A simple graphical quantitative analysis of ultrasonography images to decide when to perform fine needle aspiration biopsy in diagnosing malignancy in solid thyroid nodules? A two centres prospective study

Abstract
Objective: Since the prevalence of thyroid nodules is high and ultrasonography (USG)-guided fine-needle aspiration biopsy (FNAB) as a diagnostic means cannot be performed in all cases, we aimed to evaluate the feasibility and applicability of simple graphical analysis of USG two-dimensional images, to identify patients with suspicious thyroid nodules who would benefit from FNAB.

Subjects and methods: We studied prospectively 211 consecutive patients with thyroid nodules: 122 from the University Clinical Centre (UCC) of Maribor and 89 from the University Medical Centre (UMC) of Ljubljana who underwent USG-guided FNAB from January 2011 to October 2013. The cytology report was categorized as benign or suspicious/malignant. Blind to cytology reports, we later performed graphical analysis of USG images using ImageJ (version 1.48r) which is a public domain Java image processing and analysis programme. We compared the average gray value and standard deviation (SD) of the gray values used to generate the mean gray within the selection, with cytology reports.

Results: According to cytology reports, 24 thyroid nodules were suspicious/malignant (14/10) and 187 benign. Graphical analysis of USG images performed with ImageJ demonstrated significantly higher values of SD of the gray values used to generate the mean gray value in suspicious/malignant thyroid nodules as compared to unsuspicious nodules in both UCC Maribor and in UMC Ljubljana (P<0.001 and P=0.002, respectively). A higher value of the SD of gray value used to generate the mean gray value meant variation or dispersion from the average value and was correlated by the presence of micro-calcifications. By applying a cut-off level of the quotient between the SD value of an examined thyroid nodule and the SD value of normal/reference thyroid tissue of 1.20, we found that 21/24 nodules were classified as true positive and 114/187 as true negative.

Conclusion: Our results showed that our graphical quantitative analysis of USG images had a negative predictive value of more than 90% and was able to suggest which thyroid nodules were potentially malignant and needed further investigation.

Introduction

In adults, thyroid nodules are clinically diagnosed by a prevalence of 4%-7% [1]. It has been reported that non-palpable thyroid nodules have been found by ultrasonography (USG) in 67 of 100 studied subjects [2]. Most of the thyroid nodules are smaller than 1 cm and thus not clinically important [3]. The risk of malignancy in thyroid nodules is generally between 5%-15% [4, 5]. Out of all malignancies, thyroid cancer is relatively rare, about 1% [6] and so the high prevalence of thyroid nodules needs a cost-benefit diagnostic strategy [3]. Current management guidelines of the American Thyroid Association state that diagnostic USG should be performed in all patients with thyroid nodules and a fine needle aspiration biopsy (FNAB) should follow in potentially malignant nodules [7]. The mainstay of treatment in the vast majority of differentiated thyroid carcinoma patients is surgery, ideally gamma probe technetium-99m pertechnetate (99mTcO₄⁻) assisted completion thyroidectomy [8]. Neither thyroid USG nor thyroid scintigraphy can accurately predict the nature of thyroid nodules. The USG examination is highly available and effective in detecting nodules as small as several millimetres. Various USG findings such as marked hypoechogenicity, microcalcifi-
cations, height exceeding width, irregular margins, internal vascularity, extrathyroidal extension and lymphadenopathy are more frequent in malignant than in benign thyroid nodules. However, the sensitivities, specificities, negative and positive predictive values for these criteria are extremely variable from study to study, and no USG feature has both high sensitivity and high positive predictive value for thyroid malignancy. Diagnostic accuracy of FNAB increases when guided by USG [9, 10].

Furthermore, many of the above characteristics are highly observer-dependent and various studies attempted to make them more objective. Some studies tried to semiquantify echogenicity by the extraction of parameters based on gray-level histograms such as mean gray-scale level, standard deviation (SD) and the highest frequency gray-level [11]. Others compared mean thyroid density with that of the corresponding sternocleidomastoid muscle [12]. Other studies used the gray level histograms and compared them to different echogenicities [13]. Others tried to quantify the USG characteristics of masses or cysts of the thyroid gland [14]. Recently, it has been reported that computerized heterogeneity index can support the diagnosis of thyroid malignancy better than traditional USG evaluated by a human investigator [15]. Furthermore, an automatic system that classified the thyroid images and segmented the thyroid gland using machine learning algorithms was developed [16, 17]. Finally, a biometric system based on features extracted from thyroid tissue and accessed through two-dimensional (2D) USG was proposed [18, 19]. Several studies [11-19] were suggestive of deciding in which thyroid nodules we should perform FNAB rather than replace FNAB with picture graphical analysis of USG. Unfortunately, all these studies have not yet been applied in everyday practice.

In this prospective study we aimed to evaluate the feasibility and applicability of using a simple graphical analysis of two-dimensional USG images in the diagnostic procedure of solid thyroid nodules of any size, in order to avoid unnecessary FNAB.

Subjects and methods

Patients
Our prospective clinical study was approved by the Republic of Slovenia National Medical Ethics Committee. In the University Clinical Centre of Maribor (UCC) we studied 122 consecutive patients with thyroid nodules who were biopsied in accordance with the American Thyroid Association guidelines by the same internal medicine specialist from January 2011 to October 2013. In the University Medical Centre of Ljubljana (UMC), 89 patients with thyroid nodules biopsied between January 2013 and May 2013 were also included. All patients had solid or partly solid focal nodules. Patients with cysts or with nodules without solid component, patients with USG characteristics typical of autoimmune thyroid disorders and patients who were taking thyroid hormones or antithyroid drugs were excluded from the study.

Thyroid scintigraphy
For thyroid scintigraphy, $^{99m}$TcO$_4^-$ was used in both centres. Planar images of the thyroid were obtained within 15-30min after the intravenous (i.v.) injection of 200MBq of $^{99m}$TcO$_4^-$. In all patients FNAB under USG guidance was performed by one internal medicine specialist at the UCC Maribor and by four different internal medicine specialists at the UMC Ljubljana using a 21-gauge needle and a 20mL syringe. We used USG guidance to confirm the placement of the needle in the targeted nodule. For partially cystic nodules, sampling was directed towards the solid portion. Specimens were smeared on slides and dried in the air.

For thyroid scintigraphy, $^{99m}$TcO$_4^-$ was used in both centres. Planar images of the thyroid were obtained within 15-30min after the intravenous (i.v.) injection of 200MBq of $^{99m}$TcO$_4^-$. In all patients FNAB under USG guidance was performed by one internal medicine specialist at the UCC Maribor and by four different internal medicine specialists at the UMC Ljubljana using a 21-gauge needle and a 20mL syringe. We used USG guidance to confirm the placement of the needle in the targeted nodule. For partially cystic nodules, sampling was directed towards the solid portion. Specimens were smeared on slides and dried in the air.

Slides were examined by two experienced cytopathologists, one in each institution. The cytology report was descriptive and later categorized by two internal medicine specialists as benign or suspicious/malignant. The cytology report was considered as the final outcome of the diagnostic procedure.

Thyroid ultrasonography
In UCC Maribor, all USG investigations were performed with Sonosite Micromaxx 3.3 ultrasound and HLF38 linear transducer (6-13MHz). In UMC Ljubljana, all USG investigations were performed with Aloka Prosound C3 and linear transducer. The size and the echogenicity of thyroid nodules were evaluated. The volume of each thyroid nodule was calculated by the ellipsoid model (width x length x thickness x 1/6). The ratio between the length of the longest and the shortest axis of the thyroid nodule (the L/S axis ratio) was calculated. To minimize the inter-observer USG setting differences, we have chosen auto gain option which automatically adjusts gain on USG scanner.

Before FNAB we obtained two-dimensional images (2D mode) of the area where the FNAB was performed. Images of normal thyroid tissue or reference thyroid tissue at the same depth on the opposite side were also obtained. Images were saved in bitmap image (BMP) format as red green and blue (RGB) colour vector (colour images that can display 256 values in the red, green and blue channel).

Laboratory tests
In UCC Maribor, serum thyrotropin (TSH) concentration was measured using Roche Elecsys System with electrochemiluminescence immunoassay. Reference values for TSH were between 0.27 and 4.2mU/L. The within-run precision data expressed as coefficients of variation (CV%) was 3.3.

In UMC Ljubljana, serum concentration of TSH was measured using the ADVIA Centaur Immunoassay System (Siemens Medical Solutions Diagnostics). Reference values for TSH were between 0.35 and 5.5mU/L. The intraassay and interassay coefficients of variation ranged from 2.67% to 8.95% and from 3.75% to 12.81%, respectively.

Graphical analysis
After the acquirement of all USG images, we performed the ImageJ graphical analysis. Analysis was carried out by the same investigator who was unaware of cytology reports. ImageJ (version 1.48r) is a public domain Java image processing...
and analysis programme [21]. It runs on any computer with a Java 1.5 or later virtual machine. It can read many image formats including BMP. It is also able to calculate area and pixel value statistics of user-defined selections.

For each patient we selected the same size and depth regions of interest (ROI) in an ultrasound image of a thyroid nodule where the FNAB was taken, and in an image of normal or reference thyroid tissue (Figure 1). The following parameters were analyzed: a) Mean gray value (mean): Average gray value within the selection. This is the sum of the values of all the pixels in the selection divided by the number of pixels. For RGB images, the mean was calculated by converting each pixel to gray scale (black:0 and white:255) using the formula gray: (red+green+blue)/3 [21]. b) Standard deviation (SD): Standard deviation of the gray values used to generate the mean [21].

In order to diminish the inter-individual differences in the estimation of echogenicity using the mean value of thyroid gland as well as the differences in the USG machine properties we transformed the above-mentioned measurements into the quotients between the value obtained at the level of the patient’s thyroid nodule and the value obtained at the level of the patient’s normal or reference thyroid tissue.

**Statistical analysis**

Parameters specific for the thyroid nodule and those specific for normal or reference thyroid tissue were compared with the cytology reports. With the Student’s two-tailed t test or with the Mann Whitney’s U test, when the distribution was not normal, we compared measured values between the study groups.

Frequencies of different parameters were compared with the χ^2 test. Correlation was calculated with the Spearman’s test. Statistical analysis was performed with the Predictive Analytics SoftWare (PASW) Statistics 18.0 software package. P value below 0.05 was considered statistically significant.

**Results**

**Characteristics of patients with thyroid nodules**

As presented in Table 1, patients from UCC Maribor and UMC Ljubljana did not differ with respect to the ratio between benign and suspicious/malignant cytology reports.

With respect to gender, age and nodule volume the difference between the two centres was not statistically significant (Table 1).

The TSH level from UCC Maribor was significantly higher in patients with suspicious/malignant cytology reports than in patients with benign cytology reports. The same level of significance was not present in patients from UMC Ljubljana. With respect to TSH, neither the patients with benign cytology reports nor the patients with suspicious/malignant cytology reports differed significantly between UCC Maribor and UMC Ljubljana (Table 1).

Mean nodule L/S axis ratio in UCC Maribor was significantly lower in patients with suspicious/malignant cytology reports than in patients with benign cytology reports, the same was not true for patients from UMC Ljubljana. Additionally, in patients with benign cytology reports, the mean nodule L/S axis ratio was significantly larger in UCC Maribor than in UMC Ljubljana. However, in patients with suspicious/malignant cytology reports, L/S axis ratio did not differ between UCC Maribor and UMC Ljubljana (Table 1).

As for thyroid scintigraphy, in 168 (79.6%) patients the uptake in the nodule was lower than in the surrounding tissue, so they were considered as cold nodules. In 25 patients scintigraphy was not performed. In 16 patients the uptake in the nodule was similar as in the surrounding tissue. In 2 patients the uptake in the nodule was higher than in the surrounding tissue. All 24 nodules classified as suspicious/malignant on cytology report were among the first two groups.

**Graphical analysis of thyroid nodules and thyroid tissue**

Mean value obtained by graphical analysis of USG picture, of suspicious/malignant thyroid nodules was in 22/24 cases lower and in 2/24 cases higher than the value of the corresponding normal/reference thyroid tissue (Table 2). Similarly, 78.1% of benign nodules had higher and 21.9% lower mean value than normal/reference thyroid tissue. As evident from Table 2, mean values of benign thyroid nodules did not differ significantly from these of suspicious/malignant thyroid nodules in both centres. The same holds true for the mean values of normal/reference thyroid tissue in both groups of patients. However, patients from UCC Maribor had significantly lower mean values of normal/reference thyroid tissue than patients from UMC Ljubljana.
from UMC Ljubljana and statistically had not significantly lower mean values of thyroid nodules (Table 2). Nevertheless, the quotients between mean values of the thyroid nodules with suspicious/malignant cytology reports and normal/reference thyroid tissues were not significantly different when compared with the quotients between mean values of the thyroid nodules with benign cytology reports and normal or reference thyroid tissues in both centres.

Standard deviation values of suspicious/malignant thyroid nodules were significantly higher than SD values of benign thyroid nodules in both centres (Table 2). Between suspicious and malignant thyroid nodules we found no significant differences in graphical analysis parameters. Furthermore, in both centres, SD values of all 24 suspicious/malignant thyroid nodules were higher than SD values of the corresponding normal/reference thyroid tissue within the same gland. As for normal/reference thyroid tissue, SD values did not differ between patients with benign thyroid nodules and patients with suspicious/malignant thyroid nodules in both centres. Accordingly, the quotient between SD values of suspicious/malignant thyroid nodules and SD values of normal/reference thyroid tissue was significantly higher than the quotient between SD values of benign thyroid nodules and SD values of normal/reference thyroid tissue in both centres (Table 2).

Additionally, patients with benign as well as patients with suspicious/malignant cytology reports from UCC Maribor had significantly higher SD values of thyroid nodules than patients from UMC Ljubljana (Table 2). In patients from UCC Maribor, we also found significantly higher SD values of normal/reference thyroid tissue in patients with benign cytology reports but not in patients with suspicious/malignant cytology reports when compared with patients from UMC Ljubljana (Table 2). However, quotients between SD values of suspicious/malignant thyroid nodules and SD values of normal/reference thyroid tissue were similar in both centres. Similarly, quotients between SD values of benign thyroid nodules and SD values of normal/reference thyroid tissue did not differ between both centres.

When applying a cut-off level of the quotient between SD value of thyroid nodule and SD value of normal/reference thyroid tissue of 1.20, we found that 21/24 nodules were classified as true positive and 114/187 as true negative. The sensitivity and the specificity of the quotient as a predictor for the suspicious/malignant or benign cytology report were 87.5% and 61.0%, respectively. Furthermore, the positive and the negative predictive values for the quotient as a predictor were 22.3% and 97.4%, respectively.

**Discussion**

Graphical analysis of USG images of thyroid nodules performed with ImageJ demonstrates significantly higher value of SD of the gray values used to generate the mean gray value in thyroid nodules suspicious for thyroid carcinoma as compared to non-suspicious by cytology report nodules. Although the average mean values and SD values in both centres were different as the result of different USG scanners and their settings, the quotients between SD values of thyroid nodules and thyroid tissue in both centres were similar. Therefore, our method is robust and almost independent of

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>UCC Maribor</th>
<th>UMC Ljubljana</th>
</tr>
</thead>
<tbody>
<tr>
<td>All pts with FNAB</td>
<td>122</td>
<td>89</td>
</tr>
<tr>
<td>Cytology report</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benign</td>
<td>109</td>
<td>78</td>
</tr>
<tr>
<td>Suspicious/malignant</td>
<td>7/6</td>
<td>7/4</td>
</tr>
<tr>
<td>N of women</td>
<td>92</td>
<td>9</td>
</tr>
<tr>
<td>N of men</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>Age (Years) (mean±SD)</td>
<td>53.7±14.9</td>
<td>53.9±14.4</td>
</tr>
<tr>
<td>TSH (mU/L) (median, range)</td>
<td>1.21</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>0.01-8.81</td>
<td>0.01-8.50</td>
</tr>
<tr>
<td>Nodule volume (mL) (median, range)</td>
<td>3.56</td>
<td>3.98</td>
</tr>
<tr>
<td></td>
<td>0.11-69.08</td>
<td>0.10-33.49</td>
</tr>
<tr>
<td>L/S axis ratio (mean±SD)</td>
<td>b1.83±0.51</td>
<td>1.61±0.43</td>
</tr>
</tbody>
</table>

**UCC**: University Clinical Centre; **UMC**: University Medical Centre; **Pts**: patients; **FNAB**: fine-needle aspiration biopsy; **N**: number; **SD**: standard deviation; **TSH**: thyrotropin; **L/S axis ratio**: ratio between the length of the longest and the shortest axis of the thyroid nodule.

*P<0.05 when compared with benign, *b*P<0.001 when compared with patients from UMC Ljubljana
the level of normal or reference thyroid tissue. a P<0.001 when compared with benign, b P<0.05 when compared with benign, c P<0.001 and mean at the level of normal or reference thyroid tissue; SD quotient: calculated ratio between SD at the level of nodule and SD at the level of tissue used to generate the mean gray value; Mean quotient: calculated ratio between mean at the level of nodule and the value of SD of the normal/reference thyroid tissue; SD quotient: calculated ratio between SD at the level of nodule and SD at the level of normal or reference thyroid tissue. a P<0.001 when compared with benign, b P<0.05 when compared with benign, c P<0.001 when compared with patients from UMC Ljubljana, d P<0.05 when compared with patients from UCC Maribor.

The incidence of potentially malignant thyroid nodules in our study was 11.4% which is in accordance with other reports [4, 5]. Out of the 24 cases 10 were categorized by cytology report as almost certainly malignant and 14 as suspicious. Between these two groups there were no significant differences in graphical analysis parameters.

There is increasing evidence that higher serum TSH concentrations are found in patients with thyroid nodules harbouring malignancy [22]. Our results from UCC Maribor, but not from UMC Ljubljana, were in accordance with these observations. The discrepancy could be explained with the lower number of patients from UMC Ljubljana and with the different methods for the determination of TSH concentration.

Spherical shape was reported to be associated with an increased rate of malignancy in solid thyroid nodules [23]. Accordingly, in our study, the nodules long-to-short axis ratio was significantly lower in suspicious/malignant thyroid nodules than in benign thyroid nodules in UCC Maribor but not in UMC Ljubljana. Therefore, in UCC Maribor, more round nodules were associated with suspicious/malignant cytological features. The discrepancy between the two centres could be associated with the inter-individual differences in measurements between four internal medicine specialists in UMC Ljubljana, while in UCC Maribor all measurements were performed by the same internal medicine specialist.

Most studies have revealed that the majority of thyroid malignancies demonstrate a marked hypoechogenic USG pattern, yet most hypoechogenic nodules are benign in view of the high prevalence of benign lesions [24]. Our study showed that nodules categorized as suspicious/malignant in cytology report were statistically not significantly different in terms of hypoechogenicity estimated by the mean gray value as compared to nodules categorized as benign. The main reason for this is that hypoechogenicity is a predictor of malignancy and all specialists involved in the study predominantly performed FNAB of visually hypoechogenic nodules which have darker USG pattern than the normal thyroid tissue in the same gland.

Standard Deviation of the gray values quantifies the existence of variation or dispersion from the average value. This characteristic is almost impossible to accurately spot visually and shows a good correlation with the presence of micro-calcifications. Micro-calcifications are the most predictive USG characteristic of thyroid malignancy and are associated with a cancer risk of 82 per 1000 [25]. According to our data, graphical analysis of USG images performed with ImageJ showed that the SD of thyroid nodules with suspicious/malignant cytological features was significantly higher than SD of thyroid nodules with benign cytological features, in both centres. Therefore, suspicious thyroid nodules probably had more micro-calcifications than benign thyroid nodules. It was probably safe to assume that nodules with SD equal to or lower than those of the corresponding thyroid normal/reference tissue have a higher likelihood of being benign. Additionally, the SD value was proven to be a much better marker of thyroid malignancy than the mean value.

By using quotients between the value of SD of the thyroid nodule and the value of SD of the normal/reference thyroid tissue we were able to overcome variations in properties of different USG scanners since quotients between both centres...
did not differ significantly while they differed between benign and suspicious/malignant thyroid nodules. Quotient was shown to be especially valuable as a negative predictive factor with a higher probability of not having a suspicious/malignant cytology report if its value was below 1.20. We need to stress that our results for the suspicious/malignant nodules derive from a limited number of 24 cases compared to 187 benign nodules, which makes predictive values statistically weak.

In comparison with other recently published methods [15], our graphical analysis of USG images is based on open source software and does not require any preloaded database of similar cases. From the practical point of view, the studied graphical analysis can be easily and quickly performed. Image analysis of USG only takes five minutes with no extra cost, which makes the method also very cost-effective compared with FNAB.

In conclusion, with a simple, fast, cost-effective and user-friendly quantification of the USG image, as described in this paper, it was possible to predict benign thyroid cytology with more than 90% negative predictive value, and thus differentiate the potentially malignant nodules which needed USG guided FNAB. This type of USG image quantification is feasible in routine clinical practice.

The authors declare that they have no conflicts of interest.

Bibliography