<tr>
<td>Han [89]</td>
<td>2015</td>
<td>2.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SE – sensitivity, SP – specificity, PPV – positive predictive value, NPV – negative predictive value
the scanned structures (shear waves do not propagate in fluids) or due to technical limitations. It is recommended to acquire at least three cineloops lasting 10 seconds each, for each lesion, with the transducer held still during acquisition [98]. It is also important to set the machine for optimal image acquisition. The interval of values for the EI on the thyroid preset protocol is 0–180 kPa. Monpeyssen et al [61] recommended that the maximum value of the elastography scale should be less than 80 kPa. Another important technical factor for a good image is the intensity of the elastographic signal. If the signal is too weak, color signal will not be displayed. The gain should be increased until the limit of occurrence of noise. The operator should be aware of the pressure exerted by the transducer. Increased pressure produces superficial artifacts of hardening and variations in stiffness in homogenous tissues [42]. The size of the ROI and the standard deviation of the measured values are inversely related.

**Diagnostic value**

There are only a few studies published regarding the normal values of the elasticity index in the thyroid parenchyma, measured by the SSE technique (table XIV).

**Table XIV. The values of elasticity index in normal thyroid tissue**

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>EI (kPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sebag [97]</td>
<td>2010</td>
<td>15.9 ± 7.6 (range: 5-35)</td>
</tr>
<tr>
<td>Arda [99]</td>
<td>2011</td>
<td>10.97 ± 3.1 (range: 1–24)</td>
</tr>
<tr>
<td>Bathia [98]</td>
<td>2012</td>
<td>9.0 ± 4.0 (pre-compression of 5%) 23.9 ± 7.4 (pre-compression of 20%)</td>
</tr>
<tr>
<td>Kim [100]</td>
<td>2014</td>
<td>23.4 ± 10.8</td>
</tr>
</tbody>
</table>

Some studies assessed the thyroid nodules using SSE, trying to define a cut-off value of the EI for differentiating benign vs. malignant. All studies found increased values for malignant nodules (table XV).

The diagnostic value of SSE was assessed by different groups and is presented in table XVI. Bathia et al [98] calculated different cut-off values for EI, with the ROI kept at 2 mm, and found that increasing the cut-off value (>10.3 kPa – >132 kPa) is associated with increasing specificity (8.9 – 100%), and decreasing sensibility (100 – 7.7%). Szczepanek et al [42] reported that SSE is predictor of malignancy if the cutoff value of maximum EI is set at 50 kPa. Park et al [103] found that a mean value for EI ≥85 kPa or a maximum value ≥94 kPa are independent predictors of thyroid malignancy.

**Limitations and causes of false results**

The pressure applied on the neck of the patient may alter the values of EI [98]. The elasticity of the structures are influenced by the external pressure applied, the stiffness rising with increasing pressure due to the nonlinear elastic effects. In the isthmus region, the stiffness is increased due to the subjacent trachea. To avoid this artifact while performing SSE, a paracoronal plane for evaluating a nodule located in this area should be used [105].

Another limitation of performing SSE in thyroid pathology is represented by the presence of calcifications [101]. The accuracy of SSE measurements might be altered by the small nodules [102].

Papillary carcinoma is harder that follicular or medullary carcinoma, therefore the optimal cut-off value for

**Table XV. The mean elasticity index in thyroid nodules**

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Elasticity index (kPa)</th>
<th>Malignant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Benign</td>
<td></td>
</tr>
<tr>
<td>Sebag [97]</td>
<td>2010</td>
<td>36 ± 30 (range: 0200)</td>
<td>150 ± 95 (range: 30-365)</td>
</tr>
<tr>
<td>Bathia [98]</td>
<td>2012</td>
<td>15.0 ± 5.3 (pre-compression 5%) 53.0 ± 30.5 (pre-compression 20%)</td>
<td>44.5±27.8 (5% pre-compression) 37.4±219.0 (20% pre-compression)</td>
</tr>
<tr>
<td>Veyrieres [101]</td>
<td>2012</td>
<td>115 ± 60.4</td>
<td></td>
</tr>
<tr>
<td>Kim [102]</td>
<td>2013</td>
<td>51.46 ± 22.75</td>
<td>85.52 ± 41.94</td>
</tr>
<tr>
<td>Park [103]</td>
<td>2013</td>
<td>77.2 ± 47.1</td>
<td></td>
</tr>
</tbody>
</table>

**Table XVI. The diagnostic value of the SSE in the assessment of thyroid nodules**

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Cutoff value (kPa)</th>
<th>SE (%)</th>
<th>SP (%)</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sebag [97]</td>
<td>2010</td>
<td>65</td>
<td>85.2</td>
<td>93.9</td>
<td>92.3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>42.1</td>
<td>52.9</td>
<td>77.8</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Veyrieres [101]</td>
<td>2012</td>
<td>66</td>
<td>80</td>
<td>90.5</td>
<td>52.8</td>
<td>99.3</td>
</tr>
<tr>
<td>Carneiro [104]</td>
<td>2013</td>
<td>34.5</td>
<td>77</td>
<td>71</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2013</td>
<td>42.1</td>
<td>53</td>
<td>78</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kim [102]</td>
<td>2013</td>
<td>62</td>
<td>66.6</td>
<td>71.6</td>
<td>40.6</td>
<td>85.7</td>
</tr>
<tr>
<td>Park [103]</td>
<td>2014</td>
<td>85</td>
<td>43.6-54.4</td>
<td>71.1-87.7</td>
<td>88-93.8</td>
<td>-</td>
</tr>
</tbody>
</table>

SE – sensitivity, SP – specificity, PPV – positive predictive value, NPV – negative predictive value
all nodules is not suitable for predicting microcarci-
oma. This leads to lower sensibility but higher specificity
[103]. One study showed that the majority of microcarci-
nomas were missed by the SSE examination [104].

Proven clinical utility

This method represented an additional parameter for
characterization of thyroid cancer together with conven-
tional ultrasound [106]. SSE could guide the FNAB,
avoiding puncture of the benign nodules [105]. Liu et al
[107] found that SSE is useful for the characterization
of nodules with diameter lower than 1 cm. SSE could
evaluate the elasticity of thyroid nodules separate from
the coexistence of autoimmune thyroiditis [100,105].

Nodular thyroid disease other than papillary
carcinoma.

Medullary carcinoma. More than 50% of the nodules
with this pathology appear soft, type 1 or 2 (of 4). Quali-
tative USE was found ineffective in differentiating these
nodules from benign lesions [108]. USE may, however,
suggest, through hardness, malignity in case of small,
multiple, synchronous carcinomas [109].

Treated nodules, either with laser or radiofrequency,
become hard over time, with the maximal stiffness in-
stalled at scarring [110].

Thyroid adenomas have SR of 2.37±1.21 [45] or
1.76±1.25 [38], lower than malignant nodules.

Benign hyperfunctional nodules are hard, with RTE
specificity of 30%, independent of the scoring type used.
Therefore, the accuracy of RTE in diagnosing hot thyroid
nodules is highly questionable [111].

Benign nodules in nodular goiter show score 2 or
3 (of 4), with SR of 2.05 ± 1.33 [45] or 1.53±1.23 [38].

In acromegaly, USE detects a significantly higher
number of stiff nodules, when compared to nonacromeg-
alic goiter. These nodules are fibrotic, nonmalignant [112].

II. Diffuse thyroid disease and thyroiditis.

RTE has been used to differentiate diffuse thyroid dis-
ease from normal parenchyma, by the analysis of eleven
color parameters of the image and the “elasticity index”
derived from them. The average relative value and the
elastic index were found useful for this purpose [113].

ARFI quantification. The diagnosis of diffuse thyroid
disease is based on clinical, laboratory and usually 2D
and Doppler ultrasound. The changes in the elasticity of
the parenchyma depend of the stage of inflammation and
sclerosis, leading to increase the stiffness of the gland.
While studying the use of ARFI quantification, Sporea
and coworkers [86,114] found a statistically significant
difference in velocity values in autoimmune pathology
(Graves’ disease vs. chronic autoimmune thyroiditis)
with a value of 2.07±0.44 m/s in Graves’ disease com-
pared with 2.68±0.50 m/s in chronic autoimmune thy-
roiditis. Bahn et al found a value of 2.82±0.47 m/s in
Graves vs. 2.49±0.48 m/s in chronic autoimmune thy-
roiditis [94]. Sporea and coworkers reported a cut-off
value >2.53 m/s for differentiation between normal thy-
roid and diffuse thyroid diseases, with a sensibility and
positive predictive value >90% [86].

SSE. Magri et al [106] studied the EI in patients with
chronic autoimmune thyroiditis and found that in Hashi-
moto thyroiditis the elasticity of extra nodular tissue is
decreased according to the thyroid antibody titre and the
degree of thyroid function damage. Kim et al [100] found
a cut-off mean value of the EI, in cases of diffuse thyroid
disease, of 27.6 kPa and a maximum value of 41.3 kPa,
with a sensibility of 40.9% and a specificity of 82.9%.

Subacute (granulomatous, de Quervain) thyroiditis
presents stiff inflammatory areas [23,115]. These areas
should be included in the differential diagnosis of car-
icoma. They are harder than the ones encountered in
chronic thyroiditis [10]. The appearance is significantly
different from multinodular goiter, but not from cancer.
There is no correlation between the level of T4 and the
RTE score [23]. USE may contribute to the differential
diagnosis with Hashimoto thyroiditis. It may be used to
direct FNAB into inflamed areas unapparent on the grey-
scale image [23].

Riedel (chronic) thyroiditis is characterized by ex-
tremely stiff parenchyma, with values of 143-281 kPa.
The hypoechoic areas are the stiffest, representing fibro-
sis. The evolution of the stiff areas under cortisone treat-
ment is not known [116].

Hashimoto thyroiditis often produces pseudonodular
images. For the differentiation of true nodules from pseu-
donodules, RTE has sensibility of 92.9%, with specificity
95%, positive and negative predictive values of 86.7% and
97.4% respectively and accuracy of 94.4%, better
than the combination grey-scale and Doppler US [117].
There is no correlation between the ES or SR and the lev-
el of antibodies, STH, the size of the nodule or the pre-
ence of calcification. SR proved to be slightly better than
ES for the diagnosis of the nodules in this disease [35].

III. Value and limitations of elastography in the
diagnosis of thyroid nodules.

The existing body of knowledge allowed the elabo-
ration of several metaanalyses which offer a more com-
prehensive view on the usefulness of the elastographic
techniques. Some of the values computed in these studies
are presented in table XVII.
Table XVII. Diagnostic value of elastography in metaanalysis studies.

<table>
<thead>
<tr>
<th>Author/ Reference</th>
<th>Year</th>
<th>Technique</th>
<th>SE (%)</th>
<th>SP (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Acc (%)</th>
<th>AUROC</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bojunga [118]</td>
<td>2010</td>
<td>RTE scoring</td>
<td>92</td>
<td>90</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Razavi [119]</td>
<td>2013</td>
<td>RTE Score 4 + 5</td>
<td>82</td>
<td>82</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td></td>
<td></td>
<td>89</td>
<td>82</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Ghajarzadeh [120]</td>
<td>2014</td>
<td>RTE score 1</td>
<td>98.3</td>
<td>19.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Veer [121]</td>
<td>2014</td>
<td>Overall (RTE + shear wave)</td>
<td>87</td>
<td>80.6</td>
<td>84.9</td>
<td>96.7</td>
<td>81.7</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>RTE (external compression)</td>
<td>86.5</td>
<td>79.7</td>
<td>46.9</td>
<td>96.6</td>
<td>80.9</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shear wave</td>
<td>85.9</td>
<td>89.1</td>
<td>60.3</td>
<td>97</td>
<td>88.6</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Carotid pulsation overall</td>
<td>100</td>
<td>77.2</td>
<td>52.2</td>
<td>100</td>
<td>81.7</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Sun [122]</td>
<td>2014</td>
<td>RTE scoring</td>
<td>79</td>
<td>77</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.894</td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td></td>
<td></td>
<td>85</td>
<td>80</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.928</td>
<td></td>
</tr>
<tr>
<td>Trimboli [123]</td>
<td>2014</td>
<td>RTE scoring</td>
<td>69</td>
<td>75</td>
<td>-</td>
<td>-</td>
<td>73</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

SE – sensitivity, SP – specificity, PPV – positive predictive value, NPV – negative predictive value

A significant heterogeneity was observed for specificity in various studies [118]. Elastographic signs were found to have higher diagnostic value than grey-scale signs such as hypoechogenicity, microcalcifications, irregular margins, no halo, vertical growth and intranodular vessels [119]. The likelihood to exclude malignancy based on negative elastography is much higher than based on the grey-scale appearance [119]. An elastogram positive for malignancy is more informative than a positive grey scale exam [119].

USE has the potential to contribute in the differentiation of degenerated less than 1 cm nodules from papillary carcinoma [63]. The association of grey scale US and elastography allows mainly to define the nodules with indeterminate cytology that do not need operation [53]. The technique also plays a major role in selecting the nodules that do not require FNAB [52]. Soft nodules have a great likelihood of being benign and, if no other suspicion element is present, FNAB may be avoided or delayed [7]. The use of TSI might reduce the number of punctures with 60% [78] while concomitant application of TIRADS and elastography might reduce the number of FNABs with 33.8% [3]. A nodule with score 1 may be considered benign and FNAB or surgery may be avoided [120]. In a low risk case with indeterminate or nondiagnostic FNAB, a negative elastographic result may represent a reassuring information to avoid FNAB or surgery and institute follow-up, freeing resources for patients with suspicious nodules [121]. The constantly high negative predictive values of the elastography techniques infer that elastography is useful in determining which nodule may be followed up without FNAB [10]. On the other hand, hard nodules need to be biopsied, even if some of them will show no malignancy [7]. The combination grey scale and elastography is useful for the selection of the nodules that will undergo puncture, as it lowers the risk of overlooking nonsuspicious grey-scale appearing nodules [10].

However, it must be stressed that elastography is just a part of the US armamentarium. Its results should always be integrated with the grey-scale appearance [124]. The technique increases the specificity of grey-scale US in the identification of malignant nodules and in the identification of nodules that require FNAB [104]. It must be regarded as an important complement and not a substitute for grey-scale US [10].

Not all published studies agree on the above conclusions. Carneiro-Pla and coworkers [104] found that USE does not have enough sensibility to indicate which nodules may be followed without FNAB. Other authors [123] also found that the accuracy of USE is insufficient to indicate or dismiss the operation in patients with indeterminate cytology, and it should not be used alone to establish the indication. Some studies see a role for USE to complement the grey-scale characterization of nodules, especially for less experienced practitioners [61] and infer that the value of this technique may be limited in centers where high diagnostic performances are attained with conventional US by dedicated physicians [124]. Some authors even consider “it less likely that the use of this imaging technique to study thyroid nodules will expand significantly in the near future” [33].

Beyond the limitations presented for each of the USE techniques, above, the general limitations should be acknowledged. The many technological solutions, scores, values and cutoff data make the results less predictable.
The population based selection of nodules, exclusion of calcified nodules, small sample sizes, inconstant selection criteria and appealing to the surrogate definite diagnosis provided by FNAB instead of the true “gold standard” of surgery are, all, possible explanations of still persisting discrepancies in the reported results [96]. However, the ever increasing interobserver agreement reported in recent studies may be due to the application of quality scales and the use of shear-wave techniques, which produce more reproducible results [124].

Proven usefulness and still unanswered questions

Most of the recent studies agree that the main role of USE is to indicate which nodules may be followed up without resorting to FNAB or surgery. The high negative predictive value was obvious even in early studies [29]. The body of knowledge about thyroid elastography has grown dramatically, even in the short time span from the publication of the EFSUMB guidelines [125].

The use of USE in multinodular goiter or in patients with previous radioiodine therapy or surgery has not been studied enough. The role of USE in the management of papillary microcarcinomas still needs to be defined. It is not known to what extent this technology may aid in the management of positron emission tomography/computed tomography-positive incidental thyroid tumors [96].

It was stated that “to be a good adjunctive diagnostic tool to conventional US, additional elastography should improve the diagnostic performances of conventional US, rationalizing the consequently longer US time and the additional cost of elastographic software that comes with its use” [124]. The information gathered in this review cast an optimistic light on the perspectives of the clinical use of ultrasound elastography in thyroid disease.

References

Sorin M. Dudea et al. Ultrasound elastography in thyroid disease


110. Andrioli M, Valcavi R The peculiar ultrasonographic and elastographic features of thyroid nodules after treatment with laser or radiofrequency: similarities and differences. Endocrine 2014; 47: 967-968.


