Thyroid nodule shape suggests malignancy

Carlo Cappelli, Maurizio Castellano, Ilenia Pirola, Elena Gandossi, Elvira De Martino, Davide Cumetti, Barbara Agosti and Enrico Agabiti Rosei

Department of Medical and Surgical Sciences, Internal Medicine and Endocrinology Unit, University of Brescia, c/o 2 Medicina Spedali Civili di Brescia, Piazzale Spedali Civili no. 1, 25100 Brescia, Italy

(Correspondence should be addressed to C Cappelli; Email: cappi.tsh@libero.it)

Abstract

Objective: To evaluate if a nodule with shape taller than wide (anteroposterior/transverse diameter ratio, A/T $\geq$ 1) is a good predictor of malignancy independent of the size.

Methods: We retrospectively examined the cytological and histological results of 7455 nodules (5198 patients) referred for ultrasound-guided-fine needle aspiration cytology (US-FNAC) in our hospital from January 1991 to September 2004.

Results: A suitable FNAC was obtained from 6135 nodules (4495 patients); 34.6% were less than 1 cm in diameter (small nodules, SN). A diagnosis of carcinoma was histologically confirmed in 284/349 suspicious lesions after FNAC. The size of carcinoma nodules was not significantly associated with the occurrence of extracapsular growth (large nodules (LN): 10.5%, SN: 4.9%, NS) and lymph node metastasis (LN: 23.6%, SN: 25.0%, NS). Malignant lesions showed microcalcifications more frequently than benign nodules (72.2 vs 28.7%; $P < 0.001$; (odds ratio, OR(confidence intervals, CI) $= 9.9(7.2–13.4)$). Similarly, A/T $\geq$ 1 (76 vs 40%; $P < 0.001$; OR(CI) $= 8.6(5.5–13.1)$), blurred margins (52.8 vs 18.8%; $P < 0.001$; OR(CI) $= 7.7(5.6–10.2)$), solid hypo-echoic appearance (80.6 vs 52.4%; $P < 0.001$; OR(CI) $= 3.2(2.2–4.3)$) and intranodular vascular pattern (type 2) (61.6 vs 49.7%; $P < 0.001$; OR(CI) $= 1.7(1.3–2.3)$) were significantly more frequent in malignant than in benign nodules.

Conclusions: Our data show that no single parameter, including nodule size, satisfactorily identifies a subset of patients to be electively investigated by FNAC. We concluded that A/T $\geq$ 1 with at least two of US features (microcalcification, blurred margins, hypo-echoic pattern) is today the best compromise between missing cancers and cost–benefit.

Introduction

It has been estimated that palpable thyroid nodules are present in 4–7% of the population (1–3), but when examined by ultrasound (US), as many as 50–70% of subjects with no history of thyroid disease have been found to have incidentally discovered thyroid nodules, many of which are not palpable (4–7). In addition, nodular thyroid disease is more common in the elderly, a population subgroup, which is steadily increasing (8). Many studies have been published on the risk of malignancy in patients with thyroid nodules; these studies show that the risk of malignancy is low, approximately 5%, unless the patient has an underlying risk factor, such as a history of external neck irradiation (1–3).

Fine needle aspiration cytology (FNAC) is considered as the most reliable test for the diagnosis of thyroid nodules (9–11), but it appears not to be cost-effective to submit all these lesions to FNAC (12–15). In agreement with Kim et al. (22), we have previously demonstrated that a lesion with anteroposterior and transverse diameter ratio $\geq$ 1 is strongly associated with the malignancy in non-palpable thyroid nodules (23).

The aim of this study was to evaluate if a nodule with shape taller than wide (anteroposterior/transverse diameter ratio, A/T $\geq$ 1) is a good predictor of malignancy independent of the size.

Subjects and methods

We retrospectively examined the cytological and histological characteristics of thyroid nodules investigated by US-FNAC in 5198 patients referred to our hospital from January 1991 to September 2004 for thyroid US evaluation. The reasons for referral were quite heterogeneous, including clinical finding of palpable neck mass, follow-up controls of thyroid diseases, incidental finding of thyroid nodules at
previous echographic neck investigation (supra-aortic vessel echo Doppler).

No patient included in this series had a history of external or ionizing radiation or previous diagnosis of thyroid malignancy.

Out of the 7455 nodules examined, 2865 (38.4%) were less than 1 cm in diameter.

US investigation was performed using an ultrasonographic scanner (Siemens Elegra or ATL 5000) equipped with a 10–12 MHz linear transducer for morphological study and 4.7 MHz for colour-Doppler evaluation. Examinations were conducted and recorded by two skilled sonographers according to a standard procedure. The following US parameters were assessed: (i) echographic dimension (anteroposterior and transverse diameter); (ii) echogenicity (isoechoic, hyperechoic or hypoechoic); (iii) presence/absence of calcification; (iv) aspect of lesion margins (well defined or blurred); (v) vascular pattern as evaluated by sagittal and transverse scans performed along the maximum diameter of the nodule: type 0, absence of flow signals; type 1, vascular images in peripheral position; type 2, intranodular flow with multiple vascular images (24, 25).

FNAC was performed by using 25-gauge needles and capillarity action with a free-hand technique. Cytological specimens were evaluated by an experienced cytopathologist and classified as benign, suspicious or malignant. Suspicious cytological results included follicular or Hurthle cell neoplasms. Samples were defined as sufficient when six or more clusters of follicular cells with each group containing at least 10 follicular cells were present on the slides (12, 26).

All patients with suspicious or malignant cytology underwent surgery. All the subjects with benign cytology were re-examined by US after a 6-month period. Only one individual presented a significant increase of nodule volume: he was submitted to surgery and histology confirmed the benign nature of his lesion. Specimens were fixed in buffered formalin, embedded in paraffin and stained with haematoxylin and eosin for histological study. The final histological diagnosis on surgical specimens (according to WHO classification) was considered, the goal standard (27).

Neoplastic lesions outside the nodule examined by FNAC were considered incidental and not included in the analysis.

Statistical analysis

Clinical, US, cytological and histological findings were separately recorded and blind-processed for statistical evaluation (SPPS for Windows). Comparison of frequency distributions was performed by \( \chi^2 \)-test. Univariate and multivariate (logistic regression analysis) odds ratio (OR), with relative 95% confidence intervals (CI) were calculated to assess the relevance of US criteria to predict histological outcome.

The diagnostic value of US criteria was also assessed in terms of sensitivity, specificity, likelihood ratio, positive/negative predictive value and efficiency (28).

Results

Out of 7455 nodules, 1320 (17.7%) presented inadequate cytological specimens and were excluded from the study. A total of 6135 nodules with valid cytological specimens were obtained from 4495 subjects (3118 females, 1377 males). Multiple nodules were investigated in 1351 subjects (1118 with two, 182 with three, 46 with four, 5 with five nodules, respectively). Nodule size, as evaluated by US, ranged from 6 to 100 mm (mean ± s.d.: 15.5 ± 9.0 mm); approximately, two-thirds \( (n = 4015) \) were large nodules (LN) and the remaining \( (n = 2120) \) were small ones.

A diagnosis of carcinoma was histologically confirmed in 284 (242 papillary, 37 follicular and 5 medullary; 4.6%) of the 349 subjects with suspicious cytology after FNAC; no patient presented more than one neoplastic nodule. According to the size, the prevalence of cancer was less in small nodules (SN) than in LN (3.0 vs 5.5%, \( P < 0.001 \)). On the other side, the relative frequency of follicular and medullary histotypes was significantly greater among SN than in LN (20.8 and 1.9% vs 6.7 and 1.2%, \( P = 0.01 \)).

Malignant features at histological examination were slightly more frequent in nodules from multinodular goitre (147/2770; 5.3%) than in solitary nodules (137/3365; 4.1%) at univariate analysis, with borderline statistical significance (\( P = 0.022 \)). Neoplastic lesions were not associated with gender (5.0% in males, 4.5% in females; NS).

Local extension of thyroid neoplasia with extracapsular growth (pT3 at TNM staging) was present in 26/284 carcinomas (9.2%), whereas nodal involvement was found in 67 (23.6%). Extracapsular growth was predictable (on the basis of absence of a well-defined distinction between the margins of the lesion and the inner profile of the thyroid capsule, and/or sonographic evidence of involvement of the anterior cervical muscles) in 7/26 carcinomas (26.9%). Nodal involvement was diagnosed at preliminary US evaluation in 7/67 (10.4%) histologically proven cases of pN1 malignancy. Quite surprisingly, the size of the carcinoma nodules was not significantly associated with the occurrence of extracapsular growth (LN: 10.5%, SN: 4.9%, NS) and lymph node metastasis (LN: 23.6%, SN: 25.0%, NS).

Microcalcifications were significantly more frequent in malignant lesions than in benign nodules (72.2 vs 28.7%; \( P < 0.001 \); Odds ratio, OR(confidence intervals, CI) = 9.9(7.2–13.4)). Similarly, A/T≥1 (76 vs 40%; \( P < 0.001 \); OR(CI) = 8.6(5.5–13.1)), blurred margins (52.8 vs 18.8%; \( P < 0.001 \); OR(CI) = 7.7(5.6–10.2)), solid hypo-echoic appearance (80.6 vs 52.4%);
P < 0.001; OR(CI) = 3.2(2.2–4.3) and intranodular vascular pattern (type 2) (61.6 vs 49.7%; P < 0.001; OR(CI) = 1.7(1.3–2.3)) were significantly more frequent in malignant than in benign nodules.

Multivariate logistic regression analysis confirmed that all the above US criteria were significantly associated with the histological diagnosis of malignancy. The associations of US features with malignancy, reported above, were similar when LN and SN were examined separately. The analysis also showed that age, gender and uniqueness of lesion were not significant predictors of cancer.

The predictive value of US features is reported in Table 1. The results show that no single parameter, including A/T ≥ 1, satisfactorily identifies a subset of patients to be electively investigated by FNAC, as in any case, a relevant fraction of histologically confirmed carcinomas would be missed (from 19 up to 47%, depending on the US criterion).

Discussion

It has been estimated that palpable thyroid nodules are present in 4–7% of the population (1–3), but when examined by US, as many as 50–70% of subjects with no history of thyroid disease have been found to have incidentally discovered thyroid lesions, many of which are not palpable (4–7). Many studies have been published on the risk of malignancy in patients with thyroid nodules; these studies show that the risk of malignancy is low, approximately 5%, unless the patient has an underlying risk factor, such as a history of external neck irradiation (1–3).

FNAC is the most reliable test for the diagnosis of thyroid nodules (29–31), but it is not cost-effective to submit all these lesions to FNAC (13–15, 21). The definition of sonographic features as risk of malignancy should help to determine optimal management of these lesions, by reducing the number of FNAC and the cost of healthcare.

In our series, the percentage of non-diagnostic specimens (17.7%) was similar to literature data (15.8–18.6%) (6, 30); US features of nodules with unsuccessful FNAC were not different from those supplying adequate material, with the relevant exception of size (26% of failure in SN, about twice as compared to LN). Similarly, high percentages of inadequate cytology have been previously reported for infracentimetric nodules (6, 13, 30). It is not clear whether and how such a difference in diagnostic evaluation of SN vs LN may affect the interpretation of our data. In any case, the main finding of a rather high prevalence of malignant lesions in SN should not be modified by this limitation.

The overall percentage of malignancy in our series is lower (4.6%) than previously reported in studies of smaller size (32, 33) but similar to that observed in recent study by Lin et al. (31) in a very large series of more than 20 000 patients (3.6%). Accordingly, it is highly unlikely that a relevant number of malignant nodules have not been identified in our study. On the other hand, we reasonably exclude that a bias toward inclusion of nodules with higher risk of malignancy may have occurred, since no a priori selection criteria for patient’s referral were adopted from peripheral centres, and all patients were submitted to FNAC, irrespective of their US features.

As far as thyroid nodule size is concerned, our results show that the frequency of histologically malignant lesions is slightly less in SN vs LN, but still clinically relevant in absolute terms. Noteworthy, follicular and medullary carcinomas, i.e. histological types with a less favourable prognosis, are even more frequent in SN. On the same line of evidence, our study shows that extracapsular growth and nodal involvement at surgical staging were similar in SN and LN.

An increasing number of studies support our finding that a similar fraction of nodules over and under 10 mm present evidence of extrathyroidal spread and/or nodal involvement. The predictive value of US features is reported in Table 1. The results show that no single parameter, including A/T ≥ 1, satisfactorily identifies a subset of patients to be electively investigated by FNAC, as in any case, a relevant fraction of histologically confirmed carcinomas would be missed (from 19 up to 47%, depending on the US criterion).

Table 1 Diagnostic value of different ultrasound features and their combination for the identification of malignant thyroid nodules.

<table>
<thead>
<tr>
<th>Ultrasound features</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>Performed FNAC (%)</th>
<th>Missed carcinomas (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Size ≥ 10 mm</td>
<td>77</td>
<td>35</td>
<td>5.4</td>
<td>97</td>
<td>65</td>
<td>19</td>
</tr>
<tr>
<td>(b) Hypo-echoic*</td>
<td>81</td>
<td>47</td>
<td>7.0</td>
<td>98</td>
<td>53</td>
<td>19</td>
</tr>
<tr>
<td>(c) Blurred margins*</td>
<td>53</td>
<td>81</td>
<td>12.0</td>
<td>97</td>
<td>19</td>
<td>47</td>
</tr>
<tr>
<td>(d) Calcifications*</td>
<td>72</td>
<td>71</td>
<td>10.8</td>
<td>98</td>
<td>29</td>
<td>28</td>
</tr>
<tr>
<td>(e) Vascularity type 2*</td>
<td>62</td>
<td>50</td>
<td>5.6</td>
<td>96</td>
<td>50</td>
<td>38</td>
</tr>
<tr>
<td>(f) A/T ≥ 1</td>
<td>76</td>
<td>60</td>
<td>8.3</td>
<td>98</td>
<td>39</td>
<td>24</td>
</tr>
<tr>
<td>Derived criteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/T ≥ 1 + at least two US features (b–d)</td>
<td>99</td>
<td>57</td>
<td>6.0</td>
<td>99</td>
<td>72</td>
<td>0.9</td>
</tr>
<tr>
<td>Hypo-echoic pattern + at least one US feature (c–e) (ref. 6)</td>
<td>79</td>
<td>61</td>
<td>8.8</td>
<td>98</td>
<td>39</td>
<td>23</td>
</tr>
</tbody>
</table>

PPV, positive predictive value; NPV, negative predictive value; A/T: anteroposterior/transverse diameter ratio. *see Methods for detail.
metastasis (7, 18, 34–37), and that no valid dimensional cut-off for neoplastic aggressive behaviour can be identified.

Many studies examined the question whether ultrasonographic characteristics of thyroid nodules may be useful indicators of histological malignancy. Overall, these investigations point out to a few ultrasonographic features that are significantly more frequent in malignant than in benign thyroid nodules, and some of them try to define a set of characteristics that identify those nodules which are at higher risk of malignancy.

Our data, obtained from a large series of nodules, confirmed that the presence of microcalcifications (multivariate OR = 9.9), blurred margins (OR = 7.7), hypo-echoic pattern (OR = 3.2) and intranodular vascular flow (OR = 1.7) were useful criteria of malignancy, as already reported by many, even if not all authors (6, 12–17). Moreover, as first suggested by Kim et al. and recently confirmed in a series of non-palpable thyroid nodules (22, 23), we showed that a lesion with a shape taller than wide is another useful criterion for malignancy (OR = 8.6), independently from the size.

Our study presents a large collection of thyroid nodules retrospectively investigated by US-FNAC for potential malignancy, regardless of size. It clearly indicates that the nodule size is only a weak predictor of histological malignancy. We suggested this a few years ago in an interim analysis of a subset of the present data (38) and it was subsequently confirmed by other authors (6, 16).

Consequently, the decision of submitting a thyroid nodule to FNAC just because it is “palpable”, or larger than 10 mm, cannot be considered a proper strategy.

Reasoning in terms of pure risk/benefit ratio for the individual patient, US-FNAC should be probably considered in all nodules, also for nodules <10 mm in diameter, taking into account (a) the remarkable prevalence of malignant lesions in SN, (b) the frequency of nodal metastasis, extracapsular growth and aggressive histological types found in microcarcinomas and (c) the limited incidence of minor side effects related to the procedure.

However, since cost considerations have always been important and have recently become even more relevant for clinical guidelines in many countries, it is desirable to outline a rational approach to the decision of performing FNAC when a thyroid nodule is detected at US examination.

Papini et al. (6), aiming to limit the number of lesions to submit to FNAC, suggested that FNAC on non-palpable nodules characterized by hypo-echoic appearance and at least one of the three additional US features (intranodular vascular images, blurred margins and microcalcifications), would have missed only 13% of carcinomas, by performing FNAC in just 31% of patients. The same criterion applied to our series of nodules gives a slightly worse performance, requiring FNAC in 39% of cases, but missing 23% of carcinomas.

Our study differs from the previous one (6) mainly because we included a very large set of nodules of any size, whereas Papini selected non-palpable nodules; in this respect, our set of data may give a better assessment of the general validity of US criteria.

According to our results, as recently reported in a series of non-palpable thyroid nodules (22), we confirmed that the best compromise between the risk of missing carcinomas of potential clinical aggressiveness and need for reducing the number of procedures would consist of submitting to FNAC those nodules presenting a shape taller than wide, with at least two of the following features: microcalcifications, blurred margins or hypo-echoic pattern. Such a criterion, applied to our large series of nodules, dramatically reduces the number of missed carcinoma to 0.9% while performing FNAC in 72% of patients.

We conclude that A/T ≥ 1 with at least two additional US features may currently represent the best compromise between missing cancers and cost-benefit considerations.

References
