

Quantitative assessment of shear-wave ultrasound elastography in thyroid nodules: diagnostic performance for predicting malignancy

Hana Kim · Jeong-Ah Kim · Eun Ju Son · Ji Hyun Youk

Received: 8 January 2013 / Revised: 18 February 2013 / Accepted: 25 February 2013 / Published online: 19 April 2013
© European Society of Radiology 2013

Abstract

Objectives To evaluate the predictability of shear-wave ultrasound elastography (SWE) for thyroid malignancy and to compare the diagnostic performances of SWE and B-mode US.

Methods Retrospective review of 99 patients who underwent SWE before US-guided fine-needle aspiration of thyroid nodules was performed. SWE elasticity indices of the mean (E_{mean}), maximum (E_{max}), and minimum (E_{min}) of nodules were measured. Diagnostic performance of SWE was compared with that of B-mode US.

Results Among a total of 99 nodules, 21 were papillary thyroid carcinoma (PTC) and 78 were benign. E_{mean} , E_{max} , and E_{min} were significantly higher in PTCs than in benign nodules ($P < 0.001$). Sensitivity and specificity for predicting malignancy were 76.1 %, 64.1 % with E_{max} (65 kPa), 61.9 %, 76.1 % with E_{min} (53 kPa), and 66.6 %, 71.6 % with E_{mean} (62 kPa). Areas under the ROC curves (A_z) of SWE values were not significantly different from those of US categories on B-mode US. However, combining E_{mean} and E_{min} with B-mode US of probably benign lesions improved the specificity ($P = 0.02, 0.007$) for predicting PTC. **Conclusions** The quantitative parameter of SWE was significantly higher in PTC than in benign nodules, and combined use of quantitative SWE and B-mode US provided higher specificity for predicting malignancy.

Key Points

- Quantitative shear-wave elastography (SWE) helps differentiate benign from malignant nodules
- SWE and conventional ultrasound have comparable diagnostic performance for predicting thyroid malignancy
- Combined quantitative SWE and B-mode ultrasound is highly specific for thyroid malignancy

Keywords Shear-wave elastography · Thyroid nodule · Elasticity · Ultrasound · Papillary thyroid carcinoma

Introduction

Thyroid ultrasonography (US) and fine-needle aspiration (FNA) biopsy play an important role in differentiating benign from malignant thyroid lesions. In expert centres, FNA provides helpful results in 65–75 % of examined nodules [1]. Approximately 60–70 % of aspirates are proven to be benign and 5 % are positive for papillary thyroid carcinoma (PTC); 5–15 % of aspirates are persistently non-diagnostic. The remaining 15–25 % of aspirates is indeterminate or suspicious [2]. FNA is limited by non-diagnostic aspirates and difficulty in interpretation of cytology due to the overlap of morphological signs between benignity and malignancy [2]. The poor quality of FNA specimens may be the source of diagnostic errors, with false-negative and false-positive results that reached from 0 % to 13 % and 10 %, respectively [3, 4]. Therefore, there is a demand for another tool for evaluating thyroid nodules.

Elastography is a new dynamic tool that uses ultrasound to evaluate tissue stiffness by measuring the degree of distortion under the application of an external force [5–7]. Elastography utilises tissue strain caused by compression,

H. Kim · J.-A. Kim · E. J. Son · J. H. Youk
Department of Radiology, Gangnam Severance Hospital,
Yonsei University College of Medicine, Seoul, South Korea

J.-A. Kim (✉)
Department of Radiology, Gangnam Severance Hospital,
211, Eonjuro, Gangnam-Gu, Seoul 135-720, South Korea
e-mail: chrismd@hanmail.net

which is estimated by pre- and post-compression ultrasonic signals [5]. Elastography with freehand compression has several limitations because it is highly dependent on the organ's compressibility limits or on the extent of tissue compression and is operator dependent. Recently, a new method has been developed that uses tracking of shear-wave propagation through tissue to obtain the elastic modulus [8]. This shear-wave elastography (SWE) is operator-independent, reproducible, and quantitative. Previous reports suggested that SWE may add a new dimension to ultrasound evaluation of thyroid nodules [9]. They suggested that SWE can overcome the limitations of strain elastography with high reproducibility and quantitative elasticity measurement.

The purposes of this study were to evaluate the predictability of quantitative SWE for thyroid malignancy and to compare the diagnostic performance of SWE and B-mode US for differentiating benign and malignant thyroid nodules.

Materials and methods

This study was conducted with institutional review board approval and a waiver of patient informed consent.

Patients

From July 2011 to February 2012, 113 thyroid nodules underwent SWE before US-FNA for thyroid lesions visible on greyscale US. Fourteen thyroid nodules were excluded because of indeterminate cytology and lack of follow-up FNA cytology ($n=9$) or lack of surgical pathologic results ($n=5$). Thyroid nodules with satisfactory cytological evaluation or pathological diagnosis made by surgery were included. A total of 99 thyroid nodules in 99 patients (mean age, 45.7 years; range, 25–77 years) were included in this study.

US examination

Thyroid US examinations including B-mode US and SWE were performed with the Aixplorer US system (SuperSonic Imagine, Aix-en-Provence, France) equipped with a 15–4-MHz linear-array transducer. B-mode US features of thyroid nodules were prospectively recorded and assessed in US categories before the time of FNA by the radiologists who performed the US examinations and FNA at our institution. Suspicious US features included marked hypoechogenicity (more hypoechoic than the surrounding anterior strap muscles of neck), poorly defined margins, microcalcification, and taller than wide shape [9, 10]. When thyroid nodules showed one or

more of these suspicious US features, they were assessed as suspicious for malignancy, and the other nodules were assessed as probably benign if they had no suspicious US features [10, 11].

After obtaining B-mode US, SWE images were obtained for the thyroid nodules that were scheduled to be aspirated. The built-in region of interest (ROI) (Q-box; Super Sonic Imagine) of the system was set to include the lesion and the surrounding normal tissue, which demonstrated the semitransparent colour map of the tissue stiffness overlaid B-mode image with a range from dark blue, indicating the lowest stiffness, up to red, indicating the highest stiffness (0–180 kPa). Fixed 2×2 or 1×1 -mm ROIs were placed by an investigator over the stiffest part of the lesions noted on colour overlay images, including the immediately adjacent stiff tissue or halo, selecting one of the two fixed ROIs depending on the lesion size or extent of the SWE colour map. The system calculated the mean (E_{mean}), minimum (E_{min}), and maximum (E_{max}) elasticity indices (EI) in kPa for the lesions. A second ROI of the same size was placed in the normal thyroid parenchyma and anterior strap muscle. The elastic ratios of mean stiffness for the lesion-to-normal parenchyma ($E_{\text{mean-p}}$) and the lesion-to-strap muscle ($E_{\text{mean-m}}$) were calculated.

Statistical analysis

The elasticity values of all lesions were correlated with the pathologic diagnosis of nodules. The EIs of E_{mean} , E_{max} , E_{min} , $E_{\text{mean-p}}$, and $E_{\text{mean-m}}$ for lesions were compared between benign and malignant thyroid nodules using Student's *t*-test or Mann–Whitney U-test for unpaired data. For all analyses, two-tailed *P* values of less than 0.05 were considered to be statistically significant. Correlations between SWE indices and lesion size were analysed using Spearman's rank correlation coefficients. Receiver-operating characteristic (ROC) curve analysis was performed for each elasticity value to predict malignancy, and optimal SWE cutoff values yielding the maximal sum of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were calculated. ROC curves for the B-mode US categories and EI of SWE were analysed to compare the diagnostic performances of US categories with EI for predicting malignancy. The logistic regression model was used to perform ROC analysis of a combined set of US categories and parameters on SWE. Areas under the ROC curves (*Az*) were calculated and compared between the B-mode US categories and the parameters of SWE images, and we evaluated the potential effect of adding SWE EI to categories of B-mode US. The McNemar test was used for paired comparison of proportions

Table 1 Comparison of elasticity indices of papillary thyroid carcinoma (PTC) and benign thyroid nodules

SWE Elasticity indices (kPa)	PTC	Benign	<i>P</i> value
E_{mean}	85.52±41.94 (21–166)	51.46±22.75 (10–128)	<0.001
E_{max}	100.00±51.07 (23–218)	59.71±28.10 (17–174)	<0.001
E_{min}	60.04±30.04 (18–140)	40.56±20.28 (1–86)	<0.001
$E_{\text{mean-p}}$	3.70±2.31 (1.41–10.33)	3.53±2.62 (1.22–9.32)	0.841
$E_{\text{mean-m}}$	6.34±17.3 (0.89–77.5)	1.87±0.93 (0.46–4.9)	0.275

E_{mean} : mean elasticity index of the stiffest portion of mass or surrounding tissue

E_{max} : maximum elasticity index of the stiffest portion of mass or surrounding tissue

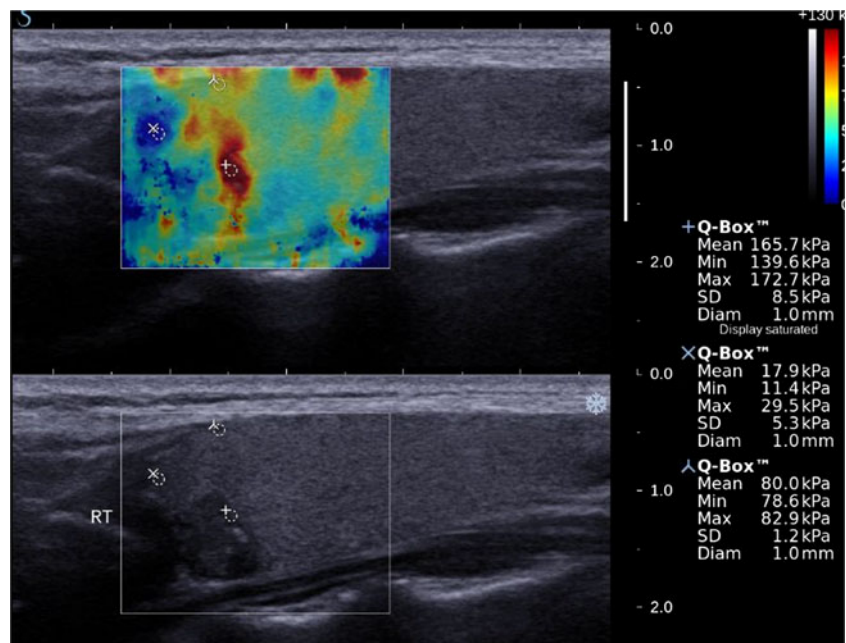
E_{min} : minimum elasticity index of the stiffest portion of mass or surrounding tissue

$E_{\text{mean-p}}$: the ratio of mean elasticity score of the lesion and parenchyma

$E_{\text{mean-m}}$: the ratio of mean elasticity score of the lesion and strap muscle

(sensitivity, specificity, and accuracy) (PASW Statistics, version 18.0.0; SPSS, Chicago, IL, USA; MedCalC, version 10.3.0.0, MedCalC Software, Mariakerke, Belgium; MAGREE SAS Macro, SAS Institute Inc., Cary, NC, USA). Statistically significant differences between Az values were reported at 95 % confidence intervals. The mean differences were regarded as being statistically significant at the 5 % level when the corresponding confidence interval did not encompass zero.

Fig. 1 A 43-year-old woman with surgery-proven papillary thyroid carcinoma. Thyroid sonography (US) shows a markedly hypoechoic nodule with irregular margins that was assessed as a suspicious nodule on greyscale US. Shear-wave elastography (SWE) displays heterogeneous colour elasticity signal with high SWE areas, such as E_{mean} of 165.7 kPa, E_{min} of 139.6 kPa, and E_{max} of 172.7 kPa



Results

Pathologic result

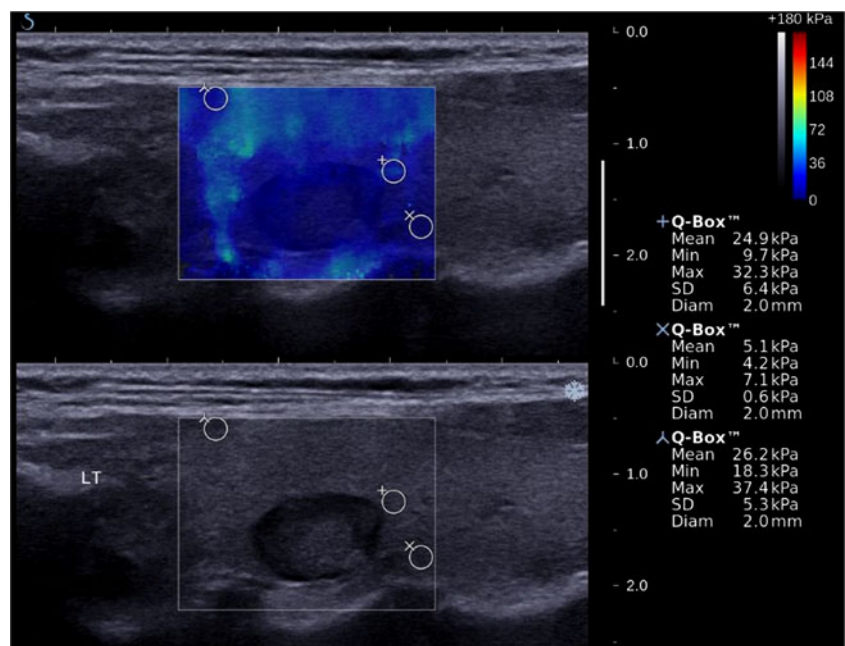
Pathologic diagnoses of thyroid nodules were obtained by FNA or surgery. Twenty-one nodules out of 99 were malignant and 78 were benign. Twenty-one nodules were confirmed as PTC by surgery; among them, 20 nodules were conventional PTC and 1 nodule was a follicular variant PTC. Seventy-eight benign lesions were confirmed by two consecutive FNA cytology tests with 3–6 month (mean 4.5 month) interval periods between the two consecutive FNAs.

SWE findings

E_{mean} , E_{max} , and E_{min} were significantly higher in malignant nodules than in benign nodules ($P<0.001$) (Table 1) (Figs. 1 and 2). The differences of $E_{\text{mean-p}}$ and $E_{\text{mean-m}}$ between malignant and benign nodules were not statistically significant ($P=0.841$ and 0.275) (Table 1). Nodule size was not correlated with SWE indices including benign and malignant thyroid nodules (Spearman's correlation coefficient, 0.01–0.27; $P=0.29$ –0.95).

The optimal sensitivity and specificity cutoff values for E_{mean} , E_{max} , and E_{min} are presented and the diagnostic performances for predicting malignancy are listed using PPV, NPV, and accuracy in Table 2. The AUC (Az) values are listed in Table 3. The Az values of all SWE parameters were not significantly different from that of US categories on B-mode US ($P>0.05$) (Table 3).

Fig. 2 A 54-year old woman with FNA-proven benign follicular thyroid nodule. Thyroid sonography (US) shows a well-defined and oval-shaped nodule with hypoechoogenicity that was assessed as a probably benign nodule on B-mode US category. The majority of the nodule displays low stiffness with mean elasticity (E_{mean}) of 24.9 kPa on SWE



There were no significant differences in sensitivity, accuracy, and AUC (Az) values between B-mode US and combined use of SWE EI and B-mode US concerning probably benign thyroid nodules (Table 4). However, upgrading the probably benign lesions on B-mode US to suspicious lesions if they have a greater value than the cutoff value in E_{mean} (=62 kPa), E_{min} (=65 kPa), or E_{max} (=53 kPa) has increased the specificity for predicting malignancy significantly (Table 4). The specificity increased from 59.7 % with the US category on B-mode US to 82.5 % adding E_{min} with the US category.

Discussion

SWE is a newly developed method that uses tracking of shear-wave propagation through tissue to obtain the elastic modulus [8]. Previous report suggested that SWE may add a new dimension to ultrasound evaluation of thyroid nodules [9]. There have been only a few studies about SWE evaluation of thyroid nodules, and reference standards of SWE indices have not been established. Sebag et al. reported that significantly higher EI was noted in malignant thyroid nodules than benign nodules, and they reported the sensitivity and specificity of SWE were 85.2 and 93.9 % using a cutoff

Table 2 Diagnostic performance of elasticity indices for predicting papillary thyroid carcinoma on SWE

Elasticity indices (kPa)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
Cutoff values					
E_{mean} 62	66.6	71.6	40.6	85.7	69.3
E_{max} 65	76.1	64.1	37.8	86.2	65.9
E_{min} 53	61.9	76.1	44.4	85.2	72.7

E_{mean} : mean elasticity index of the stiffest portion of the mass or surrounding tissue
 E_{max} : maximum elasticity index of the stiffest portion of the mass or surrounding tissue
 E_{min} : minimum elasticity index of the stiffest portion of the mass or surrounding tissue
 PPV: positive predictive value
 NPV: negative predictive value

Table 3 Comparison of diagnostic performances of B-mode US and SWE for predicting thyroid malignancy

Parameters	AUC (95 % CI)	P value (compared with B-mode)
E_{mean}	0.768 (0.619-0.917)	0.114
E_{max}	0.759 (0.631-0.887)	0.136
E_{min}	0.695 (0.557-0.833)	0.131
US category	0.823 (0.786-0.924)	

E_{mean} : mean elasticity index of the stiffest portion of the mass or surrounding tissue
 E_{max} : maximum elasticity index of the stiffest portion of the mass or surrounding tissue
 E_{min} : minimum elasticity index of the stiffest portion of the mass or surrounding tissue
 AUC: areas under the ROC curves (numbers in parentheses are 95 % confidence intervals)

Table 4 Diagnostic performance of combined use of elasticity indices of SWE with probably benign thyroid nodules on US category for predicting malignancy

Parameters	Specificity	<i>P</i> -value [#]	Sensitivity	<i>P</i> -value [#]	Accuracy	<i>P</i> -value [#]	AUC	<i>P</i> -value [#]
US category	59.7 %		90.5 %		67 %		0.823	
US+E _{mean} *	80.0 %	0.02	50.0 %	0.26	78.6 %	0.153	0.650	0.693
US+E _{max} **	72.5 %	0.167	50.0 %	0.26	71.4 %	0.609	0.613	0.589
US+E _{min} ***	82.5 %	0.007	50.0 %	0.26	81.0 %	0.077	0.663	0.730

*Probably benign nodules on US category were upgraded to suspicious nodules if the mean elasticity value was 62 kPa or more

**Probably benign nodules on US category were upgraded to suspicious nodules if the maximum elasticity value was 65 kPa or more

***Probably benign nodules on US category were upgraded to suspicious nodules if the minimum elasticity value was 53 kPa or more

[#]*P*-value was that to test the null hypothesis that there is no change in sensitivity/specificity/accuracy/AUC with addition of SW elasticity value

E_{mean}: mean elasticity index of the stiffest portion of the mass or surrounding tissue

E_{max}: maximum elasticity index of the stiffest portion of the mass or surrounding tissue

E_{min}: minimum elasticity index of the stiffest portion of the mass or surrounding tissue

level of 65 kPa [12]. They reported higher diagnostic performance with a combined score (SWE+B-mode US) than that of B-mode US only. Other SWE studies [12, 13] showed variable cutoff values yielding a maximal sum of diagnostic performance to predict thyroid malignancy.

We found that the E_{mean} was significantly higher in PTC (85.52 kPa±41.94) than in benign thyroid nodules (51.46 kPa±22.75; *P*<0.001). The E_{max} and E_{min} were also significantly higher in malignant (100.00 kPa±51.07, 60.04 kPa±30.04) than in benign (59.71±28.10, 40.56±20.28) nodules (*P*<0.001). Our result was compatible with the previous reports [11, 13]. The reported sensitivity and specificity ranged between 15.7 % and 65.4 % and between 58.2 % and 95.3 % with real-time elastography [14]. In this study, the sensitivity and specificity of SWE for predicting malignancy were 66.6 % and 71.6 using a cutoff value of 62 kPa in E_{mean}, and the sensitivity and specificity were 76.1 %, 64.1 % with the cutoff value of 65 kPa in E_{max}. Our study provided higher sensitivity for predicting thyroid malignancy than the reports with real-time elastography [14]. However, our data showed lower sensitivity and specificity compared with the previous study that reported 85.2 % sensitivity and 93.9 % specificity using SWE [12]. In our study, 49.4 % of nodules were less than 1 cm and 16.2 % of nodules were associated with macro-calcification. The accuracy of EI measurement might be altered if the thyroid nodules are small or are associated with macro-calcification or egg shell calcification [12, 14]. It was reported that the nodules associated with macro-calcifications showed a high false-positive rate for malignancy on SWE [12]. It was also reported that higher elasticity indices were noted in larger lesions for both malignant and benign lesions using real-time elastography [14, 15]. In another study, size was not correlated with SWE indices for papillary cancer [13]. In this study, lesion size was not correlated with SWE values

for either benign or malignant thyroid nodules. Further prospective studies with larger case scales about SWE values' correlation with the size of thyroid nodules are necessary.

A previous study [13] about SWE for thyroid nodules also reported that SWE ratios (calculated by dividing the SWE of the lesion by the SWE of normal parenchyma) were higher in papillary cancers than in benign thyroid lesions. We measured the elasticity at the thyroidal nodules and surrounding parenchyma to calculate the mean stiffness ratios for lesion-to-normal parenchyma (E_{mean-p}) and lesion-to-strap muscle elasticity (E_{mean-m}), but there were no statistically significant differences between the ratios of malignant and benign thyroid nodules. We assume the reasons for these observations are that the width and stiffness of strap muscles and the condition of thyroid parenchyma surrounding the nodules are different according to the individual patients, and we did not standardise these factors for comparison.

It has been reported that static elastography is not appropriate for differentiating between benign and malignant thyroid nodules when underlying diffuse thyroidal disease such as multinodular goitre or thyroiditis is associated with the thyroidal nodule [2, 7]. Lymphocytic infiltration and fibrosis, which modify thyroidal structure, may result in a change of thyroidal stiffness [16]. As the estimation of elasticity is altered by the presence of a nearby hard area, the elasticity of a nodule on static elastography may not be clearly distinguishable from nearby thyroid parenchyma [7, 16]. This could be a limitation of elastographic US for the diagnosis of thyroid nodules.

Previous studies have reported that static elastography alone or the combined use of elastography and B-mode US showed inferior diagnostic performance to B-mode US features in distinguishing benign and malignant thyroid nodules [14]. Our study reported that the diagnostic

performances of E_{mean} , E_{max} , and E_{min} on SWE were comparable to that of US categories on B-mode US for predicting thyroid malignancy. Furthermore, we did not compare the value of SWE for predicting PTC with US-FNA but only with B-mode US. A review of recently published data regarding thyroid cancer detection at US-FNA indicates a sensitivity of 76 %–98 %, specificity of 71 %–100 %, false-negative rate of 0 %–5 %, false-positive rate of 0–5.7 %, and overall accuracy of 69 %–97 % with the use of this method by Kim et al. [4]. Further studies on the correlation of diagnostic performance between SWE or combined SWE with B-mode US and US-FNA could be needed.

Although SWE did not provide superior diagnostic performance for predicting malignancy, SWE still provides adjunctive information about the lesions as a quantitative and more objective method to conventional US. In our study, combined use of quantitative SWE EI based on the greyscale US category provided significantly higher specificity for predicting PTC (Table 4). Therefore, with the quantitative EI information, we could reassess the US category for the probably benign lesions on B-mode US and we expect unnecessary FNAB in benign thyroid nodules could be avoided. However, combined use of quantitative SWE EI based on greyscale US for probably benign thyroid nodules in our study had relatively low sensitivity and Az values even though it had a higher specificity. Therefore, further prospective studies are necessary for suggesting a clinical guideline. In this study, we observed that malignant nodules showed higher EI than the benign nodules and we suggest that SWE can guide targeting the thyroid nodules during US-FNA by providing the quantitative information about the tissue elasticity and composition of lesions.

This study had several limitations. First, selection bias may exist because patients included in our study were scheduled for US-FNA for known thyroid nodules with suspicious US features, which may have affected the diagnostic performance on B-mode US. Second, SWE was performed by one radiologist and the interobserver variability for SWE was not assessed. Third, elastic indices of lesions were not correlated with the histopathological findings. In our study, all malignant thyroid nodules were PTC and follicular carcinoma was not included. A previous study reported that elasticity values show differences according to the histologic subtypes of thyroid malignancy [12]. For example elasticity values of follicular neoplasm were higher than those of papillary carcinoma, which are composed of small micro-follicles with variable amounts of colloid, affecting their elasticity and echogenicity. Therefore, further prospective studies with larger case scales on the correlation between elasticity values on SWE and various histologies of thyroid nodules are necessary. Finally, long-term follow-up

for the benign lesions with US-FNAB was not performed in this study.

In conclusion, quantitative EI of SWE was significantly higher in PTC than in benign thyroid nodules, and combined use of quantitative SWE and B-mode US provided higher specificity for predicting malignancy.

References

- Gharib H, Goellner JR (1993) Fine-needle aspiration biopsy of the thyroid: an appraisal. *Ann Intern Med* 118:282–289
- Alexander EK (2008) Approach to the patient with a cytologically indeterminate thyroid nodule. *J Clin Endocrinol Metab* 93:4175–4182
- Lewis CM, Chang KP, Pitman M, Faquin WC, Randolph GW (2009) Thyroid fine needle aspiration biopsy: variability in reporting. *Thyroid* 19:717–723
- Kim MJ, Kim EK, Park SI et al (2008) US-guided fine-needle aspiration of thyroid nodules: indications, techniques, results. *Radiographics* 28:1869–1886
- Park SH, Kim SJ, Kim EK, Kim MJ, Son EJ, Kwak JY (2009) Interobserver agreement in assessing the sonographic and elastographic features of malignant thyroid nodules. *AJR Am J Roentgenol* 193:416–423
- Lyshchik A, Higashi T, Asato R et al (2005) Thyroid gland tumor diagnosis at US elastography. *Radiology* 237:202–211
- Rago T, Vitti P (2008) Role of thyroid ultrasound in the diagnostic evaluation of thyroid nodules. *Best Pract Res Clin Endocrinol Metab* 22:913–928
- Bercoff J, Tanter M, Fink M (2004) Supersonic shear imaging: a new technique for soft tissue elasticity mapping. *IEEE Trans Ultrason Ferroelectr Freq Control* 51:396–409
- Slapa RZ, Piwowonski A, Jakubowski WS et al (2012) Shear wave elastography may add a new dimension to ultrasound evaluation of thyroid nodules: case series with comparative evaluation. *J Thyroid Res* 2012:657147
- Kim EK, Park CS, Chung WY et al (2002) New sonographic criteria for recommending fine-needle aspiration biopsy of nonpalpable solid nodules of the thyroid. *AJR Am J Roentgenol* 178:687–691
- Moon HJ, Kwak JY, Kim MJ, Son EJ, Kim EK (2010) Can vascularity at power Doppler US help predict thyroid malignancy? *Radiology* 255:260–269
- Sebag F, Vaillant-Lombard J, Berbis J et al (2010) Shear wave elastography: a new ultrasound imaging mode for the differential diagnosis of benign and malignant thyroid nodules. *J Clin Endocrinol Metab* 95:5281–5288
- Bhatia KS, Tong CS, Cho CC, Yuen EH, Lee YY, Ahuja AT (2012) Shear wave elastography of thyroid nodules in routine clinical practice: preliminary observations and utility for detecting malignancy. *Eur Radiol* 22:2397–2406
- Moon HJ, Sung JM, Kim EK, Yoon JH, Youk JH, Kwak JY (2012) Diagnostic performance of gray-scale US and elastography in solid thyroid nodules. *Radiology* 262:1002–1013
- Tan GH, Gharib H (1997) Thyroid incidentalomas: management approaches to nonpalpable nodules discovered incidentally on thyroid imaging. *Ann Intern Med* 126:226–231
- Magri F, Chytiris S, Capelli V et al (2012) Shear wave elastography in the diagnosis of thyroid nodules: feasibility in the case of coexistent chronic autoimmune Hashimoto's thyroiditis. *Clin Endocrinol (Oxf)* 76:137–141