Three-Dimensional Sonographic Morphologic Assessment of Adnexal Masses

A Reproducibility Study

Juan Luis Alcázar, MD, Manuel García-Manero, MD, Rosendo Galván, MD

Objective. The purpose of this study was to assess the reproducibility of 3-dimensional (3D) sonography for classifying adnexal masses. ***Methods.*** Eighty-two consecutive women with the diagnosis of an adnexal mass on 2-dimensional transvaginal sonography were reevaluated by 3D sonography, and 3D volume data from each mass were stored. Two different examiners (6 years and 1 year of experience in 3D sonography, respectively) reviewed 3D sonograms 1 month after the last patient was recruited and then 1 week later again. Masses had to be classified as benign or malignant. Criteria suggestive of malignancy were the presence of a thick wall, gross papillary projections, solid areas, and solid echogenicity. A definitive histologic diagnosis was obtained in every case. Intraobserver and interobserver agreement was estimated by calculating the Cohen κ index. ***Results.*** Twenty-seven (33%) tumors were malignant, and 55 (67%) were benign. Intraobserver agreement for both examiners was good ($\kappa = 0.78$ and 0.72, respectively). Interobserver agreement was also good ($\kappa = 0.70$). ***Conclusions.*** Three-dimensional sonography is a reproducible technique for morphologic assessment of adnexal masses. ***Key words:*** ovarian tumor; reproducibility; sonography; 3-dimensional sonography.

Abbreviations
CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value; 3D, 3-dimensional; 2D, 2-dimensional

Two-dimensional (2D) B-mode morphologic sonography constitutes the basis for discriminating between benign and malignant adnexal masses. This diagnosis is based on "pattern recognition"[1] scoring systems.[2] A Doppler technique has also been proposed, although its usefulness remains controversial.[3]

Recently, 3-dimensional (3D) sonography has been introduced in clinical practice. This technique overcomes some limitations of conventional 2D sonography, allowing a more detailed assessment of morphologic features of the object studied, with no restriction to the number and orientation of the scanning plane.[4] Several studies have evaluated the role of 3D transvaginal sonography in assessing adnexal masses, reporting controversial results.[5–8]

The issue of reproducibility is essential when proposing a new technique to be introduced into clinical practice. The objective of this study was to evaluate the reproducibility of 3D sonographic morphologic assessment of adnexal masses.

Received March 27, 2007, from the Department of Obstetrics and Gynecology, Clínica Universitaria de Navarra, School of Medicine, University of Navarra, Pamplona, Spain. Revision requested April 11, 2007. Revised manuscript accepted for publication April 18, 2007.

Address correspondence to Juan Luis Alcázar, MD, Department of Obstetrics and Gynecology, Clínica Universitaria de Navarra, Avenida Pio XII 36, 31008 Pamplona, Spain.

E-mail: jlalcazar@unav.es

Materials and Methods

Between March 2004 and February 2005, 82 consecutive women with the diagnosis of an adnexal mass on conventional 2D sonography were asked to participate in this study. All patients gave oral informed consent. Institutional Review Board approval was also obtained.

The patients' mean age was 43.8 years, ranging from 15 to 82 years. Fifty-five (67%) women were premenopausal, and 27 (33%) were postmenopausal.

All patients were evaluated first by 2D sonography using a Voluson 730 system (GE Healthcare, Milwaukee, WI) with a 5- to 7-MHz endovaginal probe. Immediately after 2D sonography was performed, all patients were rescanned by 3D sonography. Transabdominal sonography (3.5–5 MHz) was also performed in large tumors. The 3D volume was activated to obtain a 3D box. Once a 3D volume box was obtained, it was stored (Sonoview; GE Kretztechnik, Zipf, Austria). The volume acquisition time lasted from 6 to 10 seconds depending on the size of the volume box. In some adnexal masses, more than 1 volume box of different areas of interest were obtained and analyzed. Stored 3D volumes were to be analyzed later.

All patients had a presumptive diagnosis based on 2D sonographic findings and underwent surgery, and a definitive histologic diagnosis was obtained in every case. Three-dimensional sonographic findings were not used for clinical decisions, and examinations were at no cost to the patients. All 2D and 3D sonographic examinations were performed by 1 author (J.L.A.).

Intraobserver and interobserver agreement was assessed by the κ index according to the method of Kundel and Polansky.[9] To assess intraobserver and interobserver agreement, 1 month after the last patient was recruited, 2 examiners (J.L.A., examiner A, with 6 years of experience in 3D sonography; and M.G.-M., examiner B, with 1 year of experience) reviewed all stored 3D volume data using surface-rendering and multiplanar modes to assess the morphologic features of the tumors. This was done in that way to avoid possible bias for examiner A, who had performed 2D sonography immediately before than 3D scanning. The 3D analyzing time lasted from 5 to 10 minutes depending on the complexity of the adnexal mass.

On the basis of International Ovarian Tumor Analysis Group criteria,[10] an adnexal mass was considered malignant when at least 1 of the following features was present: a thick wall, a thick septum, thick papillary projections, solid areas, and mostly solid echogenicity (Figures 1–3).

Then each examiner reviewed all 3D volumes again 1 week later to assess interobserver agreement.[11] After this second review, both examiners had to establish a presumptive diagnosis of malignancy or benignity for each adnexal mass according to the above-mentioned criteria. This second review was also used for assessing interobserver agreement. We chose the second review arbitrarily.

Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy for 2D and 3D sonography were calculated. Sensitivity and specificity were compared by the McNemar test. The SPSS version 13.0 statistical package (SPSS Inc, Chicago, IL) was used for analysis. $P < .05$ was considered statistically significant.

Results

Intraobserver agreement for both examiners was good for examiner A (κ = 0.78; 95% confidence interval [CI], 0.63–0.92) and examiner B (κ = 0.72; 95% CI, 0.56–0.88) (Table 1). Interobserver agreement was also good (κ = 0.72; 95% CI, 0.56–0.88) (Table 2).

After surgical removal, 27 (33%) tumors proved to be malignant (17 primary invasive cancers, 2 low–malignant potential tumors, and 8 metastatic tumors to the ovary), and 55 (67%) proved to be benign (15 endometriomas, 11 hemorrhagic cysts, 8 serous cysts, 6 simple cysts, 5 dermoid cysts, 4 mucinous cysts, 2 luteal cysts, 2 peritoneal cysts, 1 hydrosalpinx, and 1 cystadenofibroma).

Overall agreement for both examiners between diagnostic impressions and definitive histologic diagnoses was very good for examiner A (κ = 0.92; 95% CI, 0.82–1.00) and good for examiner B (κ = 0.70; 95% CI, 0.55–0.86).

Sensitivity, specificity, PPV, NPV, and accuracy for 2D and 3D sonography for each examiner are shown in Tables 3 and 4. No differences were found in terms of sensitivity and specificity between 2D and 3D sonography for both examiner A (McNemar test, $P > .99$) and examiner B (McNemar test, $P = .109$). When 3D sonography was used, the diagnostic performance of examiner A was better than examiner B, but the difference did not reach statistical significance (McNemar test, $P > .99$).

Discussion

Three-dimensional sonography is being increasingly used in gynecology. Several reports have shown that this technique may be very useful in the detection of congenital uterine abnormalities,[12] assessment of uterine cavity disease,[13] and ovarian and endometrial volume estimation.[14,15]

The differential diagnosis of adnexal masses still represents a challenge despite the tremendous efforts that have been made to improve the sonographically based diagnosis. This diagnosis is based on 2D sonography.[1] An examiner's subjective impression has been shown to be reproducible.[16]

Some researchers have evaluated the role of 3D sonography for discriminating between benign and malignant adnexal masses. Bonilla-Musoles and colleagues[5] and Kurjak and coworkers[6] found that 3D sonography was more sensitive than 2D sonography. Conversely, Hata and coworkers[7] found that 3D sonography was more specific than 2D sonography. In our experience, 3D sonography was not statistically better than 2D sonography.[8] The results of this new study confirms this finding.

However, none of the above-mentioned studies assessed the reproducibility of the technique. In this study, we have addressed that question. To the best of our knowledge, no previous study has assessed the reproducibility of 3D sonography for evaluating adnexal masses morphologically. Therefore, no comparison can be made with any other study.

According to our data, this technique is reproducible. Interobserver agreement was good ($\kappa = 0.72$), as was intraobserver agreement for both examiners. However, intraobserver agreement was better for the more experienced examiner ($\kappa = 0.78$ versus 0.72).

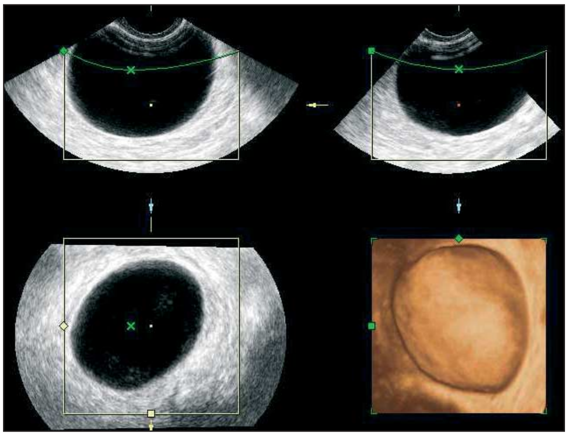

Figure 1. Three-dimensional multiplanar and surface-rendering displays of an ovarian cyst considered benign by both examiners. Histologic examination revealed a serous cystadenoma.

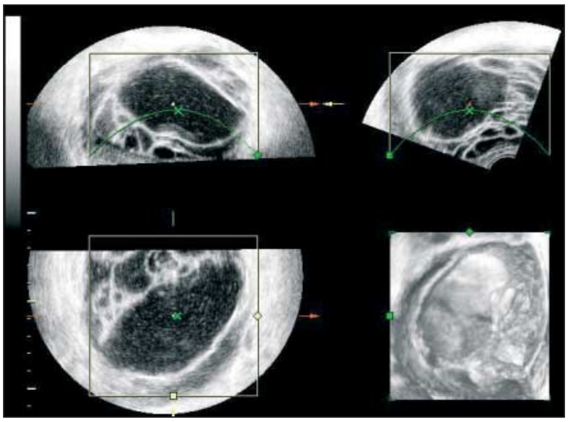

Figure 2. Three-dimensional multiplanar and surface-rendering displays of a multilocular cyst considered benign by examiner A and malignant by examiner B. Histologic examination revealed a mucinous cystadenoma.

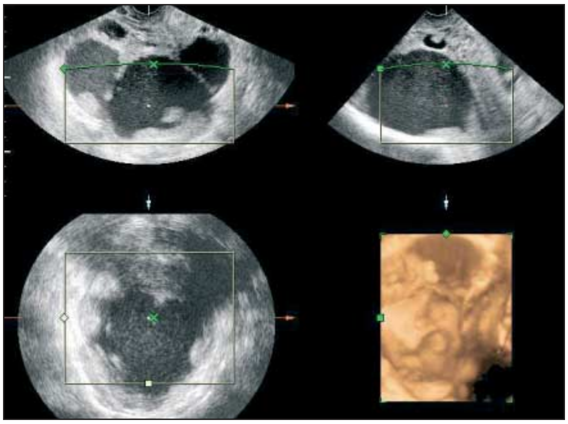

Figure 3. Three-dimensional multiplanar and surface-rendering displays of a multilocular cyst with solid components considered malignant by both examiners. Histologic examination revealed a primary invasive ovarian carcinoma.

These findings may have relevance when proposing this technique to be introduced in clinical practice. Furthermore, we have shown that concordance between a presumptive diagnosis and a final histologic diagnosis is high. Our findings may also be relevant because 3D sonographic evaluation has been advocated to be used offline and to facilitate telemedicine.[4]

Although our primary goal was not to assess the diagnostic performance of this technique for discriminating between benign and malignant adnexal masses, our results show that 3D sonography has good performance in terms of sensitivity and specificity. These figures were better for the more experienced examiner. However, the difference did not reach statistical significance. Perhaps a larger number of cases are necessary to address this question definitively.

In conclusion, according to our data, 3D sonography is a reproducible technique for morphologic assessment of adnexal masses.

Table 1. Intraobserver Agreement for Both Examiners

		Examiner A, Review 1		Examiner B, Review 1	
		Benign	Malignant	Benign	Malignant
Review 2	Benign	51	3	50	4
	Malignant	5	23	6	22

Table 2. Interobserver Agreement Between Both Examiners

		Examiner B	
		Benign	Malignant
Examiner A	Benign	51	5
	Malignant	5	21

Table 3. Diagnostic Performance of 2D and 3D Sonography for Examiner A

		Histologic Diagnosis	
		Benign	Malignant
Examiner A, 3D sonography*	Benign	54	2
	Malignant	1	25
Examiner A, 2D sonography†	Benign	52	3
	Malignant	3	24

*Sensitivity, 92.6%; specificity, 98.2%; PPV, 96.2%; NPV, 96.4%; accuracy, 96.3%.
†Sensitivity, 89%; specificity, 94.5%; PPV, 89%; NPV, 94.5%; accuracy, 92.7%.

Table 4. Diagnostic Performance of 2D and 3D Sonography for Examiner B

		Histologic Diagnosis	
		Benign	Malignant
Examiner B, 3D sonography*	Benign	50	6
	Malignant	5	21
Examiner B, 2D sonography†	Benign	48	2
	Malignant	7	25

*Sensitivity, 77.7%; specificity, 90.9%; PPV, 80.7%; NPV, 89.3%; accuracy, 86.6%.
†Sensitivity, 92.6%; specificity, 87.3%; PPV, 78.1%; NPV, 96%; accuracy, 89.3%.

References

1. Valentin L. Pattern recognition of pelvic masses by gray-scale ultrasound imaging: the contribution of Doppler ultrasound. Ultrasound Obstet Gynecol 1999; 14:338–347.
2. Alcázar JL, Mercé LT, Laparte C, Jurado M, López-García G. A new scoring system to differentiate benign from malignant adnexal masses. Am J Obstet Gynecol 2003; 188: 685–692.
3. Tekay A, Jouppila P. Controversies in assessment of ovarian tumors with transvaginal color Doppler ultrasound. Acta Obstet Gynecol Scand 1996; 75:316–329.
4. Bega G, Lev-Toaff AS, O'Kane P, Becker E Jr, Kurtz AB. Three-dimensional ultrasonography in gynecology: technical aspects and clinical applications. J Ultrasound Med 2003; 22:1249–1269.
5. Bonilla-Musoles F, Raga F, Osborne NG. Three-dimensional ultrasound evaluation of ovarian masses. Gynecol Oncol 1995; 59:129–135.
6. Kurjak A, Kupesic S, Sparac V, Bekavac I. Preoperative evaluation of pelvic tumors by Doppler and three-dimensional sonography. J Ultrasound Med 2001; 20:829–840.
7. Hata T, Yanagihara T, Hayashi K, et al. Three-dimensional ultrasonographic evaluation of ovarian tumours: a preliminary study. Hum Reprod 1999; 14:858–861.
8. Alcázar JL, Galán MJ, García-Manero M, Guerriero S. Three-dimensional ultrasound morphologic assessment in complex adnexal masses: a preliminary experience. J Ultrasound Med 2003; 22:249–254.
9. Kundel HL, Polansky M. Measurement of observer agreement. Radiology 2003; 228:303–308.
10. Timmerman D, Valentin L, Bourne TH, Collins WP, Verrelst H, Vergote I. International Ovarian Tumor Analysis (IOTA) Group. Terms, definitions and measurements to describe the sonographic features of adnexal tumors: a consensus

opinion from the International Ovarian Tumor Analysis (IOTA) Group. Ultrasound Obstet Gynecol 2000; 16:500–505.

11. Center for Biologics Evaluation and Research. CBER Guidances/Guidelines/Points to Consider. Rockville, MD: Center for Biologics Evaluation and Research, US Food and Drug Administration; 2004. Available at: http://www.fda.gov/cber/guidelines.htm. Accessed December 5, 2004.

12. Raga F, Bonilla-Musoles F, Blanes J, Osborne NG. Congenital Müllerian anomalies: diagnostic accuracy of three-dimensional ultrasound. Fertil Steril 1996; 65: 523–528.

13. Ayida G, Kennedy S, Barlow D, Chamberlain P. Contrast sonography for uterine cavity assessment: a comparison of conventional two-dimensional with three-dimensional transvaginal ultrasound—a pilot study. Fertil Steril 1996; 66:848–850.

14. Alcázar JL, Mercé LT, Manero MG, Bau S, López-Garcia G. Endometrial volume and vascularity measurements by transvaginal 3-dimensional ultrasonography and power Doppler angiography in stimulated and tumoral endometria: an interobserver reproducibility study. J Ultrasound Med 2005; 24:1091–1098.

15. Mercé LT, Bau S, Barco MJ, et al. Assessment of the ovarian volume, number and volume of follicles and ovarian vascularity by three-dimensional ultrasonography and power Doppler angiography on the HCG day to predict the outcome in IVF/ICSI cycles. Hum Reprod 2006; 21:1218–1226.

16. Timmerman D, Schwarzler P, Collins WP, et al. Subjective assessment of adnexal masses with the use of ultrasonography: an analysis of interobserver variability and experience. Ultrasound Obstet Gynecol 1999; 13:11–16.