The diagnostic efficiency of ultrasound in characterization for thyroid nodules: how many criteria are required to predict malignancy?

Alper Ozel¹, Sukru Mehmet Erturk¹, Alkin Ercan¹, Banu Yılmaz², Tulay Basak³, Vito Cantisani³, Muzaffer Basak¹, Zeki Karpat¹

¹ Sisli Etfal Research and Training Hospital, Radiology Department, Etfal Sokak 34377, Istanbul, TURKEY
² Sisli Etfal Research and Training Hospital, Pathology Department, Etfal Sokak 34377, Istanbul, TURKEY
³ University of Rome “La Sapienza”, Department of Radiology, Rome, Italy

Abstract

Objective: The purpose of this study was to define the criteria for use in differentiating benign and malignant nodules with the help of the receiver operating characteristic analysis and to increase the objective diagnostic accuracy of ultrasonography.

Materials and methods: A total of 363 patients (307 women, 56 men) with 363 nodules (22 malignant and 341 benign nodules) were included in the study. The presence and absence of each US feature of the evaluated nodule – shape taller than wide, irregular margin, hypoechogenicity, microcalcification, and intranodular vascularity – were scored 1 and 0, respectively. The total ultrasound score was obtained by the summing up of each positive ultrasound findings for malignancy. The effect of the total US score in the discrimination of benign and malignant nodules was analysed using ROC analysis.

Results: The cut off values of US score at maximum sensitivity and specificity for nodules larger and smaller than one centimeter were two (Az: 0.783) and three (Az: 0.935), respectively. For nodules greater than one centimeter, the calculated diagnostic performances including sensitivity, specificity, positive predictive value and negative predictive value, and accuracy were 62.5%, 91.5%, 30.3%, 97.7%, and 89.9%, respectively. For nodules smaller or equal to one centimeter; the sensitivity, specificity, positive predictive value, negative predictive value and accuracy were 83.3%, 94.9%, 62.5%, 98.2% and 93.8%, respectively.

Conclusion: Using ultrasound, thyroid nodules can be characterized effectively. The number of the US features used in this distinction varies in respect to the nodule size.

Keywords: thyroid nodule, thyroid carcinoma, benign nodule, ultrasonography, fine-needle aspiration biopsy

Introduction

Ultrasonography (US) is an excellent imaging technique for identifying thyroid nodules, which are very common in clinical practice. The incidence of thyroid nodules detected by US ranges from 10% to 67% [1-4]. US has been also widely used to differentiate between benign and malignant nodules using several sonographic characteristics. US features predictive of malignancy include taller than wide shape, irregular margin, hypoechogenicity, the presence of microcalcifications and intranodular vascularity [5-12]. However, some studies have found overlaps in the presence of these characteristics between benign and malignant nodules [13,14].

Additionally, it is known that no single ultrasound feature has the adequate diagnostic accuracy for diagnosing malignant nodules. Therefore, fine-needle aspiration biopsy (FNAB) is considered to be the best preoperative triage test for preoperative evaluation of the thyroid nodules.

Currently, different guidelines have been used for increasing the diagnostic accuracy of US [4,7,15,16]. Recently, in a study of Ahn et al [17], it was demonstrated that two sets of guidelines have come into prominence:
the American Association of Clinical Endocrinologists (AACE) criteria and Kim criteria. The common characteristic of both guidelines is that these guidelines use a priori criteria to diagnose malignant nodules.

In this study, we aimed to define the criteria for use in differentiating benign and malignant nodules with the help of the receiver operating characteristic (ROC) analysis and to increase the objective diagnostic accuracy of US.

**Materials and Methods**

The retrospective data analysis was performed on 439 palpable or nonpalpable thyroid nodules in 439 patients referred for FNAB in the Radiology Department between January 2009 and May 2010. Patient inclusion criteria were as follows: a) patients with an initial benign cytology and US follow up (>12 months after FNAB); b) patients who underwent surgery after FNAB for malignant nodules c) patients who underwent surgery after indeterminate cytology with FNAB. Patients with nodules with nondiagnostic cytology and patients without a final tissue diagnosis after FNAB were excluded from the study. A total of 363 patients (307 women, 56 men) with 363 nodules (22 malignant and 341 benign nodules) were included. The mean patient age was 47.6 years.

The final tissue diagnosis for malignant nodules included papillary carcinoma (n=21) and follicular carcinoma (n=1). Diagnosis of benign lesions at histologic examination included nodular hyperplasia (n=10), follicular adenoma (n=2) and thyroiditis (n=2).

**US examination and FNAB technique**

Thyroid ultrasound examinations were performed with an Aplio XV (Toshiba Tokyo Japan) machine equipped with a 7-14 MHz linear array transducer. Each biopsied nodule was documented in both transverse and longitudinal plane, while the patient lay supine with the head hyperextended. The following sonographic features were assessed for each nodule: shape, margin, echogenicity, echostucture, presence of calcifications, and vascularity on color Doppler. The shape of the nodule was classified as taller than width measured in transverse dimension or as round and wider than tall. Margins of nodules were categorized as well circumscribed when clear demarcation with normal thyroid was noted, and as not well circumscribed, which included irregular and microlobulated margins. The echogenicity of each nodule was classified as hypo-, iso- or hyperechoic in comparison with the normal background thyroid tissue. A nodule was defined as marked hypoechoic, when a nodule was hypoechoic relative to adjacent strap mus-
The diagnostic efficiency of ultrasound in characterization for thyroid nodules

**Results**

The size of the nodules ranged from 4 mm to 49 mm (mean size 17.3 mm ± 7.9). Regarding the size of the nodules, there was no significant difference between benign (17.3 mm ± 7.9) and malignant tumors (16.3 mm ± 8.1) (p>0.05).

A taller than wide shape was found more frequently in malignant nodules (22.8%) than in benign nodules (3.3%). Hypoechoogenicity (including the subgroup of markedly hypoechoic nodules) was a sonographic feature to be found in a substantial number of malignant nodules (72.7%). The frequency of hypoechoogenicity in benign nodules was low (13%). The presence of microcalcifications and intranodular vascularity on Doppler examination in malignant nodules were significantly higher than in benign ones (Table I).

A ROC analysis was performed to evaluate the effect of US scores in the discrimination of benign and malig-

**Table I. Frequency Analysis of Sonographic Features of Benign and Malignant Thyroid Nodules**

<table>
<thead>
<tr>
<th>US Feature</th>
<th>Benign Nodules (n=341)</th>
<th>Malign Nodules (n=22)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shape</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ovoid to round</td>
<td>330 (96.7)</td>
<td>17 (77.2)</td>
</tr>
<tr>
<td>Taller than wide</td>
<td>11 (3.3)</td>
<td>5 (22.8)</td>
</tr>
<tr>
<td><strong>Margin</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smooth</td>
<td>312 (91.4)</td>
<td>16 (72.7)</td>
</tr>
<tr>
<td>Irregular</td>
<td>29 (8.6)</td>
<td>6 (27.3)</td>
</tr>
<tr>
<td><strong>Echogenicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marked hypoechoic</td>
<td>1 (0.2)</td>
<td>7 (31.8)</td>
</tr>
<tr>
<td>Hypoechoic (incl. marked hypoechoic)</td>
<td>45 (13.1)</td>
<td>16 (72.7)</td>
</tr>
<tr>
<td>Isoechoic</td>
<td>286 (83.8)</td>
<td>6 (27.3)</td>
</tr>
<tr>
<td>Hyperechoic</td>
<td>10 (2.9)</td>
<td>–</td>
</tr>
<tr>
<td><strong>Calcification</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>310 (90.9)</td>
<td>10 (45.4)</td>
</tr>
<tr>
<td>Microcalcification</td>
<td>6 (1.7)</td>
<td>10 (45.4)</td>
</tr>
<tr>
<td>Amorph</td>
<td>25 (7.4)</td>
<td>2 (9.1)</td>
</tr>
<tr>
<td><strong>Vascularity on Doppler US</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>122 (35.7)</td>
<td>4 (18.2)</td>
</tr>
<tr>
<td>Peripheral</td>
<td>94 (27.5)</td>
<td>4 (18.2)</td>
</tr>
<tr>
<td>Intranodular</td>
<td>125 (36.6)</td>
<td>14 (63.6)</td>
</tr>
</tbody>
</table>
nant nodules in respect to nodule size. The cut off values of US score at maximum sensitivity and specificity for nodules larger and smaller than one centimeter were two (Az: 0.783) and three (Az: 0.935), respectively.

For the nodules greater than one centimeter, the calculated diagnostic performances including sensitivity, specificity, positive predictive value and negative predictive value were 62.5%, 91.5%, 30.3% and 97.7%, respectively. The diagnostic accuracy of US score for nodules larger than one centimeter was found; 89.9%.

For the nodules smaller or equal to one centimeter the corresponded values for the sensitivity, specificity, positive predictive value, negative predictive value and accuracy were 83.3%, 94.9%, 62.5%, 98.2% and 93.8%, respectively (table II).

Discussion

In our study, we found that different numbers of the sonographic features should be used for the discrimination of the thyroid nodules greater or smaller than one centimeter. For nodules that were greater than 10 mm in maximal dimension, as two of the sonographic features were used as a cut off, the diagnostic accuracy was 89.9%. For the nodules that were equal or less than 10 mm in size, the cut off value for sonographic features was three and diagnostic accuracy of US was 93.8%. In our opinion, the most important result we achieved was that the number of the used ultrasound features for diagnosing malignant nodules would be changed in respect to nodule size.

In the study of Kim et al, the criteria for FNAB of nonpalpable solid thyroid nodules was proposed [7]. Suspicious sonographic features were defined as irregular or microlobulated margin, marked hypoechoogenicity, microcalcifications and a shape that was more tall than it was wide. In the presence of even one of these sonographic findings the sensitivity, specificity, positive predictive value, negative predictive value and accuracy were 93.8%, 66%, 56.1%, 95.9% and 74.8%, respectively. That study was performed on nonpalpable thyroid nodules (3 mm-28 mm) and intranodular spot vascular-

Table II. Diagnostic accuracy of the total US score for malignant nodules according to size

<table>
<thead>
<tr>
<th>Size</th>
<th>Az</th>
<th>Cut off</th>
<th>Sensitivity(%)</th>
<th>Specificity(%)</th>
<th>Positive Predictive Value(%)</th>
<th>Negative Predictive Value(%)</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;10 mm</td>
<td>0.783</td>
<td>2</td>
<td>62.5 (10/16)</td>
<td>91.5 (258/282)</td>
<td>30.3 (10/33)</td>
<td>97.7 (259/265)</td>
<td>89.9 (268/298)</td>
</tr>
<tr>
<td>≤10 mm</td>
<td>0.935</td>
<td>3</td>
<td>83.3 (5/6)</td>
<td>94.9 (56/59)</td>
<td>62.5 (5/8)</td>
<td>98.2 (56/57)</td>
<td>93.8 (61/65)</td>
</tr>
</tbody>
</table>

Accordingly, due to the lack of intranodular vascular criterion and the use of only one of the suspicious criteria, the specificity was low.

Moon et al [18] evaluated the diagnostic accuracy of US for the depiction of benign and malignant thyroid nodules and found that the US criteria including a shape taller than wide, a spiculated margin, marked hypoecho-ogenicity, microcalcification and macrocalcification were helpful for discrimination of malignant nodules from benign ones. According to their results, the diagnostic accuracy for the nodules one centimeter or less in size was 77% when one of the five malignant findings was used. The relative low diagnostic accuracy in that study could be due to the lack of intranodular vascularity into suspicious diagnostic US criteria. When compared to our results, we believe that the inclusion of intranodular spot vascularity in the malignant US findings has an important role in the diagnosis of malignant thyroid nodules.

Recently, Ahn et al [17] compared the different sets of guidelines including AACE criteria for discriminating benign and malignant thyroid nodules. According to AACE guidelines, FNAB should be performed on all hypoechoic nodules with at least one of the following additional ultrasound features: taller than wide, irregular margins, microcalcifications, or intranodular vascular spots. For nodules of one centimeter or more in diameter, Ahn et al calculated the diagnostic accuracy with receiver operating characteristic analysis and found an Az value of 0.842. In that study, the diagnostic accuracy for nodules under one centimeter was not reported. Additionally, the lack of intranodular vascularity criterion is a limitation of that study.

Regarding the clinical impact, the ultrasonographic assessment of thyroid nodules is problematic and a subject on which a consensus has not been yet reached. In general, it can be said that four gray scale ultrasonographic features including a shape taller than wide, irregular margin, marked hypoechoogenicity and microcalcifications, are the most accepted criteria for diagnosing malignant nodules. In our study, we evaluated the four gray scale ultrasound features and intranodular vascular...
spot on Doppler US in respect to nodule size. In contrast to other previous studies, we found that at least three and two ultrasound features have to be taken into consideration for nodules under one centimeter and equal or greater than one centimeter, respectively. The increase of diagnostic accuracy of US in that way could play an important role in respect to clinical management of thyroid nodules.

Our study had some limitations. The first limitation is the retrospective design of the study. Therefore, the ultrasonographic features of the biopsied nodules were evaluated. However, in recent literature almost all studies have the same limitation [17,18]. Secondly, the small number of the malignant nodules was another limitation. Nevertheless, the reported malignancy rate of incidental detected thyroid nodules in the literature is 5-15% [19-22]. The rate of malignant nodules in our study was 6.0% and it was in accordance with the literature.

In conclusion, benign and malignant discrimination of thyroid nodules with US could be made with high diagnostic accuracy. The number of the US features used in this distinction varies in respect to nodule size.

Conflict of interest: none

References