

Interobserver and intraobserver variations in sonographic measurement of thyroid volume in children

Ali Özgen, Cengiz Erol, Ayten Kaya, Mustafa N Özmen, Deniz Akata and Okan Akhan

Department of Radiology, Hacettepe University School of Medicine, Sıhhiye, 06100, Ankara, Turkey

(Correspondence should be addressed to A Özgen, Sporcular Sitesi, İlkyerlesim Mahallesi, 438. sokak No:12 Batikent 06370, Ankara, Turkey)

Abstract

Objective: To determine the interobserver and intraobserver variations in sonographic measurement of thyroid volume in children.

Design: Thyroid volumes of 30 healthy children were measured by three separate observers. Additionally each observer measured thyroid volumes of ten separate children three times.

Methods: The data were used to assess intraobserver and interobserver variations in measurement of thyroid volume. Interobserver and intraobserver variations in measuring each diameter of the thyroid gland were also determined. The effect of thyroid size on interobserver and intraobserver variations was analyzed.

Results: Intraobserver variation in measurement of thyroid volume was $8.4 \pm 6.7\%$ (mean \pm s.d.). Interobserver variation was $13.3 \pm 8.2\%$. The widest interobserver variation was encountered in determining the craniocaudal diameter of the thyroid gland. No correlation was found between thyroid volume and interobserver variation ($r = -0.12$, $P = 0.27$), whereas a slight but statistically significant correlation was found between thyroid volume and intraobserver variation ($r = -0.26$, $P = 0.012$).

Conclusion: Significant interobserver and intraobserver variation occurs in sonographic measurement of thyroid volume in children.

European Journal of Endocrinology 140 328–331

Introduction

High resolution real-time ultrasound (US) is widely used to image the thyroid as a non-invasive and easy-to-perform method. This method is also used to determine thyroid size in both diagnosis and follow-up of patients with diseases that could affect the thyroid gland (1–10). Normal ranges of thyroid volumes in pediatric age groups and adults have been published in many references, and may be a guide to decide whether thyroid size is normal or not (11–15). The aim of this study was to determine the interobserver and intraobserver variation in US measurement of thyroid volume in children.

Materials and methods

Patients

Sixty children between the ages of 7 and 16 years (mean, 12 years) were examined. The study was performed in a 1 week period. All patients were free of any complaint that could be related to the thyroid. The patients were examined by US whether the clinical examination was normal or not. Informed consent was obtained from the families of the patients.

Methods

High resolution real-time scanners (Toshiba SSA 270-A, Tokyo, Japan) with 7.5 MHz linear probes were used to measure thyroid size. Patients were in the supine position with mild extension of the neck. For each lobe, mediolateral, anteroposterior and craniocaudal diameters of the thyroid gland were measured. The volume of each lobe was calculated using the formula for an ovoid ($\text{width} \times \text{depth} \times \text{length} \times \pi/6$). The isthmus was not included in volume calculation. The volume of the gland was calculated by adding the volume of each lobe.

Thirty separate children were examined three times by three of us (Observers A, B and C). The mean of the three measurements was accepted as the true volume of the thyroid gland for each patient. A total of 90 examinations was performed. Error was calculated by subtracting each volume measured by one observer from the volume measured by the other for each patient. Therefore, measurements of Observer A were subtracted from those of Observers B and C, and measurements of Observer B were subtracted from those of Observer C (A–B, B–C and A–C). The error in percentages was calculated by dividing each value by the true thyroid volume for each patient. The absolute values of the

Table 1 Interobserver variation in sonographic measurement of thyroid volume in children. All numbers are percentages.

	Thyroid volume			Right lobe diameters			Left lobe diameters		
	Mediolateral	Anteroposterior	Craniocaudal	Mediolateral	Anteroposterior	Craniocaudal	Mediolateral	Anteroposterior	Craniocaudal
Mean ± s.d.	9.6 ± 5.7	10.6 ± 7.5	11.5 ± 7.8	9.8 ± 7.1	11.4 ± 8.1	15.9 ± 12.8	9.8 ± 7.1	11.4 ± 8.1	15.9 ± 12.8
Ninety-five percent confidence interval (mean ± 2 s.d.s)	21.0	25.6	27.1	24.0	27.7	41.5	24.0	27.7	41.5
	13.4 ± 8.3								
	30.1								

errors in percentages were used in statistics to assess the interobserver variation since we were concerned only with the magnitude of the difference.

Each observer measured thyroid volumes of 10 separate children three times and a total of 90 examinations were performed on 30 children. The error was determined in the same way. Therefore, the first measurement was subtracted from the second and the third measurements, and the second measurement was subtracted from the third measurement for each observer (1–2, 2–3 and 1–3). The error in percentages was calculated by dividing each value by the true thyroid volume for each patient. The absolute values of the errors in percentages were used in statistics to assess the intraobserver variation. To assess the interobserver and intraobserver variation in measuring any diameter of the thyroid gland the same procedure was performed for each diameter of each thyroid lobe.

This investigation was performed within a study approved by the Ministry of Health and which included 450 children to determine the normal ranges of thyroid volume in pediatric age groups in an area with near-normal iodine uptake to compare them with those in areas with low iodine uptake. The observers did not know which patients were included in the study. They also had no knowledge of one another's results. An interval of at least five patients was left before re-measuring a patient to determine intraobserver variation. All measurements of each patient used in this study were performed within the same day.

Statistics

Means and standard deviations (s.d.s) of the errors were calculated. Ninety-five percent confidence intervals for each variation were computed by adding 2 × s.d. to the mean. Repeated measures ANOVA was used to compare both the mean thyroid volume measurement of the three observers and the mean thyroid volume measurements of the same observers. We calculated correlations using Pearson's correlation. All statistics in this study were done using SPSS for Windows (SPSS Inc., Chicago, IL, USA).

Results

The distribution of ages of the patients were similar for each observer. No patient was recorded as 'uncooperative'. Volumes of the lobes were not separately evaluated since the volume of the whole gland is commonly used in diagnosis and follow-up of the patients.

Repeated measures ANOVA showed no significant differences in thyroid volume measurements by the three separate observers (P = 0.631) or by the same observers (P = 0.689).

The interobserver variation is given in Table 1. The widest interobserver variation was encountered in determining the craniocaudal diameter of the thyroid

Table 2 Intraobserver variation in sonographic measurement of thyroid volume in children. All numbers are percentages.

	Right lobe diameters			Left lobe diameters			
	Thyroid volume	Mediolateral	Anteroposterior	Craniocaudal	Mediolateral	Anteroposterior	Craniocaudal
Mean \pm s.d.	8.4 \pm 6.8	7.9 \pm 6.2	7.5 \pm 5.7	7.5 \pm 5.9	8.9 \pm 6.8	11.4 \pm 8.8	6.9 \pm 5.8
Ninety-five percent confidence interval (mean \pm 2 s.d.s)	21.9	20.3	18.9	19.4	22.5	29.0	18.4

gland. No correlation was found between the thyroid volume and the interobserver variation ($r = -0.12$, $P = 0.27$).

The intraobserver variation is given in Table 2. A slight but statistically significant correlation was found between the thyroid volume and the intraobserver variation in determining the thyroid volume ($r = -0.26$, $P = 0.012$). That is, intraobserver variation in determination of the thyroid volume tends to decrease slightly as the volume of the gland increases.

Discussion

US is a safe, easy-to-perform method and is used as a non-invasive tool to image the thyroid. US measurement of thyroid volume is used in diagnosis and follow-up of patients, and normal US thyroid volumes in adult and pediatric age groups are published in many references and may be used as a guide to decide whether the thyroid size is normal or not (11–15). However, to our knowledge, no study has been performed to determine the interobserver and intraobserver variation in US measurement of thyroid volume in children.

The thyroid volume of a child varies with age. The gland enlarges as the patient grows. Our patient population included children from different age groups that have statistically different thyroid volumes. We then calculated and presented our data in percentages to assess the results more reliably. To calculate the ranges for interobserver and intraobserver variations, we added $2 \times$ s.d. to the mean values, since this range would result in a 95% confidence interval.

Our data show that up to 30% interobserver error may occur in determination of thyroid volume by US. That means that there is a 95% probability that a measurement of thyroid volume performed by one observer will be within $\pm 30\%$ of the measurement performed by another observer. This value may seem to be high at first sight, but the volume of the thyroid gland is calculated using the formula for the volume of an ovoid ($\text{width} \times \text{depth} \times \text{length} \times \pi/6$). That means that any difference between each diameter of the thyroid may increase the error separately between the two calculated volumes of the gland by two separate observers. When each measurement by an observer is 10% more than that of another, the total volume of the gland calculated from the measurements of the first observer would be 33% more than the total volume of the gland calculated from the measurements of the other. Compared with a study of interobserver variation in US measurement of renal volume in children (16), our result for interobserver variation is smaller. We think that this difference may be due to the fact that imaging of the thyroid is easier because the gland is more superficially located than the kidney and is not susceptible to the motion caused by respiration of the patient.

Intraobserver variation in our study was calculated to be $8.4 \pm 6.8\%$. That is, a measurement of the thyroid volume by a single observer will be within approximately $\pm 22\%$ of a second measurement performed by the same observer with a probability of 95%. As one might expect, intraobserver variation is smaller than interobserver variation.

In conclusion, differences equal to or smaller than 30% in US determination of thyroid volume in children by two separate observers will be within the 95% confidence interval and may be regarded as statistically not significant. The variation was independent of the thyroid volume. Thus, this value for interobserver variability can be applied to all age groups and thyroid volumes. Caution is suggested in measuring the craniocaudal diameter of the gland since the widest variation was encountered in estimation of this diameter. In the same way, differences smaller than 22% between two separate measurements performed by the same observer in US estimation of the thyroid volume may be regarded as statistically not significant. We believe that these data are of value in evaluating patients with diseases that could affect the thyroid gland.

Acknowledgements

The authors wish to thank Prof. Dr N Yordam and her colleagues in the Department of Pediatric Endocrinology in Hacettepe University.

References

- Hintze G, Emrich D & Kobberling J. Treatment of endemic goitre due to iodine deficiency with iodine, levothyroxine or both: results of a multicentre trial. *European Journal of Clinical Investigation* 1989 19 527–534.
- Hintze G, Windeler J, Baumert J, Stein H & Köbberling J. Thyroid volume and goitre prevalence in the elderly as determined by ultrasound and their relationships to laboratory indices. *Acta Endocrinologica* 1991 124 12–18.
- Hintze G & Kobberling J. Treatment of iodine deficiency goiter with iodine, levothyroxine or a combination of both. *Thyroidology* 1992 4 37–40.
- Rzepka AH, Cissewski K, Olbricht T & Reinwein D. Effectiveness of prophylactic therapy on goiter recurrence in an area with low iodine uptake – a sonographic follow-up study. *Clinical Investigation* 1994 72 967–970.
- Cheung NW, Lou JC & Boyages SC. Growth hormone does not increase thyroid size in the absence of thyrotropin: a study in adults with hypopituitarism. *Journal of Clinical Endocrinology and Metabolism* 1996 81 1179–1183.
- Liesenkötter KP, Kiebler A, Stach B, Willgerodt H & Grüters A. Small thyroid volumes and normal iodine excretion in Berlin schoolchildren indicate full normalization of iodine supply. *Experimental Clinical Endocrinology and Diabetes* 1997 105 46–50.
- Nygaard B, Faber J, Veje A & Hansen JEM. Thyroid volume and function after ^{131}I treatment of diffuse non-toxic goitre. *Clinical Endocrinology* 1997 46 493–496.
- Cheung NW & Boyages SC. The thyroid gland in acromegaly: an ultrasonographic study. *Clinical Endocrinology* 1997 46 545–549.
- Smyth PPA, Hetherington AMT, Smith DF, Radcliff M & O'Herlihy C. Maternal iodine status and thyroid volume during pregnancy: correlation with neonatal iodine intake. *Journal of Clinical Endocrinology and Metabolism* 1997 82 2840–2843.
- Todd CH & Dunn J. T Intermittent oral administration of potassium iodide solution for the correction of iodine deficiency. *American Journal of Clinical Nutrition* 1998 67 1279–1283.
- Hegedüs L, Perrild H, Poulsen LR, Andersen JR, Holm B, Schnohr P *et al.* The determination of thyroid volume by ultrasound and its relationships to body weight, age and sex in normal subjects. *Journal of Clinical Endocrinology and Metabolism* 1983 6 260–263.
- Berghout A, Wiersinga MW, Smits NJ & Touber JL. Determinants of thyroid volume as measured by ultrasonography in healthy adults in a non-iodine deficient area. *Clinical Endocrinology* 1987 26 273–280.
- Chanoine JP, Toppet V, Lagasse R, Spehl M & Delange F. Determination of thyroid volume by ultrasound from the neonatal period to late adolescence. *European Journal of Pediatrics* 1991 150 395–399.
- Vitti P, Martino E & Aghini-Lombardi F. Thyroid volume measurement by ultrasound in children as a tool for the assessment of mild iodine deficiency. *Journal of Clinical Endocrinology and Metabolism* 1994 79 600–603.
- Delange F, Benker G & Caron P. Thyroid volume and urinary iodine in European schoolchildren: standardization of values for assessment of iodine deficiency. *European Journal of Endocrinology* 1997 136 180–187.
- Sargent MA, Long G, Karmali M & Cheng SM. Interobserver variation in the sonographic estimation of renal volume in children. *Pediatric Radiology* 1997 27 663–666.

Received 17 September 1998

Accepted 16 November 1998