WFUMB GUIDELINES AND RECOMMENDATIONS ON THE CLINICAL USE OF ULTRASOUND ELASTOGRAPHY: PART 4. THYROID

DAVID COSGROVE,∗ RICHARD BARR,†‡ JOERG BOJUNGA,§ VITO CANTISANI,¶ MARIA CRISTINA CHAMMAS,‖ MANJIRI DIGHE,** SUDHIR VINAYAK,†† JUN-MEI XU,**‡‡ and CHRISTOPH F. DIETRICH∥∥

∗Division of Radiology, Imperial and Kings Colleges, London, UK; †Department of Radiology, Northeastern Ohio Medical University, Rootstown, Ohio, USA; ‡Southwoods Imaging, Youngstown, Ohio, USA; §Department of Internal Medicine 1, Endocrinology and Diabetology, Goethe University Hospital, Frankfurt am Main, Germany; ¶Department of Radiological Sciences, Oncology and Pathology, Policlinico Umberto I, University Sapienza, Rome, Italy; ‖Ultrasound Division, Department of Radiology, Hospital das Clínicas, School of Medicine, University of São Paulo, São Paulo, Brazil; **Department of Radiology, University of Washington, Seattle, Washington, USA; ††Department of Imaging and Diagnostic Radiology, Aga Khan University Hospital, Nairobi, Kenya; ‡‡Department of Medical Ultrasound, Shanghai Tenth People’s Hospital, Ultrasound Research and Education Institute, Tongji University School of Medicine, China; ∥∥Department of Internal Medicine 2, Caritas Krankenhaus, Bad Mergentheim, Germany; and **Sino-German Research Center of Ultrasound in Medicine, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

Abstract—The World Federation for Ultrasound in Medicine and Biology (WFUMB) has produced guidelines for the use of elastography techniques including basic science, breast and liver. Here we present elastography in thyroid diseases. For each available technique, procedure, reproducibility, results and limitations are analyzed and recommendations are given. Finally, recommendations are given based on the level of evidence of the published literature and on the WFUMB expert group’s consensus. The document has a clinical perspective and is aimed at assessing the usefulness of elastography in the management of thyroid diseases. (E-mail: christoph.dietrich@ckbm.de) © 2016 World Federation for Ultrasound in Medicine & Biology.

Key Words: World Federation for Ultrasound in Medicine and Biology, Guidelines, Elastography, Shear wave elastography, Strain elastography.

INTRODUCTION

Epidemiology

Thyroid nodules are a common finding in the general population, and their detection is increasing with the widespread use of ultrasound (US). The prevalence of thyroid nodules varies from 19% to 67% and increases with age, affecting about 50% of the population older than 40 years of age (Remonti et al. 2015). A thyroid nodule is a discrete lesion within the thyroid gland that is distinct from the surrounding thyroid parenchyma. Some palpable lesions may not correspond to distinct imaging abnormalities and so do not meet the strict definition for thyroid nodules (Haugen et al. 2015). Non-palpable nodules detected on US or other anatomic imaging studies are termed incidentally discovered nodules or “incidentalomas.” Some authors report that ultrasound of the neck performed for reasons other than thyroid disease has revealed that 13% to 88% of patients have thyroid nodules (Brander et al. 2000; Chammas et al. 2005; Tramalloni et al. 1999). Non-palpable nodules have the same risk of malignancy as palpable nodules of the same size (Haugen et al. 2015).

Epidemiologic studies have reported the prevalence of palpable thyroid nodules to be approximately 5% in women and 1% in men living in iodine-sufficient parts of the world. In contrast, high-resolution US has detected thyroid nodules in 19%–68% of randomly selected individuals, with higher frequencies in women and the elderly (Haugen et al. 2015). However, less than 15% of these nodules are malignant. Nodules that have a higher risk of malignancy include those found in the young (<14 y) and elderly (>70 y), in females (Ferraioli et al. 2015) and in those with a family history of thyroid cancer, a personal history of exposure to ionizing radiation in childhood, a prior history of thyroid cancer or nodules positive on [18F]fluorodeoxyglucose positron emission tomography (Haugen et al. 2015; Kwak et al. 2011; Sipos 2009).

Patients with multiple nodules have the same likelihood of malignancy (14%) as those with a solitary nodule, a finding that contradicts traditional teaching that the risk of malignancy decreases with increasing
numbers of nodules (Sipos 2009). Most nodules of the thyroid are not true neoplasms but are benign hyperplastic nodules that form as a result of cycles of hyperplasia and involution of the thyroid parenchyma (Ahuja et al. 1996; Reading et al. 2005).

It is currently recommended that US-guided fine-needle aspiration (FNA) biopsy be performed on up to four nodules, preferentially sampling those with the most suspicious findings on US (Sipos 2009). The clinical importance of thyroid nodules rests with the need to exclude thyroid cancer. Information on the probability of each US feature to be associated with malignancy would help in making the clinical decision to perform FNA biopsy (Remonti et al. 2015).

Examination technique (B-mode and elastography)

Thyroid ultrasound (B-mode and elastography) should be performed using a high-resolution scanner, equipped with a 12- to 15-MHz linear probe. During B-mode US, thyroid nodules are identified and a region of interest for elastography is identified (Lyshchik et al. 2005). For shear wave elastography (SWE), some systems require the use of a lower-frequency transducer (9 MHz). In these cases, the high-frequency transducer should be used to evaluate the B-mode, and the lower-frequency probe should be used to perform SWE.

The stiffness of the gland depends on the structural properties of the matrix of tissues (cells, membranes, extravascular matrix, microvessels, etc.), whereas in conventional US, it is the microscopic structure that determines reflectivity. Thus, elastograms contain contrast based on histologic tissue structure, enabling the differentiation of normal gland from nodules and parenchymal diseases. Tissue stiffness is a feature that reflects the nature of the thyroid nodule; neoplasia and inflammation alter the tissue composition and structure and increase the parenchymal stiffness.

Anatomical considerations regarding the use of elastography

The thyroid gland is located relatively superficially. It is surrounded by a sheath derived from the pre-tracheal layer of the deep fascia, which attaches the gland to the trachea and the larynx. Anterior and lateral to the lobes lie the strap muscles (sternohyoid, sternothyroid, sternomastoid and omohyoid). Posterior and lateral are the common carotid arteries, internal jugular veins and vagus nerves. More medial and posterior to the lobes are the larynx, trachea, esophagus and longus colli muscles (Ahuja 2000).

Strain elastography is especially problematic in two circumstances: First, the combination of a relatively superficial protuberant mass and sparse overlying subcutaneous and deep cervical connective tissue produces difficulties in applying uniform stress over the entire region of interest; second, lesions adjacent to the carotid arteries have lateral displacements from the pulsations, which can result in mistracking artifacts. For the latter, orientating the transducer in the longitudinal plane parallel to the long axis of the adjacent artery appeared to improve the consistency and hence the quality of elastograms (Bhatia et al. 2010).

Ultrasound criteria, B-Mode and color Doppler US

Ultrasound criteria (B-mode, Doppler) are important but are not part of this guideline. We refer to other guidelines (Frates et al. 2006) and reviews (Frates et al. 2005).

Consensus reports

The American Thyroid Association has defined as suspicious sonographic findings nodules with microcalcifications, increased intranodular vascularity, hypo-echogenicity and irregular, infiltrative margins (Haugen et al. 2015). Additional suspicious features include a taller-than-wide shape, marked hypo-echogenicity, extension beyond the thyroid capsule and cervical lymph node metastases. Less specific US features that may raise suspicion include lack of a hypo-echoic halo and solid composition (Nachiapppan et al. 2014).

Thyroid ultrasound has been widely used to stratify the risk of malignancy in nodules and aid decision making on whether FNA is indicated (Haugen et al. 2015). The probability of malignancy increases with the number of suspicious US features. The American Association of Clinical Endocrinologists states that the coexistence of at least two suspicious US features greatly increases the likelihood of thyroid cancer (Gharib et al. 2010).

Thyroid Imaging Reporting and Data System

Appropriate criteria are needed to avoid a further increase in unnecessary biopsies of thyroid nodules. Thyroid Imaging Reporting and Data System (TI-RADS) was conceived to improve understanding and communication among specialists, facilitating US reports (American Thyroid Association Guidelines Taskforce on Thyroid et al. 2009). This classification (Horvath et al. 2009) was modified by Kwak et al. (2011). It is modeled on the Breast Imaging Reporting and Data System and is based on the likelihood of malignancy (Dietrich and Bojunga 2015; Hao et al., 2015; Kwak et al. 2011).

The TI-RADS classification (Stoian et al. 2015) incorporates the following categories:

- **TI-RADS 1**: Normal thyroid gland
- **TI-RADS 2**: Benign conditions (0% malignancy)
- **TI-RADS 3**: Probably benign nodules (<5% malignancy)
- **TI-RADS 4**: Suspicious nodules (5%–80% malignancy rate) (subdivision into 4a [malignancy between 5% and
10% and 4b [malignancy between 10% and 80%] is optional)

- TI-RADS 5: Probably malignant nodules (>80% malignancy)
- TI-RADS 6: Biopsy-proven malignant nodules (Horvath et al. 2009)

Current TIRADS category suggestions include 3 (no suspicious US features), 4a (one suspicious US feature), 4b (two suspicious US features), 4c (three or four suspicious US features) and 5 (five suspicious US features), using the likelihood of malignancy from the TIRADS categorization (Kwak et al. 2011).

Recommendation 1. The TIRADS reporting system for conventional ultrasound should be used (Kwak et al. 2011). Level of evidence (LoE): 3a, grade of recommendation (GoR): B; 57% consensus.

Reporting recommendations

Standardized US reporting criteria should be followed (Dietrich and Bojunga 2015; Gharib et al. 2010). The report should convey the nodule’s size (in three dimensions) and location and a description of the nodule’s sonographic features, including composition (solid, cystic proportion or spongiform), echogenicity, margins, presence and type of calcifications, shape if taller than wide and vascularity. The pattern of sonographic features in a nodule confers a risk of malignancy and, combined with nodule size, guides FNA decision making (Haugen et al. 2015).

Recommendation 2. Standardized US reporting criteria should be followed, indicating localization, shape, size, margins, content and echogenic and vascular patterns of the nodule (90%) (Gharib et al. 2010; Haugen et al. 2015). LoE: 1b, GoR: A; 100% consensus.

PUBLISHED GUIDELINES

To date the only published guidelines on thyroid US elastography are the European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) Guidelines and Recommendations, which state that both strain elastography (SE) and SWE elastography may be performed with no patient preparation; dedicated equipment that supports elastography is required. Elastography is recommended as an additional tool to conventional ultrasound and to guide follow-up of lesions previously diagnosed as benign at FNA biopsy.

STRAIN ELASTOGRAPHY

Introduction to strain elastography

Strain elastography indicates the stiffness in the tissues, defined as the change in length during compression divided by the length before compression. Young’s modulus, the relationship between compression (or stress) and strain, is defined as \( E = \frac{\text{stress}}{\text{strain}} \). Commercial ultrasound elastography equipment cannot measure the applied stress, so direct quantification is not possible (Carlsen et al. 2015), and strain elastograms depict only relative stiffness. The stress in SE is usually applied externally either by manual compression with the transducer (Fig. 1) or by acoustic radiation force impulse (ARFI) (Fig. 2). Alternatively, physiologic shifts within the patient, such as carotid artery pulsations, can be used (Fig. 2). Stress with the transducer is applied by continuously and uniformly compressing and decompressing the skin of the patient a few millimeters at a time. The elastogram is calculated from the change in signals from before to after compression and is displayed in a split-screen mode with both the conventional B-mode image and the elastogram on the monitor. In an alternative display, the elastogram is a color overlay on the B-mode image. The tissue stiffness is displayed either in gray scale (Fig. 3) or, more commonly, in a continuum of colors from red to green to blue, designating soft (high strain), intermediate (equal strain) and hard (no strain) (Fig. 4). However, at present there is no color standard and the display preference varies, with some systems having a color scale inverse to the one mentioned (Fig. 5). Details have been published as Part I of the WFUMB guidelines on elastography (Shiina et al. 2015) and by EFSUMB (Bamber et al. 2013) and are not repeated here.

Strain imaging: review of literature

Literature reviews were carried out using the following key words: thyroid and SWE, thyroid and shear wave and thyroid nodules and ARFI (Table 1) (Ghajarzadeh et al. 2014; Razavi et al. 2013; Sun et al. 2014).

There are several ways of providing semi-quantitative elastographic measures which are especially useful in evaluating focal lesions. These include visual scoring systems, such as the Tsukuba system (Itoh et al. 2006) for breast tumors, and the strain ratio.

Visual scoring system. The Tsukuba scoring system for breast ultrasound was revised using a 5-point scale (Itoh et al. 2006) (Fig. 6): Score 1 indicates deformability of the entire lesion; score 2, deformability of most of the lesion with some small stiff areas; score 3, deformability of the peripheral portion of the lesion with stiff tissue in the center; score 4, stiffness of the entire lesion; score 5, stiffness of the entire lesion and surrounding tissue. If a lesion is classified between 1 and 3, it is considered benign; if classified 4 or 5, it is considered to be malignant. This has been adapted and used in thyroid elastography in multiple studies.
Recommendation 3. The Tsukuba five-pattern visual scoring system can be used for thyroid nodules (Itoh et al. 2006). LoE: 2a, GoR: B; 86% consensus.

The Tsukuba score was used in a study of 92 patients, in which 49 patients were scored 1 and 2, all benign; 13 patients were scored 3, one carcinoma and 12 benign lesions; and 30 patients were scored 4 and 5, all carcinomas. Thus, scores 4 and 5 were highly predictive of malignancy ($p < 0.0001$), with a sensitivity of 97%, specificity of 100%, positive predictive value of 100% and negative predictive value of 98% (Rago and Vitti 2008; Rago et al. 2007).

The 4-point system for scoring strain histograms was modified from the Tsukuba classification (Fig. 7). ES 1 was assigned to nodules that are soft throughout the whole region of interest (ROI); ES 2, to nodules that are soft in a large portion of the ROI; ES 3, to nodules with stiffness in a large portion of the ROI; and ES 4, to entirely stiff nodules (Asteria et al. 2008). Sensitivities and specificities of US elastography for thyroid cancer diagnosis were 94.1% and 81%, with positive and negative predictive values of 55.2% and 98.2%, respectively.

Recommendation 4. The four-pattern scoring system for RTE of thyroid nodules is an alternative visual system (Rago et al. 2007; Rago and Vitti 2008). LoE: 2a, GoR: B; 100% consensus.

A two-pattern scoring system has also been used by combining the Tsukuba 1 + 2 + 3 into score 1 and the Tsukuba 4 + 5 into score 2 with excellent sensitivity and specificity (Rago and Vitti 2009). Score 1 was defined as <50% blue and hence soft nodules and score 2 as ≥50% blue (Chong et al. 2013).

Although most systems have used a color-based evaluation scoring system, some have used gray-scale scoring systems (Fig. 3) (Ding et al. 2011; Lyshchik et al. 2005). In one study, the elastograms were displayed by the machine using a 4-point gray scale with score 1 being a very dark lesion, score 2 being markedly darker than surrounding parenchyma, score 3 being slightly darker than surrounding parenchyma and score 4...
being as bright as or brighter than surrounding parenchyma (Lyshchik et al. 2005). The original color thyroid elastograms from a red–green–blue color space to the hue–saturation–value color space and extracted texture features were converted in the other study (Ding et al. 2011). They had a classification accuracy of 93.6% in differentiating between benign and malignant nodules. Histogram analysis of the color elastograms (elasticity index and mean) were used for diagnosis and management of patients with diffuse thyroid disease (Cantisani et al. 2015a, 2015b; Yoon et al. 2014), and another study indicated excellent inter-observer agreement (Lim et al. 2012).

![Fig. 2. Benign nodule. This benign nodule was part of a multinodular goiter. The color scale is coded with red as hard, so the nodule is soft to intermediate in stiffness. Carotid pulsation is used as the driving force in this system. Note also the vertical bar alongside the color scale: This is the quality indicator, which should be full green. Image courtesy of Paul Sidhu, King’s College Hospital, London.](image1)

Recommendation 5. SE can be used in combination with conventional US to improve specificity (Cantisani et al. 2015a, 2015b; Cosgrove et al. 2013). LoE: 1b, GoR: A; 86% consensus.

Strain ratio. The strain ratio, a relatively new method for assessing thyroid nodules, is the ratio between the strain in adjacent thyroid tissue and strain in the nodule and is an objective and semi-quantitative method of analyzing tissue stiffness (Figs. 8 and 9). Two types of strain ratios have been defined: parenchyma-to-nodule strain ratio (PNSR), which is the mean strain in the normal thyroid parenchyma divided by the mean strain

![Fig. 3. Adenoma. Longitudinal scan. This well-defined nodule in a 60-y-old man was benign on fine-needle aspiration. The coding used is in gray scale, with black as hard. The strain ratio, using adjacent strap muscle as the reference, is 3.93, consistent with the cytology.](image2)
within the thyroid nodule, and muscle-to-nodule strain ratio (MNSR), which is the mean strain in an adjacent strap muscle divided by the strain in the nodule (Ciledag et al. 2012). No significant difference were found between PNSR and MNSR in the distinction between benign and malignant thyroid nodules, suggesting that MNSR

![Figure 4](image1.png)

**Fig. 4.** Papillary carcinoma: lymph node metastasis. In this patient with a biopsy-proven papillary carcinoma, right-sided cervical lymphadenopathy was palpable. The node visualized on B-mode has a rounded shape with effaced internal architecture (no hilum indicated) and appeared as medium stiff on elastography. The color scale codes hard as blue.

![Figure 5](image2.png)

**Fig. 5.** Papillary carcinoma. An incidentally found 25-mm nodule on B-mode in this 28-y-old male has heterogeneous, predominantly low-level echoes. On strain elastography its hard nature is displayed as predominantly red shades, the coding used on this system. It was scored as Tsukuba 5 and fine-needle aspiration cytology was classified as Bethesda category VI. Image courtesy of Pingtong Huang, Second Zhejiang University Hospital, Hangzhou, China.
could be used in situations in which PNSR could not be used, for example, when a large nodule occupies the whole thyroid gland or in the presence of abnormal thyroid parenchyma (e.g., thyroiditis) (Aydin et al. 2014).

Several studies have assessed the strain ratio, but agreement on the critical cutoff point to differentiate between benign and malignant nodules has not been reached (Cantisani et al. 2012a, 2012b, 2016; Ding et al. 2011; Kagoya et al. 2010; Xing et al. 2011).

A cutoff >4.22 resulted in a sensitivity of 81.8%, specificity of 82.9% and accuracy of 88% (Ning et al. 2012), whereas 3.79 was the best cutoff point in another study, with a sensitivity of 97.8% and specificity of 85.7% (Xing et al. 2011). Cutoff points as low as 1.5 and as high as 5 have been suggested.

**Semi-quantitative approach using carotid artery pulsation.** Another semi-quantitative approach similar to the strain ratio exploits the pulsation of the adjacent carotid artery to provide the stress. It is known as the elasticity contrast index (Lim et al. 2012). The transducer is held still, with light contact on the skin over the thyroid. Patients are asked to hold their breath for a few seconds, and elastography data are acquired over 3–4 s. After elastography data are acquired, two ROIs are drawn, one close to the area of highest strain adjacent to the carotid artery in the surrounding muscle and another within the thyroid nodule at the region of lowest strain (highest stiffness), to calculate the thyroid stiffness index (Dighe et al. 2008). The elasticity contrast index is calculated by co-concurrence matrix comparison of benign and malignant features in a nodule using complex calculations (Lim et al. 2012).

**Display of strain.** The default display used by most manufacturers depicts stiff areas (little or no strain) in

---

**Table 1. Meta-analyses of strain elastography**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Methods</th>
<th>No. of studies</th>
<th>No. of nodules</th>
<th>No. of patients</th>
<th>Sens</th>
<th>95% CI</th>
<th>Spec</th>
<th>95% CI</th>
<th>AUROC</th>
<th>Positive likelihood ratio</th>
<th>Negative likelihood ratio</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ghajarzadeh et al. 2014</td>
<td>Elasticity score 4- or 5-point scoring scale</td>
<td>12</td>
<td>1180</td>
<td>86</td>
<td>81.9–89.4</td>
<td>66.7</td>
<td>63.4–69.9</td>
<td>3.82</td>
<td>0.16</td>
<td>27.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sun et al. 2014</td>
<td>Strain ratio</td>
<td>31</td>
<td>5481</td>
<td>4468</td>
<td>77–81</td>
<td>77</td>
<td>76–79</td>
<td>0.8941</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Razavi et al. 2013</td>
<td>Strain ratio</td>
<td>24</td>
<td>3,531</td>
<td>2,624</td>
<td>82</td>
<td>82</td>
<td>80</td>
<td>77–83</td>
<td>0.89</td>
<td>4.52</td>
<td>0.22</td>
<td></td>
</tr>
</tbody>
</table>

Sens = sensitivity; Spec = specificity; CI = confidence interval; AUROC = area under the receiver operating characteristic curve.

---

Fig. 6. Five-pattern scoring system for strain elastography. Score 1: the nodule is entirely soft (green); score 2: the nodule is mostly soft (green, with some blue areas); score 3: the nodule is only soft at the periphery (blue core, green periphery); score 4: the nodule is entirely rigid (blue); score 5: the nodule and the surrounding tissue is rigid (blue) (Itoh et al. 2006).

Fig. 7. Four-pattern scoring system for strain elastography of thyroid nodules. Score 1: the nodule is entirely green; score 2: the nodule is mostly green, with some blue areas; score 3: the nodule is mostly blue, with some green areas; score 4: the nodule is entirely blue (Rago and Vitti 2008; Rago et al. 2007).
contrast to the better inter-observer agreement for conventional ultrasound (Park et al. 2009b). They speculated that different external compression cycles and carotid artery pulsations caused the poor agreement in elastography. However, in that study, the machine did not have an indicator of the compressive force applied by the transducer. Newer machines have a quality display that gives real-time feedback to the operator about the stress applied to the tissue. Recent reports indicate substantial or almost perfect agreement within multiple operators, as outlined in Table 2.

Novices were taught how to perform elastography and were found to obtain results similar to those of an expert after scanning four to seven patients. It was found that the strain ratio is more easily learned than interpretation of the elasticity scores, which is more subjective and hence requires more expertise (Tatar et al. 2013).

Recommendation 6. The SR has lower inter-observer variability and is more easily learned than SE (Cantisani et al. 2015a, 2015b; Tatar et al. 2013). LoE: 1b, GoR: A; 100% consensus.

Intra-observer variability. No in vivo studies of intra-observer variability have been reported.

Recommendation 7. There are insufficient data to make a recommendation on intra-observer variability. LoE: 5, GoR: D; 87% consensus.

Influence of pathology of the nodule on elastographic appearance

Thyroid malignancy. Although most of the more common papillary carcinomas are stiff, it is well documented in multiple studies that follicular carcinomas may appear soft on elastography (Friedrich-Rust et al. 2010; Lippolis et al. 2011; Ning et al. 2012). A 44% false-negative result for follicular carcinoma on SE was reported (Oliver et al. 2011). Medullary, undifferentiated and metastatic carcinomas may also appear soft (Hong et al. 2009; Unlütürk et al. 2012) (Figs. 10 and 11).

Elastography is not useful in detecting malignancy in a low-risk population because of its low sensitivity and specificity when used alone (Vidal-Casariego et al. 2012). When partially cystic nodules are considered, the diagnostic value of SE is limited, as only the solid component should be assessed (Liu et al. 2014; Rago et al. 2007, 2010) and the cystic component may produce no data or artifacts (Oliver et al. 2011).

Recommendation 8. Tumors other than papillary carcinomas may be misleadingly soft (Hong et al. 2009; Unlütürk et al. 2012). LoE 1b, GoR: A; 100% consensus.

Diffuse thyroid diseases (thyroiditis). Elastography in thyroiditis was first reported in patients with de Quervain disease as increased thyroid stiffness (Ruchala et al. 2011) (Fig. 12). SE was not able to differentiate between subacute thyroiditis and malignancy as both appear stiff (Xie et al. 2011). A follow-up study reported that the stiffness in acute thyroiditis resolved to normal on follow-up, whereas the stiffness in subacute thyroiditis (and, to a lesser extent, in chronic thyroiditis) was higher than normal. Stiffness in subacute thyroiditis decreased at 4 and 10 weeks from starting treatment (Ruchala et al. 2012). Similar findings were reported in other studies (Menzilcioğlu et al. 2014; Yang et al. 2015). Cappelli
et al. (2015) found elastography to be of limited value in detecting thyroid malignancies in patients with Hashimoto’s disease, whereas Sahin et al. (2014) found that they needed a lower SR cutoff level in patients with nodules and Hashimoto’s disease, although the sensitivity and specificity were lower. The MNSR was used in patients with coexistent Hashimoto’s thyroiditis and was found to be more specific and accurate (Wang et al. 2015). However, Wang et al. used a higher cutoff value of 5.03 for the SR to differentiate between benign and malignant nodules.

** Recommendation 9. In patients without adjacent normal thyroid tissue, a surrounding muscle can be used for the strain ratio (Aydin et al. 2014). LoE: 2a, GoR: B; 100% consensus.**

Examination technique

**Depth penetration.** With SE imaging, as long as an adequate B-mode image can be obtained, a diagnostic elastogram should be possible. However, optimizing the displacement of deep tissue of interest may require a change in the technique with more stress applied. This may result in poor-quality elastograms in the near field.

**Focus zone.** The focus zone should be placed to optimize the B-mode image in the ROI as it is the changes in the B-mode data with applied stress that generate the SE image.

**Tissue to compare.** Semi-quantitative measurements can be obtained by comparing the stiffness of the nodule

---

**Table 2. Inter-observer variability**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Year published</th>
<th>Type</th>
<th>No. of nodules</th>
<th>Statistical method</th>
<th>Statistical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Park et al. 2009a, 2009b</td>
<td>2009</td>
<td>SE</td>
<td>52</td>
<td>Spearman correlation coefficient</td>
<td>0.08–0.22</td>
</tr>
<tr>
<td>Park et al. 2009a, 2009b</td>
<td>2009</td>
<td>SR</td>
<td>52</td>
<td>Spearman correlation coefficient</td>
<td>0.03–0.23</td>
</tr>
<tr>
<td>Merino et al. 2011</td>
<td>2011</td>
<td>SE</td>
<td>106</td>
<td>Cohen’s κ statistic</td>
<td>0.82</td>
</tr>
<tr>
<td>Ragazzoni et al. 2012</td>
<td>2012</td>
<td>SE</td>
<td>132</td>
<td>Cohen’s κ statistic</td>
<td>0.64</td>
</tr>
<tr>
<td>Kim et al. 2013</td>
<td>2012</td>
<td>SE</td>
<td>99</td>
<td>Cohen’s κ statistic</td>
<td>0.738</td>
</tr>
<tr>
<td>Calvete et al. 2014</td>
<td>2013</td>
<td>SE</td>
<td>89</td>
<td>Cohen’s kappa statistic</td>
<td>0.838</td>
</tr>
<tr>
<td>Cantisani et al. 2014a, 2014b</td>
<td>2014</td>
<td>SR</td>
<td>344</td>
<td>Cohen’s kappa statistic</td>
<td>0.95</td>
</tr>
<tr>
<td>Cantisani et al. 2015a, 2015b</td>
<td>2015</td>
<td>ECI index</td>
<td>154</td>
<td>Cohen’s kappa statistic</td>
<td>0.71–0.79</td>
</tr>
</tbody>
</table>

SE = strain elastography.
with that of normal thyroid or muscle (Cantisani et al. 2015a, 2015b).

**Image characteristics.** Each nodule should undergo at least two SE acquisitions. The patient should be placed supine with a pillow under the neck, as with routine thyroid ultrasound. The ultrasound probe is placed over the nodule with gel interposed. If manual compression is used, light manual vibration is applied with the transducer to create uniform compression over the thyroid nodule.

**Slice thickness.** The slice thickness may be a source of misleading results, especially in smaller lesions, because thicker slices cause averaging around small nodules (Hong et al. 2009). This is less of a problem with highly focused transducers.

**Region of interest size and content.** As SE displays relative elasticity within an ROI, it is advisable to set the ROI as large as possible, preferably covering the whole nodule together with adjacent normal thyroid tissue. One should avoid the inclusion of vessels, bones

---

Fig. 10. Anaplastic carcinoma. This large and heterogeneous mass is color coded in *blue* (for hard) and *green*, indicating that it is stiff. Histology on the resection specimen revealed an anaplastic carcinoma.

Fig. 11. Metastasis. A rare and aggressive skin malignancy, Merkel tumor, metastatic to the thyroid gland. Though fairly homogeneous on gray scale, it has extensive *blue* regions (coding for hard) on strain elastography.
Recommendation 10. The ROI for SE should be as large as possible, covering the nodule and some adjacent thyroid tissue (Dudea and Botar-Jid 2015). LoE: 2a, GoR: B; 86% consensus.

Nodule position. Isthmic nodules are difficult to assess, being compressed between two hard planes (transducer and trachea) and lacking reference tissue and, in some cases, being too superficial and thus affected by the near-field artifact (Wang et al. 2010). Deeply located nodules are subject to the stress decay phenomenon because of their distance from the transducer. Nodules located adjacent to the common carotid artery are the most susceptible to pulsation interference (Tranquart et al. 2008).

Recommendation 11. Nodules in the isthmus, close to the carotid artery, and those that are deeply located may be more difficult for SE (Wang et al. 2010). LoE: 1b, GoR: B 100% consensus.

Nodule size. Nodules larger than 3 cm or large enough to replace an entire lobe of the thyroid cannot be assessed with SE because of the lack of reference tissue and the depth of their deeper portions (Tranquart et al. 2008). Coalescent nodules are also unsuitable for SE (Wang et al. 2012).

Recommendation 12. Large nodules are difficult for SE as there may be no surrounding reference thyroid tissue (Tranquart et al. 2008). LoE: 1b, GoR: B; 100% consensus.

Nodule characteristics. Calcifications within a nodule are associated with increased stiffness, irrespective of the underlying pathology, and this can produce unreliable results (Fig. 13). Fibrosis inside longstanding benign nodules or associated with subacute or Hashimoto thyroiditis may also induce stiffness within nodules (Cantisani et al. 2014a, 2014b; Shuzhen 2012). Intranodular colloid alters the SE appearance, making the nodules unsuitable for elastography. For partially cystic nodules, only the solid component should be assessed.

Operator experience. Operators need to be trained and to have experience in performing thyroid elastography before interpreting or documenting the exams (Cantisani et al. 2014a, 2014b).


Strain quality indicator. A strain quality indicator is helpful to provide real-time feedback on the quality of compression. Various manufactures have different
recommendations; however, a quality factor of 3–4 (Hitachi) or 50 (Siemens) has been recommended (Calvete et al. 2013; Cappelli et al. 2012). A visual equivalent is sometimes provided (Fig. 8).

Patients should be asked to hold their breath for a short period. If carotid artery pulsation is used for compression, the probe is placed such that the image includes the nodule and carotid artery in transverse section, and an additional few minutes of elastographic acquisition is performed. The displayed elasticity scale should be adapted according to the stiffness of the tissue (Cantisani et al. 2014a, 2014b). For a very stiff nodule, the scale can be changed to increase the upper limit of display to allow differentiation between relatively stiff and soft areas; similarly, in a very soft nodule, the scale should be lowered to enable display of the relatively hard regions in a nodule.

Recommendation 14. A strain quality indicator is very helpful in obtaining good strain elastograms (Calvete et al. 2013; Cappelli et al. 2012) LoE: 1b, GoR: B; 100% consensus.

Practical points. Strain elastography can be performed as an extension to routine thyroid scanning and adds 3–5 min to the exam time; SE should always be interpreted alongside B-mode imaging because US has been found to have excellent sensitivity to detect suspicious nodules (Moon et al. 2007). However, the clinical utility of SE must be carefully understood. Suspicious US findings in patients should trump the SE information; that is, if the nodule is suspicious on US, FNA should be recommended. In case of the opposite findings, current literature does not support active intervention. Prospective studies are needed to understand if FNA can be avoided in patients with nodules that appear benign on SE. SE information may be particularly helpful in patients with non-diagnostic or indeterminate cytology to dictate the next steps (repeat biopsy or follow-up). Operator experience is essential for SE because significant false-positive results or inadequate samples can occur with lack of experience. Performance of at least five cases under guidance with the proper technique is essential (Tatar et al. 2013). In addition, the quality display should be used to optimize the compression–relaxation cycle. Static images and cine clips should be saved for retrospective review.

Limitations. Several factors can affect the results of elastography, including nodule characteristics (calcifications and cystic components), the experience of the operator and motion artifacts from respiration and carotid pulsations. Recognition of these pitfalls and limitations is important when evaluating thyroid nodules. Transverse scans are more susceptible to interference from carotid artery pulsations (Oliver et al. 2011; Park et al. 2009a, 2009b) and, therefore, less suitable for elastography with external compression. Longitudinal scans are less susceptible to carotid pulsations and also offer more thyroid reference tissue (Bhatia et al. 2011; Ning et al. 2012).

Recommendation 15. Longitudinal scans are useful for SR; however, if carotid pulsation is used, transverse scans should be employed (Bhatia et al. 2012; Hou et al. 2013). LoE: 1b, GoR: B; 100% consensus.

SHEAR WAVE ELASTOGRAPHY

Introduction

In SWE, ultrasound push pulses (ARFI, similar to color Doppler pulses) are sent into the tissue, and the minute motion produced by the acoustic radiation force they generate sets up transverse shear waves that travel away from the push pulse lines. The induced shear wave speed (SWS) is related to the stiffness of the tissue, stiffer tissue conducting shear waves faster. Through use of pulse-echo US, the ultrasound machine is able to measure the shear wave speed (in m/s) from which Young’s modulus can
be calculated (in kPa), making some assumptions regarding the tissue properties. Details have been published as Part I of the WFUMB guidelines on elastography (Shiina et al. 2015) and by EFSUMB (Bamber et al. 2013) and are not repeated here.

SWE can be performed using the point shear wave technique (pSWE), in which a small ROI of fixed size is placed at the location of the desired measurement. When pSWE is activated, the system provides a numerical result of the SWS of the tissue in the ROI box (in either m/s or kPa) (Fig. 14). In 2-D SWE, a larger ROI that can be controlled by the operator is placed, and when it is activated, a color-coded map of the SWS is displayed in the field of view. One or more measurement ROIs can then be placed in the field of view (Figs. 15 and 16). Two-dimensional SWE can be performed as a “one-shot” technique or as a “real-time” technique. With both pSWE and 2-D SWE techniques, it is important that there is no motion during acquisition. With 2-D SWE, the patient must remain still for several seconds to obtain a stable elastogram.

**Shear wave elastography: review of the literature**

A literature review carried out using the key words acoustic radiation force impulse technology and thyroid (32 results) and shear wave elastography and thyroid (69 articles, but including several on strain US elastography), and reported four meta-analyses.

The meta-analysis results are listed in Table 3. The four meta-analyses include more than 6,000 nodules; however, there is some overlap of the studies included. The meta-analyses include both pSWE and 2-D SWE, with the majority of papers evaluating pSWE. The sensitivities for SWE ranged from 0.80 to 0.86 (95% confidence interval [CI]: 0.73–0.92), whereas the specificities ranged from 0.84 to 0.90 (95% CI: 0.80–0.94). The area under the receiver operating characteristic curve (AUROC) values ranged from 0.91 to 0.94. Two studies (Dong et al. 2015; Liu et al. 2015a, 2015b, 2015c, 2015d) also reported positive likelihood ratios of 5.21 and 7.04 (95% CI: 3.56–11.4), negative likelihood ratios of 0.17 and 0.23 (95% CI: 0.10–0.32) and positive odds ratios of 46.7 and 27.5 (95% CI: 14.0–111.8). Significant heterogeneity was noted in both sensitivity and specificity (p < 0.001) (Zhan et al. 2015). All the meta-analyses concluded that SWE (pSWE and 2-D SWE) are useful complements to B-mode ultrasound in differentiating between benign and malignant nodules. It was concluded that SWE can be useful in selecting patients with thyroid nodules for surgery (Zhan et al. 2015).

Several articles not included in the meta-analyses are listed in Table 4. The sensitivities of these studies ranged from 0.35 to 1.00, with the majority between 0.68 and 0.95. The specificities ranged from 0.71 to 0.97.

The diagnostic performance of pSWE for differentiating benign from malignant thyroid nodules was evaluated, as was the change in diagnostic confidence compared with conventional US (Zhang et al. 2015).
One hundred seventy-four pathologically proven thyroid nodules (139 benign, 35 malignant) in 154 patients were included. US, Virtual Touch tissue imaging (VTI) and Virtual Touch quantification (VTQ) were performed and interpreted by two blinded readers with different experience who independently scored the likelihood of malignancy using a 5-point scale in three different image-reading sets (Zhang et al. 2014a, 2014b, 2014c). The specificity of both readers improved significantly after viewing the VTI/VTI and VTQ images (all \( p \), 0.05). After review of the results of both VTI and VTQ, the numbers of correctly diagnosed nodules increased for nodules \(<1.0\) cm for both readers and in both nodular goiters and papillary thyroid carcinomas for the junior reader (\( p \), 0.05). The confidence in characterizing a nodule as benign or malignant increased after review of VTI and VTQ images versus conventional US for the senior reader (\( p \), 0.05). The authors concluded that adding pSWE improved the specificity in diagnosing malignant thyroid nodules compared with conventional US on its own.

**Recommendation 16.** SWE (pSWE and 2-D SWE) has good sensitivity and specificity for identification of thyroid nodules (Dong et al. 2015). LoE: a, GoR: B; 100% consensus.

**Recommendation 17.** SWE (pSWE and 2-D SWE) may be useful in selecting patients with thyroid nodules for surgery (Zhan et al. 2015). LoE: 2a, GoR: B; 100% consensus.

**Recommendation 18.** Point SWE is useful in evaluating the stiffness of thyroid nodules and differentiating between malignant and benign nodules. SWS can be considered as a useful complement to conventional ultrasound (Dong et al. 2015). LoE: 2a, GoR: B; 100% consensus.

**Recommendation 19.** Compared with conventional US, pSWE has improved specificity in diagnosing malignant thyroid nodules, particularly nodules smaller than 1.0 cm. Point SWE increases the diagnostic confidence of the readers (Zhang et al. 2014a, 2014b, 2014c). LoE: 2a, GoR: B; 100% consensus.

**Inter-observer and intra-observer variability.** Three studies have evaluated the inter-observer variability of SWE. One study reported a concordance rate of \( k = 0.75 \) and an inter-class correlation of 0.97 (95% CI: 0.96–0.98) in interpretation of 51 lesions with pSWE (Grazhdani et al. 2014). The second study (Veyrieres et al. 2012) obtained an interclass correlation of 0.97 (95% CI: 0.96-0.98) in the evaluation of 102 nodules using 2-D SWE. The third study (Zhang et al. 2012) reported an intra-observer variability of 0.90 and inter-observer variability of 0.86 in interpretation of 80 lesions with pSWE. Regarding intra-observer
variability, a low value of 0.35 was reported when the means and standard deviations of 10 measurements at one location were considered, whereas a better value, 0.82, was obtained when all measurements were included. Friedrich-Rust et al. (2012) concluded that this variability did not affect the differentiation between benign and malignant nodules (Friedrich-Rust et al. 2012).

**Recommendation 20.** The inter-observer and intra-observer repeatability of both pSWE and 2-D SWE are high (Friedrich-Rust et al. 2012; Grazhdani et al. 2014; Veyrieres et al. 2012; Zhang et al. 2012). LoE: 2a, GoR: B; 100% consensus.

**Interpretation of the results of SWE. Normal thyroid.** Three studies report the stiffness values of the normal thyroid gland (Friedrich-Rust et al. 2012; Fukuhara et al. 2015a, 2015b; Veyrieres et al. 2012). These range from 1.60 ± 0.18 m/s for pSWE to 2.6 ± 1.8 m/s for 2-D SWE.

**Nodules.** Table 3 lists studies performed with both pSWE and 2-D SWE with suggested cutoff values to distinguish benign from malignant lesions. Studies using SWE to differentiate benign from malignant thyroid nodules reported cutoff values ranging from 3.65 to 4.70 m/s (34.5–66 kPa). Increasing the cutoff value from 1.8 to 6.7 m/s (10.3 kPa to 132 kPa) increased the specificity from 8.9% to 100% (Bhatia et al. 2012).

Most studies have evaluated SWE for the differentiation of thyroid nodules in a general population with promising results (Liu et al. 2015a, 2015b, 2015c, 2015d; Park et al. 2015; Sebag et al. 2010). In one, a mean stiffness value of >85 kPa or a maximum value >94 kPa was an independent predictor of malignancy (Park et al. 2015). One study (Zhang et al. 2012) reported SWSs for benign and malignant thyroid nodules of 2.34 ± 1.17 m/s (range: 0.61–9.00 m/s) and 4.82 ± 2.53 m/s (range: 2.32–9.00 m/s), respectively ($p < 0.001$). These results were similar to those of other studies in which cutoff values in the range 2.55–2.75 m/s were reported (Bojunga et al. 2012; Friedrich-Rust et al. 2012; Gu et al. 2012; Han et al. 2015; Hou et al. 2013). In another study (Zhang et al. 2012), the SWS was found to be significantly higher in malignant than benign lesions, with a value higher than 2.87 m/s strongly suggestive of malignancy (Cantisani et al. 2013). In their meta-analysis, Zhang et al. found that a total of 15 studies reported the SWS cutoff to be in the range 2–3 m/s, which they called the “gray zone.” Furthermore, 11 studies reported the best SWS cutoff as ≥2.5 m/s. One possible solution would be to perform FNA in all patients with a range 3 m/s vs. ≥ 2.5 m/s.

**Table 3.** Results of meta-analyses of SWE for characterization of thyroid nodules.

<table>
<thead>
<tr>
<th>Reference</th>
<th>No. of studies</th>
<th>No. of patients</th>
<th>Sens. 95% CI</th>
<th>Spec. 95% CI</th>
<th>AUROC</th>
<th>PPV</th>
<th>NPV</th>
<th>Positive likelihood ratio</th>
<th>Negative likelihood ratio</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhan et al. 2015</td>
<td>16</td>
<td>2456</td>
<td>0.80 (0.73–0.87)</td>
<td>0.85 (0.80–0.90)</td>
<td>0.91</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94 (0.90–0.98)</td>
<td>0.87 (0.83–0.91)</td>
<td>0.94</td>
</tr>
<tr>
<td>Dong et al. 2015</td>
<td>13</td>
<td>1451</td>
<td>0.86 (0.80–0.92)</td>
<td>0.90 (0.83–0.94)</td>
<td>0.91</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94 (0.90–0.98)</td>
<td>0.87 (0.83–0.91)</td>
<td>0.94</td>
</tr>
<tr>
<td>Liu et al. 2015a, 2015b, 2015c, 2015d</td>
<td>13</td>
<td>1641</td>
<td>0.81 (0.77–0.86)</td>
<td>0.84 (0.81–0.87)</td>
<td>0.91</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94 (0.90–0.98)</td>
<td>0.87 (0.83–0.91)</td>
<td>0.94</td>
</tr>
<tr>
<td>Lin et al. 2014</td>
<td>15</td>
<td>1525</td>
<td>0.84 (0.77–0.90)</td>
<td>0.88 (0.84–0.92)</td>
<td>0.93</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95 (0.92–0.98)</td>
<td>0.89 (0.85–0.93)</td>
<td>0.95</td>
</tr>
</tbody>
</table>

SWS = shear wave elastography; pSWE = point shear wave elastography; Sens = sensitivity; Spec = specificity; CI = confidence interval; PPV = positive predictive value; NPV = negative predictive value; AUROC = area under the receiver operating characteristic curve.
In a study that included 146 nodules in 93 patients (Sebag et al. 2010) a scoring system for US was compared with SWS features. The sensitivity and the specificity for malignancy were 51.9% and 97.0% for gray-scale US and 81.5% and 97.0% for the combination of gray-scale US and SWE, respectively.

Normal thyroid tissue has a follicular structure with a low cell density; chronic autoimmune thyroiditis (CAT) has a diffuse fibrotic structure; benign thyroid nodules have a high cell density; and papillary thyroid carcinoma (PTC) has a mixed pathologic structure with cells, fibrosis and adipose tissue. Five hundred ninety-nine thyroid tissue samples were divided into four groups based on their pathologic structure: 254 normal controls, 128 cases of CAT with diffuse fibrosis, 165 benign nodules that had high cell density and 52 PTCs that exhibited high cell density and fibrosis (Fukuhara et al. 2015a, 2015b). The mean SWS in each group was 1.60 ± 0.18 m/s in normal thyroid, 2.55 ± 0.28 m/s in CAT, 1.72 ± 0.31 m/s in benign nodules and 2.66 ± 0.95 m/s in PTCs. The SWSs of CAT and PTC were significantly higher than those of normal thyroid (p < 0.001). SWS was significantly increased by fibrosis.

**Recommendation 21.** Cutoff values for discriminating benign from malignant nodules vary from 2.4 to 4.7 m/s (Bhatia et al. 2012). LoE: 2a, GoR: B; 100% consensus.

**Size.** In two studies, the correlations of SWE indices with nodule size were inconsistent (Hou et al. 2013). Another reported that the AUROC of SWE in nodules ≤10 mm was significantly lower than that in nodules >20 mm (Zhang et al. 2012). Sensitivity analysis indicated that the performance of SWE was not improved by excluding three studies with nodules of mean diameter <10 mm. Limited information prevented a conclusion as to whether these factors significantly influence the performance of SWE. The effects of the elasticity of normal thyroid tissue on the diagnostic performance of ultrasound elastography should be evaluated in future studies.

One study reported that the performance of pSWE quantification is better for nodules >20 mm (Zhang et al. 2014a, 2014b, 2014c). Measurements of SWS in nodules <20 mm in diameter are not stable (Fukuhara et al. 2014). The variability of the five measurements in PTC (>0.29 m/s) was larger than those in the other groups, suggesting that fibrosis and pathologic heterogeneity are significant factors affecting the measurement of SWS in the thyroid. In cases of markedly heterogeneous histopathology within a given ROI, the SWS varies between the region’s different tissues. The velocity cannot be calculated when the SWS is not constant within the ROI, as is the case in PTC, which is composed of solid cells, fibrosis, adipose tissue and other components (Fukuhara et al. 2015a, 2015b).

**Recommendation 22.** The influence of nodule size on SWS is uncertain (Hou et al. 2013; Zhang et al. 2012). LoE: 5, GoR: D; 100% consensus.

**Calcifications and cysts.** Sometimes no SWS can be measured, and the display reads “X.XX” or “0.00” because these nodules contain cysts and calcifications.

---

**Table 4. Results of peer-reviewed articles evaluating SWE of thyroid nodules**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Method</th>
<th>No. of nodules</th>
<th>No. of patients</th>
<th>Malignant proportion</th>
<th>Cutoff m/s (kPa)</th>
<th>Sens</th>
<th>Spec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang et al. 2015</td>
<td>pSWE + ARFI SE</td>
<td>174</td>
<td>154</td>
<td>35/174</td>
<td>2.66 (21)</td>
<td>1.00</td>
<td>0.82</td>
</tr>
<tr>
<td>Hamidi et al. 2015</td>
<td>pSWE</td>
<td>95</td>
<td>95</td>
<td>52/599</td>
<td>2.66 (21)</td>
<td>1.00</td>
<td>0.82</td>
</tr>
<tr>
<td>Fukuhara et al. 2015a,</td>
<td>pSWE</td>
<td>599</td>
<td>599</td>
<td>52/599</td>
<td>2.66 (21)</td>
<td>1.00</td>
<td>0.82</td>
</tr>
<tr>
<td>Fukuhara et al. 2015b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sebag et al. 2010</td>
<td>2-D-SWE</td>
<td>146</td>
<td>93</td>
<td>4.65 (65)</td>
<td>0.82</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>Bhatia et al. 2012</td>
<td>2-D-SWE</td>
<td>81</td>
<td>74</td>
<td>3.45 (34.5)</td>
<td>0.77</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>Veyrieres et al. 2012</td>
<td>2-D-SWE</td>
<td>151</td>
<td>148</td>
<td>4.70 (66)</td>
<td>0.80</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>Slapa et al. 2012</td>
<td>2-D-SWE</td>
<td>6</td>
<td>4</td>
<td>4.65 (65)</td>
<td>0.80</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>Liu et al. 2015a, 2015b</td>
<td>2-D-SWE</td>
<td>331</td>
<td>271</td>
<td>3.65 (39.3)</td>
<td>0.66</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>Liu et al. 2015c, 2015d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Park et al. 2015</td>
<td>2-D-SWE</td>
<td>476</td>
<td>453</td>
<td>Mean: 5.3 (85)</td>
<td>0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang and Han 2013</td>
<td>pSWE</td>
<td>155</td>
<td>155</td>
<td>62/155</td>
<td>2.84</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>Hou et al. 2013</td>
<td>pSWE</td>
<td>173</td>
<td>142</td>
<td>2.87</td>
<td>0.66</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>Zhang et al. 2012</td>
<td>pSWE</td>
<td>158</td>
<td>138</td>
<td>21/158</td>
<td>2.57</td>
<td>0.35</td>
<td>0.79</td>
</tr>
<tr>
<td>Bojunga et al. 2012</td>
<td>pSWE</td>
<td>98</td>
<td>72</td>
<td>22/98</td>
<td>2.56</td>
<td>0.86</td>
<td>0.93</td>
</tr>
<tr>
<td>Gu et al. 2012</td>
<td>pSWE</td>
<td>118</td>
<td>140</td>
<td>2.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Han et al. 2015</td>
<td>pSWE</td>
<td>60</td>
<td>55</td>
<td>3/60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friedrich-Rust et al. 2012</td>
<td>pSWE</td>
<td>183</td>
<td>159</td>
<td>66/183</td>
<td>0.68</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Xu et al. 2014</td>
<td>pSWE + ARFI SE</td>
<td>64</td>
<td>49</td>
<td>19/64</td>
<td>3.60 (38.3)</td>
<td>0.68</td>
<td>0.87</td>
</tr>
<tr>
<td>Liu et al. 2014</td>
<td>2-D-SWE</td>
<td>173</td>
<td>142</td>
<td>2.87</td>
<td>0.66</td>
<td>0.67</td>
<td></td>
</tr>
</tbody>
</table>

SWE = shear wave elastography; pSWE = point shear wave elastography; ARFI = acoustic radiation force impulse; SE = strain elastography; Sens = sensitivity; Spec = specificity.

WFUMB guidelines on US elastography of thyroid ● D. COSGROVE et al. 19
This phenomenon is likely to be influenced by the size of the ROI. Pathologic imaging of the PTC samples indicated that 18 of 21 unmeasurable samples had heterogeneous pathology, and 9 had coarse calcification: One or both of these characteristics were observed in 20 of the 21 (95%) cases. On the other hand, 25 of 29 measurable samples had homogeneous pathology, and only one had coarse calcification. These results corroborate the above hypothesis. One study reported no difference in SWE indices between completely solid lesions and the solid component of partially cystic lesions or between calcified and non-calcified benign nodules (Bhatia et al. 2012). In contrast, another study reported that the rate of false classification was significantly higher in calcified than non-calcified nodules (Vorlander et al. 2010).

**Recommendation 23.** SWS may be difficult to measure in heterogeneous nodules (Bhatia et al. 2011; Vorlander et al. 2010). LoE: 2b, GoR: B; 100% consensus.

**Nodules in autoimmune disease.** Two studies found that SWE can differentiate thyroid nodules, even in the presence of autoimmune thyroiditis. ARFI elastography, including VTI (ARFI SE) and VTQ (pSWE) (Han et al. 2015; Liu et al. 2015a, 2015b, 2015c, 2015d), can be performed for the differential diagnosis of malignant from benign thyroid nodules independently of the coexistence of CAT, with promising diagnostic accuracy. Additionally, the diagnostic performance of ARFI was better than that of SE (Liu et al. 2015a, 2015b, 2015c, 2015d). However, the optimal cutoff values of the three SWE indices were higher in the Hashimoto’s thyroiditis (HT) group than in a general population. The diagnostic performance of SWE was not satisfactory. The majority of the microcarcinomas were missed by SWE (Liu et al. 2014).

Conventional ultrasound and SWE were performed in 243 patients with 286 thyroid nodules with histologic proof (Fig. 17) (Liu et al. 2015a, 2015b, 2015c, 2015d). The HT group consisted of 93 patients with 117 nodules. The non-HT group consisted of 140 patients with 169 nodules. In the benign and malignant nodules, there were no significant differences in SWE values between the HT and non-HT groups. However, SWE values of extra-nodular thyroid parenchyma were significantly higher in the HT group. In the HT group, the maximum SWE value had the highest AUROC (0.82; 95% CI: 0.74–0.90), and there were no significant differences when compared with other SWE features. In the multivariate analysis, hypo-echogenicity (odds ratio = 9.86, p = 0.002), microcalcification (odds ratio = 3.98, p = 0.046) and maximum SWE value (odds ratio = 40.71, p = 0.001) were independent predictors of thyroid malignancy.

**Influence of pathology of the nodule on elastographic appearance.** Thyroid carcinoma. In distinguishing benign from malignant thyroid nodules, we need to recognize that not all malignant nodules are stiff: Some are soft or heterogeneous. Follicular carcinomas, in particular, can be soft and difficult to distinguish from benign nodules, although some good results have been obtained with SWE (Samil et al. 2015). A cutoff value of 22.30 kPa helps differentiate malignant from benign follicular thyroid lesions with a sensitivity of 82%, specificity of 88% and positive and negative predictive values of 75% and 91%, respectively (Samil et al. 2015). To date, only a few articles on medullary thyroid tumors have been published, and these have involved only SE (Andrioli and Persani 2014).

**Recommendation 24.** SWS readings may be low in non-papillary thyroid carcinomas (Samil et al. 2015). LoE: 1b, GoR: B; 100% consensus.

**Diffuse thyroid diseases.** Chronic autoimmune thyroiditis (Hashimoto’s thyroiditis), Basedow-Graves’ disease and multinodular goiter are usually diagnosed on the basis of clinical and laboratory findings, supported by ultrasound. Their typical fibrosis and inflammation increase the stiffness of the gland, and although differences in the extent of inflammation and scarring may affect the precise values, the general tendency is that of a diffusely stiff gland. A group in two studies reported a statistically significant differences in ARFI stiffness between normal subjects and patients with autoimmune pathology (Graves’ disease).

![Fig. 17. Thyroiditis. This very stiff nodule in a patient with thyroiditis reveals the extreme values associated with the intense fibrosis that is typical. The shear wave elastography scale is set to a maximum of 100 kPa, and most of the nodule exceeds this value.](image-url)
Sue was greater in the HT than in the non-HT group (Magri et al. 2012). The stiffness of the extranodular tissue was greater in the HT than in the non-HT group (24.0 ± 10.5 kPa vs. 20.8 ± 10.4 kPa), but this difference was not statistically significant (p = 0.21). In subacute (granulomatous, de Quervain) thyroiditis, two studies found the inflammatory areas to be stiff (Menzilcioglu et al. 2014) (Fig. 17). Any focal inflammatory areas should be included in the differential diagnosis of carcinoma because they are stiffer than those encountered in Hashimoto thyroiditis (Andrioli and Persani 2014; Dudea and Botar-Jid 2015).

SWE in 75 patients was used with benign thyroid nodules on cytology, 33 with Hashimoto’s thyroiditis (HT group) and 42 with uni- or multinodular goiters (Magri et al. 2012). The stiffness of the extranodular tissue was greater in the HT than in the non-HT group (24.0 ± 10.5 kPa vs. 20.8 ± 10.4 kPa), but this difference was not statistically significant (p = 0.21). In subacute (granulomatous, de Quervain) thyroiditis, two studies found the inflammatory areas to be stiff (Menzilcioglu et al. 2014) (Fig. 17). Any focal inflammatory areas should be included in the differential diagnosis of carcinoma because they are stiffer than those encountered in Hashimoto thyroiditis (Andrioli and Persani 2014; Dudea and Botar-Jid 2015).

SWE was evaluated for diagnosing CAT (Fukuhara et al. 2015a, 2015b) in 229 patients with 253 normal thyroid lobes (controls) and 150 CAT lobes. The SWS in CAT (2.47 ± 0.57 m/s) was significantly higher than that in controls (1.59 ± 0.41 m/s) (p < 0.001). The AUROC for CAT was 0.899, and the SWS cutoff value was 1.96 m/s. The sensitivity, specificity and diagnostic accuracy were 87.4%, 78.7% and 85.1%, respectively. Anti-thyroglobulin antibody levels and thyroid isthmus thickness were correlated with tissue stiffness in CAT. However, there was no correlation between levels of anti-thyroglobulin antibodies and tissue stiffness. They concluded that quantitative SWE is useful for diagnosing CAT, and it is possible that SWE can be used to evaluate the degree of fibrosis in patients with CAT.

Riedel’s (chronic) thyroiditis is characterized by extremely stiff parenchyma, with values of 143–281 kPa (Dudea and Botar-Jid 2015) obtained using ARFI quantification (Carneiro-Pla 2013; Kim et al. 2014a, 2014b). Significantly higher values were obtained in autoimmune pathology (CAT and Graves’ disease): 2.82 ± 0.47 m/s for Graves’ disease compared with 2.68 ± 0.50 m/s for CAT (Sporea et al. 2011, 2012; Vlad et al. 2015). A cutoff value >2.53 m/s was reported for differentiation between normal thyroid and diffuse thyroid diseases, with a sensitivity and positive predictive value >90% (Kim et al. 2014a, 2014b).

In Hashimoto thyroiditis, the stiffness of extra nodular tissue increases with the thyroid antibody titer and degree of thyroid function damage (Magri et al. 2012). A mean cutoff value for the elasticity index, in cases of diffuse thyroid disease, of 3.0 m/s (27.6 kPa) and a maximum value of 3.7 m/s (41.3 kPa), were obtained with a sensitivity of 40.9% and a specificity of 82.9% (Kim et al. 2014a, 2014b). These areas should be included in the differential diagnosis of carcinoma. They are harder than those encountered in chronic thyroiditis (American Thyroid Association Guidelines Taskforce on Thyroid et al. 2009).

**Recommendation 25. SWS is increased in many diffuse thyroid diseases** (Carneiro-Pla 2013, Dudea and Botar-Jid 2015, Kim et al. 2014a, 2014b). **LoE: 2b, GoR: B; 100% consensus.**

**Point and 2-D shear wave elastography**

**Technical factors in SWE.** The intensity of the elastographic signal determines the quality of the shear wave speed measurement. If the signal is too weak, no signal will be displayed. The gain should be increased until noise limits the image quality. The operator should be aware of any pressure exerted by the transducer: Excess pressure produces superficial hardening artifacts and increases the stiffness. The size of the ROI and the standard deviation of the measured values are inversely related.

The color (or, optionally, gray-scale) overlay indicates the SWS (or the converted kPa values), and ROIs can be placed on the stiffest part of a lesion and on adjacent thyroid or muscle to obtain quantitative readouts as well as ratios. Liberal use of ultrasound coupling gel can help (Grazhdani et al. 2014). The ROI of pSWE is not modifiable and is available in two sizes: 5 × 6 mm and 20 × 20 mm. For this reason, in small nodules, the ROI includes adjacent normal thyroid, and this may lower the measurements. As in other elastography techniques, cystic regions and calcifications may yield false results (Grazhdani et al. 2014). However, peculiar to ARFI, the use of the standard-size ROI sometimes makes it impossible to exclude fluid or calcified portions of a nodule. In addition, there are technical limitations to VTQ measurements. Whether thyroid nodules smaller than the VTQ measurement box should be excluded remains controversial. Shear wave speed up to 9 m/s can be measured; higher speeds are displayed as “X.XX m/s” (Bojunga et al. 2012; Zhang et al. 2012). The ARFI penetration depth is limited to about 5.5 cm, so large thyroids or very large and deeply located nodules cannot be properly assessed using ARFI quantification (Goertz et al. 2011).

**Recommendation 26. Limitations of pSWE elastography include very large and small nodules and calcifications and cystic regions that cannot be**
excluded (Bojunga et al. 2012; Grazhdani et al. 2014; Zhang et al. 2012). LoE: 1b, GoR: A; 100% consensus.

**Examination technique. Depth penetration.** As a general rule, both pSWE and 2-D SWE, when performed with a linear high-frequency transducer that would be used for thyroid imaging, obtain accurate measurements up to 4–5 cm in depth. Deeper than this, the ARFI pulses are attenuated, and the tissue displacement is too small to track the shear waves accurately.

**Focal zone.** For all vendors, the SWE focal zone cannot be changed by the operator. The system automatically selects the correct focus zone for the ROI (pSWE) or field of view (2-D SWE) that is placed on the image for SWS measurement.

**Tissue to compare.** Most studies using SWE (pSWE or 2-D SWE) usually use a quantitative stiffness value of the maximum or mean stiffness value of the nodule for characterization. A strain ratio can be calculated using either normal thyroid tissue or muscle as the reference, similar to SE. However, it is important to recognize that stiffness ratios obtained using SWE are not the same as those in elasticity index, and stiffness ratios using m/s differ from those calculated with kPa.

**Quality indicator.** The propagation mode is a quality control map that displays the shear wave propagation as lines in the selected window; it guides the optimum position of the ROI very precisely (Fig. 16). Other vendors also display a quality map.

**Size of the ROI.** When pSWE is used, a fixed ROI is placed inside the nodule, taking care to avoid cystic or calcified areas, while patients hold their breath and the operator exerts only minimal pressure with the probe (Bojunga et al. 2012; Dudea and Botar-Jid 2015; Friedrich-Rust et al. 2010). The software displays the shear wave speed together with the depth of the ROI.

**Recommendation 27.** A standard ROI of $5 \times 6$ mm should be placed inside the nodule, taking care to avoid cystic or calcified areas (Bojunga et al. 2012; Dudea and Botar-Jid 2015). LoE: 1b, GoR: A; 100% consensus.

**How many measurements?** In the experience of some researchers, 10 measurements are needed to obtain a median that yields reliable accuracy (Bojunga et al. 2012, Dudea and Botar-Jid 2015). Another research group obtained good results with five measurements (Sporea et al. 2012).

**Recommendation 28.** Five to ten measurements are needed to obtain a median value yielding reliable accuracy (Bojunga et al. 2012; Dudea and Botar-Jid 2015; Sporea et al. 2012). LoE: 1b, GoR: A; 100% consensus.

**Limitations.** In SWE, as in other elastography techniques, pressure applied by the probe increases the tissue’s stiffness (Bhatia et al. 2012; Cantisani et al. 2013; Lyshchik et al. 2005). Therefore, experience is required to perform reliable examinations (Shao et al. 2015). Nodules in the isthmus are difficult because they are trapped between the skin and the stiff trachea.

**COMPARISON OF THE METHODS**

Ultrasound elastography, using either SE or SWE, is a valid and useful additional tool to gray-scale and color Doppler US in thyroid evaluation, as evidenced by the literature and the EFSUMB guidelines. However, to achieve reliable SE, adequate training, suitable cutoff values for both strain and SWE, adequate equipment and clinically appropriate examinations are necessary. Our suggestions are to minimize pre-compression, check the ROI size and positions, avoid areas with artifacts or with gross calcifications or cystic areas and instruct patients to cooperate properly. Future technical developments to reduce inter-observer and intra-observer variability will be helpful.

SWE appears to have the advantage of being less operator dependent; in addition, the learning curve for strain elastography seems to be short (Tatar et al. 2013) and has been improved significantly by the availability of real-time operator feedback. Some authors claim that SWS is an operator-independent and reproducible technique (Hegedus 2010; Sebag et al. 2010). However, only a small number of studies have been published; future multicenter studies are needed.

Only a few articles have been published comparing the value of SE and SWE. SWE and qualitative five-grade, color-coded SE were compared to discriminate 64 focal thyroid nodules in 49 patients with surgical pathology (Liu et al. 2014). Of the 64 nodules, 19 were papillary thyroid carcinomas and 45 were benign. Using the most accurate cutoff of a mean $\geq 38.3$ kPa to predict malignancy, the diagnostic specificity, sensitivity, accuracy, positive predictive value and negative predictive value of SWE and SE were 68.4% versus 79.0%, 86.7% versus 84.4%, 81.3% versus 78.1%, 68.4% versus 64.7% and 86.7% versus 83.3%, respectively. Liu et al. concluded that SWE is a promising tool for differentiating thyroid nodules with results comparable to those for real-time elastography, having slightly lower sensitivity and slightly higher specificity. ARFI and qualitative SE of 158 nodules were compared in 138 patients (Bojunga et al. 2012). There was no significant difference in diagnostic accuracy for the diagnosis of malignant thyroid nodules between SE and ARFI imaging (0.74 vs. 0.69, $p = 0.54$), and the combination of RTE with ARFI did not improve diagnostic accuracy.


Tatar IG, Kurt A, Yilmaz KB, Akinci M, Kalacoglu H, Hekimoglu B. The learning curve of real time elastosonography: A preliminary


Zhang FJ, Han RL, Zhao XM. The value of Virtual Touch tissue image (VTI) and Virtual Touch tissue quantification (VTQ) in the differential diagnosis of thyroid nodules. Eur J Radiol 2014a;83: 2033–2040.


