

DOI: <http://dx.doi.org/10.12996/gmj.2026.4561>

## Comparison of Two and Three-Dimensional Ultrasound in the Diagnosis of Polycystic Ovary Morphology: A Cross Sectional Study in Unselected Women

Polikistik Over Morfolojisinin Tanısında İki ve Üç Boyutlu Ultrasonografinin Karşılaştırılması: Seçilmemiş Kadınlarda Kesitsel Bir Çalışma

① Zuhâl Yapıcı Coşkun<sup>1</sup>, ② Gürkan Bozdağ<sup>2</sup>, ③ Bülent Okan Yıldız<sup>3</sup>, ④ İbrahim Esinler<sup>4</sup>, ⑤ Hakan Yaralı<sup>4,5</sup>

<sup>1</sup>Department of Obstetrics and Gynecology, Ankara Bilkent City Hospital, Ankara, Türkiye

<sup>2</sup>Bahçeçi Health Group IVF Center, İstanbul, Türkiye

<sup>3</sup>Division of Endocrinology and Metabolism, Department of Internal Medicine, Hacettepe University Faculty of Medicine, Ankara, Türkiye

<sup>4</sup>Department of Obstetrics and Gynecology, Hacettepe University Faculty of Medicine, Ankara, Türkiye

<sup>5</sup>Anatolia IVF Center, Ankara, Türkiye

### ABSTRACT

**Objective:** To compare the visual appearance of the ovaries with two dimensional and three dimensional (3D) ultrasound (US) modalities in unselected women of reproductive age and to evaluate whether 3D changes the likelihood of polycystic ovary syndrome (PCOS) diagnosis.

**Methods:** This cross-sectional study was conducted in patients diagnosed within an unselected population of women (n = 115). The primary outcome measures were antral follicle count (AFC) and ovarian volume (OV), assessed using two different sonographic modalities. The agreement of these two sonographic modalities was tested using a Bland–Altman plot.

**Results:** Whereas the mean AFC with both sonographic modalities was  $10.3 \pm 4.8$  vs.  $11.2 \pm 5.5$  ( $p = 0.61$ ), the OV estimated were  $7.5 \pm 3.0$  vs.  $9.6 \pm 4.0$  ( $p < 0.001$ ) mL respectively. The mean bias, representing the upper and lower limits of agreement between the modalities for estimation of AFC, was  $-0.17$  ( $-8.45$  to  $8.11$ ). The respective figure for OV was  $-1.99$  mL ( $-7.72$  to  $3.75$ ). Therefore, 3D revealed OV values that were 22.5% larger and identified 2 additional cases of PCOS according to the Rotterdam criteria.

**Conclusion:** 3D US yields larger OV values and classifies slightly more women as having PCOS. Although concerns related to the diagnosis of PCOM mainly focus on AFC, those findings also underscore the reliability of OV as a diagnostic parameter for the syndrome. These

### Öz

**Amaç:** Bu çalışmanın amacı fertil çağda olan seçilmemiş kadın popülasyonunda iki boyutlu (2D) ve üç boyutlu (3D) ultrasonografi (US) ile yumurtalıkların görünümünü karşılaştırmak ve 3D'nin polikistik over sendromu (PKOS) tanı olasılığını değiştirip değiştirmediğini değerlendirmektir.

**Yöntemler:** Bu kesitsel çalışma seçilmemiş bir kadın popülasyonunda PKOS tanısı alan hastalar arasında (n = 115) yapılmıştır. Birincil sonuç ölçümleri 2 boyutlu ve 3 boyutlu US ile antral folikül sayısı (AFC) ve over hacmi (OV) idi. 2D ve 3D US'nin uyumu Bland-Altman grafiği ile test edilmiştir.

**Bulgular:** 2D ve 3D ile ortalama AFC  $10,3 \pm 4,8$  ve  $11,2 \pm 5,5$  ( $p = 0,61$ ) iken, 2D ve 3D ile tahmin edilen OV  $7,5 \pm 3,0$  ve  $9,6 \pm 4,0$  ( $p < 0,001$ ) mL idi. AFC tahmini için 2D ve 3D arasındaki ortalama sapma (uyumun üst ve alt sınırları)  $-0,17$  ( $-8,45$  ila  $8,11$ ) idi. OV için ilgili rakam  $-1,99$  ( $-7,72$  ila  $3,75$ ) mL idi. Dolayısıyla, 3D OV için %22,5 daha büyük değerler ortaya koymuş ve Rotterdam kriterlerine göre ek olarak 2 PKOS vakası daha tespit etmiştir.

**Sonuç:** 3D US daha büyük OV değerleri vermekte ve hafifçe daha fazla sayıda kadına PKOS tanısı konmasına yardımcı olmaktadır. PKOM tanısı ile ilgili endişeler esas olarak antral folikül sayımına odaklansa da, bu bulgular sendromun tanısında OV'nin güvenilirliğini de vurgulamaktadır. Bu bulgular ovaryan ölçümlerin değerlendirilmesi ve tanısal kriterlerin

**Cite this article as:** Yapıcı Coşkun Z, Bozdağ G, Yıldız BO, Esinler İ, Yaralı H. Comparison of two and three-dimensional ultrasound in the diagnosis of polycystic ovary morphology: A cross sectional study in unselected women. Gazi Med J. 2026;37(2):177-182

**Address for Correspondence/Yazışma Adresi:** Zuhâl Yapıcı Coşkun, Department of Obstetrics and Gynecology, Ankara Bilkent City Hospital, Ankara, Türkiye

**E-mail / E-posta:** [zuhalyapici@yahoo.com](mailto:zuhalyapici@yahoo.com)

**ORCID ID:** [orcid.org/0000-0002-9338-4669](https://orcid.org/0000-0002-9338-4669)

**Received/Geliş Tarihi:** 02.10.2025

**Accepted/Kabul Tarihi:** 17.02.2026

**Publication Date/Yayınlanma Tarihi:** 31.03.2026



©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Gazi University Faculty of Medicine. Licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

©Telif Hakkı 2026 Yazar(lar). Gazi Üniversitesi Tıp Fakültesi adına Galenos Yayınevi tarafından yayımlanmaktadır. Creative Commons Atf-GayriTicari-Türetilemez 4.0 (CC BY-NC-ND) Uluslararası Lisansı ile lisanslanmaktadır.

**ABSTRACT**

findings suggest that the imaging modality should be considered when interpreting ovarian measurements and applying diagnostic criteria.

**Keywords:** Ultrasonography, polycystic ovary syndrome, antral follicle count, three dimensional ultrasonography

**ÖZ**

uygulanmasında görüntüleme tekniği ve modalitesinin göz önünde bulundurulmasını önermektedir.

**Anahtar Sözcükler:** Ultrason, polikistik over sendromu, antral folikül sayımı, üç boyutlu ultrasonografi

**INTRODUCTION**

The role of assessment of ovarian morphology for the diagnostic value in polycystic ovary syndrome (PCOS) has been a debate (1). Following the consensus meeting held by the National Institutes of Health (NIH), it has been stated that clinical and/or biochemical hyperandrogenism accompanied by chronic anovulation are required for the diagnosis (2). PCOS Consensus Workshop Group sponsored by European Society of Human Reproduction and Embryology /American Society for Reproductive Medicine (ESHRE/ASRM) suggested that polycystic ovary morphology (PCOM) should be one of the three criteria for the confirmation of PCOS diagnosis (3). Alternatively, Androgen Excess-PCOS Society (AE-PCOS Society) recommended that PCOM would not have diagnostic value for PCOS without any finding of hyperandrogenism accompanying it (4).

As of today, not only the importance but also the definition of PCOM remains controversial (1). Initially, it was reported that antral follicle count (AFC) of more than or equal to 12 or ovarian volume (OV) of  $\geq 10 \text{ cm}^3$  is essential for PCOM in Rotterdam (3). However, a revision of the morphological criteria for PCOM is required after this initial definition. The first reason is that PCOM is observed among otherwise healthy women whose AFCs overlap with those of women with PCOS. Therefore, it might be sensible to implement a higher cut-off, primarily for AFC, which would classify fewer otherwise healthy women as PCOM. On the other hand, the metabolic/ovarian states of women with PCOM-only are unclear, and it may not be straightforward to determine their health risk, if any, when they are classified as "normal". Secondly, due to technological improvement of ultrasound (US) evaluation and its resolution, we might be assigning more ovaries as "polycystic" than early ages. As noted previously (5), follicle enumeration with advanced US and grid-based methods necessitates a higher threshold across the ovary. And lastly, data about the validity of three dimensional (3D) features of US while counting antral follicles and OV in the diagnosis of PCOM is not clear. Although 3D US is increasingly used in gynecology, its concordance with two dimensional (2D) US for PCOM warrants investigation and clarification.

Our study aims to compare the visual properties of the ovaries using 2D and 3D US in unselected women of reproductive age, and to evaluate whether 3D changes the likelihood of a PCOS diagnosis based on either AFC or OV.

**MATERIALS AND METHODS****Ethical Approval**

This study was approved by the Institutional Ethics Review Board of Hacettepe University Faculty of Medicine on 19 November 2009 (approval number: 2545, date: 19.11.2009) and conducted in accordance with the principles of the Declaration of Helsinki.

Written informed consent was obtained from all participants and/or their legal guardians.

**Participants**

This study was conducted on an unselected population of women, evaluating them with both 2D and 3D US at the Institute of Mineral Research and Exploration (6). Unselected women describes volunteers from the general women population who responded to an open invitation. No inclusion criteria were based on clinical symptoms or suspicion of PCOS. All participants provided informed consent before participating in the study.

The female subjects were aged 18-45 year. Menopausal status, a history of hysterectomy or bilateral oophorectomy, pregnancy, and use of oral contraceptive pills for any reason were exclusion criteria (n = 43). A total of 115 women undergoing both 3D and 2D US are enrolled in the current study.

**Study Protocol**

For initial evaluation, a standardized medical form was used to obtain interview-based information on women's age, obstetric history, medical conditions, medications, menstrual regularity, gynecological and family history. Menstrual cycles with an interval  $\geq 35$  or  $\leq 23$  days were supposed to define ovulatory dysfunction.

The amount and distribution of terminal hair on the designated body areas were evaluated using a modified Ferriman Gallwey (mFG) scoring system, including 9 areas of the body, as described previously (7,8), to describe hirsutism. In case of mFG score being  $\geq 6$  participant was accepted to have hirsutism regardless of alopecia or acne existence. Biochemical hyperandrogenism was evaluated by analysing blood samples an overnight fast on the second to fifth day of menstruation between 8:00 and 10:30 AM. For participants using oral contraceptives, samples were collected during the interval. To define biochemical hyperandrogenism (hyperandrogenemia) in this study, otherwise healthy, non-hirsute women without PCOM and with regular menstrual cycles were considered the reference group. This group corresponds to 216 of the 392 women. Biochemical hyperandrogenemia is defined as increased levels of at least one of the androgens over 95<sup>th</sup> percentile such as total testosterone (tT), androstenedione, (A4), dehydroepiandrosterone sulfate and/or free androgen index (FAI).

If any of the androgen levels exceeds 95<sup>th</sup> percentile of healthy, non-hirsute women having regular menstrual cycles with no PCOM features (n = 216), then the patient assigned as having hyperandrogenemia.

**Hormonal and Biochemical Analyses**

After blood sampling, the specimen was transferred to a central laboratory by 11:00 AM. Following 20 minutes of centrifugation, specimen has been stored in polypropylene tubes at -70 °C until

final analysis. The hormonal analyses included TSH, prolactin, 17-OH progesterone, sex hormone binding globulin (SHBG), and previously defined androgens. The FAI was calculated using tT and SHBG levels as follows:  $(FAI = tT \times 100 / SHBG)$ .

### Ultrasonography

All examinations with 2 and 3D US were performed on second to seventh day of participant's menstrual bleeding with Voluson e (GE Healthcare, İstanbul, Türkiye) were performed by a single physician. Based on marital status and patient preference, either an abdominal (2–7 MHz) or a transvaginal (5–9 MHz) probe was used.

PCOM was defined as an AFC  $\geq 12$  follicles measuring 2–9 mm and/or an OV  $\geq 10 \text{ cm}^3$  in either ovary. If one ovary could not be assessed, PCOM classification was based on the measurable ovary. When a persistent cyst or dominant follicle prevented accurate volume measurement on one ovary, AFC from both ovaries and OV from the contralateral ovary were used to determine PCOM status.

AFC and OV were determined using both US modalities. Under 2D US, antral follicles that have a diameter between 2 and 9 mm were counted in the transverse section at each site. OV was calculated from three diameters –anterior-posterior (a), maximum longitudinal (b), and transverse (c) –using the formula  $(a \times b \times c \times 0.5)$ . Under 3D US, the antral follicles were counted at the same time. The OV was processed using the Virtual Organ Computer-aided Analysis imaging program, employing Plane A and 60-degree rotational steps.

### Definiton of PCOS

Depending on NIH criteria, PCOS had been defined as biochemical and/or clinical hyperandrogenism accompanied by ovulatory dysfunction as recommended (2). Rotterdam criteria revised the definition of PCOS by the presence of at least two of the findings/symptoms: 1) clinical and/or biochemical hyperandrogenism, 2) ovulatory dysfunction and, 3) PCOM (9). Regarding the AE-PCOS Society criteria, PCOS was diagnosed as biochemical and/or clinical hyperandrogenism associated with ovulatory dysfunction or PCOM (4,10). For the definitions of PCOS, only 2D evaluations were considered.

### Statistical Analysis

The independent numerical parameters were analyzed with paired-samples t-tests. Reliability and consistency between parameters were assessed using Spearman correlation analysis. The term "mean bias" is defined as the average difference between the true OV and the estimated volume obtained using either 2D or 3D US. Bland-Altman plot of differences was drawn as defined previously (11). Parameters were noted as mean  $\pm$  SD, unless stated otherwise. The SPSS 13.0 package (SPSS Inc., Chicago, IL) was used for statistical analysis. The figures of agreement were generated with GraphPad Prism 6.0 (trial version).

## RESULTS

In this study, we compared 2D and 3D US measurements of AFC and OV in 115 women to determine whether the imaging method affects the classification of PCOM and PCOS. As depicted in the Figure 1,

when 392 women were referred (80.2% of whole population), the prevalence of PCOS due to subsets of NIH (6.1%), Rotterdam ESHRE/ASRM (19.9%) and AE-PCOS Society (15.3%) criteria were calculated (6). According to 2D US, the rate of PCOM in women was 36.5% (143/392). Among 143 women with PCOM, 59 had bilateral and 84 had unilateral PCOM. Notably, PCOM was diagnosed in 95 (66.4%) based on AFC, 6 (4.2%) based on OV and 42 (29.4%) women based on the presence of both.

Of the 392 participants, we evaluated 115 with both 2D and 3D sonographic modalities in at least one ovary (Figure 1). Whereas the mean AFC with these modalities were  $10.3 \pm 4.8$  vs.  $11.2 \pm 5.5$  ( $p = 0.61$ ), the OV estimated were  $7.5 \pm 3.0$  vs.  $9.6 \pm 4.0$  ( $p < 0.001$ ) mL respectively.

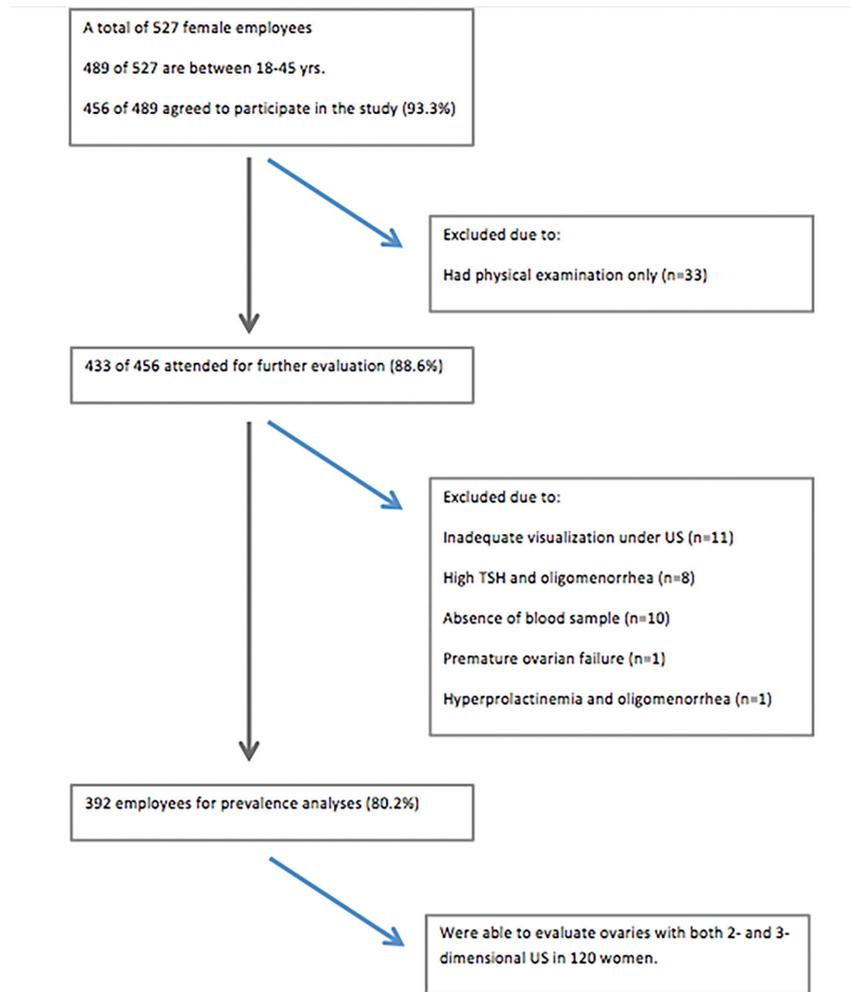
Sixty-six ovaries were classified as PCOM when US was performed using 2D imaging (Table 1). Sixty-five (93.4%) of them were confirmed as PCOM under 3D US. However, of the 49 ovaries who were not noted to be PCOM on 2D, 34 were diagnosed as PCOM with 3D US. That revision revealed 2 more cases of PCOS according to the Rotterdam criteria, when OV and hyperandrogenism were considered. However, this numerical increase is small, has not been statistically tested, and is unlikely to represent a statistically significant difference. The findings suggest that 3D US may shift the classification in borderline cases, but may not substantially alter PCOS prevalence in a population of this size.

The coefficients between these sonographic modalities, analysed for AFC and OV calculations, were 0.783 and 0.626, respectively. When the limit of agreement was tested, the mean bias between 2D and 3D for the estimation of AFC was  $-0.17$  ( $-8.45$  to  $8.11$ ). The respective figure for OV was  $-1.99$  ( $-7.72$  to  $3.75$ ) mL. Therefore, 3D yielded values that were 22.5% larger than the OV estimated by 2D (Figures 2 and 3).

## DISCUSSION

Our main finding is that 3D US yields consistently larger ovarian-volume measurements than 2D US, resulting in more women being classified with PCOM and, in some cases, PCOS. For the AFC, both the agreement and the consistency of 2D and 3D US were superior to those of OV. However, despite good agreement in OV between the two methods, 3D estimated OV was 22.5% larger, which might be clinically important in the diagnosis of PCOM. That finding once again supports the type of statistical analysis used in comparing modalities. This highlights that OV may be less stable and more sensitive to technical differences than AFC; therefore, diagnostic thresholds may need to be adjusted for different US technologies. The major strength of this study is that it reports US findings using sonographic modalities of different dimensionalities in a subset of a large unselected population.

In the available literature, the agreement among sonographic modalities of different dimensionalities regarding AFC and OV is inconclusive. According to our data, in spite of good agreement between 2D and 3D US regarding AFC and OV across the whole population, 3D US detected a slightly higher number of cases among patients with PCOS according to the Rotterdam criteria. The technical aspects of transabdominal US with poor spatial resolution and low-frequency probes limit clarity compared to transvaginal probe that provides 3D evaluation more clearly (12).



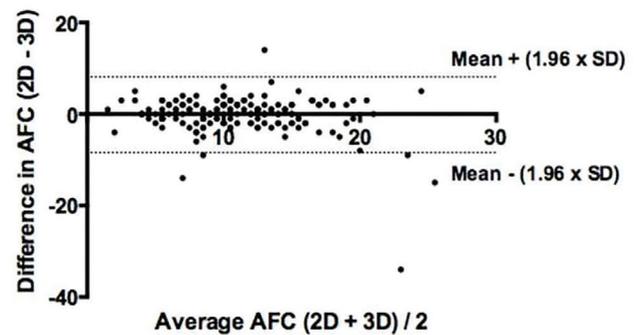
**Figure 1.** Study population.

TSH: Thyroid-stimulating hormone, US: Ultrasound.

**Table 1.** The relation of 2D and 3D US regarding polycystic ovarian morphology.

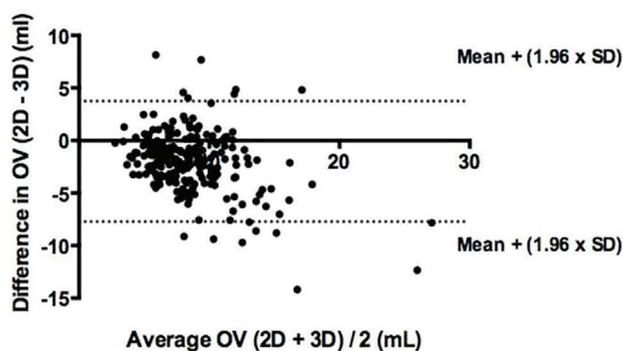
	2D, PCOM (+)	2D, PCOM (-)
3D, PCOM (+)	65	34
3D, PCOM (-)	5	16

PCOM: Polycystic ovarian morphology, 2D: Two dimensional, 3D: Three dimensional, US: Ultrasound.



**Figure 2.** Bland-Altman plot of 2D and 3D US for AFC.

AFC: Antral follicle count, 2D: Two dimensional, 3D: Three dimensional, US: Ultrasound, SD: Standard deviation.



**Figure 3.** Bland-Altman plot of 2D and 3D US for OV.

SD: Standard deviation, 2D: Two dimensional, 3D: Three dimensional, US: Ultrasound, OV: Ovarian volume, SD: Standard deviation.

Although a significant correlation is observed between 2D and 3D US in our study similar to previous literature (13), the level of agreement might present some diversions. According to Scheffer et al (14), the coverage intervals for the difference between two methods were -5.3 to 8.3 follicles which indicate a moderate agreement. Of note, the agreement gets worse as the number of antral follicles increase (14). In another study that assesses the reliability of measurements made by manual 3D or with sono-automatic volume count software, the respective mean  $\pm$  standard deviation of AFCs were  $6.5 \pm 4.8$  vs.  $19.4 \pm 10.9$  (15). We observed better agreement between the two methods for AFC, but agreement worsened for OV estimation. Nevertheless, previously we had reported that whereas 2D US brought 18% larger, 3D US revealed 11% smaller values compared to the absolute OV that was calculated according to Archimedes' principles following oophorectomy in patients undergoing surgery (16). Therefore, one may hypothesize that different thresholds might be required depending on US specifications and the availability of software to define PCOM, which would not be reasonable to apply in clinical practice.

Since defining universally accepted morphological criteria for the appearance of polycystic ovaries is unlikely, two main strategies might be pursued. Those are revising the threshold levels with parallel to the improvement in imaging or secondly, replacing it with a biochemical marker that is less physician and cycle dependent. However, the former strategy requires revisiting the thresholds regularly as technology improves (1). Nevertheless, a recent task force by AE-PCOS Society (17) recommended using higher threshold levels of AFC for the definition of PCOM, particularly when using newer technology (i.e. transducer frequency  $\geq 8$  MHz) which provides better resolution for ovarian follicles. Therefore, the technological evolution of US does not allow the use of the same threshold for AFC or OV over time. In this context the diagnostic thresholds for US evaluation of follicle number per ovary has been revised and increased in the International Evidence-based Guideline for the assessment and management of PCOS 2023 (18).

A recent study has shown that artificial intelligence, using the backpropagation algorithm, accurately identifies the three-dimensional ovarian structure and measures both OV and the AFC.

US parameters, in addition to endocrine and metabolic parameters, represent objective diagnostic tools and have clinical importance. Comprehensive studies on this issue with larger participant population needs to be performed to better diagnose and evaluate PCOS and direct clinical management (19).

### Study Limitations

As highlighted earlier, the sampling methodology represents one of the principal limitations of this study. Although women working at the institution participated in the study with a high response rate, potential selection bias due to undetermined differences between the study sample and the background community cannot be excluded. The second limitation is the low proportion of women who had at least one ovary eligible for evaluation by both modalities. The unusually high rate of missing data may be attributable to certain issues. Initially, all patients declined the endocavitary probe because of virginal status and discomfort with the vaginal probe ( $n = 174$ ). Nevertheless, the technical aspects of transabdominal US with poor spatial resolution and low-frequency probes limit clarity (12). Potentially, image quality may be reduced secondary to central obesity, which is commonly seen in women with PCOS (12). Among the whole group, the percentages of overweight women [body mass index (BMI): 25.0–29.9 kg/m<sup>2</sup>] and obese women (BMI  $\geq 30$  kg/m<sup>2</sup>) were 24.0 and 10.2%, respectively. We did not include subjects in the final analysis unless one ovary was properly visualized. The physician's meticulous policy might also have caused the lower number of cases included. The effect of these missing data on the outcome and conclusions of the study is unknown, but the interpretation of the results using the Bland-Altman plot might be more valuable than comparing the mean values. These factors should be considered when interpreting the results.

### CONCLUSION

We suggest that 3D US estimates larger OV when compared with 2D, resulting in a slightly higher number of women being diagnosed with PCOS. A difference of 22.5% might be clinically important for women whose values are close to current cut-offs for the definition of PCOM and, in turn, for assigning them to the syndrome. Although previous debates have focused on detecting antral follicles, those findings also underscore the reliability of OV as a diagnostic parameter for PCOS. These differences suggest that diagnostic thresholds for PCOM may need to be adjusted for newer imaging technologies. In the following studies, the concordance and consistency of 2D and 3D sonographic modalities for sonographic parameters such as OV and AFC needs to be investigated with larger study population.

### Ethics

**Ethics Committee Approval:** The study was approved by the Hacettepe University Faculty of Medicine Institutional Review Board on 19 November 2009 (approval number: 2545, date: 19.11.2009).

**Informed Consent:** Prior to inclusion in the study, informed consent was obtained from the participants.

**Footnotes****Authorship Contributions**

Surgical and Medical Practices: Z.Y.C., G.B., B.O.Y., İ.E., H.Y., Concept: Z.Y.C., G.B., B.O.Y., İ.E., H.Y., Design: Z.Y.C., G.B., B.O.Y., İ.E., H.Y., Data Collection or Processing: Z.Y.C., G.B., B.O.Y., İ.E., H.Y., Analysis or Interpretation: Z.Y.C., G.B., B.O.Y., İ.E., H.Y., Literature Search: Z.Y.C., G.B., B.O.Y., İ.E., H.Y., Writing: Z.Y.C., G.B., B.O.Y., İ.E., H.Y.

**Conflict of Interest:** No conflict of interest was declared by the authors.

**Financial Disclosure:** This study was partially sponsored by Merck Serono.

**REFERENCES**

- Dewailly D, Gronier H, Poncelet E, Robin G, Leroy M, Pigny P, et al. Diagnosis of polycystic ovary syndrome (PCOS): revisiting the threshold values of follicle count on ultrasound and of the serum AMH level for the definition of polycystic ovaries. *Hum Reprod.* 2011; 26: 3123-6.
- Zawadzki JK, Dunaif A. Diagnostic criteria for polycystic ovary syndrome. In: Dunaif A, Givens JR, Haseltine F, Merriam GR, editors. *Polycystic ovary syndrome.* Boston: Blackwell Scientific Publications; 1992. p. 377-84.
- Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril.* 2004; 81: 19-6.
- Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, et al. Positions statement: criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: an Androgen Excess Society guideline. *J Clin Endocrinol Metab.* 2006; 91: 4237-28.
- Lujan ME, Jarrett BY, Brooks ED, Reines JK, Peppin AK, Muhn N, et al. Updated ultrasound criteria for polycystic ovary syndrome: reliable thresholds for elevated follicle population and ovarian volume. *Hum Reprod.* 2013; 28: 1361-7.
- Yildiz BO, Bozdag G, Yapici Z, Esinler I, Yarali H. Prevalence, phenotype and cardiometabolic risk of polycystic ovary syndrome under different diagnostic criteria. *Hum Reprod.* 2012; 27: 3067- 6.
- Ferriman D, Gallwey JD. Clinical assessment of body hair growth in women. *J Clin Endocrinol Metab.* 1961; 21: 1440-7.
- Yildiz BO, Bolour S, Woods K, Moore A, Azziz R. Visually scoring hirsutism. *Hum Reprod Update.* 2010; 16: 51-64.
- Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum Reprod.* 2004; 19: 41-7.
- Azziz R, Woods KS, Reyna R, Key TJ, Knochenhauer ES, Yildiz BO. The prevalence and features of the polycystic ovary syndrome in an unselected population. *J Clin Endocrinol Metab.* 2004; 89: 2745-9.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986; 1: 307-10.
- Yoo RY, Sirlin CB, Gottschalk M, Chang RJ. Ovarian imaging by magnetic resonance in obese adolescent girls with polycystic ovary syndrome: a pilot study. *Fertil Steril.* 2005; 84: 985-10.
- Nardo LG, Buckett WM, Khullar V. Determination of the best-fitting ultrasound formulaic method for ovarian volume measurement in women with polycystic ovary syndrome. *Fertil Steril.* 2003; 79: 632-1.
- Scheffer GJ, Broekmans FJ, Bancsi LF, Habbema JD, Looman CW, Te Velde ER. Quantitative transvaginal two- and three- dimensional sonography of the ovaries: reproducibility of antral follicle counts. *Ultrasound Obstet Gynecol.* 2002; 20: 270-5.
- Deb S, Jayaprakasan K, Campbell BK, Clewes JS, Johnson IR, Raine-Fenning NJ. Intraobserver and interobserver reliability of automated antral follicle counts made using three-dimensional ultrasound and SonoAVC. *Ultrasound Obstet Gynecol.* 2009; 33: 477-83.
- Bozdag G, Salman MC, Mumusoglu S, Yapici Z, Gunalp S. Is ovarian volume estimation reliable when compared with true volume? *Am J Obstet Gynecol.* 2012; 206: 44.e1-4.
- Dewailly D, Lujan ME, Carmina E, Cedars MI, Laven J, Norman RJ, et al. Definition and significance of polycystic ovarian morphology: a task force report from the Androgen Excess and Polycystic Ovary Syndrome Society. *Hum Reprod Update.* 2014; 20: 334-52.
- Teede HJ, Tay CT, Laven J, Dokras A, Moran LJ, Piltonen TT, et al. Recommendations from the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome†. *Hum Reprod.* 2023; 38: 1655-79.
- Wei L, Wu F, Zhang J, Li J, Yang D, Wen G. Evaluation of endocrine and metabolic changes in polycystic ovary syndrome by ultrasonic imaging features under an intelligent algorithm. *Comput Math Methods Med.* 2022; 2022: 1411943.