Assessing the accuracy and reliability of ultrasonographic three-dimensional parathyroid volume measurement in a patient with secondary hyperparathyroidism: a comparison with the two-dimensional conventional method

Sung-Hye You¹, Sang-il Suh², In-Seon Ryoo³, Gyu-Ri Son¹, Young-Hen Lee³, Hyung Suk Seo⁵, Nam Joon Lee¹, Hae Young Seol²

¹Department of Radiology, Korea University Anam Hospital, Seoul; ²Department of Radiology, Korea University Guro Hospital, Seoul; ³Department of Radiology, Korea University Ansan Hospital, Ansan, Korea

Purpose: The purpose of this study was to investigate the accuracy and reliability of the semi-automated ultrasonographic volume measurement tool, virtual organ computer-aided analysis (VOCAL), for measuring the volume of parathyroid glands.

Methods: Volume measurements for 40 parathyroid glands were performed in patients with secondary hyperparathyroidism caused by chronic renal failure. The volume of the parathyroid glands was measured twice by experienced radiologists by two-dimensional (2D) and three-dimensional (3D) methods using conventional sonograms and the VOCAL with 30° angle increments before parathyroidectomy. The specimen volume was also measured postoperatively. Intraclass correlation coefficients (ICCs) and the absolute percentage error were used for estimating the reproducibility and accuracy of the two different methods.

Results: The ICC value between two measurements of the 2D method and the 3D method was 0.956 and 0.999, respectively. The mean absolute percentage error of the 2D method and the 3D VOCAL technique was 29.56% and 5.78%, respectively. For accuracy and reliability, the plots of the 3D method showed a more compact distribution than those of the 2D method on the Bland-Altman graph.

Conclusion: The rotational VOCAL method for measuring the parathyroid gland is more accurate and reliable than the conventional 2D measurement. This VOCAL method could be used as a more reliable follow-up imaging modality in a patient with hyperparathyroidism.

Keywords: Imaging, three-dimensional; Parathyroid glands; Organ size; Ultrasonography
Materials and Methods

Patients
This prospective study was approved by the Institutional Review Board of Korean University Guro Hospital, and all participants provided written informed consent at the beginning of the study. Between January 2011 and January 2013, a total of 73 patients with SHPT who failed medical treatment underwent US examinations for assessing the parathyroid glands. The volume of enlarged parathyroid glands in patients with SHPT was measured by an experienced neuroradiologist A (1st measurement) using the conventional 2D method. At the same time, the same radiologist A scanned the hypertrophied parathyroid glands by using the 3D method. The scanning dataset was transferred to a personal computer, and the radiologist A created a reconstructed 3D image and measured the parathyroid volume. The 2D and 3D volumes of 203 parathyroid glands in 73 patients were measured, and these volume data were provided to the nephrologist. Ultimately, 15 patients with 40 parathyroid glands who decided to undergo parathyroidectomy were included in this study. The day before the operation, the same neuroradiologist A re-measured (second measurement) the volume of the target parathyroid glands by using the 2D and 3D methods, and provided the volume data to the surgeon. The interval between the first and the second measurements was not more than 2–4 weeks. Immediately after the resection of the parathyroid glands, the specimen volume was measured by another experienced radiologist B.
as the reference. The contour of the parathyroid gland on the axial plane was drawn, and two calipers were located on the superior and inferior extremities of the target lesion in order to designate the rotational axis. After the radiologist selected the 30° rotational VOCAL method for measuring the parathyroid volume, the program automatically extracted the six sequential planes perpendicular to the reference axial plane. After the manual drawing of the contour of the parathyroid gland on the six acquired planes, the computer calculated the volume of each parathyroid gland automatically and then displayed a reconstructed 3D image of the parathyroid gland with its volume (Fig. 1B). In all, it took approximately 10 minutes to acquire the US data and calculate the parathyroid gland volume by using the 30° rotational VOCAL technique.

Postoperatively, the specimen volume was measured using the 3D method by radiologist B. The specimen was buried in jelly in order to avoid air and was scanned using the 3D probe (Fig. 1C). The scanning dataset was transferred to a personal computer, and the volume was calculated. The schematic representations of the 2D and 3D measurements and the gold standard of a parathyroid gland are shown in Fig. 2.

Statistical Analysis
All statistical analyses were carried out using SAS ver. 9.2 (SAS Institute Inc., Cary, NC, USA). The intraobserver reliability of the 2D and 3D measurements was evaluated by using the intraclass correlation coefficient (ICC) and Bland-Altman graphical approach. The absolute percentage error of each imaging technique was used for comparing the accuracy of the 2D and 3D techniques. The absolute percentage error was defined as the percentage difference between the ultrasonographically measured parathyroid volume and the specimen volume: 100×(|measured vol.-specimen vol.|)/specimen vol. The Bland-Altman graphical approach was also used for comparing the accuracy of the 2D and 3D techniques. The absolute percentage error was defined as the percentage difference between the ultrasonographically measured parathyroid volume and the specimen volume: 100×(|measured vol.-specimen vol.|)/specimen vol. The Bland-Altman graphical approach was also used for comparing the accuracy of the 2D and 3D techniques.

Fig. 1. Example of the two-dimensional (2D), three-dimensional (3D), and specimen volume measurements in a 61-year-old man with a hemodialysis history for 22 years, showing the enlargement of all four parathyroid glands.
A. The volume of the right lower parathyroid gland calculated using the 2D method was 0.40 cm$^3$. B. The volume of the same parathyroid gland calculated using the 3D method was 0.42 cm$^3$. C. The specimen volume was 0.42 cm$^3$. 

3D US for parathyroid glands

used for an easy comparison of the accuracy of the 2D and 3D measurements.

Results

In all, 15 patients (4 men and 11 women) with sHPT caused by CRF were enrolled (median age, 52.9 years; mean dialysis time before parathyroidectomy, 12.3 years). The mean serum iPTH levels immediately prior to surgery were 213.0 pg/mL (Table 1). The mean volume of the parathyroid glands on the first and the second measurement conducted using the 2D ultrasound were 1.22±1.66 (mean±SD, cm$^3$) and 1.23±1.40 (mean±SD, cm$^3$), respectively. There was no significant difference between the two measurements (P=0.915) (Table 2). The mean volume of the parathyroid glands on the first and second measurement conducted using the 3D ultrasound were 1.26±1.62 (mean±SD, cm$^3$) and 1.25±1.58 (mean±SD, cm$^3$), respectively. There was no significant difference between the two measurements (P=0.411) (Table 2). The mean specimen volume of the 40 enlarged parathyroid glands measured by the 3D VOCAL method was 1.23±1.57 (mean±SD, cm$^3$).

The ICC value of the 2D measurement was 0.956, indicating a reproducibility of 95.6% (Table 3). The Bland-Altman graph shows that the difference between the two measurements was −0.01±0.63 (mean±SD, cm$^3$), and the limits of agreement (mean±1.96SD, cm$^3$) were 1.23 and −1.25 (Fig. 3A). The ICC value of the 3D measurement was 0.999, indicating a reproducibility of 99.9% (Table 3). The Bland-Altman graph showed that the difference between the two measurements was 0.01±0.08 (mean±SD, cm$^3$), and the limits of agreement (mean±1.96SD, cm$^3$) were 0.15 and −0.17 (Fig. 3B).

Table 1. Patient characteristics

| Characteristic | Value |
|----------------|-------|
| No. of patients | 15 |
| No. of target nodules | 40 |
| Sex (M:F) | 4:11 |
| Age (yr)$^d$ | 52.9 (19–82) |
| Dialysis type (PD:HD) | 3:12 |
| Dialysis duration (yr) | 12.3 (HD 13.2±4.7, PD 9.0±2.0) |
| PTH before surgery (pg/mL) | 1,213.0±827.4 |
| PTH after surgery (pg/mL) | 142.2±333 |

Values are presented as number or mean±SD.

PD, peritoneal dialysis; HD, hemodialysis; PTH, parathyroid hormone.

$^d$Median age: male 53.3 yr, female 52.8 yr.
The mean absolute percentage difference of the 2D ultrasound measurement of the enlarged parathyroid glands was 29.56±17.37 (mean±SD, %), indicating that the mean measurement error of the 2D method was approximately 30% (Table 3). The Bland-Altman graph for the accuracy of the first 2D measurement showed that the difference between the first 2D measurement and the specimen volume was −0.00±0.41 (mean±SD, cm³), and the limits of agreement (mean±1.96SD, cm³) were 0.81 and −0.80 (Fig. 4A). The difference between the second 2D measurement and the specimen volume was 0.00±0.62 (mean±SD, cm³), and the limits of agreement (mean±1.96SD, cm³) were 1.22 and −1.21 (Fig. 4B).

The mean absolute percentage difference of the 3D ultrasound measurement of the enlarged parathyroid glands was 5.78±6.84 (mean±SD, %), indicating that the mean measurement error of the 3D method was approximately 6%. The Bland-Altman graph for the accuracy of the first 3D measurement showed that the difference between the first 3D measurement and the specimen volume was 0.04±0.08 (mean±SD, cm³), and the limits of agreement (mean±1.96SD, cm³) were 0.19 and −0.15 (Fig. 4C). The difference between the second 3D measurement and the specimen volume was 0.02±0.06 (mean±SD, cm³), and the limits of agreement (mean±1.96SD, cm³) were 0.13 and −0.18 (Fig. 4D).

Table 3. ICC of and mean absolute percentage difference of the 2D and 3D measurements

| Parameter                  | ICC (P-value) | Mean absolute percentage difference (mean±SD, %) | P-value |
|----------------------------|---------------|-----------------------------------------------|--------|
| 2D measurement             |               |                                               |        |
| 1st measurement            | 0.956 (<0.001)| 29.56±17.37                                   | <0.001 |
| 2nd measurement            | 0.999 (<0.001)| 5.78±6.84                                     | <0.001 |
| 3D measurement             |               |                                               |        |
| 1st measurement            | 0.956 (<0.001)| 29.56±17.37                                   | <0.001 |
| 2nd measurement            | 0.999 (<0.001)| 5.78±6.84                                     | <0.001 |
| Specimen volume (cm³)      |               |                                               |        |

Both two-dimensional (2D) and three-dimensional (3D) methods showed good intraclass correlation coefficient (ICC), but the 3D measurement showed slightly better ICC. The 3D measurement exhibited significantly higher accuracy than the 2D measurement.

Table 2. Mean calculated volume and specimen volume of the hypertrophic parathyroid glands

| Parameter                  | Mean±SD | Paired t test (P-value) |
|----------------------------|---------|-------------------------|
| 2D volume measurement (cm³)|         |                         |
| 1st measurement            | 1.22±1.66| 0.915                   |
| 2nd measurement            | 1.23±1.40|                         |
| 3D volume measurement (cm³)|         |                         |
| 1st measurement            | 1.26±1.62| 0.411                   |
| 2nd measurement            | 1.25±1.58|                         |
| Specimen volume (cm³)      | 1.23±1.57|                         |

There was no significant difference between the two measurements in the two-dimensional (2D) and the two-dimensional three-dimensional (3D) methods.

Fig. 3. The Bland-Altman graph for the reliability of the two-dimensional (2D) and the three-dimensional (3D) methods for measuring the parathyroid volume.

A. The plots show the two measurements using 2D ultrasound of one operator against their differences. B. The plots show the two measurements using 3D ultrasound of one operator against their differences. The solid line indicates the mean of the differences between the measurements, whereas the dotted lines indicate the limits of agreement of the measurements.
Discussion

In the medical field, accurate and reliable volume measurement of internal organs is very important. Traditionally, volume has been measured using the 2D method and such measurement requires several assumptions. The basic assumption is that the target object has an ellipsoidal shape, but most of the organs have an irregular shape. Therefore, 3D volume measurement was introduced and has been actively used in computed tomography (CT) scans. For example, 3D volume measurements of the liver and the renal cortex provide useful clinical information [15–17]. However, the main drawbacks of 3D measurement using CT scans are radiation hazards and adverse

![Fig. 4. The Bland-Altman graph for the accuracy of the two-dimensional (2D) and the three-dimensional (3D) methods for measuring the parathyroid volume.](image)

A, B. The plots show the first/second measurements using 2D ultrasound and the specimen volumes against their differences. C, D. The plots show the first/second measurements using 3D ultrasound and the specimen volumes against their differences. The solid line indicates the mean of the differences between the 2D measurements and the specimen volume, whereas the dotted lines indicate the limits of agreement of the measurements.
effects from iodine contrast agents. The 3D ultrasonographic volume measurement technique was developed recently, and several studies have shown that 3D ultrasonographic volume measurement is accurate and reliable in many organs [4-12]. Further, because of the easy applicability of US examinations without the use of radiation and contrast media injection, a 3D US volume measurement is more attractive than a CT-based measurement.

There are several 3D US volume measurement techniques. We adopted the 30° rotational VOCAL technique on the basis of the following previous studies. Kot et al. [12] compared the 3D parallel planes technique and freehand technique using three phantoms. Both techniques showed high accuracy and reliability, but the 3D multiplanar technique (3D parallel planes technique) was more accurate (accuracy, 81.5%-83.4% vs. 74.7%-84.2%). Another phantom study compared multiplanar (5-mm interval), VOCAL (30°) and extended imaging VOCAL (5, 10, 15, and 20 planes) methods [11]. All three methods were reliable and accurate (average accuracy, 97.11%). Raine-Fenning et al. [18] showed that rotational VOCAL was superior to the multiplanar method for irregularly shaped organs, but there were no significant differences in measurements among the different rotation angles (VOCAL, 6°, 9°, 15°, and 30°). The relatively high reliability and accuracy of the rotational technique can be explained by the fact that in the multiplanar technique, it is difficult to define the starting and the finishing points for measurement. This problem is overcome in the rotational technique. Further, Peixoto-Filho et al. [9] revealed that the use of a 30° rotation step as compared to a 15° rotation step resulted in a significantly faster evaluation while maintaining excellent reliability in both cases. To summarize, among the several 3D measurement techniques, the rotational VOCAL technique is more accurate and reliable than other 3D measurement techniques such as the 3D parallel planes technique, and the 30° rotational step is time-efficient for measuring the organ volume while maintaining sufficient accuracy and reliability. On the basis of these findings, we selected the 30° rotational VOCAL technique.

In our study, both the 2D and 3D methods showed good reliability, but the 3D ultrasound volume measurement method was more reliable than the 2D method. The ICC of the 2D and 3D measurements was 0.956 and 0.999, respectively. The lower reproducibility of the 2D ultrasound may be explained by the fact that the volume was calculated by using the ellipsoid equation. In this equation, subtle variations in the measurements of each of the three dimensions lead to a large difference in the calculated total volume. Further, in the 2D measurement, it is difficult to obtain the same and exact axial and longitudinal planes of the parathyroid gland on each US examination. However, because the 3D ultrasound volume measurement outlines the parathyroid glands and calculates the volume, it provides more reliable volume data. In practice, when the clinician selects the follow-up imaging modality, he/she considers the two important points of the modality, namely reliability and convenience such as time. Thus, our results have clinical implications, described in the two following points. First, the 3D ultrasonic volume measurement could be a better follow-up imaging modality in patients with sHPT with respect to high reliability. Clinically, high reliability is very important in patients with sHPT for the purpose of evaluating the response to medical therapy. Depending on the change in the parathyroid volume, clinicians will adjust the medication dose in order to give the minimum dose to minimize side effects while ensuring response to the treatment. Second, the 2D method may be a sufficient tool for the follow-up of the parathyroid volume with respect to time because it has a relatively high ICC (95.6%) and takes less time to perform. Approximately an additional 10 minutes is required for measuring the volume by using the 3D method.

This study showed that the 3D ultrasound volume measurement is more accurate than the 2D method. The mean absolute percentage difference of the 3D and the 2D ultrasound measurement of the enlarged parathyroid glands is approximately 6% and 30%, respectively. These results suggest that the volume data from the 3D measurement could help determine the optimal time for parathyroidectomy. As mentioned above, if the volume of one or more parathyroid glands exceeds 0.5 cm³ and medical treatment fails, the clinician considers parathyroidectomy [3]. Further, the reconstructed 3D images and the volume information of each hyperplastic parathyroid gland may aid the surgeon to build an image of the target parathyroid gland. Finally, it could provide more accurate clinical information about the response to medical treatment.

Our study has some limitations. First, the number of cases that we examined was small, which prevents us from drawing firmer conclusions from this preliminary result. Second, there is a problem with respect to the gold standard. Several previous in vivo studies revealed the high accuracy of the 3D rotational VOCAL method using the phantoms that have the gold standard of the volume [11,12,18]. Ideally, the gold standard for measuring the volume of the parathyroid gland is the water displacement method by dropping the specimen into the water. However, in reality, we did not adopt this method because the specimen could not be perfectly separated from the fat and the adjacent connective tissues. If the volume was measured by the water displacement method, it would include the volume of the non-parathyroid soft tissues. Instead, we measured the specimen volume after immersing the specimen in jelly, which provides a different tissue contrast. Although we recognize that this is not the ideal gold standard, we accept the accuracy of this technique by using this indirect method.
In conclusion, the rotational VOCAL method for measuring the parathyroid gland is more reliable and accurate than the 2D method, and relatively large-scale studies are required for investigating the value of this new approach.

ORCID: Sung-Hye You: http://orcid.org/0000-0001-5005-9383; Sang-il Suh: http://orcid.org/0000-0001-8933-0492; In-Seon Ryoo: http://orcid.org/0000-0001-6866-7926; Gyu-Ri Son: http://orcid.org/0000-0001-0614-071X; Young-Hen Lee: http://orcid.org/0000-0001-7739-4173; Hyung Suk Seo: http://orcid.org/0000-0003-2303-7415; Nam Joon Lee: http://orcid.org/0000-0003-4231-5409

Conflict of Interest
No potential conflict of interest relevant to this article was reported.

References
1. Indridason OS, Heath H 3rd, Khosla S, Yohay DA, Quarles LD. Non-suppressible parathyroid hormone secretion is related to gland size in uremic secondary hyperparathyroidism. Kidney Int 1996;50:1663-1671.
2. Druke TB. Cell biology of parathyroid gland hyperplasia in chronic renal failure. J Am Soc Nephrol 2000;11:1141-1152.
3. Schomig M, Ritz E. Management of disturbed calcium metabolism in uremic patients: 3. Potential perspectives: calcimimetics. Nephrol Dial Transplant 2000;15 Suppl 5:30-31.
4. Boito SM, Struijk PC, Ursem NT, Stijnen T, Wladimiroff JW. Assessment of fetal liver volume and umbilical venous volume flow in pregnancies complicated by insulin-dependent diabetes mellitus. BJOG 2003;110:1007-1013.
5. Kuno A, Hayashi Y, Akiyama M, Yamashiro C, Tanaka H, Yanagihara T, et al. Three-dimensional sonographic measurement of liver volume in the small-for-gestational-age fetus. J Ultrasound Med 2002;21:361-366.
6. Araujo Junior E, Nardozza LM, Pires CR, Filho HA, Moron AF. Comparison of the two-dimensional and multiplanar methods and establishment of a new constant for the measurement of fetal lung volume. J Matern Fetal Neonatal Med 2008;21:81-88.
7. Kalache KD, Espinoza J, Chaiworapongsa T, Londono J, Schoen ML, Treadwell MC, et al. Three-dimensional ultrasonic fetal lung volume measurement: a systematic study comparing the multiplanar method with the rotational (VOCAL) technique. Ultrasound Obstet Gynecol 2003;21:111-118.
8. Roelfsema NM, Hop WC, Boito SM, Wladimiroff JW. Three-dimensional sonographic measurement of normal fetal brain volume during the second half of pregnancy. Am J Obstet Gynecol 2004;190:275-280.
9. Peixoto-Filho FM, Sa RA, Lopes LM, Velarde LG, Marchiori E, Ville Y. Three-dimensional ultrasound fetal urinary bladder volume measurement: reliability of rotational (VOCAL) technique using different steps of rotation. Arch Gynecol Obstet 2007;276:345-349.
10. Sur SD, Clewes JS, Campbell BK, Raine-Fenning NJ. Embryo volume measurement: an intraobserver, intermethod comparative study of semiautomated and manual three-dimensional ultrasound techniques. Ultrasound Obstet Gynecol 2011;38:516-523.
11. Barreto EQ, Milani HJ, Araujo Junior E, Haratz KK, Rolo LC, Nardozza LM, et al. Reliability and validity of in vitro volume calculations by 3-dimensional ultrasonography using the multiplanar, virtual organ computer-aided analysis (VOCAL), and extended imaging VOCAL methods. J Ultrasound Med 2010;29:767-774.
12. Kot BC, Sin DM, Ying M. Evaluation of the accuracy and reliability of two 3-dimensional sonography methods in volume measurement of small structures: an in vitro phantom study. J Clin Ultrasound 2009;37:82-88.
13. Zimmermann P, Takala T, Poyhonen L, Punnonen R. Ultrasonography of the thyroid gland in pregnancies complicated by autoimmune thyroid disease. J Clin Ultrasound 1993;21:109-113.
14. Vitti P, Martino E, Aghini-Lombardi F, Rago T, Antonangeli L, Maccherini D, et al. Thyroid volume measurement by ultrasound in children as a tool for the assessment of mild iodine deficiency. J Clin Endocrinol Metab 1994;79:600-603.
15. Shin CI, Kim SH, Rhim JH, Yi NJ, Suh KS, Lee JM, et al. Feasibility of commercially available, fully automated hepatic CT volumetry for assessing both total and territorial liver volumes in liver transplantation. J Korean Soc Radiol 2013;68:125-136.
16. Goumard C, Perdigao F, Cazejust J, Zalinski S, Soubrane O, Scatton O. Is computed tomography volumetric assessment of the liver reliable in patients with cirrhosis? HPB (Oxford) 2014;16:188-194.
17. Muto NS, Kamishima T, Harris AA, Kato F, Onodera Y, Terae S, et al. Renal cortical volume measured using automatic contouring software for computed tomography and its relationship with BMI, age and renal function. Eur J Radiol 2011;78:151-156.
18. Raine-Fenning NJ, Clewes JS, Kendall NR, Bunkheila AK, Campbell BK, Johnson IR. The interobserver reliability and validity of volume calculation from three-dimensional ultrasound datasets in the in vitro setting. Ultrasound Obstet Gynecol 2003;21:283-291.