นิพนธ์ต้นฉบับ

Original Article

ปัจจัยที่เกี่ยวข้องกับการไม่สอดคล้องกันของขนาดของต่อมไทรอยด์ที่วัดโดยการคลำและการทำอัลตราซาวด์เพื่อการรับการรักษาโรคเกรฟส์โดยการกลืนแร่

อลิสรา วงศ์สุทธิเลิศ พ.บ., วัลลภ ใจดี ปร.ด.**, สุธาสินี คงพร้อมสุข พ.บ.****

บทคัดย่อ

วัตถุประสงค์: ศึกษาหาปัจจัยที่เกี่ยวข้องกับความไม่สอดคล้องกันของการวัดขนาดของต่อมไทรอยด์โดยวิธีการคลำและการทำอัลตราซาวด์รวมถึงหาจุดตัดของขนาดของต่อมไทรอยด์ที่ทำให้เกิดความไม่สอดคล้องกัน

วิธีการศึกษา: การศึกษานี้เป็นการศึกษาแบบย้อนหลังโดยเก็บข้อมูลจากผู้ป่วยที่มารับการกลืนแร่ด้วยโรคเกรฟส์ที่โรงพยาบาลมหาวิทยาลัยบูรพาจำนวน 477 รายในช่วงมกราคม พ.ศ. 2562 ถึง กันยายน พ.ศ. 2565 ผู้ป่วยจะได้รับการวัดขนาดของต่อมด้วยการคลำและการทำอัลตราซาวด์ การศึกษานี้ใช้สถิติ one-way ANOVA, Chi-square, univariate regression, multivariate logistic regression และ linear regression analyses เราหาจุดตัดของขนาดของต่อมไทรอยด์ที่ทำให้เกิดความไม่สอดคล้องกัน รวมถึงดู sensitivity, specificity, PPV, NPV, and accuracy ของจุดตัดนี้ กำหนดนัยสำคัญทางสถิติเป็น $p<0.05$

ผลการศึกษา: ขนาดของต่อมไทรอยด์ที่คลำได้เล็กกว่าที่ควรจะเป็นเกี่ยวข้องกับการสูบบุหรี่ (aOR: 2.37; 95%CI: 1.16–4.84) และอาการคอพอก (aOR: 2.42; 95%CI: 1.53–3.82) และพบจุดตัดที่ 30 กรัมมี 91.21% sensitivity, 68.14% specificity, 63.85% PPV, 92.63% NPV, and 76.94% accuracy.

สรุป: การสูบบุหรี่ อาการคอพวก รวมถึงขนาดของต่อมไทรอยด์ที่มากกว่า 30 กรัม สัมพันธ์กับขนาดของต่อมไทรอยด์ที่คลำได้เล็กกว่าที่ควรจะเป็น และผู้ป่วยเหล่านี้ควรได้รับการทำอัลตราซาวด์

คำสำคัญ : ขนาดของต่อมไทรอยด์, การคลำ, อัลตราซาวด์, โรคเกรฟส์, การกลืนแร่

Factors correlated with Discordance between Thyroid Volume Measurement by Palpation and Ultrasound Methods for Radioiodine Treatment in Graves' Disease

Alisara Wongsuttilert M.D.*, Wanlop Jaidee Ph.D**, Sutasinee Kongpromsuk M.D.***

Objective: To evaluate the related factors of thyroid volume (TV) discordance between palpation (P-TV) and ultrasound (US-TV) methods and the TV's cut-off point that separates the discordance and concordance groups.

Material and Method: In this retrospective study, we included 477 patients with Graves' disease (GD) who were referred for radioiodine treatment from January 2019 to September 2022. The TVs of these patients were measured via P-TV and US-TV. The data were analyzed using one-way ANOVA, Chi-square, univariate regression, multivariate logistic regression, and linear regression analyses. We evaluated the cut-off point of the TV that separates the discordance and concordance groups. We also tested the accuracy of this cut-off value and reported the

,* สาขาวิชารังสีวิทยาและเวชศาสตร์นิวเคลียร์ คณะแพทยศาสตร์ มหาวิทยาลัยบูรพา*

*** คณะสาธารณสุขศาสตร์ มหาวิทยาลัยบูรพา*

Corresponding Author: Sutasinee Kongpromsuk

,* Division of Radiology and Nuclear medicine, Faculty of medicine, Burapha university*

*** Department of Fundamentals of Public Health, Faculty of Public Health, Burapha University*

Received : Mar 13, 2023 Revised : Aug 23, 2023 Accepted : Aug 31, 2023

sensitivity, specificity, PPV, NPV, and accuracy. The threshold for statistical significance was set at $p < 0.05$.

Results: P-TV underestimation was positively correlated to smoking (aOR: 2.37; 95%CI: 1.16–4.84) and goiter symptoms (aOR: 2.42; 95%CI: 1.53–3.82). Regarding P-TV overestimation, there was only the pre-tracheal thickness that was negatively correlated (aOR: 0.82; 95%CI: 0.71–0.95). The cut-off point value of the TV was 30 grams with 91.21% sensitivity, 68.14% specificity, 63.85% PPV, 92.63% NPV, and 76.94% accuracy.

Conclusions: Smoking and goiter symptoms, as well as a thyroid volume of >30 grams, correlate with P-TV underestimation; thus, such patients should undergo ultrasonography.

Keywords : Thyroid volume, palpation, ultrasound, Graves's disease, radioiodine

Introduction

Radioiodine treatment (RIT) is highly efficient and has minimal side effects for the treatment of Graves' disease (GD). The precision volume of the thyroid is essential for the calculation of the therapeutic radioiodine (RAI) activity in each patient. According to the American Thyroid Association in 2016, thyroid volume (TV) can be determined via manual palpation (P-TV) or ultrasonography (US-TV). The equation of the therapeutic RAI activity is RAI activity [μCi] = gland weight [g] * 50–200 μCi/g * [1/24-hour uptake in % of administered activity][1]. Based on the data of 232 patients who received RIT with the calculated dose regimen in our institution between 2015 and 2018, the median value of 24-hour RAI uptake is 69.6% (IQR: 57.5%–78.7%). We modified the equation for the fixed-dose regimen as RAI activity [μCi] = gland weight [g] * 200 μCi/g * [1/69.6]. Then, we discovered that the 3.5-grams error of the TV will take the error in the calculation of the therapeutic RAI activity to about 1 mCi. This error may increase the risk of the patient either receiving inadequate therapeutic RAI activity or suffering excessive unnecessary radiation exposure.

Previous studies have shown that the US-TV method has a high correlation with volume measurements from 3D MRI[2,3], CT[4], planar scintigraphy[3,5], and SPECT[6]. Although the US-TV method has higher precision in the evaluation of TV than the P-TV method, the patient may require extra visits and more expenses, which may cause delays in the performance of RIT. The P-TV method is easy, fast, and cost effective for estimating RAI activity. In the past, two studies have revealed that the P-TV method, when performed by experienced physicians, has a strong correlation with the US-TV method[7,8]; however, other studies reported a poor correlation between these two methods[9,10]. Both methods have their advantages and disadvantages. If we know the factors that are correlated with the discordance between the two methods, we can develop a new appropriate approach for TV measurement in this situation.

In this study, we aim to evaluate the factors associated with TV discordance between the P-TV and US-TV methods. Therefore, we assessed the cut-off point of the TV that separates the discordance group from the concordance group.

Materials and Methods

Population

In this retrospective study, we enrolled 477 consecutive patients both male and female who were aged ≥18 years, had a diagnosis of GD, and underwent the first session of RIT at our institution from January 2019 to September 2022. The TV of each patient was assessed by palpation followed by US testing at our

institution. We retrospectively collected each patient's data from the well-organized database of the Nuclear Medicine Unit of our institution. We also collected thyroid US images of all participants from the picture archiving and communication system and then remeasured the sizes of the thyroid glands and nodules of all images. This study was approved by the Institutional Review Board of our institution.

P–TV method

Clinical assessment of thyroid gland size was performed by a nuclear medicine physician with eight years of experience in thyroid palpation. The palpation technique was performed using standard protocol by palpating the patient's thyroid from behind, with the patient in a sitting position. We estimated the size of the thyroid gland when the patient swallows and his/her thyroid gland moves up and down in the examiner's hands like a ball. We estimated the size of each lobe and isthmus should be approximately the size of the distal phalanx of the thumb (approximately 5 grams)[11].

US–TV method

All patients underwent thyroid US at Burapha University Hospital, and US was performed using two devices; the first was an Aplio 300 with a 7–14 MHz linear transducer while the second is an Apilo 200 with a 7–14 MHz linear transducer (Toshiba, Tokyo, Japan). Our protocol required the patient to lay down in a supine position and slightly overextend his/her neck with a pillow behind the shoulders. We scanned the whole gland at the upper, mid, and lower poles, and also at the isthmus[12]. Then, we measured each lobe in three dimensions: the maximum width (W), maximum depth (D), and maximum length (L; Fig. 1). The US-TV was the sum of the volumes of both lobes and the isthmus. The TV of each lobe was calculated using the standard formula for ellipsoid volumes: TV = 0.52 (W*D*L)[12]. The limits of normal US-TV are 10–15 grams for females and 12–18 grams for males[13]. We also evaluated the appearance and size of the thyroid nodules and cervical lymph nodes. We hypothesized that the soft tissue of the anterior neck may have influenced the P-TV measurements, particularly the strap muscles. Then, we also measured the pre-tracheal soft tissue thickness, the anterior soft tissue thickness of both thyroid lobes, and the strap muscle thickness of both thyroid lobes (Fig. 2). Ultrasound was done by radiologists who were blinded to the P-TV result.

Determination of TV discordance between the two methods

Per the P-TV method, the size of the distal phalanx of the thumb represents about 5 grams of the TV. So, we classified the participants into three groups. If the difference in the size of the thyroid gland between the P-TV and US-TV methods was less than 5 grams, we assigned the patient to the TV concordance group. If the size of the gland measured via P-TV were smaller than that measured via US-TV by ≥5 grams, we assigned the patient to the P-TV underestimation group. If the size of the gland determined via P-TV exceeded that determined via US-TV by ≥5 grams, we assigned the patient to the P-TV overestimation group.

Statistical analysis

For continuous data, age was presented using mean values with standard deviations. Median values with interquartile ranges (IQRs) were used for P-TV, US-TV, and the soft tissue thickness of the anterior neck. As for categorical data, sex, smoking, goiter symptoms, a bulge in the thyroid or a changed thyroid shape (with nodules present) were presented as counts and percentages. Smoking included both current and past smoking. Goiter symptom included feeling of neck fullness, difficulty swallowing or breathing. cough, hoarseness and snoring.

We used the one-way ANOVA to compare

continuous data and the Chi-square test to compare categorical data to compare clinical data between the TV concordance, P-TV underestimation, and P-TV overestimation groups. Univariate and multivariate logistic regression analyses were used to determine independent risk factors for P-TV underestimation and P-TV overestimation. The linear regression analysis was used to find the correlation between P-TV and US-TV and to show this correlation using scatter plots and the R^2 coefficient. Then, we evaluated the cut-off point of the TV that separated the discordance and concordance groups. We also tested the accuracy of this cut-off point and reported it in terms of sensitivity, specificity, PPV, NPV, and accuracy. The threshold for statistical significance was set at $p < 0.05$.

Results

Among a total of 477 patients, 31.0% were male and 69.0% were female. Their mean age was 38.4 years (SD: 11.9 years; Supplementary Table 1). The mean TV assessed by US was 34.7 grams (SD: 23.7 ml) with a range of 7.2–249.6 grams. Only 43.4% of our study participants belonged to the TV concordance group (Table 1).

P-TV underestimation was positively correlated to smoking (aOR: 2.37; 95%CI: 1.16–4.84) and goiter symptoms (aOR: 2.42; 95%CI: 1.53–3.82)(table 2). In contrast, P-TV overestimation, only the pre-tracheal thickness was negatively correlated (aOR: 0.82; 95%CI: 0.71–0.95)(table 3).

The P-TV method showed a moderate correlation with US-TV (R^2 = 0.4846; Fig. 3). The cut-off value of the TV was 30 grams. When we used this cut-off value for the P-TV method, it showed 91.21% sensitivity, 68.14% specificity, 63.85% PPV, 92.63% NPV, and 76.94% accuracy.

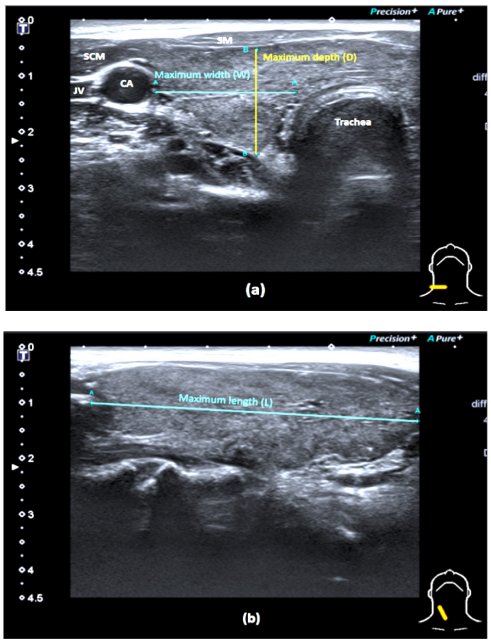

Figure 1 Transverse (a) and longitudinal (b) images of the right thyroid lobe. Abbr: common carotid artery (CA); jugular vein (JV); sternocleidomastoid (SCM); and strap muscles (SM).

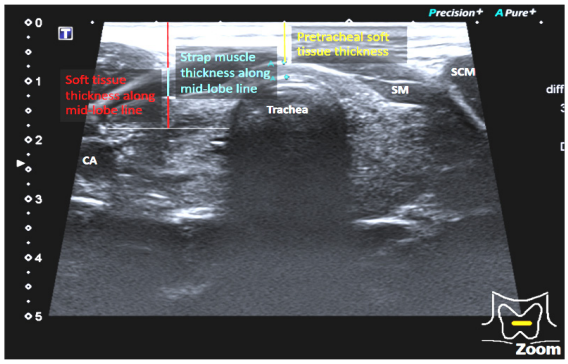

Figure 2 Transverse image of the thyroid. Abbr: common carotid artery (CA); sternocleidomastoid (SCM); and strap muscles (SM).

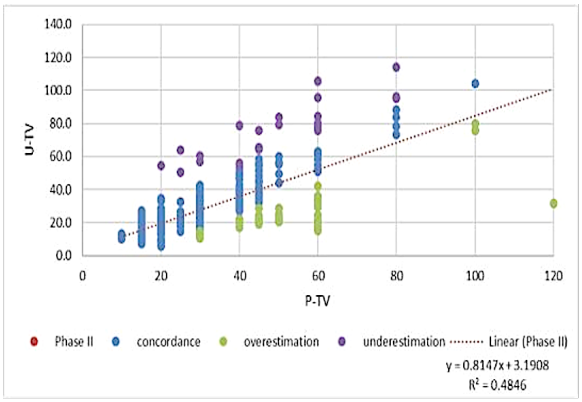

Figure 3 Scatter plot of TV distribution between P-TV and US-TV methods.

Table 1 Clinical data of this study (n = 477).

Variable	TV concordance (n = 207)	P-TV underestimation (n = 154)	P-TV overestimation (n = 116)	*p*-value
Age (year)	39.9 (12.9)	37.4 (11.1)	37.2 (11.0)	0.072
Gender				0.038*
Male	52 (25.1)	52 (33.8)	44 (37.9)	
Female	155 (74.9)	102 (66.2)	72 (62.1)	
Smoking	26 (12.6)	39 (25.3)	27 (23.3)	0.045*
Goiter symptom	75 (36.2)	89 (57.8)	48 (41.4)	0.002*
Thyroid nodule(s)	41 (19.8)	31 (20.1)	18 (15.5)	0.568
Bulge in the thyroid or changed thyroid shape	15 (7.2)	11 (7.1)	5 (4.3)	0.546
US-TV (gram)	24.4	49.5	22.1	<0.001*
	(17.8, 32.2)	(35.3, 62.2)	(16.1, 30.8)	
Isthmus	0.9	2.0	0.8	<0.001*
	(0.5, 1.5)	(1.2, 3.2)	(0.5, 1.4)	
Right lobe	12.7	26.1	12.0	<0.001*
	(9.7, 17.3)	(17.1, 34.6)	(8.6, 16.6)	
Left lobe	10.3	19.5	8.9	<0.001*
	(7.1, 13.7)	(14.0, 27.0)	(6.7, 12.3)	
Soft tissue thickness of anterior neck				
Pre-tracheal (mm)	4.7	5.0	4.2	0.005*
	(3.6, 6.3)	(3.7, 6.3)	(3.3, 5.3)	
Mid-lobe soft tissue (mm)				
Right lobe	6.4	6.0	6.6	0.122
	(5.1, 8.0)	(4.5, 7.7)	(5.1, 7.8)	
Left lobe	7.5	7.0	7.2	0.487
	(6.0, 9.2)	(5.7, 9.0)	(5.9, 8.8)	
Mid-lobe strap muscle (mm)				
Right lobe	2.2	2.3	2.5	0.465
	(1.2, 3.4)	(0.0, 4.1)	(1.6, 3.7)	
Left lobe	3.2	3.6	2.9	0.398
	(2.0, 4.7)	(1.8, 5.3)	(2.1, 4.9)	

Data are in mean(SD), n(%) and median (IQR)

Table 2 Univariate and multivariate logistic regression analyses of factors association with P-TV underestimation (n = 154).

Factors	Crude OR	*p*-value	Adjusted OR	*p*-value
Age (year)	0.98		0.99	
	(0.97–1.00)	0.060	(0.97–1.01)	0.421
Male vs Female	1.52		1.13	
	(0.96–2.40)	0.074	(0.62–2.09)	0.688
Smoking	2.36		2.37	
	(1.36–4.09)	0.002*	(1.16–4.84)	0.017*
Goiter symptom	2.41		2.42	
	(1.57–3.69)	<0.001*	(1.53–3.82)	<0.001*
Thyroid nodules(s)	1.02		1.24	
	(0.61–1.72)	0.939	(0.64–2.39)	0.523
Bulge in the thyroid or changed thyroid shape	0.98		0.87	
	(0.44–2.21)	0.970	(0.32–2.38)	0.781
Pre-tracheal thickness (mm)	1.09		1.08	
	(0.97–1.22)	0.137	(0.96–1.22)	0.192

Table 3 Univariate and multivariate logistic regression analyses of factors association with P-TV overestimation (n = 116).

Factors	Crude OR	*p*-value	Adjusted OR	*p*-value
Age (year)	0.98		0.98	
	(0.96–1.00)	0.058	(0.96–1.01)	0.127
Male vs Female	1.82		1.44	
	(1.12–2.97)	0.016*	(0.77–2.67)	0.252
Smoking	2.11		1.65	
	(1.16–3.83)	0.014*	(0.79–3.47)	0.186
Goiter symptom	1.24		1.22	
	(0.78–1.98)	0.361	(0.74–2.01)	0.444
Thyroid nodules(s)	0.74		1.04	
	(0.40–1.37)	0.340	(0.50–2.16)	0.919
Bulge in the thyroid or changed thyroid shape	0.58		0.70	
	(0.20–1.63)	0.299	(0.20–2.41)	0.573
Pre-tracheal thickness (mm)	0.82		0.82	
	(0.71–0.94)	0.005*	(0.71–0.95)	0.008*

Discussion

The key to the success of RIT for patients with GD is precise therapeutic RAI activity. Presently, thyroid ultrasonography is still a high-precision method of measuring the TV for this purpose[1-6,14,15]. However, this US-TV method is not available in all local hospitals because of the paucity of local staff and diagnostic tools, or even prolonged ultrasound waiting lists. These problems may cause a delay in RIT and may change a patient's therapeutic outcome.

In our study, the P-TV method was moderately correlated to the US-TV method (R^2 = 0.4846; Fig. 3), and the rate of TV concordance was only 43.4% (Table 1). This result differed from the findings of Nordmeyer et al., who showed a good correlation between the P-TV and US-TV methods (R^2 = 0.856–0.872). Although the mean TV measurement via US did not differ from that in our study (35.9–37.3 grams vs. 34.7 grams), the US-TV was calculated using a different formula, which may have influenced the correlation analysis. In our study, we used an ellipsoid formula (f = 0.524); however, the study conducted by Nordmeyer et al. used Brunn et al.'s formula. (f = 0.479)[7].

To the best of our knowledge, our study is the first to investigate the factors associated with the TV discordance between the P-TV and US-TV methods. The rate of P-TV underestimation was 32.3%, and the underestimated P-TV was positively correlated with smoking (aOR: 2.37; 95%CI: 1.16–4.84) and the experience of goiter symptoms (aOR: 2.42; 95%CI: 1.53–3.82; Table 2). We also found that the median US-TV of the patients in this group was significantly larger than that of patients in the TV concordance group (49.5 grams vs. 24.4 grams; Table 1). A large thyroid gland would make TV estimation via palpation difficult, unlike a normal-sized or slightly-enlarged thyroid. Interestingly, the nodules produced a bulge in the thyroid, and the soft tissue thickness of the anterior neck was not significantly correlated to the underestimation of the P-TV method (Table 2).

On the contrary, the pre-tracheal thickness was only one factor that was correlated with P-TV overestimation (aOR: 0.82; 95%CI: 0.71–0.95). In this group, the median pre-tracheal thickness was significantly thinner than in the TV concordance group (4.2 millimeters vs. 4.7 millimeters; Table 1). We also found that the median US-TV did not differ significantly between the P-TV overestimation and TV concordance groups (22.1 grams vs. 24.4 grams; Table 1).

We discovered that the cut-off value of 30 grams can sufficiently differentiate the TV discordance and TV concordance groups. When we used this cut-off value, the accuracy of TV estimation by palpation increased from 43.4% to 76.9%. From the practical point of view, we suggest that patients with GD should use ultrasonography to estimate their TVs when physicians palpate their thyroid glands and estimate that they weigh ≥30 grams. Based on the findings of our study, there were 260 patients (54.5%) whose thyroid glands weighed less than 30 grams, and we can reduce the cost associated with thyroid ultrasonography testing by nearly half (about 8,000 USD). However, in some patients, US-TV estimates of <30 grams were still inaccurate (11.9% in the TV underestimation group and 32.7% in the TV overestimation group). This result will increase the risks of RIT failure in the TV underestimation group and will increase the rate of unnecessary radiation exposure in the TV overestimation group. In our next study, we will search for an innovative, more precise formula for TV measurement for RIT. We will apply the factors associated with TV discordance to the development of the new formula. The main limitation of this study is the lack of treatment outcomes for the study participants.

In conclusion, the P-TV method was not sufficiently accurate in measuring TVs of glands weighing ≥30 grams; thus, physicians should consider thyroid ultrasound testing for GD patients with such glands who plan to opt for treatment with RAI.

References

1. Ross DS, Burch HB, Cooper DS, Greenlee MC, Laurberg P, Maia AL, et al. 2016 American thyroid association guidelines for diagnosis and management of hyperthyroidism and other causes of thyrotoxicosis. Thyroid. 2016; 26: 1343-421.
2. Reinartz P, Sabri O, Zimny M, Nowak B, Cremerius U, Setani K, et al. Thyroid volume measurement in patients prior to radioiodine therapy: comparison between three-dimensional magnetic resonance imaging and ultrasonography. Thyroid .2002; 12: 713-7.

3. van Isselt JW, de Klerk JM, van Rijk PP, van Gils AP, Polman LJ, Kamphuis C, et al. Comparison of methods for thyroid volume estimation in patients with Graves' disease. Eur J Nucl Med Mol Imaging. 2003; 30: 525-31.
4. Nygaard B, Nygaard T, Court-Payen M, Jensen LI, Søe-Jensen P, Gerhard Nielsen K, et al. Thyroid volume measured by ultrasonography and CT. Acta Radiol. 2002; 43: 269-74.
5. Plesniak J, Urbanski S. Comparative thyroid gland volume by two methods: ultrasonography and planar scintigraphy. Pol J Radiol. 2012; 77: 19-21.
6. Chen Y, Xie SP, He F, Chen J. Evaluation of the efficacy of standardized uptake value (suv)-shape scheme for thyroid volume determination in graves' disease: a comparison with ultrasonography. Asia Ocean J Nucl Med Biol. 2017; 5: 44-8.
7. Nordmeyer JP, Simons M, Wenzel C, Scholten T. How accurate is the assessment of thyroid volume by palpation? A prospective study of 316 patients. Exp Clin Endocrinol Diabetes. 1997; 105: 366-71.
8. Lucas KJ. Use of thyroid ultrasound volume in calculating radioactive iodine dose in hyperthyroidism. Thyroid. 2000; 10: 151-5.
9. Berghout A, Wiersinga WM, Smits NJ, Touber JL. The value of thyroid volume measured by ultrasonography in the diagnosis of goitre. Clin Endocrinol (Oxf) 1988; 28: 409-14.
10. Sheikh M, Doi SA, Sinan T, Al-Shoumer KA. Technical observations on the assessment of thyroid volume by palpation and ultrasonography. J Ultrasound Med. 2004; 23: 261-6.
11. Hennessey JV. Physical examination of the thyroid gland. In Werner & Ingbar's the thyroid: a fundamental and clinical text, edn 9, ch. 18, pp 393-7. Philadelphia: Lippincott Williams & Wilkins; 2005.
12. Dighe M, Barr R, Bojunga J, Cantisani V, Chammas MC, Cosgrove D, et al. Thyroid ultrasound: state of the art part 1-thyroid ultrasound reporting and diffuse thyroid diseases. Med Ultrason. 2017; 19: 79-93.
13. Brunn J, Block U, Bos I, Kunze WP, Scriba PC. Volumetrie der schilddrüsenlappen mittels real-time-sonographie. Dtsch Med Wochenschr. 1981; 106: 1338-40.
14. Hegedus L, Perrild H, Poulsen LR, Andersen JR, Holm B, Schnohr P, et al. The determination of thyroid volume by ultrasound and its relationship to body weight, age, and sex in normal subjects. J Clin Endocrinol Metab 1983;56:260-3.
15. Hegedus L. Thyroid size determined by ultrasound: influence of physiological factors and non-thyroidal disease. Dan Med Bull 1990;37:249-63.

Supplementary Table

Supplementary Table 1 Characteristics of this study (n = 477)

Variable	Results
Age (yr; *mean, SD*)	38.4 (11.9)
Male *(n, %)*	148 (31.0)
Female *(n, %)*	329 (69.0%)
Smoking *(n, %)*	92 (19.3)
Goiter symptom *(n, %)*	212 (44.4)
Thyroid nodule (s; *n, %*)	90 (18.9)
Bulge in the thyroid or changed thyroid shape *(n, %)*	31 (6.5)
P-TV (g) Median (IQR)	30.0 (20.0, 40.0)
US-TV (g)	28.0 (20.1, 41.1)
Soft tissue thickness of anterior neck	
Pre-tracheal (mm; *median, IQR)*	4.7 (3.5, 6.0)
Mid-lobe soft tissue (mm; *median, IQR)*	
Right lobe	6.3 (4.9, 7.9)
Left lobe	2.3 (1.1, 3.6)
Mid-lobe strap muscle (mm; *median, IQR)*	
Right lobe	7.3 (5.8, 9.0)
Left lobe	3.2 (2.0, 4.9)
Volume concordance *(n, %)*	
P-TV underestimation	154 (32.3)
TV concordance	207 (43.4)
P-TV overestimation	116 (24.3)