Which Method is Gold Standard for Determination of Thyroid Volume?

¿Cual es el Método Gold Standard para Determinar el Volumen de la Glándula Tiroides?

Nurcan Erçikti1; Niyazi Acer2; Nihal Apaydi3; İnanç Güven4 & Gökmen Zararsız5


SUMMARY: Change of the thyroid gland volume is often the symptom of most common pathological conditions some thyroid diseases. The exact calculation for the thyroid volume is very important for the assessment and management of thyroid disorders. The volume of thyroid gland, using computed tomography (CT), ultrasound (USG) and magnetic resonance imaging (MRI) has been accessed in few studies published; however a gold standard method has not yet been determined. The purpose of this study was to estimate the volume of normal thyroid gland to define an optimal correction factor therefore was to compare different techniques using the CT. We used computed tomography images obtained from 8 cadavers (2 females, 6 males) to calculate the thyroid volumes. In the present study, the actual thyroid volumes were measured using the water-displacement method as a gold standard, point-counting as a stereology, and ellipsoid methods. Mean squared errors and correction factors were calculated and modeled for each model to find an optimal correction factor and from 0.450 to 0.600 in steps of 0.001 separately for thyroid volume estimation. The average volume of the thyroid glands were 14.58 ± 9.84, 15.28 ± 9.38, and 14.97 ± 8.35 cm³ by fluid displacement, stereology and ellipsoid formula, respectively. No significant difference was found among the methods (P >0.05). The results of this study suggested that the volume of thyroid gland can be measured on CT scans stereologically for diagnosis, as well as provide reliable measure of thyroid volume, management and follow-up of thyroid diseases and for preoperative planning.

KEY WORDS: Thyroid volume; Actual volume; Correction factor; Stereology.

INTRODUCTION

Thyroid volume (TV) determination by using planar scintigraphy is a routine procedure performed in nuclear medicine clinics. Assessment of TV is also used as a method for follow-up after treatment of thyroid gland with radioactive iodine (Lucas, 2000).

Ultrasound (USG) estimation of TV is reported as a safe, cheap and easily available method. However, determination of the TV by using 2D USG-TV has reported to cause some problems like over-estimation or under-estimation of the TV (Yokoyama et al., 1986; Zaidi, 1996; van Isselt et al., 2003; Ruggieri et al., 2008). Magnetic resonance imaging (MRI) has also been reported as a gold standard technique for measuring TV. But it is relatively more expensive than USG (van Isselt et al.). Volume determination by means of MRI and USG are time-consuming and biased by where the operator decides to draw the ROI on a blurred and noisy image. In addition to this; high cost, availability and capacity of MRI limits the clinical application of MRI for diagnosis and follow up of thyroid diseases such as Graves and hyperthyroidism.

Analysis of the thyroid gland based on 3D USG data which uses a semi-automatic segmentation approach for classification (Kollorz et al., 2008). Ruggieri et al. and Trimboli et al. (2008) used an elliptic formula by 2D-USG and compared it with the postsurgical total TV measured by Archimedes’ principle. They found that the mean USG volume was significantly lower than the mean postsurgical total thyroid volume and they developed mathematical for-

1 Health Sciences University, Gulhane Medical Faculty, Department of Anatomy, Ankara, Turkey
2 University School of Medicine, Department of Anatomy, Kayseri, Turkey
3 Ankara University Faculty of Medicine, Department of Anatomy, Turkey.
4 Medical Park, Ankara Hospital, Ankara, Turkey.
5 Erciyes University School of Medicine, Department of Biostatistics, Kayseri, Turkey.
mulas in order to reduce USG volume underestimation and to predict the real TV using a linear model. Vurdem et al. (2012) compared USG-TV with the actual TV and reported that USG resulted in a 10.62 % underestimation of TV. The Archimedean principle is highly accurate for determination of volume so it is widely accepted as Gold Standard technique. However it is not possible application in routine practice because it is needed to remove the thyroid gland to make this measurement.

Computed Tomography (CT) is also one of the routinely used methods for evaluation of thyroid diseases. However, to our knowledge, there is only study which compares the TV determined by CT and by the gold standard method which was previously suggested as fluid displacement based on Archimedes’ principle (Lee et al., 2014). So, the aim of the present study is to compare these techniques and to define an optimal correction factor for to determine the volume of thyroid gland, specifically by CT, linear measurement and fluid displacement method.

MATERIAL AND METHOD

The TV was calculated on 8 cadavers (16 thyroid lobes) with no known pathology or surgery related with thyroid or their neck region. The mean age of the cadavers (2 females, 6 males) at the time of their death was 51 years (range, 27-75 years).

The TV were evaluated using three different techniques which are:

1) Fluid displacement technique
2) Volumetric ellipsoid method
3) CT by using Cavalieri principle

Fluid displacement technique. The thyroid glands of the cadavers were dissected and extracted from the neck carefully without disturbing the gland and its capsule. The exact thyroid gland volume was calculated using fluid displacement technique, which also known as the ‘Archimedes’ principle’. For this purpose, each gland was immersed in a 500 ml graduated cylinder filled with distilled water at room temperature. And the volume of the displaced fluid was recorded. Each measurement was performed twice, and the mean of the measurements was accepted as the actual volume (Acer et al., 2011).

Volumetric ellipsoid method. The width, height, and depth of each thyroid lobe were calculated by a digital caliper with 2 decimals sensitivity. Finally, we used the volumetric ellipsoid method (width × height × depth × 0.52) for to determine the TV as previously described by Shabana et al. (2003)

CT by using Cavalieri principle. All cadavers were scanned in anatomic position using a high-resolution scanner before dissection. All examinations were performed with a 16-slice scanner (Somatom ARC; Siemens). The scanning parameters were 1.25-mm collimation, 3-mm slice thickness, 20-cm field of view, 512 × 512-pixel matrix, 120 kVp, and 240 mAs. For all CT scans, the pixel dimensions were 0.40x 0.40 mm².

Measurement of the thyroid volume using Cavalieri method. A section series with 3-mm thickness for CT images of were used to estimate the TV. These images transferred to a personal computer and transparent square grid test system (with d = 0.15 cm between test points) was transposed on the film to cover the entire image frame. The points hitting the thyroid sectioned surface area for each section were counted, and the TV was estimated by using a previously described formula (Acer et al., 2007; Acer et al., 2008; Acer et al., 2009; Acer et al., 2010; Acer et al., 2011) (Fig. 1).

\[ V = T \times \left[ d^2 \right] \times \sum p \]

Where ‘T’ is the section thickness, ‘d’ the distance between the test points of the grid, ‘P’ is the total number of points hitting the sectioned cut surfaces area of the structure. According to this volumetric technique, a square grid of test points was positioned on each CT images, and all points hitting the structure were counted (Fig. 1).

Error prediction and confidence interval formulae. Coefficient of error (CE) and confidence interval (CI) calculations were performed according to García-Fiñana (García-Fiñana et al., 2003; García-Fiñana, 2006; García-Fiñana et al., 2009). We make the calculation a lower and upper bound values for the TV by applying to the Cavalieri sample.

To estimate Var(QT) via Eq. (2) we have to calculate first α(q), C0, C1, C2 and C4 (Ertekin et al., 2011). From Eq. (3), we have:

\[ \text{var}(Q_T) = a(q)(3C_0 - 4C_1 + C_2)T^2 q \in [0,1] \]

\[ C_k = \sum_{i=1}^{n-k} f_i f_{i+k}, k = 1,2,...,n-1 \]
Equation (3) is an extended version of the variance estimator according to Garcia Finana (2006). The quantities \( C_0, C_1 \) and \( C_2 \) computed from the systematic data sample of as follows:

The smoothness constant \( q \) and coefficient \( \alpha(q) \) can be estimated from Eq. (4) and (5) as follows:

\[
q = \left\{ \begin{array}{cl}
0, & \frac{1}{2 \log 2} \log \left[ \frac{3C_0 - 4C_1 + C_2}{3C_0 - 4C_1 + C_2} \right] - \frac{1}{2} \\
\end{array} \right.
\]

\[
\alpha(q) = \frac{\Gamma(2q + 2)\zeta(2q + 2)\cos(\pi q)}{(2\pi)^{2q+2}} (1 - 2^{2q-1}) \quad q \in [0,1]
\]

where \( \Gamma \) and \( \zeta \) show the gamma function and the Riemann Zeta function, respectively.

We applied Eq. (5) with \( \alpha = 0.92 \).

The coefficient \( \alpha(q) \) has the following expression:

\[
\alpha(0.92) = \frac{\Gamma(2 \times 0.92 + 2)\zeta(2 \times 0.92 + 2)\cos(\pi 0.92)}{(2\pi)^{2 \times 0.92+2}} (1 - 2^{2 \times 0.92-1}) = 0.0056
\]

Therefore, the estimate of \( \text{Var}(Q^T) \) obtained via Eq. (2) is:

\[
\text{Var}(Q^T) = \alpha(q)(3C_0 - 4C_1 + C_2)\gamma^2
\]

\[
\text{Var}(Q^T) = 0.0056 \times (3 \times 51076 - 4 \times 49448 + 45221) \times (0.3)^2
\]

\[
\text{Var}(Q^T) = 0.33
\]

Therefore, the estimate of \( \text{Var}(Q^T) \) obtained via Eq. (2) is:

We predicted coefficient of error:

\[
CE(Q^T) = \sqrt{2.26 / 13.61} = 0.042 = 4.2\%
\]

The upper and lower CI interval for the TV is obtained by applying Eq. (6). We have:

\[
\left( Q_T mT \lambda_q \times \sqrt{\alpha(q)} \left( 3C_0 - 4C_1 + C_2 \right) \right)
\]

\[
\left( 13.61 - 3.3 \times \sqrt{0.33}, 13.61 + 3.3 \times \sqrt{0.33} \right)
\]

\[
= (11.7 - 15.1) \text{cm}^3
\]

We have used the identity \( \lambda_{0.92} = 3.3 \) according to Garcia-Finana (2006). We found a satisfactory predicted interval and it provides a relatively narrow confidence interval.

**Statistical analyses.** The results of the measurements were recorded as mean, standard deviation, minimum and maximum values. The statistical differences of the estimated volumes obtained by 3 different approaches were compared using Friedman Analysis, and Turkey test. Bland–Altman analysis was also applied to determine the difference between the results of two different approaches. Pearson correlation test was applied to evaluate the relation between the results of volume estimates. Mean squared errors and correction factors were calculated and modeled for each model to find an optimal correction factor and from 0.450 to 0.600 in steps of 0.001 separately for thyroid volume estimation. We used R 3.0.2 software (www.r-project.org) by considering a p value less than 5% as statistically significant.
RESULTS

The mean thyroid height, width and depth by linear measurements were 44.87 mm ± 7.59 mm, 45.81 mm ± 12.37 mm and 13.16 mm ± 4.02 mm, respectively. The mean ± SD thyroid volumes used by the fluid displacement technique, by CT using Cavalieri principle and by volumetric ellipsoid method were 14.58 ± 9.84 cm³, 15.28 ± 9.38 cm³, and 14.97 ± 8.35 cm³, respectively. We didn’t found statistically difference among the methods of calculating thyroid volume (P >0.05).

Table I. Calculation of the constants C0, C1, C2, C4.

<table>
<thead>
<tr>
<th>Section, i</th>
<th>Pi</th>
<th>Pi, Pi+1</th>
<th>Pi, Pi+2</th>
<th>Pi, Pi+4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>14</td>
<td>33</td>
<td>65</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>196</td>
<td>910</td>
<td>1190</td>
</tr>
<tr>
<td>3</td>
<td>33</td>
<td>1089</td>
<td>2145</td>
<td>2805</td>
</tr>
<tr>
<td>4</td>
<td>65</td>
<td>4225</td>
<td>5525</td>
<td>6110</td>
</tr>
<tr>
<td>5</td>
<td>85</td>
<td>7225</td>
<td>7990</td>
<td>7225</td>
</tr>
<tr>
<td>6</td>
<td>94</td>
<td>8836</td>
<td>7990</td>
<td>8084</td>
</tr>
<tr>
<td>7</td>
<td>85</td>
<td>7225</td>
<td>7310</td>
<td>6885</td>
</tr>
<tr>
<td>8</td>
<td>86</td>
<td>7396</td>
<td>6966</td>
<td>5676</td>
</tr>
<tr>
<td>9</td>
<td>81</td>
<td>6561</td>
<td>5346</td>
<td>3888</td>
</tr>
<tr>
<td>10</td>
<td>66</td>
<td>4356</td>
<td>3168</td>
<td>2112</td>
</tr>
<tr>
<td>11</td>
<td>48</td>
<td>2304</td>
<td>1536</td>
<td>1104</td>
</tr>
<tr>
<td>12</td>
<td>32</td>
<td>1024</td>
<td>736</td>
<td>320</td>
</tr>
<tr>
<td>13</td>
<td>23</td>
<td>529</td>
<td>230</td>
<td>69</td>
</tr>
<tr>
<td>14</td>
<td>10</td>
<td>100</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>726</td>
<td>51076</td>
<td>49448</td>
<td>45221</td>
</tr>
<tr>
<td>Co</td>
<td>C1</td>
<td>C2</td>
<td>C4</td>
<td></td>
</tr>
</tbody>
</table>

Table II. Confidence interval (CI) and CE of thyroid volume parameters with % 95 confidence interval value for Cavalieri.

<table>
<thead>
<tr>
<th>Case number</th>
<th>Cavalieri volume</th>
<th>Cl lower</th>
<th>Cl upper</th>
<th>CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27.89</td>
<td>24.72</td>
<td>31.06</td>
<td>3.4</td>
</tr>
<tr>
<td>2</td>
<td>32.1</td>
<td>30.34</td>
<td>33.85</td>
<td>1.7</td>
</tr>
<tr>
<td>3</td>
<td>13.61</td>
<td>11.71</td>
<td>15.51</td>
<td>4.2</td>
</tr>
<tr>
<td>4</td>
<td>9.72</td>
<td>7.20</td>
<td>12.12</td>
<td>5.9</td>
</tr>
<tr>
<td>5</td>
<td>6.84</td>
<td>5.20</td>
<td>8.40</td>
<td>6.1</td>
</tr>
<tr>
<td>6</td>
<td>9.93</td>
<td>5.90</td>
<td>13.90</td>
<td>5.5</td>
</tr>
<tr>
<td>7</td>
<td>13.31</td>
<td>11.14</td>
<td>15.61</td>
<td>4.8</td>
</tr>
<tr>
<td>8</td>
<td>9.44</td>
<td>7.50</td>
<td>12.41</td>
<td>5.5</td>
</tr>
</tbody>
</table>

There was a correlation between three methods (r=0.952, 0.990, 0.963) (Fig. 2AB). When we compared the volumetric ellipsoid method with the fluid displacement method a difference of (-6.6) - 5.9 cm³ was found between the measurements. This was a 13 % underestimation of thyroid gland volume by ellipsoid formula when compared with fluid displacement (Fig. 3A). When we compared Cavalieri with the fluid displacement technique as the gold standard, we found a difference of (-3.6) - 2.2 cm³ between the measurements. This was again a 13 % underestimation of thyroid gland volume by Cavalieri principle (Fig. 3B).

We didn’t found statistically significant difference between a real volume using a correction factor of 0.523 (p = 0.748) according to a paired student’s t test (Fig. 4).

The mean CEs for the thyroid gland volume estimates derived from the technique of the Cavalieri principle was under 5 %. We found a satisfactory predicted interval and this provided a relatively narrow confidence interval. The reference values of TV parameters with 95 % confidence interval (CI) are listed in Table II.
DISCUSSION

According to World Health Organization and the International Council or the Control of Iodine Deficiency Disorders, USG has been accepted method for the estimation of thyroid volume. Because of the irregular profiles of the gland, US thyroid volume may disagree with the in vivo volume (Shabana et al., 2003; Rago et al., 2006). Calculation of thyroid volume is relevance to establish the prevalence of thyroid diseases such as iodine deficiency areas, endemic and sporadic goiter, nodules, thyroiditis, and Graves disease (Lee et al.).

Vurdem et al. compared the volumes of thyroidectomy specimens with different approaches and revealed that the mean ± SD thyroid volumes of the fluid displacement, Cavalieri and ellipsoid were 82.75 ± 48.87, 80.45 ± 48.96, and 75.50 ± 46.59 cm³, respectively. The authors found no significant difference among the methods of calculating thyroid volume (P > 0.05). They found that the 2D-USG volume is a 10.62 % underestimation of the thyroid gland volume compared with the actual volume. Miccoli et al. (2006), reported that 88 % of cases were underestimated by ultrasound, Trimboli et al. found that the volume was underestimated by ultrasound in 77% of cases and Lee et al., reported that was underestimated by ultrasound in 38 %. We found that was 13 % underestimation of thyroid gland volume by ellipsoid formula when compared with fluid displacement.

Berghout et al. (1987) found a mean adult TV of 10.7±4.6 ml in healthy adult subjects. Hegedüs et al. (1983) found the mean male TV was 19.6±4.7 ml and the mean female TV was 17.5 ± 4.2. Nygaard et al. (1993) found that TV’s were 12 ml, 18 ml, 18 ml and 18 ml for 4 age groups lines as 15, 30, 45 and 60 respectively in 391 women from Denmark. Oberhofer et al. (1989) found the mean thyroid volume as 13.35 ml (men 14.94 ml, women 12.09 ml in 500 healthy adults at 1989 in Austria. Gómez et al. (2002), found the thyroid volume as 9.8 ± 4.6 ml for man and 6.5±2 ml for women. Seker & Tas (2010) reported the mean thyroid volume 15.87±7.18 ml and 10.94±4.53 ml respectively. The mean value of the thyroid gland volume measured by the fluid displacement, the Cavalieri and ellipsoid methods were 14.58 ± 9.84, 15.28 ± 9.38, and
14.97 ± 8.35 cm³, respectively. Our mean thyroid gland volume was smaller than Hegedüs et al. Studies, bigger than other studies. A possible reason for this difference may be the use of different imaging techniques. So we used fluid displacement technique for thyroid gland volume estimation and compared CT scan result.

There are a few reports about correction factor for thyroid gland volume using ellipsoid formula. Brunn et al. (1981) suggested the use of a correction factor of 0.479 instead of the accepted 0.524. Shabana et al. (2003) proposed a correction factor of 0.529 when using the ellipsoid formula. Their study has got some limitations. Their studies were based on volume measurement by MDCT. We found our correction factor as 0.524 (π / 6) which also is acceptable according to literature.

Shabana et al. (2006), Nygaard et al. (2002) and Lee et al. found that thyroid volumes using ultrasound did not different significantly from CT-estimated thyroid volumes. They suggested that the CT volumes were reproducible and that CT volumetry could also be used in goiters with substernal extension.

We used different correction factors for calculating actual TV using fluid displacement method to compare volume measurements using CT images. We suggest the use of a mean value of 0.523 to calculate the volume of the thyroid lobe according to real volume. The stereological approach we used provides an opportunity for the investigator making appropriate changes on their sampling or estimating procedures. Therefore, the Cavalieri method gives a coefficient of error (CE) of estimation for each volume assessments. If you use too few slices or too few points are taken for volume estimation, such problems may occur. The investigator can change the spacing of points in the grid or the number of slices available in any MRI or CT studies to provide a reasonable CE value (11, 27). The coefficient of error and confidence interval predicted the recent literature (Sahin et al., 2003; Acer et al., 2009). CE value is acceptable in a range of less than 5 % according to the literature. An acceptable grid size and the number of slices required for volume estimation of an object is important at the beginning.

This study has some limitations: small sample size and other accurate modalities for thyroid volume estimation like 2D and 3D ultrasound were not used in our study. CT application in medicine has got some disadvantages such as ionizing radiation applied on the body in CT scans is energetic enough to directly or not damage to DNA.

We confirmed that there is no significant difference in calculation of thyroid volume by gold standard methods and Cavalieri’s principle. So the TV can be measured on CT scans stereologically for diagnosis, management and follow-up of thyroid diseases and for preoperative planning.


RESUMEN: El cambio de volumen de la glándula tiroides es a menudo el síntoma de las condiciones patológicas más comunes de algunas enfermedades de dicha glándula. El cálculo exacto del volumen tiroides es muy importante para la evaluación y el tratamiento de los trastornos tiroides. El volumen de la glándula tiroides, utilizando la tomografía computarizada (TC), el ultrasonido (USG) y la resonancia magnética (RM) ha sido presentado en varias publicaciones. Sin embargo, aún no se ha determinado un gold standard. El propósito de este estudio fue estimar el volumen de la glándula tiroides normal para definir un factor de corrección óptimo, por lo que se compararon diferentes técnicas utilizando TC. Para calcular los volúmenes tiroides se utilizaron imágenes de tomografía computarizada obtenidas de 8 cadáveres (dos mujeres y seis hombres). En el presente estudio, los volúmenes reales de la glándula tiroides se midieron utilizando como gold standard los métodos esterológicos de desplazamiento de agua y conteo de puntos y el método volumétrico elipsoide. Se calcularon y modelaron los errores cuadráticos medios y los factores de corrección para cada modelo con el objetivo de encontrar un factor de corrección óptimo y de 0.450 a 0.600 en pasos de 0,001 por separado para la estimación del volumen tiroides. El volumen medio de las glándulas tiroides fue de 14,58 ± 9,84, 15,28 ± 9,38 y 14,97 ± 8,35 cm³ calculados por desplazamiento de fluido, esterología y fórmula elipsoide, respectivamente. No se encontró diferencia significativa entre los métodos (P>0,05). Los resultados de este estudio sugieren que el volumen de la glándula tiroides puede ser medido estereológicamente por TC, estableciéndose como una medida fiable del volumen tiroides, para el diagnóstico, manejo y seguimiento de las enfermedades tiroides y la planificación preoperatoria.

PALABRAS CLAVE: Volumen; Volumen real; Factor de corrección; Estereología.
REFERENCES


Rago, T.; Bencivelli, W.; Scutari, M.; Di Cosmo, C.; Rizzo, C.; Berti, P.; Miccoli, P.; Pinchera, A. & Vitti, P. The newly developed three-dimensional (3D) and two-dimensional (2D) thyroid ultrasound are strongly correlated, but 2D overestimates thyroid volume in the presence of nodules. J. Endocrinol. Invest., 29(5):423-6, 2006.


