Thyroid Ultrasound Features and Risk of Carcinoma: A Systematic Review and Meta-Analysis of Observational Studies

Luciana Reck Remonti,1 Caroline Kaercher Kramer,2 Cristiane Bauermann Leitão,3 Lana Catani F. Pinto,4 and Jorge Luiz Gross3

Background: Thyroid nodules are a common finding in the general population, and their detection is increasing with the widespread use of ultrasound (US). Thyroid cancer is found in 5–15% of cases depending on sex, age, and exposure to other risk factors. Some US parameters have been associated with increased risk of malignancy. However, no characteristic seems sufficiently reliable in isolation to diagnose malignancy. The objective of this meta-analysis was to evaluate the diagnostic performance of US features for thyroid malignancy in patients with unslected thyroid nodules and nodules with indeterminate fine-needle aspiration (FNA) cytology.

Methods: Electronic databases were reviewed for studies published prior to July 2012 that evaluated US features of thyroid nodules and reported postoperative histopathologic diagnosis. A manual search of references of review and key articles, and previous meta-analyses was also performed. A separate meta-analysis was performed including only nodules with indeterminate cytology. Analyzed features were solid structure, hypoechogenicity, irregular margins, absence of halo, microcalcifications, central vascularization, solitary nodule, heterogeneity, taller than wide shape, and absence of elasticity.

Results: Fifty-two observational studies (12,786 nodules) were included. Nine studies included nodules with indeterminate cytology as a separate category, comprising 1851 nodules. In unselected nodules, all US features were significantly associated with malignancy with an odds ratio varying from 1.78 to 35.7, and microcalcifications, irregular margins, and a taller than wide shape had high specificities (Sp; 87.8%, 83.1%, 96.6%) and positive likelihood ratios (LHR; 3.26, 2.99, 8.07). Absence of elasticity was the single feature with the best diagnostic performance (sensitivity 87.9%, Sp 86.2%, and positive LHR 6.39). The presence of central vascularization was the most specific US feature in nodules with indeterminate cytology (Sp 96% and positive LHR 2.13).

Conclusions: US features in isolation do not provide reliable information to select nodules that should have a FNA performed. A combination of US characteristics with higher likelihood ratios and consequently with higher post-test probabilities of malignancy—microcalcifications, or a taller than wide shape, or irregular margins, or absence of elasticity—will probably identify nodules with an increased risk for malignancy. Further studies are required to standardize elastography techniques and evaluate outcomes, especially in nodules with an indeterminate cytology.

Introduction

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 nodularity varies from 19% to 67%, and increases with age, affecting about 50% of the population older than 40 years of age (1–5). The clinical significance of thyroid nodules relates to the need to exclude thyroid cancer, which is found in 5–15% of cases, depending on sex, age, and exposure to other risk factors (5–8). The incidence of thyroid cancer has increased about fivefold in the last 50 years, mostly due to small papillary thyroid cancers, the most indolent form of thyroid cancer (9).

Some US parameters, such as microcalcifications, hypoechogenicity, absence of a halo, increased intranodular vascularity, nodule shape or irregular margins, have been traditionally associated with increased risk of malignancies.
(10). However, none of these characteristics seems sufficiently reliable in isolation to diagnose malignancy. Diagnostic sensitivity ranges from 26.5% to 87.1% for hypochoicogenicity, 54.3% to 74.3% for intranodular vascularity, and 26.1% to 59.1% for microcalcifications, whereas specificity ranges from 43.4% to 94.3%, 78.6% to 80.8%, and 85.8% to 95%, respectively (4,10,11). More recently, US determination of tissue elasticity (elastography) has been suggested to detect malignancy in thyroid nodules. A meta-analysis found a sensitivity of 92% and specificity of 90% using this technique. However, only a few studies were included, and only three used histopathology of surgical specimens for final diagnosis (12). Fine needle aspiration (FNA) biopsy is considered the most accurate procedure to identify malignant nodules. To implement biopsies in all patients harboring a thyroid nodule is too burdensome, and the results of FNA have some limitations. The indications are broad and vague, and usually include patients with a family history of thyroid cancer, or those who have had significant radiation exposure, or those who have a combination of suspicious US features (5,10). However, there is no information about the probability of the US features associated with malignancy and which combination would be more clinically useful. US features may be also useful in clinical decision making for patients with FNA specimens insufficient for diagnosis (10%) or where specimens are indeterminate (15–30%), the latter carrying a 20–30% risk of malignancy (4,5). A recent meta-analysis evaluating the accuracy of US to predict malignancy in thyroid nodules found sensitivities ranging from 26% to 87%, and specificities from 40% to 93%. In this study, a taller than wide shape showed the highest diagnostic odds ratio (OR) for cancer. However, that meta-analysis included studies that used cytology, instead of histology, as a final diagnosis for benign nodules. Besides, it did not evaluate the accuracy of elastography to predict malignancy (13). Moreover, there was no description of the probability—likelihood ratio—of US characteristics associated with malignancy. The likelihood ratio would provide more information to be used in the clinical decision making of thyroid nodules than just sensitivity and specificity (14).

The aim of this study was to conduct a systematic review and meta-analysis of observational studies evaluating the diagnostic performance of US features considered to be associated with thyroid malignancy in patients with unselected thyroid nodules or nodules with indeterminate FNA cytology, considering only histopathologic diagnosis of surgical specimens as the final diagnosis.

Material and Methods

Search strategy

MEDLINE was searched using the following medical subject heading terms: “Thyroid Nodule”[MeSH] AND (“Ultrasoundography”[MeSH] OR “ultrasoundography”[Subheading] OR “Ultrasoundography, Doppler”[MeSH]). EMBASE was searched using EmTree terms “Thyroid nodule” and “Ultrasound.” The search period ended in July 2012. A manual search of the references of review articles, previous meta-analyses, and key articles was also performed. All potentially eligible studies were considered for review regardless of the primary outcome or language.

Study selection

Observational studies of patients with thyroid nodules evaluated by US and submitted to thyroidectomy regardless of the reason for surgery were considered for inclusion. Only studies with histopathologic diagnosis of surgical specimens were considered. Two independent investigators (L.R.R. and C.K.K.) selected potentially eligible studies based on titles and abstracts. All the studies selected were retrieved for full-text evaluation. Disagreements were solved by a third investigator (C.B.L.).

Data extraction and quality assessment

Two investigators reviewed the selected studies for patient characteristics, US features, and histopathologic results. Any discrepancies between the data extracted were discussed until a consensus was reached. The absolute number of patients with and without the evaluated features and with and without malignancy was extracted. These data were entered into a computerized spreadsheet considering true positives, true negatives, false positives, and false negatives.

The diagnostic ability to diagnose thyroid malignancy of the following US features was evaluated: solid structure, hypochoicogenicity, irregular margins, absence of halo, microcalcifications, central vascularization, solitary nodule, heterogeneity, taller than wide shape, and absence of elasticity. The presence of these features was defined as described in the original study.

Two independent investigators (L.R.R. and L.C.F.P.) evaluated the quality of the included studies using the QUADAS-2 tool (15). Any disagreements were solved by a third investigator (CBL). The present meta-analysis was described according to proposed by Stroup et al. (16). Details are available in Supplementary Table S1 (Supplementary Data are available online at www.liebertpub.com/thy).
Table 1. Characteristics of Studies Included in Meta-Analysis

<table>
<thead>
<tr>
<th>Author (reference)</th>
<th>Year</th>
<th>Number of nodules</th>
<th>Male sex (%)</th>
<th>Age (mean)</th>
<th>Clinical background</th>
<th>US features evaluated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walker J (65)</td>
<td>1985</td>
<td>94</td>
<td>14.8</td>
<td>All patients submitted to surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggarwal S (17)</td>
<td>1989</td>
<td>36</td>
<td>Patients with cold nodules</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoX M (29)</td>
<td>1991</td>
<td>68</td>
<td>10 48.6</td>
<td>Patients submitted to surgery with no compressive symptoms or obvious malignancy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hübsch V (36)</td>
<td>1992</td>
<td>65</td>
<td>30 42</td>
<td>Patients with cold nodules, compressive symptoms, hyperthyroidism, or malignancy suspected or confirmed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brkljacic B (21)</td>
<td>1994</td>
<td>426</td>
<td>13.3 46.5</td>
<td>Patients with multinodular goiter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ousehal A (46)</td>
<td>1996</td>
<td>100</td>
<td>39</td>
<td>Unselected patients submitted to surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rago T (55)</td>
<td>1998</td>
<td>104</td>
<td>33 42.3</td>
<td>Patients with single nodule with compressive symptoms or suspicion of malignancy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kakkos S (37)</td>
<td>2000</td>
<td>188</td>
<td>All patients submitted to surgery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bozbora A (20)</td>
<td>2002</td>
<td>81</td>
<td>25 33</td>
<td>Patients with cold solitary or dominant nodule</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giammanco M (33)</td>
<td>2002</td>
<td>125</td>
<td>21.6 57</td>
<td>All patients submitted to surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Khoo M (38)</td>
<td>2002</td>
<td>361</td>
<td>Consecutive patients submitted to surgery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kountakis S (40)</td>
<td>2002</td>
<td>83</td>
<td>All patients submitted to surgery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leenhardt L (41)</td>
<td>2002</td>
<td>155</td>
<td>25.1</td>
<td>Patients submitted to surgery with no hormonal dysfunction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peccin S (47)</td>
<td>2002</td>
<td>80</td>
<td>20 45.3</td>
<td>Patients with compressive symptoms or suspicion of malignancy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Casella C (25)</td>
<td>2003</td>
<td>66</td>
<td>15.1 44.6</td>
<td>Patients submitted to surgery with no cold solitary or dominant nodule</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alexopoulou O (18)</td>
<td>2004</td>
<td>109</td>
<td>Patients with nontoxic multinodular goiter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fukunari N (31)</td>
<td>2004</td>
<td>310</td>
<td>14.19 47</td>
<td>Patients with cold solitary nodules</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Penfold A (48)</td>
<td>2004</td>
<td>83</td>
<td>13.2</td>
<td>Patients submitted to surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seiberling K (59)</td>
<td>2004</td>
<td>159</td>
<td>23.2 46</td>
<td>Patients submitted to surgery with diagnosis of follicular nodule</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kobayashi K (39)</td>
<td>2005</td>
<td>910</td>
<td>23.2 49</td>
<td>Patients submitted to surgery with diagnosis of follicular nodule</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nicola H (30)</td>
<td>2005</td>
<td>86</td>
<td>Patients with follicular neoplasms on FNA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Popowicz B (51)</td>
<td>2006</td>
<td>356</td>
<td>Patients submitted to surgery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sahin M (56)</td>
<td>2006</td>
<td>472</td>
<td>17.8 51.5</td>
<td>Patients submitted to surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang N (66)</td>
<td>2006</td>
<td>322</td>
<td>18.4 44</td>
<td>Patients submitted to surgery with diagnosis of follicular nodule</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cappelli C (24)</td>
<td>2006</td>
<td>349</td>
<td>Patients submitted to surgery</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Author (reference)</th>
<th>Year</th>
<th>Number of nodules</th>
<th>Male sex (%)</th>
<th>Age (mean)</th>
<th>Clinical background</th>
<th>US features evaluated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rago T (52)</td>
<td>2007</td>
<td>505</td>
<td>21.18</td>
<td>45</td>
<td>Cold nodules with cytological diagnosis of follicular or Hürthle cell lesion</td>
<td>Hypoechochogenicity, microcalcifications, irregular margins</td>
</tr>
<tr>
<td>Rago T (53)</td>
<td>2007</td>
<td>92</td>
<td>31.52</td>
<td>41.6</td>
<td>Patients submitted to surgery for compressive symptoms or FNA suspicious</td>
<td>Hypoechochogenicity, absence of halo, microcalcifications, central vascularization, elasticity</td>
</tr>
<tr>
<td>Sharma R (60)</td>
<td>2007</td>
<td>52</td>
<td>26.9</td>
<td></td>
<td>Patients with cold solitary nodules</td>
<td>Central vascularization</td>
</tr>
<tr>
<td>Sippel R (63)</td>
<td>2007</td>
<td>325</td>
<td>18.5</td>
<td>47</td>
<td>Patient with FNA diagnosis of follicular or Hürthle cell neoplasm or indeterminate</td>
<td>Hypoechochogenicity, central vascularization, solid, heterogeneity</td>
</tr>
<tr>
<td>Varverakis E (64)</td>
<td>2007</td>
<td>85</td>
<td>20</td>
<td></td>
<td>Patients submitted to surgery due to risk of malignancy or compressive symptoms</td>
<td>Central vascularization</td>
</tr>
<tr>
<td>Bakhshae M (19)</td>
<td>2008</td>
<td>85</td>
<td>13</td>
<td>36.88</td>
<td>Patients submitted to surgery due to FNA diagnosis or obstructive or cosmetic reasons</td>
<td>Central vascularization, solid</td>
</tr>
<tr>
<td>Choi Y (28)</td>
<td>2008</td>
<td>175</td>
<td></td>
<td></td>
<td>Patients submitted to surgery</td>
<td>Hypoechochogenicity, microcalcifications, taller than wide shape, solid, irregular margins</td>
</tr>
<tr>
<td>Gulcelik N (34)</td>
<td>2008</td>
<td>98</td>
<td>16.32</td>
<td>46.7</td>
<td>Patients with cytology reporting follicular neoplasm</td>
<td>Hypoechochogenicity, microcalcifications, solitary, solid</td>
</tr>
<tr>
<td>Salmaslioglu A (57)</td>
<td>2008</td>
<td>1926</td>
<td>19</td>
<td>46.9</td>
<td>Patients with multinodular goiter submitted to surgery</td>
<td>Hypoechochogenicity, microcalcifications, solid, irregular margins</td>
</tr>
<tr>
<td>Chen G (27)</td>
<td>2009</td>
<td>758</td>
<td>23</td>
<td>46</td>
<td>Patients submitted to surgery</td>
<td>Microcalcifications</td>
</tr>
<tr>
<td>Hong Y (35)</td>
<td>2009</td>
<td>145</td>
<td>17.7</td>
<td>46</td>
<td>Consecutive patients submitted to surgery</td>
<td>Hypoechochogenicity, microcalcifications, central vascularization, elasticity, taller than wide shape, irregular margins</td>
</tr>
<tr>
<td>Liu F (43)</td>
<td>2009</td>
<td>40</td>
<td>7.5</td>
<td>43.7</td>
<td>Patients with lymphocytic thyroiditis and nodules FNA malignant or indeterminate</td>
<td>Hypoechochogenicity, absence of halo, microcalcifications, irregular margins</td>
</tr>
<tr>
<td>Mendelson A (45)</td>
<td>2009</td>
<td>77</td>
<td>16.8</td>
<td></td>
<td>FNA reporting follicular, Hürthle cell or nondiagnostic</td>
<td>Microcalcifications, solid</td>
</tr>
<tr>
<td>Phuttharak W (49)</td>
<td>2009</td>
<td>31</td>
<td>3.3</td>
<td>41.8</td>
<td>Patients with risk of malignancy after US and FNA</td>
<td>Hypoechochogenicity, microcalcifications, central vascularization, taller than wide shape</td>
</tr>
<tr>
<td>Popowicz B (50)</td>
<td>2009</td>
<td>1141</td>
<td>49.5</td>
<td></td>
<td>Patients submitted to surgery</td>
<td>Hypoechochogenicity, microcalcifications, central vascularization, solid, taller than wide shape</td>
</tr>
<tr>
<td>Rago T (54)</td>
<td>2010</td>
<td>195</td>
<td>26.1</td>
<td>44</td>
<td>Patients with indeterminate or nondiagnostic cytology</td>
<td>Hypoechochogenicity, absence of halo, microcalcifications, central vascularization, elasticity</td>
</tr>
<tr>
<td>Schueller-Weidekamm C (58)</td>
<td>2010</td>
<td>31</td>
<td>31.42</td>
<td>55.2</td>
<td>Patients with cold nodules</td>
<td>Hypoechochogenicity, microcalcifications, central vascularization, irregular margins</td>
</tr>
<tr>
<td>Sillery J (62)</td>
<td>2010</td>
<td>102</td>
<td>35</td>
<td>53</td>
<td>Patients with diagnosis of follicular carcinoma and adenoma</td>
<td>Hypoechochogenicity, absence of halo, microcalcifications, central vascularization, heterogeneity</td>
</tr>
</tbody>
</table>

(continued)
Statistical analysis

The overall OR was calculated to assess the predictive value of each US feature for malignancy. The Cochran chi-square and the I² tests were used to evaluate statistical heterogeneity among studies, and a threshold value of \( p = 0.10 \) was considered significant. Risk estimates were obtained with a random effects meta-analysis if significant heterogeneity was found among the studies in preliminary models.

The pooled sensitivity, specificity, positive and negative likelihood ratios, and post-test probabilities (14) were calculated using a mean pretest probability of 10% based on the average of malignancy found in thyroid nodules in general (5–8). The likelihood ratio represents how many times more (or less) frequently patients with the disease present that particular result than a patient without the disease; it is a statistical means that summarizes the diagnostic accuracy of a test (14). Likelihood ratios >10 or <0.1 are considered strong evidence to, respectively, confirm or rule out the diagnosis of interest (14).

The possibility of publication bias was evaluated using a funnel plot of a trial’s effect size against the SE. Funnel plot asymmetry was analyzed by the Begg and Egger tests. Trim-and-fill computation was used to estimate the effect of publication bias.

A separate meta-analysis was performed including only patients with nodules with an indeterminate cytology. As FNA cytology classification has changed over time, indeterminate cytology was defined as reported in the original article including those classified as indeterminate or suspicious.

All statistical analyses were performed using Stata v11.0 software (StataCorp LP, College Station, TX).

Results

The initial search retrieved 1917 articles, of which 1766 were excluded based on title and abstract. Full-text assessment was performed on 151 articles, and of these, 52 were selected for the present study (Fig. 1). Therefore, 12,786 nodules were included in the analysis (17–68). Nine studies including patients with indeterminate cytology aspirates, comprising 1851 nodules, were included in a separate meta-analysis. The characteristics of the included studies are described in Table 1.

High statistical heterogeneity was identified in the analysis of all but two US features (heterogeneity and having a taller than wide shape); therefore, the random effects meta-analysis if significant heterogeneity was found among the studies in preliminary models.

Quality of studies

Included studies had, in general, a low risk of bias. The most concerning issue was the lack of description if the US assessor was blinded for the histopathologic diagnosis. As US has to be performed prior to surgery, the person who performed the US was not aware of the histopathologic diagnosis. It was also considered that some studies may have limitations due to patient selection, in most cases because they included only patients with cold nodules. Details about quality of trials are described in Supplementary Table S2.
Diagnostic performance of US features in all nodules

All the features evaluated were significantly associated with malignancy, with an overall OR ranging from 1.77 to 35.7 (Fig. 2). However, the sensitivity of US features traditionally associated with malignancy was somewhat low, ranging from 26.7% to 63%, which means that, using these features individually, 37% to 73.3% of cancers would not be diagnosed. Four of these features—microcalcifications, central vascularization, irregular margins, and a taller than wide shape—showed better specificity than the other features: 87.8%, 78%, 83.1%, and 96.6%, respectively. The positive likelihood ratio...
ranged from 1.33 to 8.07, and the negative likelihood ratio from 0.13 to 0.77 (Table 2). Considering a pretest probability of 10%, the post-test probability of malignancy ranged from 12.8% to 47.0% after a positive test, and 1.4% to 7.8% with a negative test result. Absence of elasticity was the US feature that showed the best diagnostic accuracy, with a sensitivity of 87.9%, a specificity of 86.2%, and a positive and negative LHR of 6.39 and 0.13, respectively (Table 2).

**Diagnostic performance of US features in nodules with indeterminate cytology**

Only a few of the studies reported the histopathologic diagnosis specifically for nodules with an indeterminate cytology. Because of that, only the following features were analyzed: absence of halo, absence of elasticity, hypoechogenicity, solid structure, presence of microcalcifications, solitary nodule, irregular margins, and central vascularization. Of these, pooled diagnostic accuracy statistics could be calculated only for hypoechogenicity, central vascularization, and presence of microcalcifications because more than three studies are needed in order to perform a meta-analysis of a diagnostic test. Only the presence of microcalcifications was significantly associated with malignancy (Fig. 3). However, in this subgroup of nodules, any of the US features was not able to determinate the risk of malignancy with an acceptable sensitivity (Table 3). Presence of central vascularization was the feature with the best specificity (96%).

**Table 2. Diagnostic Performance of Each US Feature in the Differentiation of Benign and Malignant Thyroid Nodules in Unselected Nodules**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Positive likelihood ratio</th>
<th>Post-test probability (%)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Negative likelihood ratio</th>
<th>Post-test probability (%)&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taller than wide</td>
<td>26.7</td>
<td>96.6</td>
<td>8.07</td>
<td>47.0</td>
<td>0.75</td>
<td>7.6</td>
</tr>
<tr>
<td>Halo absent</td>
<td>56.4</td>
<td>72.0</td>
<td>2.02</td>
<td>18.1</td>
<td>0.60</td>
<td>6.2</td>
</tr>
<tr>
<td>Absence of elasticity</td>
<td>87.9</td>
<td>86.2</td>
<td>6.39</td>
<td>41.3</td>
<td>0.13</td>
<td>1.4</td>
</tr>
<tr>
<td>Heterogeneity</td>
<td>47.5</td>
<td>70.0</td>
<td>1.58</td>
<td>14.8</td>
<td>0.74</td>
<td>7.5</td>
</tr>
<tr>
<td>Hypoechogenicity</td>
<td>62.7</td>
<td>62.3</td>
<td>1.66</td>
<td>15.4</td>
<td>0.62</td>
<td>6.3</td>
</tr>
<tr>
<td>Solid</td>
<td>72.7</td>
<td>53.2</td>
<td>1.55</td>
<td>14.6</td>
<td>0.51</td>
<td>5.3</td>
</tr>
<tr>
<td>Microcalcifications</td>
<td>39.5</td>
<td>87.8</td>
<td>3.26</td>
<td>26.4</td>
<td>0.68</td>
<td>7.0</td>
</tr>
<tr>
<td>Solitary</td>
<td>53.0</td>
<td>60.2</td>
<td>1.33</td>
<td>12.8</td>
<td>0.77</td>
<td>7.8</td>
</tr>
<tr>
<td>Central vascularization</td>
<td>45.9</td>
<td>78.0</td>
<td>2.09</td>
<td>18.7</td>
<td>0.69</td>
<td>7.1</td>
</tr>
<tr>
<td>Irregular margins</td>
<td>50.5</td>
<td>83.1</td>
<td>2.99</td>
<td>24.7</td>
<td>0.59</td>
<td>6.1</td>
</tr>
</tbody>
</table>

<sup>a</sup>Probability of malignancy after having a positive test result.
<sup>b</sup>Probability of malignancy after having a negative test result.
FIG. 3. Forest plot representing OR for malignancy of each US feature evaluated in nodules with indeterminate cytology.
positive likelihood ratio ranged from 1.12 to 2.52, and the negative likelihood ratio from 0.66 to 0.95. Considering a pretest probability of 10%, the post-test probability of malignancy ranged from 11% to 21.8% after a positive test, and 6.8% to 9.5% with a negative test result.

**Meta-regression**

In the analysis of some features, fewer than 10 studies were available, preventing a meta-regression from being performed. For analysis of hypoechogenicity, irregular margins, microcalcifications, solid structure, and central vascularization, a meta-regression was performed using the year of publication and/or the prevalence of cancer in the study sample as covariates. However, none of these variables was able to explain the high heterogeneity found significantly.

**Discussion**

In the present meta-analysis, the US features associated with a higher risk and post-test probability of malignancy were taller than wide shape, absence of elasticity, presence of microcalcifications, and irregular margins. However, none of the US features analyzed singly had a clinically relevant positive likelihood ratio (> 10) and post-test probabilities to suggest malignancy. Most likely, the use in combination may provide stronger risk and probability of malignancy. However, it was not possible to estimate the real risk of malignancy by using the combination of US features because very few studies have analyzed this aspect, and they differ regarding the selected features.

The strengths of the present meta-analysis are the large number of nodules evaluated and the fact that all the nodules included had a histopathologic diagnosis, which is the reference method for the definite diagnosis of thyroid nodules. Moreover, the performance of US in nodules with indeterminate cytology was also analyzed, which constitutes the most challenging group of patients for clinical decision making. Another relevant aspect was the calculation of the likelihood ratio statistics, which summarizes how many times more (or less) likely patients with the disease are to have that particular result than patients without the disease (14). The likelihood ratio of a diagnostic test is more useful clinically than sensitivity and specificity. To the best of the authors’ knowledge, there are no previous systematic reviews and meta-analyses that focus on histopathology only.

The present study has some limitations. First, no information was available on individual characteristics of patients regarding risk factors for malignancy, and on the reason for surgery. Also, the number of studies was insufficient for the analysis of some US features in patients with indeterminate cytology, possibly the subgroup of patients that would most benefit from the use of US as a tool to help in clinical management decision.

Our results confirm the findings of previous isolated studies. Moon et al. (69) evaluated 831 patients with thyroid nodules and found low sensitivity values for most of the US features. Hypoechogenicity was the only finding that showed a sensitivity of 87.2%. In the same study, taller than wide shape, speculated margins, marked hypoechogenicity, and micro- and macrocalcifications demonstrated a high specificity for malignancy, ranging from 90.8% to 96.1%. In one of the largest series comprising 672 patients and 1141 nodules, Popovicz et al. also found low sensitivity values for most US features for malignancy. However, microcalcifications and taller than wide shape features had high specificity (50). Moreover, in another study including 550 patients with multinodular goiter, Salmaslioglu et al. found that the presence of microcalcifications had a sensitivity of 89.3% for malignancy (57). The best diagnostic performance in the present meta-analysis was seen for absence of elasticity. Usually, elasticity is described in a scale ranging from 1 to 4 (1–2 being suggestive of a benign nodule and 3–4 of malignancy) or 1–5 (where 1–3 is suggestive of a benign lesion and 4–5 of malignancy) (53,70–72). This US feature was initially described for breast or prostate cancer, but several studies have evaluated its performance to differentiate between malignant and benign thyroid nodules, revealing high sensitivity and specificity (81.8–97% and 81.1–100%) (52,72). A recent meta-analysis including eight studies with a total of 639 nodules diagnosed by FNA cytopathology or histopathology reported a sensitivity of 90% and specificity of 92% for elasticity. However, not all studies included had a final histopathologic diagnosis of the nodule (70). A recent study, which included 498 thyroid nodules evaluated by US, color flux Doppler, and real-time elastography, concluded that the combination of elastography with US parameters increased the sensitivity for malignancy to 97% (73).

The present findings have important clinical implications. They reinforce that isolated US features on their own do not provide strong evidence to confirm (likelihood ratio > 10) or rule out (likelihood ratio < 0.1) a diagnosis of malignancy. The American Thyroid Association recommends the use of a combination of US features to select thyroid nodules that should be biopsied (5). Information about the probability of each US feature to be associated with malignancy would help the clinical decision to perform FNA biopsy.

### Table 3. Diagnostic Performance of Each US Feature in the Differentiation of Benign and Malignant Thyroid Nodules in Nodules with Indeterminate Cytology

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Positive likelihood ratio</th>
<th>Post-test probability (%)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Negative likelihood ratio</th>
<th>Post-test probability (%)&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypoechogenicity</td>
<td>49.7</td>
<td>56.0</td>
<td>1.12</td>
<td>11.0</td>
<td>0.89</td>
<td>8.9</td>
</tr>
<tr>
<td>Microcalcifications</td>
<td>45.6</td>
<td>81.9</td>
<td>2.52</td>
<td>21.8</td>
<td>0.66</td>
<td>6.8</td>
</tr>
<tr>
<td>Central vascularization</td>
<td>8.4</td>
<td>96.0</td>
<td>2.13</td>
<td>19.1</td>
<td>0.95</td>
<td>9.5</td>
</tr>
</tbody>
</table>

<sup>a</sup>Probability of malignancy after having a positive test result.

<sup>b</sup>Probability of malignancy after having a negative test result.
The present findings also suggest that more accurate criteria are needed to recommend surgery in patients with indeterminate cytology. This is an important practical matter, since it would be helpful to select better which patients should be submitted to FNA and, specially, when surgery should be indicated in those nodules with indeterminate cytology (74, 75).

Attempts have been made to improve patient selection in evaluation of thyroid nodules. Moon et al. evaluated a classification that considered as suspicious for malignancy a nodule that was solid plus having two additional risk features. Those authors found a sensitivity of 87.7%, a specificity of 97.8%, and an overall accuracy of 96.2% (76). A recent, retrospective case-control study of patients who underwent thyroid US reported three ultrasound nodule characteristics (microcalcifications, size > 2 cm, and an entirely solid composition) as the only findings associated with the risk of thyroid cancer (77). However, this study has important aspects that limit its generalization such as the low prevalence of thyroid cancer and the definition of noncancerous nodules (78).

Conclusions

The present results show that there is no isolated US feature capable of predicting malignancy in thyroid nodules with acceptable diagnostic accuracy. However, the presence of some US features, such as a microcalcifications, a taller than wide shape, irregular margins, central vascularization, or absence of elasticity probably, will identify nodules with an increased risk for malignancy. Ideally, meta-analyses should be performed with individual patient data, which would enable the creation of a risk classification for malignancy in thyroid nodules combining US features in combination and other risk factors to define better which patients should be submitted to FNA and surgery. Elastography is a new technique and may be a good tool to select patients at increased risk for thyroid malignancy. Nevertheless, more studies are required to standardize the technique and confirm its usefulness.

Acknowledgment

J.L.G. holds a grant from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)—project 558637/2008-6.

Author Disclosure Statement

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that they have no competing interests relevant to this work. J.L.G. has served on boards for Bristol-Myers Squibb, GlaxoSmithKline, Novo Nordisk, and Eli Lilly, and has received payment for the development of educational presentations for Bristol-Myers Squibb, Novo Nordisk, and Eli Lilly.

References


Address correspondence to:
Luciana Reck Remonti, MD
Hospital de Clínicas de Porto Alegre
Ramiro Barcelos 2350
Porto Alegre, RS
Brazil
E-mail: lucianalreck@gmail.com