

Research paper

Comparison of color-Doppler and qualitative and quantitative strain-elastography for differentiation of thyroid nodules in daily practice

Manuela Götzberger,¹ Susanne Krueger,² Roland Gärtner,¹
Martin Reincke,¹ Matthias Pichler,¹ Gerald Assmann,³
Thomas Mussack,⁴ Michael Zimmermann⁵, Veit Gülberg¹

¹Medizinische Klinik und Poliklinik IV, Klinikum der Universität München; ²Klinikum Bogenhausen; ³Pathologisches Institut, Klinikum der Universität München; ⁴Klinik für Viszeral- und Endokrine Chirurgie - Campus Innenstadt, Klinikum der Universität München; ⁵Maria-Theresia-Klinik; Munich, Germany

ABSTRACT

OBJECTIVE: Strain-elastography provides a new ultrasound-based method that can offer information about the stiffness of thyroid nodules as an indicator of malignancy. The aim of our study was to compare the utility of color-Doppler and strain-elastography in differentiating between benign and malignant nodules. **DESIGN AND METHODS:** 77 thyroid nodules (70 benign and 7 malignant) from 70 unselected patients (48 female/22 male, mean age 49.7±14.3 years) were evaluated with color-Doppler and elastography based on a five-scale elastogram score for qualitative elastography and strain ratio for quantitative elastography. As reference tissue we chose normal thyroid tissue [strain ratio a (SR a)] and cervical muscles [strain ratio b (SR b)]. The cytological or histological results were used as a reference standard. Diagnostic performances of qualitative and quantitative elastography were compared using ROC curves. **RESULTS:** Vascularization score 3 or 4 was associated with malignancy ($p=0.024$) as well as elastogram score 4 or 5 ($p=0.070$, n.s.s.). SR a was indicatively higher and SR b lower in the group of malignant nodules ($p=0.065$ and $p=0.246$, n.s.s.). The best cut-off points predicting malignancy were 3.32 for SR a (66.7% sensitivity, 83.3% specificity) and 0.10 for SR b (71.4% sensitivity, 67.1% specificity). **CONCLUSION:** In our study, the accuracy of elastography did not surpass other sonographic parameters in differentiating thyroid nodules. The technique can play a role as a supplementary parameter in assessment of malignancy to improve diagnostic efficacy. The best parameter is SR a, but SR b can serve as an alternative if SR a is not assessable.

Key words: Elasticity score, Strain ratio, Thyroid cancer, Ultrasound

Address for correspondence:

Manuela Götzberger, M.D., Klinikum Freising, Alois-Steinecker-Str. 18, D-85354 Freising, Germany, Tel.: 0049-8161-245316, Fax: 0049-8161-244957, E-mail: goetzbergerm@klinikum-freising.de; manuela_goetzberger@yahoo.de

Accepted: 15-01-2016, Accepted: 18-05-2016

INTRODUCTION

Ultrasound is the most important diagnostic tool to detect thyroid pathologies. Thyroid nodules occur with a prevalence of 35-45% and are therefore very frequently detected during ultrasound examination.¹ Malignant diseases however account only for 4%-6% of all thyroid nodules.² The accuracy of conventional ultrasound in the differentiation between benign and malignant nodules is not very high.³ Fine-needle-aspiration (FNA) as an additional tool to optimize cancer diagnosis has some limitations as well. The result depends on the expertise of the cytologist and a significant number of findings remain indeterminate.⁴ Therefore, evaluation of further sonographic parameters for differentiation of benign and malignant thyroid nodules is still of importance. Manual palpation is the basic approach to detect thyroid nodules since a hard consistency upon palpation is associated with malignancy. In recent years, real-time strain-elastography (SE) has become established as a new ultrasound method allowing an objective determination of tissue stiffness. The tissue deformation is produced by an external palpation with the probe. The result is coded in color-scale images displayed over the B-mode images, which are related to tissue stiffness. For classification, on the one hand a visual categorical score is used. On the other hand, regions of interest (ROIs) can be drawn over target areas and adjacent healthy thyroid tissue to calculate a semiquantitative strain ratio (SR) coefficient.⁵

In the past few years, some studies have analyzed the diagnostic accuracy of elastogram scores as well as SR in thyroid nodules. However, most published studies have used qualitative SE score analysis. Only a minority of data provide a quantitative analysis of SR. Therefore, a recent meta-analysis has stressed the need for greater use of SR measurement in future studies.⁶ In addition, there has been till now no standardized elastographic score and methods vary between different studies. Thus, the aim of our study was to characterize the efficacy of elastograms in different pathological nodules and assess the utility of qualitative and quantitative SE in differentiating benign from malignant thyroid nodules in unselected patients attending an endocrinological outpatient clinic.

SUBJECTS AND METHODS

Patients

We prospectively examined 110 patients admitted to the Department of Endocrinology and Diabetology of the University of Munich for further examination of thyroid nodules. Exclusion criteria were age below 18 years, previous surgical or radiation therapy in the neck, nodule size below 5 mm, completely cystic lesions, nodules with macrocalcifications and lack of written informed consent. In 40 patients 44 nodules showing typical benign features on ultrasound examination without compressive symptoms were identified. Therefore, these patients did not undergo FNA or surgery. Finally, 70 patients with 77 nodules were enrolled in the next stage of the study (48 female/22 male, mean age 49.7 ± 14.3 years). The volumes of the nodules ranged from 4 ml to 110 ml. Sixty-four patients were sent to surgery and in 6 patients FNA of the thyroid nodule was performed. The cytological or histological results were used as a reference standard. Patients with negative FNA were observed over a minimum of 6 months with control ultrasound to confirm the benign nature of the nodule. All patients provided written informed consent for the study protocol before enrollment. The study was conducted according to the Declaration of Helsinki (DRKS ID 00007796).

Conventional ultrasound and SE

Conventional ultrasound and SE were performed using the same linear 10-MHz probe (Hitachi Medical Systems, EUB-8500, Tokyo, Japan). SE examination was performed following the conventional B-mode session. The probe was placed on the slightly extended neck with light pressure. All examinations and documentations were performed by one skilled sonographer with 8 years of experience in ultrasound and 3 years of experience in elastography.

The target nodule was evaluated for size, position within the lobe, echogenicity, margin, internal microcalcifications, halo sign and cystic components. Nodule size was measured in three dimensions [depth (D), width (W) and length (L)] from two section scans (transverse and sagittal). Subsequent volume (V) was calculated by the formula of the ellipsoid as $V = D \times W \times L$ (in cm) $\times 0.524$. Blood flow pattern was

judged with color-Doppler on a five-point scale: 0 no detectable flow, 1 partial peripheral flow, 2 complete peripheral flow, 3 rare internal flow, 4 rich internal flow. A score of 3 and 4 showing internal flow was judged as sign of malignancy.

Evaluation based on elastogram scores

To provide an adequate elastogram image, pressure degree was illustrated on a four-point numerical scale, which was displayed laterally on the elastogram. An appropriate pressure was defined as a pressure ranging from levels 3-4 on the scale during the examinations. The ultrasound elastogram was displayed over the B-mode image on a color map that ranges from red (softest components) to blue (hardest components). Each evaluation of the elastogram score was repeated by means of compression and relaxation cycles at least three times until the findings were confirmed to be reproducible for the investigator. The elastographic score was assigned according to the color pattern of the nodule, with a classification into a five-point categorical scale (score 1: homogeneously green, score 2: predominantly green with a few blue areas at the margin of the nodule, score 3: predominantly green with a few blue areas in the center of the nodule, score 4: predominantly inhomogeneous blue with a few green areas, score 5: homogeneously blue) (Figure 1 a-b). Based on previous results, it was postulated that a score of 4 and 5 with predominantly hard areas indicates malignancy.⁷

Evaluation based on SR

A region of interest (ROI) A was placed inside the nodule. Then, a ROI B was selected at the same depth as the target lesion as reference in normal thyroid tissue SR a. As second reference tissue we chose cervical muscles SR b, which appeared homogeneously blue. Additionally, we measured SR b in 30 thyroid lobes of healthy volunteers as reference for a standard value (Figure 1c). SR coefficient was calculated in real time by the ultrasound software. SR coefficient was measured 3 times in each nodule for both reference tissues and mean value was used for analysis. Images were saved on an electronic documentation system.

Statistical analysis

Data with normal distribution are expressed as mean \pm standard deviation. We compared the distribu-

tion of patient characteristics between the groups of benign and malignant nodules using a non-parametric t-test (Mann-Whitney-U-Test) for quantitative variables and Pearson's Chi-square test in the event of categorical variables. Statistical significance was assumed for $p < 0.050$. Receiver operating characteristic (ROC) curves were analyzed and the diagnostic performance of different variables was expressed as the area under the curve (AUC). Statistical analysis was performed using SPSS software (Version 15.0; SPSS Inc., Chicago, USA).

RESULTS

In total, 70 benign and 7 malignant nodules were diagnosed. In the group of benign nodules histology showed 62 nodular goiters and 8 adenomas. Among the malignant lesions 3 follicular carcinomas, 2 medullary carcinomas, 1 papillary carcinoma and 1 metastatic breast carcinoma were detected. There was no statistically significant difference regarding the volume of the nodules between the two groups. A vascularization score of 3 or 4 in color-Doppler evaluation was significantly associated with malignancy (57% vs 16%, $p=0.024$).

Elastogram score and SR

Among the group of benign nodules 77% were scored 1-3 in the visual elastogram score; however, 16 (23%) nodules were scored 4 or 5, suggesting already present malignancy. Three out of the 7 malignant lesions (43%) were scored between 1-3, including 2 cases of follicular carcinoma and 1 metastatic carcinoma. The remaining 4 (57%) malignant cases (1 follicular carcinoma, 2 medullary carcinomas and 1 papillary carcinoma) showed hard elastogram scores of 4 or 5, indicating malignancy. The mean value of SR a was indicatively higher in the group of malignant nodules (3.80 ± 1.95 vs. 3.45 ± 5.35 , $p=0.065$). SR b was lower in malignant than in benign nodules, but the difference was not statistically significant (0.15 ± 0.09 vs. 0.20 ± 0.33 , $p=0.246$) (Table 1).

ROC analysis

The AUC of the elastography score was 0.683 and the AUC of SR a and b was 0.731 and 0.635, respectively. Comparison of ROC curves between the different parameters is shown in Figure 2. The

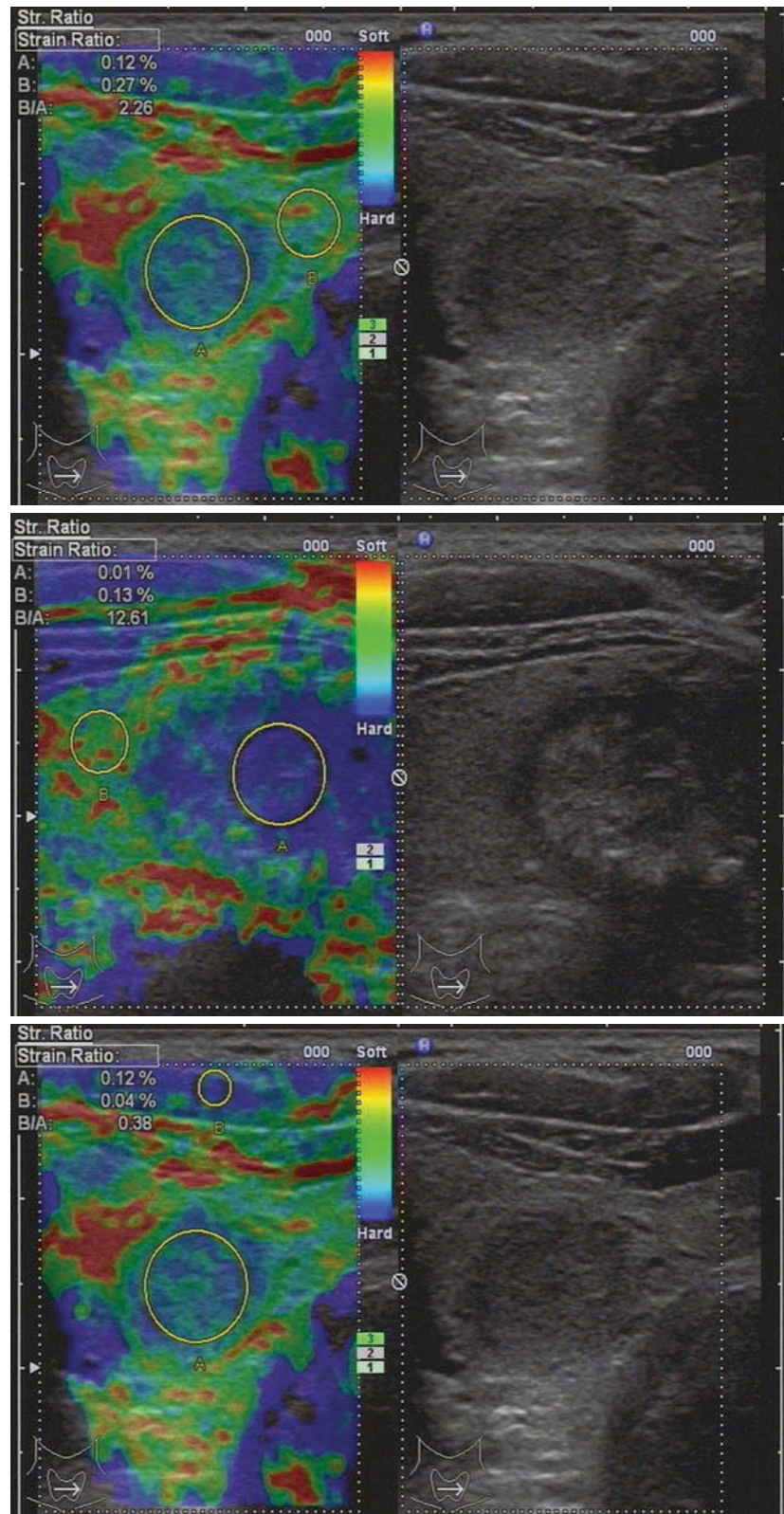
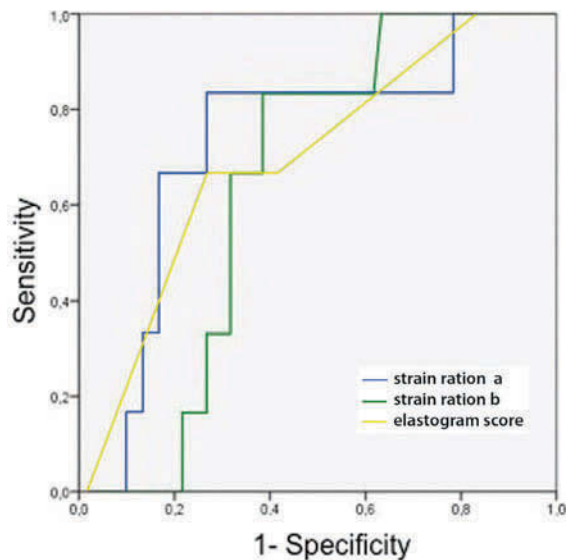


Figure 1. Examples of elastograms with representative elastogram scores. **a.** A case of nodular goiter. Elastogram was scored 2. SR a was 2.26. **b.** A case of medullary carcinoma with elasticity score 4. SR a was 12.61. **c.** Evaluation of SR b = 0.38 in a benign nodule with reference ROI in normal cervical muscle.

Table 1. Ultrasound and elastography features in the groups of benign and malignant nodules according to final histology

n=77	Benign nodules (n=70)	Malignant nodules (n=7)	p
Volume (ml)	7.8 ± 8.8	16.0 ± 22.9	0.328
Vascularization			
0-2	59 (84%)	3 (43%)	0.024
3-4	11 (16%)	4 (57%)	
Elastogram Score			
1-3	54 (77%)	3 (43%)	0.070
4-5	16 (23%)	4 (57%)	
Strain Ratio			
SR a	3.45 ± 5.35	3.80 ± 1.95	0.065
SR b	0.20 ± 0.33	0.15 ± 0.09	0.246

best cut-off point for differentiation of benign from malignant nodules was 3.5 for the elastogram score with a sensitivity of 57.1% and specificity of 77.1%. For SR a the best cut-off point was 3.32 (sensitivity

**Figure 2.** Comparison of the Receiver Operating Characteristic (ROC) curves between SR a, b and elastogram score.**Table 2.** AUC analysis for the different parameters

Parameter	AUC	Cut-off point	Sensitivity	Specificity	95% CI
Elasto-Score	0.683	3.50	57.1%	77.1%	0.487-0.878
SR a	0.731	3.32	66.7%	83.3%	0.526-0.935
SR b	0.635	0.10	71.4%	67.1%	0.474-0.796

of 66.7% and specificity of 83.3%) and 0.10 for SR (b) (sensitivity of 71.4% and specificity of 67.1%) (Table 2).

DISCUSSION AND CONCLUSION

For assessment of malignancy in thyroid nodules no conventional US parameter has a diagnostic accuracy higher than 75%.¹ Nevertheless, in national guidelines ultrasound is recommended as the main diagnostic tool in the management of thyroid nodules.^{8,9} SE provides a new ultrasound-based method that can offer information about the hardness of the nodule as an indicator of malignancy.

A recent meta-analysis showed an overall mean sensitivity for differentiation of thyroid nodules of 79% for the elastogram score and 85% for SR assessment. The overall specificity was 77% for the elastogram score and 80% for SR assessment.⁶ SR assessment achieved higher results than the elastography score, but the difference was not significant. In our study the sensitivity was only 57% and 67% for the elastogram score and SR assessment with a comparable specificity of 77% and 83%, respectively. Sun et al. observed a high heterogeneity among the data for the elastography score as well as SR.⁶ Sensitivity of the elastography score differed between 36% and 100%,^{10,11} whereas for SR assessment sensitivity was estimated between 69% and 98%.^{12,13} These divergent findings include a number of publications that report less promising results, thus raising the question of the real clinical utility of ultrasound SE.^{12,14} Numerous studies demonstrated high sensitivities in a highly selected population with a prevalence of malignancy up to 46%.¹³ In this respect, the present study is limited to a low malignancy rate of 9%, which may result in a lower sensitivity. However, it probably reflects the real situation in daily clinical practice more accurately than other studies selecting a high number of malignant tumors.

Malignant nodules in our collective had a higher elastogram score and SR a compared to benign results. However, the difference in elastogram score was not statistically significant, which is in contrast to the vascularization score. SR a was indicatively higher in the group of malignant nodules and showed a considerable trend towards significance ($p=0.069$). The probable reason for this borderline result is a relatively high number of follicular carcinomas (3/7) within the group of malignant nodules, whereas in most other studies the malignant masses were mainly papillary carcinomas.^{15,16} In the meta-analysis of Sun et al only a total of 11 follicular carcinomas were reported, leading to 10 missed diagnoses by elastography evaluation.⁶ In our patients two follicular carcinomas displayed features of benign nodules in qualitative and quantitative elastography evaluation. Only one follicular carcinoma showed conclusive malignant results in conventional ultrasound as well as in elastography evaluation, probably due to advanced tumor stage. For small follicular carcinomas elastography seems not to be an effective diagnostic tool.

The AUC of SR a was higher than that of the elastogram score (0.731 vs. 0.697), which is in accordance with other results revealing a better accuracy for SR than for the qualitative elastogram score.¹⁵ Several studies used special offline software-systems for quantitative analysis, which are not feasible in daily practice.^{10,17} Quantitative real-time elastography up to now has been investigated in four studies using SR comparing thyroid nodule to normal thyroid tissue.^{13,15,16,18} Cut-off levels were described in a range between 3.79 and 6.66. So far, no precise cut-off point for nodule-to-gland SR has been defined. Therefore, Sun et al underlined the need for further studies investigating the efficacy of SR to differentiate between thyroid nodules.⁶ Our results reveal 3.32 as the best cut-off point for nodule-to-gland SR, which is in line with most of the results proposing cut-off levels between 3.79 and 4.22.^{13,15,18} Nevertheless, further evaluation of SR measurement is needed to determine the optimal cut-off value.

To date, only two studies have calculated the nodule-to-muscle SR.^{19,20} A cut-off value of 1.5 was described as a predictor of malignancy in the study of Kagoya et al.¹⁹ The parameter showed a sensitivity of 90% and a specificity of only 50%.¹⁹ The study of

Aydin et al revealed a cut-off value of 1.85.²⁰ Diagnostic performance for SR measurement showed a sensitivity of 96% and a specificity of 93% with good agreement between parenchyma-to-nodule SR and muscle-to-nodule SR. Only one follicular carcinoma was found among the malignant cases (4%), which might explain the higher diagnostic accuracy achieved in this study compared to our work.²⁰

Our data revealed a cut-off below 0.10 to assess malignancy. The variability between these results may be due to the fact that the elasticity of cervical muscles can vary depending on the patients position. In the center of the muscles we usually find homogeneously blue areas in elastogram images. By choosing these areas as reference, we have achieved reproducible results between malignant nodules, benign nodules and normal tissue. However, measurements were not significantly different and the AUC was lower than for nodule-to-gland SR. Furthermore, in the study of Aydin et al, sonoelastography was performed using a different ultrasound system (Aplio XG, Toshiba), whereas Kagoya's and our work was conducted with the more common Hitachi system. The cut-off levels of SR may not be directly comparable between the two systems and standard values still have to be evaluated in further studies.

Since we also included multinodular goiters in our study, in 9% of the patients nodule-to-gland SR was not quantifiable because of a lack of normal thyroid tissue. Several studies included only a single solid nodule within one thyroid lobe for elastography evaluation.²¹ This selection may also account for better results in these studies. Determination of nodule-to-muscle SR as an alternative technique enables quantitative SE measurement even in multinodular goiters, which is a frequent clinical finding. Hence, it is of great interest to further evaluate this method in more cases and to optimize cut-off values.

Recently, the first results in shear wave elastography of thyroid nodules as an alternative method of elastography were described.^{22,23} This method can provide dynamic quantitative measurement of tissue stiffness. Shear wave speed is reported to be higher in malignant than in benign nodules mainly investigated in patients submitted to surgery.²³ Evaluation in a study population representing daily clinical practice as well

as correlation with SR measurement is lacking up to now. Therefore, it remains unclear whether shear wave elastography is superior to strain elastography.

In conclusion, strain elastography has been under continuous development over the past few years and is of considerable relevance in the diagnosis of thyroid nodules. It is convenient and fast in clinical use. However, in spite of the initial optimism, several studies reporting less promising results must be mentioned. The use of this technique can play a role as a supplementary parameter together with conventional ultrasound criteria in assessment of malignancy. Therefore, it can help to determine which nodule meets concordant criteria of benignancy and can be followed up and in which case FNA is needed. Furthermore, it may be used to guide follow-up of FNA negative lesions.²⁴ Some limitations need to be taken into account, e.g. the technique seems insufficient to reliably detect follicular carcinomas. Additionally, our results verify that intranodular blood-flow demonstrated on color flow Doppler is also a finding strongly predicting malignancy and should not be disregarded in favor of new developments. In our investigation, a cut-off point of 3.32 for nodule-to-gland SR and 0.10 for nodule-to-muscle SR achieved the best results for differentiation of thyroid nodules. To optimize the validity of the method in the future, a generally accepted standardized elastogram score as well as the determination of cut-off values for SR measurement are needed.

CONFLICT OF INTEREST

The authors do not have any conflict of interest.

REFERENCES

1. Moon WJ, Jung SL, Lee JH, et al, 2008 Benign and malignant thyroid nodules: US differentiation – multicenter retrospective study. *Radiology* 247: 762-770.
2. Hegedüs L, Bonnema SJ, Bennedbaek FN, 2003 Management of simple nodular goiter: current status and future perspectives. *Endocr Rev* 24: 102-132.
3. Gharib H, Papini E, Paschke R, et al, 2010 American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and European Thyroid Association Medical Guidelines for Clinical Practice for the Diagnosis and Management of Thyroid Nodules. *Endocr Pract* 16: 1-43.
4. Tee YY, Lowe AJ, Brand CA, Judson RT, 2007 Fine-needle aspiration may miss a third of all malignancy in palpable thyroid nodules: a comprehensive literature review. *Ann Surg* 246: 714-720.
5. Bamber J, Cosgrove D, Dietrich CF, et al, 2013 EFSUMB Guidelines and Recommendations on the clinical use of ultrasound elastography. Part 1: basic principles and technology. *Ultraschall in Med* 34: 169-184.
6. Sun J, Cai J, Wang X, 2014 Real-time ultrasound elastography for differentiation of benign and malignant thyroid nodules. *J Ultrasound Med* 33: 495-502.
7. Fukunari N, 2007 More accurate and sensitive diagnosis for thyroid tumors with elastography. *Medix* 20: Suppl 3: 16-19.
8. Pacini F, Schlumberger M, Dralle H, Elisei R, Smit JWA, Wiersinga W and the European Thyroid Cancer Taskforce, 2006 European consensus for the management of patients with differentiated thyroid carcinoma of the follicular epithelium. *European Journal of Endocrinology* 154: 787-803.
9. American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer, Cooper DS, Doherty GM, Haugen BR, et al, 2009 Revised American Thyroid Association Management Guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid* 19: 1167-1214.
10. Lyschchik A, Higashi T, Asato R, et al 2005 Thyroid gland tumor diagnosis at US elastography. *Radiology* 237: 202-211.
11. Mansor M, Okasha H, Esmat S, Hashem AM, Attia KA, Hussein HE, 2012 Role of ultrasound elastography in prediction of malignancy in thyroid nodules. *Endocr Res* 37: 67-77.
12. Unlütürk U, Erdogan MF, Demir O, Güllü S, Baskal N, 2012 Ultrasound elastography is not superior to grayscale ultrasound in predicting malignancy in thyroid nodules. *Thyroid* 22: 1031-1038.
13. Xing P, Wu L, Zhang C, Li S, Liu C, Wu C, 2011 Differentiation of benign from malignant thyroid lesions: calculation of the strain ratio on thyroid sonoelastography. *J Ultrasound Med* 30: 663-669.
14. Lippolis PV, Tognini S, Materazzi G, et al, 2011 Is Elastography actually useful in the presurgical selection of thyroid nodules with indeterminate cytology? *J Clin Endocrinol Metab* 96: E1826-E1830.
15. Ning CP, Jiang SQ, Zhang T, Sun LT, Liu YJ, Tian JW, 2012 The value of strain ratio in differential diagnosis of thyroid solid nodules. *Eur J Radiol* 81: 286-291.
16. Cakir B, Ersoy R, Cuhaci FN, et al, 2014 Elastosonographic strain index in thyroid nodules with atypia of undetermined significance. *J Endocrinol Invest* 37: 127-133.
17. Cantisani V, D'Andrea V, Biancari F, et al, 2012 Prospective evaluation of multiparametric ultrasound and quantitative elastosonography in the differential diagnosis of benign and malignant thyroid nodules: preliminary

- experience. *Eur J Radiology* 81: 2678-2683.
18. Wang H, Brylka D, Sun L-N, Lin Y-Q, Sui G-Q, Gao J, 2013 Comparison of strain ratio with elastography score system in differentiating malignant from benign thyroid nodules. *Clinical Imaging* 37: 50-55.
 19. Kagoya R, Monobe H, Tojima H, 2010 Utility of elastography for differential diagnosis of benign and malignant thyroid nodules. *Otolaryngology-Head and Neck Surgery* 143: 230-234.
 20. Aydin R, Elmali M, Polat AV, et al, 2014 Comparison of muscle-to-nodule and parenchyma-to-nodule strain ratios in the differentiation of benign and malignant thyroid nodules: which one should we use? *Eur J Radiology* 83: e131-e136.
 21. Ding J, Cheng H, Ning C, Huang J, Zhang Y, 2011 Quantitative measurement for thyroid cancer characterization based on elastography. *J Ultrasound Med* 30: 1259-1266.
 22. Xu JM, Xu X-H, Xu H-X, et al, 2014 Conventional US, US Elasticity Imaging, and Acoustic radiation force impulse imaging for prediction of malignancy in thyroid nodules. *Radiology* 272: 577-586.
 23. Veyrieres J-B, Albarel F, Vaillant Lombard J, et al, 2012 A threshold value in Shear Wave elastography to rule out malignant thyroid nodules: a reality? *Eur J Radiology* 81: 3965-3972.
 24. Cosgrove D, Piscaglia F, Bamber J, et al, 2013 EFSUMB Guidelines and Recommendations on the clinical use of ultrasound elastography. Part 2: Clinical applications. *Ultraschall in Med* 34: 238-253.