Diagnostic Role of Conventional Ultrasonography and Shearwave Elastography in Asymptomatic Patients with Diffuse Thyroid Disease: Initial Experience with 57 Patients

Injoong Kim, Eun-Kyung Kim, Jung Hyun Yoon, Kyung Hwa Han, Eun Ju Son, Hee Jung Moon, and Jin Young Kwak

1Department of Radiology, Research Institute of Radiological Science, Yonsei University College of Medicine, Seoul; 2Biostatistics Collaboration Unit, Medical Research Center, Yonsei University College of Medicine, Seoul; 3Department of Radiology, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, Korea.

Purpose: Thyroid ultrasonography (US) is a useful diagnostic tool in the evaluation of diffuse thyroid disease (DTD), whereas shearwave elastography is a dynamic technique that can provide information about tissue hardness by using acoustic shearwaves remotely induced by a focused ultrasonic beam. This study aims at investigating the role of conventional US and shearwave elastography in the diagnosis of asymptomatic patients with DTD. Materials and Methods: Fifty-seven patients who underwent both conventional US and shearwave elastography were included in this study. Interobserver variability of the three radiologists in assessment of underlying thyroid echogenicity on conventional US was analyzed. Diagnostic performances for diagnosing DTD on conventional US and shearwave elastography were calculated and compared. Results: Fair agreement was observed in the identification of DTD with conventional US (kappa value=0.27). The area under the receiver operating characteristic curve (Az) were 0.52-0.585 on conventional US by three radiologists. The Az values when using the mean and maximum elasticity values as a diagnostic criteria for DTD were 0.619 and 0.59 on shearwave elastography. Patients with DTD showed higher mean [24.1±10 kilo-Pascals (kPa)] and maximum (36.4±13.3 kPa) elasticity values on shearwave elastography when compared to the normal group (23.4±10.8 kPa and 33.7±12.4 kPa, respectively), although without statistical significance (p=0.802 and p=0.452, respectively). Conclusion: Conventional US did not show reliable interobserver agreement in the diagnosis of DTD. Although not statistically significant, shearwave elastography may provide additional information in the diagnosis of DTD. Therefore, larger prospective studies are needed to define the values of shearwave elastography for diagnosing DTD.

Key Words: Ultrasonography, shearwave elastography, thyroid gland, diffuse thyroid disease

INTRODUCTION

Thyroid ultrasonography (US) is a useful diagnostic tool in the evaluation of thy-
roid disease including thyroid nodules and diffuse thyroid disease (DTD).1-9 Hashimoto thyroiditis (HT) and Graves disease (GD) are included in DTD, and known US features of DTD are decreased or increased diffuse heterogeneous parenchymal echogenicity, a coarse echotexture, micronodulation, and scattered microcalcifications.4,6,7,10,11 US can provide valuable information on the course of subclinical hypothyroidism (SCH) and is also being used in the evaluation of DTD in patients with thyroid functional abnormality with thyroid autoantibodies such as thyroid peroxidase antibody (TPOAb) and thyroglobulin antibody (TgAb).5,12,13 Rosário, et al. analyzed 117 patients with SCH, and reported that both diffuse hypoechoigenicity on US and the presence of TPOAb increase the risk of SCH developing into overt hypothyroidism.15

US elastography is a newly developed dynamic technique that provides information about tissue hardness by using a degree of distortion under the application of an external force or by using shear acoustic waves remotely induced by a focused ultrasonic beam.14-16 This method has been applied in the differentiation of benign and malignant thyroid lesions.14-17 Previous studies have demonstrated the diagnostic role of US used in DTD diagnosis.5,7 However, to the best of our knowledge, no reports have shown the interobserver agreement of US or the diagnostic performances of shearwave elastography in DTD diagnosis. In this study, we investigated the role of conventional US and shearwave elastography in the diagnosis of asymptomatic patients with DTD.

**MATERIALS AND METHODS**

**Patients**

The institutional review board of our institution approved this retrospective study and required neither patient approval nor informed consent for the review of images and records.

This study was performed at our institution (a referral center) during October 2010. During this period, 94 consecutive patients underwent US-fine needle aspiration (FNA). Twenty nine patients were excluded from the original 94 patients because they did not undergo shearwave elastography. Among these patients, 28 patients made the personal choice not to undergo any additional examinations, while 1 patient did not have shearwave elastography performed due to an insufficient amount of underlying thyroid parenchyma being observed as there were multiple nodules in both thyroid glands. Eight additional patients were also excluded because they did not undergo thyroid autoantibody tests. The final study included 57 patients. All patients were initially referred not just for evaluation of thyroid function abnormality, but also for US-FNA on thyroid nodules. Patient age ranged from 26 to 72 years (mean, 49.9 years). Mean age of male patients was 50.4 years (range 44-55 years) and that of female patients was 49.8 years (range 26-72 years).

**Conventional US and shearwave elastography of the thyroid**

Real-time conventional US and shearwave elastography were performed using a linear transducer with a central frequency of 4 to 15 MHz (ShearWave™ Elastography, Supersonic Imagine, Aix-en-Provence, France). Real-time US was performed by one of three board-certified radiologists with 5 to 13 years of experience in thyroid imaging. First, a careful evaluation of the underlying echogenicity of the thyroid on conventional US was made. After examinations using conventional US, shearwave elastography was performed by the same radiologists who had performed the conventional US. Regions of interest for elasticity images were set to include more than 1/3 of a thyroid lobe for sufficient amount of the underlying thyroid parenchyma and not to include thyroid nodule or nodules on either transverse or longitudinal plane (Figs. 1 and 2). By measuring the speed of shearwave propagation at every point in the image, a quantitative elasticity value was expressed in kilo-Pascals (kPa).

We prospectively recorded the quantitatitive elasticity value of the underlying thyroid gland on shearwave elastography, without considering the US features of the thyroid nodules targeted for US-FNA. Data were entered in a computer database both for patient care and this study.

Thyroid parenchyma was retrospectively analyzed for US features of DTD when the thyroid gland showed one or more of the following suspicious findings: diffuse heterogeneous echogenicity or diffuse hypoechoogenicity with/without micronodules, thyroid enlargement [anteroposterior (AP) diameter of thyroid on longitudinal scan >2 cm], multiple linear echogenicity, and scattered microcalcifications6,7,18,19 by three radiologists, (E.J.S, H.J.M, and E.K.K.) with 5 to 13 years of experience in thyroid US independently. They were fully blinded to the serologic test results for all patients. Each blinded reader evaluated the patients’ images separately and independently.

**Measurement of anti-thyroid autoantibodies**

In all patients, thyroid function was evaluated using venous
blood samples. Thyroid autoantibodies as well as thyroid function were routinely measured at our institution to evaluate thyroid functional status. Serum thyroid peroxidase antibody (TPOAb), thyroglobulin antibody (TgAb) and thyroid-stimulating hormone (TSH)-binding inhibitory immunoglobulins (TBII) levels were measured by radioimmunoassay (Brahms, Hennigsdorf/Berlin, Germany). The existence of each TPOAb and/or TgAb was defined with a serum concentration of thyroid autoantibody >60 IU/L. Patients with HT were defined as those who were positive for TPOAb and/or TgAb. Activity of TSH receptor antibody (TRAb) was measured using a first generation TRAb kit prepared by Brahms (Brahms, Hennigsdorf/Berlin, Germany). A value of TBII exceeding 10% was considered positive. Patients with GD were defined as those who were positive for TBII. In this study, we evaluated patients with DTD including both HT and GD.

Statistical analysis

Patient medical records were reviewed for clinical characteristics. The Fisher’s exact test was used to compare gender to DTD. We also evaluated the association of age and the values of mean and maximum elasticity on shearwave elastography to DTD using the independent two-sample t-test. Interobserver variability was estimated using the generalized kappa statistic of Landis and Koch. The degree of agreement was categorized in terms of kappa values: 0 corresponds to no agreement, 1.00 to complete agreement, less than or equal to 0.20 to slight agreement, 0.21-0.40 to fair agreement, 0.41-0.60 to moderate agreement, 0.61-0.80 to substantial agreement, and 0.81-1.00 to near perfect agreement.

The diagnostic performances (sensitivity, specificity, accuracy, positive predictive value, negative predictive value) for diagnosing DTD on conventional US were calculated. The area under the receiver operating characteristic curve (Az) was calculated by the three radiologists on conventional US.

We compared the mean and maximum elasticity values between the patients of the DTD group and the patients of the normal serology group. To evaluate diagnostic performances of diagnosing DTD on shearwave elastography, the cut-off value was calculated by using the Youden’s index. The Az was calculated by using the continuous data of the mean and maximum elasticity.

All statistical analysis was performed with SAS software (version 9.1.3; SAS Institute, Cary, NC, USA). Statistical significance was assumed when the p value was less than 0.05. All reported p values are 2-sided.

RESULTS

Of the 57 patients, DTD was diagnosed in 22 patients using serum TPOAb or TBII. The mean age (50.1±11.6 years) of
patients with DTD was older than that (49.7±10.4 years) of patients without DTD ($p=0.887$). Gender was not associated with DTD ($p=0.639$).

Reproducibility of conventional US observations was assessed between three observers who analyzed the same images. If only patients diagnosed with heterogeneous echogenicity by all three radiologists were considered, the generalized kappa value was 0.27, which means a fair degree of agreement.

The sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) of the three radiologists were 18.2-59.1%, 57.1-94.3%, 57.9-66.7%, 44.4-71.4% and 62.5-69.0%, respectively, on conventional US for the identification of DTD. The $A_z$s were 0.52-0.585 on conventional US by three radiologists. Also, the $A_z$s were not statistically different between the three radiologists ($p=0.387$) (Table 1).

In the 35 (61.4%) patients consisting of the normal serology group, mean elasticity value was 23.4±10.8 kPa (range, 9.0 kPa to 54.5 kPa), and mean maximum elasticity value was 33.7±12.4 kPa (range, 11.0 kPa to 61.4 kPa). In the 22 (38.6%) patients of the DTD group, mean elasticity value was 24.1±10.0 kPa (range, 9.3 kPa to 53.6 kPa), and mean maximum elasticity value was 36.4±13.3 kPa (range, 15.1

In the box plots of mean elasticity values (A) and maximum elasticity values (B), mean values of both mean and maximum elasticity were higher in patients with diffuse thyroid disease than those in the normal group. However, there were no statistically significant differences between the two groups ($p=0.802$ and $p=0.452$). DTD, diffuse thyroid disease; Mean, mean elasticity value; Maximum, maximum elasticity value.

Table 1. Comparison of Diagnostic Performances for the Identification of Diffuse Thyroid Disease

| Total number | Gold standard for diagnosis | Diagnostic modality | TP | TN | FP | FN | Sensitivity | Specificity | Accuracy | PPV | NPV | $A_z$ |
|--------------|-----------------------------|---------------------|----|----|----|----|-------------|-------------|----------|------|-----|------|
| 57           | Anti-thyroid autoantibodies | Conventional US     | 5  | 33 | 2  | 17 | 22.7       | 94.3        | 66.7     | 71.4 | 66  | 0.585 |
| 57           | Anti-thyroid autoantibodies | Conventional US     | 4  | 30 | 5  | 18 | 18.2       | 85.7        | 59.6     | 44.4 | 62.5 | 0.52  |
| 57           | Anti-thyroid autoantibodies | Conventional US     | 13 | 20 | 15 | 9  | 59.1       | 57.1        | 57.9     | 46.4 | 69  | 0.581 |
| 57           | Anti-thyroid autoantibodies | Shearwave elastography | 9  | 29 | 6  | 13 | 40.9       | 82.9        | 66.7     | 60   | 69  | 0.619 |
| 57           | Anti-thyroid autoantibodies | Shearwave elastography | 9  | 27 | 8  | 13 | 40.9       | 77.1        | 63.2     | 52.9 | 67.5 | 0.59  |
| 322          | Histology                   | Conventional US     | 50 | 244| 21 | 7  | 87.7       | 92.1        | 91.3     | 70.4 | 97.2 |       |
| 74           | Clinical features           | Elastography using the ARFI technique | 64.7 | 91.3 | 72.9 | 94.2 | 53.8 | 0.8 |

TP, true positive; TN, true negative; FP, false positive; FN, false negative; PPV, positive predictive value; NPV, negative predictive value; ARFI, acoustic radiation force impulse.

$A_z$: the area under the receiver operating characteristic curve.
Diffuse Thyroid Disease with Shearwave Elastography

kPa to 70.6 kPa) (Fig. 3). Although the mean values of mean and maximum elasticity were higher in patients with DTD than those in the normal group, there were no statistically significant differences between the two groups ($p=0.802$ and $p=0.452$).

The cutoff values of shearwave elastography in DTD diagnosis were 27.6 kPa and 41.3 kPa using the mean and maximum elasticity, respectively. When using the mean elasticity value as a diagnostic criterion for DTD, the sensitivity and specificity were 40.9% and 82.9%, respectively. And when the maximum elasticity value was used as a diagnostic criterion for DTD, the sensitivity and specificity were 40.9% and 77.1%, respectively. Using the mean and maximum elasticity values as a diagnostic criteria for DTD, $A_1$ values were 0.619 and 0.59, respectively (Table 1).

**DISCUSSION**

Shin, et al.\(^5\) indicated that thyroid US in SCH has a significant predictive value in treatment outcome, which was comparable to other well-known prognostic parameters such as TPOAb or high TSH levels. Recently, Kim, et al.\(^7\) prospectively evaluated the efficacy of thyroid US in the identification of asymptomatic DTD, and concluded that real-time thyroid US is useful in differentiating DTD from normal thyroid parenchyma in asymptomatic patients.

However, US is very subjective and dependent on the performer, which causes interobserver variability between examiners.\(^22-23\) There are several reports on the interobserver agreement in US assessment of thyroid nodules.\(^25-27\) These reports conclude that if US is performed by experienced radiologists, relatively good agreement can be expected.\(^22-24\) However, there were no reports on the interobserver variability of US in the evaluation of echogenicity of the underlying thyroid gland. In this study, fair agreement with a kappa value of 0.26 was observed on the heterogeneous echogenicity on conventional US by three radiologists. This result showed that in diagnosing DTD, conventional US had poor reproducibility even when all images were simultaneously reviewed by three experienced radiologists. Also, diagnostic performances were variable among the observers. The sensitivity, specificity, accuracy, PPV, and NPV of three radiologists were 18.2-59.1%, 57.1-94.3%, 57.9-66.7%, 44.4-71.4% and 62.5-69.0% on conventional US when identifying DTD (Table 1). These results are different from those of other studies.\(^7-26\) We speculate that the reasons for the differences\(^7-26\) are different gold standards, patient populations, and interobserver variability among radiologists and pathologists. Especially, lymphocytic thyroiditis, of which the pathologic hallmark is lymphocytic infiltration, can show a certain pathologic spectrum, ranging from lymphocytic thyroiditis only to atrophic and eosinophilic change of thyroid cells with fibrosis.\(^27\) Depending on which stage of lymphocytic thyroiditis is being observed, there will be variations in diagnostic performances that will, in turn, affect the results. Therefore, the pathologic stage of disease is thought to be an important factor that decides what kinds of results US elastography obtains.

US elastography is based on the deformity formed by external compression,\(^17\) and the diagnostic performances of US elastography in general greatly depend on the quality of the freehand compression applied.\(^28\) Shearwave elastography is a technique which characterizes tissue structure by using shear acoustic waves remotely induced by a focused ultrasonic beam.\(^29\) Because shearwave elastography produces the radiation force in the probe, this may be more reproducible than conventional elastography, in which radiating force is produced by repetitive external compression by an operator.\(^16\) Furthermore, shearwave elastography shows quantitative tissue elasticity within a given region of interest.\(^30\) This reproducibility and quantitative information are very useful advantages of shearwave elastography. In the present study, we performed a quantitative measurement of thyroid tissue stiffness employing shearwave elastography, using mean and maximum elasticity values, and found that the mean values of mean and maximum elasticity in patients with DTD were not significantly different from those in the normal group (mean 24.1±10 kPa versus 23.4±10.8 kPa and maximum 36.4±13.3 kPa versus 33.7±12.4 kPa, respectively, $p=0.802$ and $p=0.452$ (Table 1). The specificity was higher when the mean elasticity value was used instead of the maximum elasticity value, however, the sensitivities of shearwave elastography were same when both indices were used in diagnosing DTD.

Our study has some limitations. First, the number of patients included in our study was relatively small to generalize reproducibility of conventional US and diagnostic performance of shearwave elastography. Future studies with larger populations are needed. Second, final diagnosis of the patients included in this study was based on the serologic levels of autoantibodies. False-negative and false-positive serologic results might have an effect on our results, nevertheless, they were not considered. Third, this study was a retrospective analysis to evaluate a reproducibility of
conventional US between three observers using only static captured images. To resolve and validate our result, real time US might be helpful in further studies. Lastly, this study did not evaluate diagnostic performances of US and elastography according to pathologic spectrum of DTD. For example, hepatitis has variable US results, depending on liver fibrosis stage, because stiffness changes by stage. Thus, for a more insight, studies are needed on how US elastography results vary by pathologic stages of DTD to understand the full potential of its practical use. In conclusion, conventional US did not show a reliable interobserver agreement in the diagnosis of DTD. Although not statistically significant, shearwave elastography may provide additional information for the diagnosis of DTD. Therefore, larger prospective studies are needed to define the values of shearwave elastography in diagnosing DTD.

REFERENCES

1. Kim EK, Park CS, Chung WY, Oh KK, Kim DI, Lee JT, et al. New sonographic criteria for recommending fine-needle aspiration biopsy of nonpalpable solid nodules of the thyroid. AJR Am J Roentgenol 2002;178:687-91.
2. Kim SJ, Kim EK, Park CS, Chung WY, Oh KK, Yoo HS. Ultrasound-guided fine-needle aspiration biopsy in nonpalpable thyroid nodules: is it useful in infracentimetric nodules? Yonsei Med J 2003;44:635-40.
3. Moon WJ, Baek JH, Jung SL, Kim DW, Kim EK, Kim JY, et al. Ultrasonography and the ultrasound-based management of thyroid nodules: consensus statement and recommendations. Korean J Radiol 2011;12:1-14.
4. Pedersen OM, Aaradal NP, Larsen TB, Varhaug JE, Myking O, Vik-Mo H. The value of ultrasonography in predicting autoimmune thyroid disease. Thyroid 2000;10:251-9.
5. Shin DY, Kim EK, Lee EJ. Role of ultrasonography in outcome prediction in subclinical hypothyroid patients treated with levothyroxine. Endocr J 2010;57:15-22.
6. Yeh HC, Futterweit W, Gilbert P. Micronodulation: ultrasonographic sign of Hashimoto thyroiditis. J Ultrasound Med 1996;15:813-9.
7. Kim DW, Eun CK, In HS, Kim MH, Jung SJ, Bae SK. Sonographic differentiation of asymptomatic diffuse thyroid disease from normal thyroid: a prospective study. AJNR Am J Neuroradiol 2010;31:1956-60.
8. Sohn YM, Yoon JH, Moon HJ, Kim EK, Kwav JY. Mixed echoic thyroid nodules on ultrasound: approach to management. Yonsei Med J 2012;53:812-9.
9. Lee MJ, Hong SW, Chung WY, Kwav JY, Kim MJ, Kim EK. Cytological results of ultrasound-guided fine-needle aspiration cytology for thyroid nodules: emphasis on correlation with sonographic findings. Yonsei Med J 2011;52:838-44.
10. Marconcì C, Vitti P, Cetani F, Catalano F, Concetti R, Pinchera A. Thyroid ultrasonography helps to identify patients with diffuse lymphocytic thyroiditis who are prone to develop hypothyroidism. J Clin Endocrinol Metab 1991;72:209-13.
11. Hayashi N, Tamaki N, Konishi J, Yonekura Y, Senda M, Kasagi K, et al. Sonography of Hashimoto’s thyroiditis. J Clin Ultrasound 1986;14:123-6.
12. Nys P, Cordray JP, Merceron RE. Etiologic discussion and clinical relevance of thyroid ultrasonography in subclinical hypothyroidism. A retrospective study in 1945 patients. Ann Endocrinol (Paris) 2009;70:59-63.
13. Rosário PW, Bessa B, Valadão MM, Purisch S. Natural history of mild subclinical hypothyroidism: prognostic value of ultrasound. Thyroid 2009;19:9-12.
14. Asteria C, Giovannardi A, Pizzocaro A, Cozzaglio L, Morabito A, Somalvico F, et al. US-elastography in the differential diagnosis of benign and malignant thyroid nodules. Thyroid 2008;18:523-31.
15. Rago T, Santini F, Scutari M, Pinchera A, Vitti P. Elastography: new developments in ultrasound for predicting malignancy in thyroid nodules. J Clin Endocrinol Metab 2007;92:2917-22.
16. Sebag F, Vaillant-Lombard J, Berbis J, Griset V, Henry JF, Petit P, et al. Shear wave elastography: a new ultrasound imaging mode for the differential diagnosis of benign and malignant thyroid nodules. J Clin Endocrinol Metab 2010;95:5281-8.
17. Ophir J, Alam SK, Garra B, Kallef F, Konofagou E, Krouskop T, et al. Elastography: ultrasonic estimation and imaging of the elastic properties of tissues. Proc Inst Mech Eng H 1999;213:203-33.
18. Choi JS, Chung WY, Kwak JY, Moon HJ, Kim MJ, Kim EK. Staging of papillary thyroid carcinoma with ultrasonography: performance in a large series. Ann Surg Oncol 2011;18:3572-8.
19. Tessler FN, Tublin ME. Thyroid sonography: current applications and future directions. AJR Am J Roentgenol 1999;173:437-43.
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
21. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3:32-5.
22. Kim SH, Park CS, Jung SL, Kang BJ, Kim JY, Choi JJ, et al. Observer variability and the performance between faculties and residents: US criteria for benign and malignant thyroid nodules. Korean J Radiol 2010;11:149-55.
23. Park CS, Kim SH, Jung SL, Kang BJ, Kim JY, Choi JJ, et al. Observer variability in the sonoerographic evaluation of thyroid nodules. J Clin Ultrasound 2010;38:287-93.
24. Choi SH, Kim EK, Kwak JY, Kim MJ, Son EJ. Interobserver and intraobserver variations in ultrasound assessment of thyroid nodules. Thyroid 2010;20:167-72.
25. Park CS, Kim SJ, Kim EK, Kim MJ, Son EJ, Kwak JY. Interobserver agreement in assessing the sonographic and elastographic features of malignant thyroid nodules. AJR Am J Roentgenol 2009;193:W416-23.
26. Sporea I, Vlad M, Bota S, Sirli RL, Popescu A, Danila M, et al. Thyroid stiffness assessment by acoustic radiation force impulse elastography (ARFI). Ultraschall Med 2011;32:281-5.
27. LiVolsi VA. The pathology of autoimmune thyroid disease: a review. Thyroid 1994;4:333-9.
Montaldo G, et al. Quantitative assessment of breast lesion viscoelasticity: initial clinical results using supersonic shear imaging. Ultrasound Med Biol 2008;34:1373-86.

31. Sporea I, Sirli R, Bota S, Popescu A, Sendroiu M, Jurchis A. Comparative study concerning the value of acoustic radiation force impulse elastography (ARFI) in comparison with transient elastography (TE) for the assessment of liver fibrosis in patients with chronic hepatitis B and C. Ultrasound Med Biol 2012;38:1310-6.