human reproduction update

Definition and significance of polycystic ovarian morphology: a task force report from the Androgen Excess and Polycystic Ovary Syndrome Society

Didier Dewailly[1,*], Marla E. Lujan[2], Enrico Carmina[3], Marcelle I. Cedars[4], Joop Laven[5], Robert J. Norman[6], and Héctor F. Escobar-Morreale[7]

[1]Department of Endocrine Gynaecology and Reproductive Medicine, Hôpital Jeanne de Flandre, Centre Hospitalier de Lille, University of Lille, Lille, France [2]Division of Nutritional Sciences, Human Metabolic Research Unit, Cornell University, Ithaca, NY, USA [3]Endocrinology Unit, DISMOT Department, University of Palermo, Palermo, Italy [4]Division of Reproductive Endocrinology and Infertility, Department of Obstetrics, Gynecology and Reproductive Sciences, University of California San Francisco, San Francisco, CA, USA [5]Division of Reproductive Medicine, Department of Obstetrics and Gynecology, Erasmus MC University Medical Center, Rotterdam, Netherlands [6]Robinson Institute, School of Paediatrics and Reproductive Health, University of Adelaide, Adelaide, SA, Australia [7]Diabetes, Obesity and Human Reproduction Research Group, University of Alcalá & Hospital Ramón y Cajal & Centro de Investigación Biomédica en Red Diabetes y Enfermedades Metabólicas Asociadas CIBERDEM & Instituto Ramón y Cajal de Investigación Sanitaria IRYCIS, Madrid, Spain

*Correspondence address. Department of Endocrine Gynaecology and Reproductive Medicine, Hôpital Jeanne de Flandre, C.H.R.U., 59037 Lille, France; E-mail: didier.dewailly@chru-lille.fr

TABLE OF CONTENTS

BACKGROUND: The diagnosis of polycystic ovary syndrome (PCOS) relies on clinical, biological and morphological criteria. With the advent of ultrasonography, follicle excess has become the main aspect of polycystic ovarian morphology (PCOM). Since 2003, most investigators have used a threshold of 12 follicles (measuring 2–9 mm in diameter) per whole ovary, but that now seems obsolete. An increase in ovarian volume (OV) and/or area may also be considered accurate markers of PCOM, yet their utility compared with follicle excess remains unclear.

METHODS: Published peer-reviewed medical literature about PCOM was searched using PubMed.gov online facilities and was submitted to critical assessment by a panel of experts. Studies reporting antral follicle counts (AFC) or follicle number per ovary (FNPO) using transvaginal ultrasonography in healthy women of reproductive age were also included. Only studies that reported the mean or median AFC or FNPO of follicles measuring 2–9 mm, 2–10 mm or <10 mm in diameter, or visualized all follicles, were included.

RESULTS: Studies addressing women recruited from the general population and studies comparing control and PCOS populations with appropriate statistics were convergent towards setting the threshold for increased FNPO at ≥25 follicles, in women aged 18–35 years. These studies suggested maintaining the threshold for increased OV at ≥10 ml. Critical analysis of the literature showed that OV had less diagnostic potential for PCOM compared with FNPO. The review did not identify any additional diagnostic advantage for other ultrasound metrics such as specific measurements of ovarian stroma or blood flow. Even though serum concentrations of anti-Müllerian hormone (AMH) showed a diagnostic performance for PCOM that was equal to or better than that of FNPO in some series, the accuracy and reproducibility issues of currently available AMH assays preclude the establishment of a threshold value for its use as a surrogate marker of PCOM. PCOM does not associate with significant consequences for health in the absence of other symptoms of PCOS but, because of the use of inconsistent definitions of PCOM among studies, this question cannot be answered with absolute certainty.

CONCLUSIONS: The Task Force recommends using FNPO for the definition of PCOM setting the threshold at ≥25, but only when using newer technology that affords maximal resolution of ovarian follicles (i.e. transducer frequency ≥8 MHz). If such technology is not available, we recommend using OV rather than FNPO for the diagnosis of PCOM for routine daily practice but not for research studies that require the precise full characterization of patients. The Task Force recognizes the still unmet need for standardization of the follicle counting technique and the need for regularly updating the thresholds used to define follicle excess, particularly in diverse populations. Serum AMH concentration generated great expectations as a surrogate marker for the follicle excess of PCOM, but full standardization of AMH assays is needed before they can be routinely used for clinical practice and research. Finally, the finding of PCOM in ovulatory women not showing clinical or biochemical androgen excess may be inconsequential, even though some studies suggest that isolated PCOM may represent the milder end of the PCOS spectrum.

Key words: Anti-Müllerian hormone / antral follicle count / follicle number per ovary / ovarian volume / polycystic ovaries

Introduction

The combination of oligomenorrhoea, infertility, hirsutism and bilateral enlarged polycystic ovaries was identified as an entity by Stein and Leventhal (1935), who for some time gave their name to the syndrome. The term 'polycystic ovary syndrome' and its acronym PCOS appeared in the 1960s and gradually replaced the Stein–Leventhal syndrome designation.

Meanwhile, the progressive decline in the practice of wedge resection of the ovaries deprived researchers of a source of valuable study material. The last major histological study dates back to 1982 (Hughesdon, 1982) and provides a detailed description of the 'polycystic' appearance of the ovaries as being simply an increase in the number of growing follicles measuring <10 mm in diameter.

Because the ovarian follicle may visually resemble a 'cyst', the presence of 'cystic' images in the ovary must be considered a normal event. It is only their excessive number that must be regarded as pathologic with PCOS being the major if not the exclusive cause. Therefore, it would have been more accurate to speak of 'multifollicular ovaries' but the name 'polycystic' was time-honoured and has lasted until now. This is unfortunate because this term wrongly worries patients and is sometimes misinterpreted by physicians who are not familiar with ovarian pathology.

With the advent of ultrasonography, follicle excess has become the main aspect of polycystic ovarian morphology (PCOM). An increase in ovarian volume (OV) and an increased ovarian area (OA) are also considered accurate markers of PCOM, provided the measurements are carried out on a median section of the ovaries. Histopathologic studies also confirm that both OV and OA are indeed a good reflection of

stromal hypertrophy and follicle excess, which are the anatomical hallmarks of PCOM (Hughesdon, 1982).

Nowadays there is an almost universal consensus on the choice of follicular excess and ovarian enlargement as criteria to define PCOM by ultrasound. However, establishing the normal values for follicle number per ovary (FNPO) and OV, and especially the setting of accurate thresholds for distinguishing normal ovaries from PCOM, is still the subject of great controversy. This poses a real problem because the item PCOM is included in the most commonly used classifications for the diagnosis of PCOS (The Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2004b; Azziz *et al.*, 2009) and also included in the latest recommendations from the National Institutes of Health (Johnson *et al.*, 2012). But it is clear that the threshold of FNPO currently proposed for the diagnosis of PCOM needs to be revisited, as this threshold is currently met by >50% of normal young ovulatory women in some series (Johnstone *et al.*, 2010). The aim of this systematic review was therefore to analyse the available literature and determine whether we can now achieve a new consensus on the definition of PCOM.

Methods

Panel

The Androgen Excess and PCOS (AE-PCOS) Society Board appointed a panel of experts on PCOM, selected from those researchers who had authored many original articles in the field. Panel members and the Board Director constituted the Writing Committee.

Data

Published peer-reviewed medical literature about PCOM was searched using PubMed.gov online facilities (http://www.ncbi.nlm.nih.gov/pubmed) introducing the following search terms: 'polycystic appearing ovary', 'polycystic ovarian morphology', 'polycystic ovaries', 'polycystic ovary', 'polyfollicular ovary', 'multifollicular ovary' 'multicystic ovaries', 'ovarian area', 'ovarian volume', 'subcortical stroma', 'ovarian stroma', 'ovarian stromal hypertrophy', 'AMH', 'anti-Müllerian factor', 'anti-Müllerian hormone', 'Müllerian inhibiting factor', 'Müllerian-inhibiting hormone', 'Müllerian-inhibiting substance', 'receiver operating characteristic curve', 'sensitivity and specificity', 'inter-observer agreement', 'intra-observer agreement', 'reliability', 'reproducibility', 'prevalence' and 'Rotterdam'.

In parallel, a search of studies reporting antral follicle count (AFC) or FNPO using transvaginal ultrasonography in healthy women of reproductive age was performed using as search terms: 'antral follicle count', 'transvaginal ultrasonography', 'regular menstrual cycles' and 'follicle number'. Results were limited to studies published after the year 2000 because the aim of this search was to identify the upper limits of normal of these variables in ovulatory women when using modern ultrasound equipment. Studies in which the time-span of data collection or transducer frequency could not be confirmed were excluded. Control populations were defined as healthy women with regular menstrual cycles recruited from the general population, women with regular menstrual cycles with confirmed male factor or tubal infertility, and/or those with regular menstrual cycles in which hyperandrogenism, PCOS and/or other endocrine disorders were excluded. Only studies that reported means or medians of AFC and/or FNPO of follicles measuring 2–9, 2–10, <10 mm in diameter, or reported all follicles being visualized, were included. In instances where AFC was reported, values were divided by two to generate an FNPO.

More than 300 articles were initially available for review. Some studies were eliminated because data were not related to the focus of the guidelines, were insufficient for analysis or were duplicated in several publications. All data sources were analysed recognizing positive publication bias.

Process

The review process included individual studies, systematic reviews, hand searches, abstracts, and individual databases and expert data. Each review was conducted by at least two investigators, and the criteria for inclusion/exclusion were agreed upon by at least two reviewers in each area and arbitrated by a third when necessary. The position statement applied part of the Grading of Recommendations, Assessment, Development and Evaluation group criteria (Atkins *et al.*, 2004; Swiglo *et al.*, 2008) in which the strength of a recommendation was indicated by 'recommend' or, if a weaker recommendation was indicated, by 'suggest'. Recommendations were made based on evidence that was considered appropriate in making the recommendation. The writing committee critically reviewed the manuscript for intellectual content before submitting the manuscript to the AE-PCOS Society Board for endorsement. Institutional Review Board approval was not obtained because the study reviewed publicly available medical literature.

Follicle Excess to Define PCOM

Does ultrasound afford a reliable estimate of the follicle excess of polycystic ovaries?

Validation of follicle counts

While sonohistopathological assessments confirmed a high accuracy in detecting PCOM by transabdominal ultrasonography, these conclusions were based on qualitative assessments of ovarian morphology and not on quantitative measurements such as follicle counts (Saxton *et al.*, 1990). Likewise, using ovarian laparoscopy as the gold-standard, ultrasonography was calculated to have a sensitivity of 91% and a specificity of 100% in detecting PCOM, but this did not apply specifically to FNPO (Fox and Hull, 1993).

In fact, there are few data addressing the accuracy of ultrasonographic estimates of follicle counts in polycystic ovaries. Takahashi *et al.* (1994) demonstrated a good correlation between follicle counts obtained by histopathology and transvaginal ultrasonography. However, the follicle counts were higher when estimated by ultrasound than when measured by histopathology (Takahashi *et al.*, 1994), a result that should not be unexpected because histological assessment was conducted on ovarian wedge resections and did not rely on the examination of whole ovaries (Takahashi *et al.*, 1994). Current ethical restrictions preclude the histological validation of ultrasonographic assessment of follicle counts using newer ultrasound technology because of the invasive nature of the techniques needed to obtain ovarian tissue for histopathology.

Limitations inherent to the ultrasound assessment of polycystic ovaries

There can be doubt as to whether or not every sonolucency in the ovary actually represents a follicle, given the variability in size and shape of ovarian follicles (Broekmans *et al.*, 2010). This is especially true for small follicles (≤ 2 mm in diameter) and it is questioned whether those follicles should be counted since there is lack of data supporting the reliability of ultrasound in the quantification of such small follicles. Likewise, there can be difficulty in interpreting contiguous follicles as being one or more than one (Broekmans *et al.*, 2010).

Moreover, there is also risk of recounting or overlooking follicles particularly in real-time examination, where there is no opportunity to flag follicles as they are counted. To that end, a standardized approach to quantifying follicle populations in polycystic ovaries has recently been proposed.

The method is limited to offline assessments in which medical imaging software allows for the flagging of individual follicles as they are counted (Lujan et al., 2010). The method involves compartmentalizing the ovary into grid sections and performing focused follicle counts on individual segments of the ovary to generate estimations of FNPO. The investigators demonstrated a high degree of agreement between multiple observers when a grid system was used and that little to no variation was evident when a single observer assessed the same images for FNPO (Lujan et al., 2010). Unfortunately, since the grid technique is restricted to offline analysis, it is mainly helpful for clinical research but not for clinical practice.

Follicle counting by three-dimensional ultrasonography: is it more reliable?

It is now possible to count follicles from stored three-dimensional (3D) datasets while simultaneously visualizing three perpendicular planes. The ability to cross-check follicles in multi-planar view aids in identifying follicles and this method was shown to have higher levels of reliability compared with two-dimensional (2D) methods when used to assess non-polycystic ovaries (Scheffer et al., 2002; Merce et al., 2005; Jayaprakasan et al., 2007; Deb et al., 2009). Another option for estimating follicle populations involves the use of 3D reconstruction volume calculation software (e.g. VOCAL™ and SonoAVC™), which can detect and quantify anechoic structures within an acquired 3D dataset. Use of this software showed better accuracy in determining follicular volume (Raine-Fenning et al., 2008; Lamazou et al., 2010; Salama et al., 2010) and reduced observer variation in follicle counts (Jayaprakasan et al., 2007; Deb et al., 2009) compared with manual 2D measurements. Because use of this software involves variable degrees of image processing, the reliability of follicle counts would be expected to depend heavily on image quality (Jayaprakasan et al., 2007) and on the aptitude of the interpreter for using the software. Moreover, the reliability of software in detecting small follicles (<5 mm in diameter) (Deb et al., 2010) and follicle counts >15 (Scheffer et al., 2002) was associated with substantially lower levels of agreement compared with other methods, which has important implications for their use in assessing polycystic ovaries.

To date, data attesting to the reliability of 3D ultrasonography to estimate follicle populations in polycystic ovaries are sparse as only a few studies have attempted to use 3D ultrasonography in this setting (Allemand et al., 2006; Ng et al., 2006; Lam et al., 2007, 2009; Sun and Fu, 2007; Pascual et al., 2008; Battaglia et al., 2012). In the only study to derive follicle thresholds for PCOM using 3D volume-based software, an agreement level of 0.82 among two observers was reported as part of an internal validation assessment (Allemand et al., 2006). Compared with 2D estimates, the 3D method counted more follicles in subjects with polycystic ovaries, which was the opposite to that observed when imaging normal ovaries (Jayaprakasan et al., 2007; Deb et al., 2009, 2010) and is in contrast with the findings of Battaglia et al. (2012) who reported similar follicle counts with 2D and 3D methodology in polycystic ovaries. Taken together, these data suggest that 3D ultrasonography holds promise in the evaluation of PCOM but that further studies are required before its routine use can be recommended.

Proposed thresholds for follicle excess in PCOM

Selection of the threshold value for defining follicle excess is a highly complex issue. Our review of the available literature demonstrated that three approaches have been previously used, namely: (i) arbitrary (i.e. choice of threshold based on clinical experience); (ii) use of receiver operator characteristic (ROC) curve analyses (which report the diagnostic power of a parameter to distinguish between the diseased and non-diseased conditions and propose thresholds that balance test sensitivity and test specificity) or (iii) use of the 95th percentile of age-matched control populations considered as normal.

Each approach has its drawbacks. Arbitrary choice cannot be accepted any longer as no consensus can be reached through this approach. Whether ROC curve analysis is appropriate for a condition like PCOS, in which there is controversy in defining criteria for the 'diseased condition', is debatable. Lastly, the use of a control population raises the issue as to what extent the population is truly normal. Many authors have used non-hyperandrogenic normo-ovulatory patients referred to their clinic, using various and debatable exclusion criteria. Others have used supposedly healthy women recruited from the general populations, although with variable definitions of 'healthy'.

The first set of most widely adopted criteria, proposed by Adams et al. (1985, 1986) in the 1980s, arbitrarily described PCOM as an ovary containing 10 or more follicles (measuring 2–8 mm in diameter) in one cross section of the ovary by using transabdominal ultrasonography. Since then, transabdominal approaches have been replaced by higher frequency transvaginal approaches which afford a greater likelihood of detecting the ovaries and a much better resolution for imaging small follicles. Moreover, thresholds for follicle counts now rely primarily on estimates of follicle populations throughout the entire ovary (FNPO), rather than in a single cross section (follicle number per section, FNPS), which is a highly important distinction that has led to confusion in both clinical practice and the literature.

The first study that used the ROC curve analyses for defining a threshold for FNPO was that of Jonard et al. (2003), demonstrating a 75% sensitivity and 99% specificity in distinguishing PCOS cases from controls as judged by transvaginal ultrasonography, with an FNPO threshold of ≥12 follicles measuring 2–9 mm in diameter (mean of both ovaries). The 2003 Rotterdam consensus, the most common ultrasound definition employed to date, was based on this single study and on expert agreement (Balen et al., 2003). This approach has been repeated recently in two studies comparing PCOS to controls by means of ROC curve analysis (Dewailly et al., 2011; Lujan et al., 2013). The conclusions of these studies were to raise the diagnostic threshold substantially to ≥19 and to ≥26 follicles per ovary, respectively. The different thresholds proposed by the two studies may be explained by differences in the analysis of the control populations because Dewailly et al. (2011) applied cluster analysis in order to exclude clinically normal women with PCOM (see section 7−2) prior to determining the threshold for FNPO. Had this preliminary step been omitted, their analysis would have yielded a cut-off value of 25 follicles, similar to the results obtained by Lujan et al. (2013).

However, these data differ from findings obtained in non-European or non-Caucasian populations. Chen et al. (2008) confirmed the ≥12 threshold for FNPO applying ROC analysis in the Chinese population, whereas Kosus et al. (2011) proposed an FNPO threshold of 8 follicles per ovary for Turkish women. These cut-offs are far below the newly proposed values for Western countries. Whether such a difference is solely due to ethnic variation or derived from the use of lower frequency transducers remains unclear.

The difficulty in choosing the threshold for FNPO that defines PCOM is confirmed by studies reporting on so-called 'normal' or 'general'

populations. In several recent studies of women of child-bearing age, the previous FNPO ≥12 threshold resulted in very large prevalences of PCOM in women from the general population, especially in those under 30 years old (Duijkers and Klipping, 2010; Johnstone et al., 2010; Kristensen et al., 2010, 2012; Rosen et al., 2010; Jokubkiene et al., 2012) (Table I).

In selected populations of women with regular menstrual cycles and no evidence for hyperandrogenism, three recent studies (Bentzen et al., 2013; Deb et al., 2013; Lujan et al., 2013) yielded median values of FNPO between 11 and 13 (Table I), indicating that at least half of the controls aged 20–35 years had PCOM when using an FNPO ≥12 threshold.

Therefore, these studies published in the 10 years after the 2003 Rotterdam consensus was made available strongly suggest that the FNPO ≥12 threshold is no longer valid for defining PCOM. Some investigators have interpreted these findings as evidence for the inadequacy of follicle counts as a criterion for PCOS (Johnstone et al., 2010) while others have considered these findings as indicative of the need to re-evaluate the FNPO threshold diagnostic of PCOM (Kristensen et al., 2010; Bentzen et al., 2013).

The systematic review and analysis of the literature identified several studies in which data regarding FNPO in the 'general' or 'normal' populations were sufficiently well documented (Table I). Interestingly, the FNPO 95th percentiles in most of these populations are quite similar to the FNPO thresholds proposed by the two recent studies that relied on ROC curve analysis (Dewailly et al., 2011; Lujan et al., 2013) (Table II), indicating that an FNPO threshold of around 25 follicles may be best used to distinguish normal ovarian morphology from PCOM in most populations (Table I).

Why such variability and controversy about the threshold for FNPO?

The appropriateness of proposed thresholds for FNPO can be influenced by several factors as described below (Table III).

Differences in the methods of counting follicles

Considerable variability exists in both published studies and in clinical practice in the technical methods used to count, measure and report follicles. A consortium of experts recently met with the aim of standardizing real-time methods for estimating antral follicle populations (Broekmans et al., 2010). Their recommendation for a systematic method of counting follicles included performing initially a 'scout sweep' of the ovary in two planes to discern its boundaries, and establishing then the size of the largest follicle by making orthogonal measurements of the follicular antrum. It was recommended that calliper-based measurements should be made of all follicles >10 mm in diameter before proceeding to counting all remaining follicles between 2 and 10 mm in the preferred (longitudinal) sweep of the ovary. Counts for both ovaries should be summed, and follicles >10 mm subtracted, to obtain a total AFC.

This approach differs from that proposed by Balen et al. (2003) who recommended estimating the follicles in multiple planes and reporting the mean follicle counts of the left and right ovary when assessing PCOM (i.e. to generate an FNPO). Moreover, they recommended performing these estimates only in the absence of a dominant follicle, which could interfere with obtaining accurate measurements of OV and could

influence the accurate estimation of the follicle count. Whether there has been widespread adoption of either method for estimating follicle populations is uncertain. Moreover, we are unaware of any systematic evaluation of these real-time methods to reduce observer variability.

Observer variability in assessing follicle number

In the first prospective evaluation of the variability associated with assessing PCOM by ultrasound, Amer et al. (2002) demonstrated that the ultrasonographic diagnosis was highly subjective and called into question its utility as an aid in the diagnosis of PCOS. Two recent studies (Lujan et al., 2008, 2009) have evaluated prospectively the variability associated with counting follicles in polycystic ovaries and reported poor agreement between observers when estimating FNPO. These data stand in contrast to past reports of good agreement in the assessment of FNPO in subfertile women without PCOS (Scheffer et al., 2002; Jayaprakasan et al., 2007).

Differences among studies might be best explained by differences in follicle populations among the clinical populations. Better agreement would be expected for normal ovaries since they contain fewer follicles, and those present are typically larger and not as densely packed as those found in polycystic ovaries. Groups evaluating follicle counts in normal ovaries have noted a distinct decrease in agreement when follicle counts (sum of both ovaries) exceed 15 (Scheffer et al., 2002), consistent with the notion that the higher numbers of follicles in polycystic ovaries would result in a greater inter-observer variability in FNPO. Collectively, these studies indicate that obtaining accurate estimates of FNPO in polycystic ovaries is challenging and prone to significant inter-observer variability.

Impact of recent advancements in imaging technology on the variability in follicle counts

There have been marked improvements in the level of spatial resolution afforded by newer ultrasound scanners. Some investigators credit the improved spatial resolution that has occurred in the 10 years since the Rotterdam consensus for primarily driving the need to re-evaluate criteria for PCOM (Dewailly et al., 2011) (Fig. 1). Considering the first year of patient inclusion in studies evaluating follicle populations in healthy women of reproductive age, there is a clear increase in the median values for FNPO over time, consistent with notion that the increasing transducer frequency of newer ultrasound scanners facilitates the detection of more follicles (Fig. 2). Regression analysis confirmed a significant effect of Max Transducer Frequency on FNPO ($P = 0.023$), independent of the mean age of the patients reported for each study. Post hoc analysis of FNPO at different transducer frequencies revealed a significant increase in reported FNPO when the transducer frequency was ≥8 MHz ($P < 0.0001$).

This has very important implications for when choosing the FNPO threshold diagnostic of PCOM in clinical practice and research.

Recommendations

After critical analysis of the recent and pertinent literature, this Task Force recommends setting the FNPO threshold for the definition of PCOM at ≥25 follicles for most populations.

However, because the age of equipment likely impacts the number of follicles that might be visible on ultrasound, practitioners should be cautioned to check whether this threshold fits with the technology available

Table I 5th, 50th and 95th percentiles (P5, P50 and P95) for follicle number per ovary and ovarian volume in healthy women with regular menstrual cycles and normal androgens.

Author (year)	Control population	Age	Technical aspects	FNPO	OV (ml)
Duijkers and Klipping (2010)	n = 171 General population Regular menstrual cycles Hyperandrogenism not an exclusion criterion	Min.: 18 Median: 25 Max.: 40	January − July 2003 TVUS/6 MHz Scanned Day 6−9 Real-time counts of all visible follicles Mean OV reported	P5: 4 P50: 13 P95: 27	P5: 3.9 P50: 7.1 P95: 11.8
Dewailly et al. (2011)	n = 105 Investigation for male factor or tubal infertility Regular menstrual cycles No hyperandrogenism Combined data for Groups 1A+1B	P5: 21.7 P50: 29.5 P95: 34.2	2008−2010 TVUS/5−9 MHz Scanned Day 2−5 Real-time counts of all follicles <10 mm Mean OV reported	P5: 7 P50: 15 P95: 26	P5: 2.7 P50: 5.5 P95: 9.8
Kristensen et al. (2012)	n = 44 (out of 76) for FNPO General population Regular menstrual cycles Subset of non-OCP users Not screened for hyperandrogenism	P25: 19.9 P50: 20.2 P75: 20.4	2008−2009 TVUS/4−9 MHz Scanned at random Real-time counts of 2−9 mm follicles No available data on OV	P5: 10 P50: 16 P95: 25	NA
Lujan et al. (2013)	n = 70 General population Regular menstrual cycles No hyperandrogenism	P5: 18.6 P50: 26.0 P95: 35.5	2006−2012 TVUS/6−12 MHz Scanned Day 2−5 Offline counts of 2−9 mm follicles Mean OV reported	P5: 6 P50: 13 P95: 26	P5: 2.0 P50: 6.0 P95: 13.0
Bentzen et al. (2013)	n = 228 General population Regular menstrual cycles No hyperandrogenism	Group 1 20−29 y n = 62 Group 2 30−34 y n = 166	2008−2010 TVUS/4−9 MHz Scanned Day 2−5 Real-time counts of 2−10 mm Mean OV reported	Group 1 P5: 6 P50: 13 P95: 23 Group 2 P5: 5 P50: 12 P95: 21	Group 1 P5: 2.2 P50: 5.4 P95: 8.8 Group 2 P5: 2.8 P50: 5.4 P95: 8.9
Johnstone et al. (2010)	n = 509 for FNPO n = 447 for OV General population Regular menstrual cycles No hyperandrogenism	P5: 25 P50: 30 P95: 35	2008−2011 TVUS/4−8 MHz Scanned Day 2 −4 Real-time counts of 2−10 mm follicles Max. OV in n = 447 reported	P5: 4 P50: 9 P95: 19	P5: 2.4 P50: 5.7 P95: 10.9
Combined data[a]	n = 1127 for FNPO n = 1021 for OV	See above	See above	P5: 0 P50: 11 P95: 23	P5: 0.9 P50: 5.9 P95: 10.8

FNPO, follicle number per ovary; NA, not available; OCP, oral contraceptive pill; OV, ovarian volume; TVUS, transvaginal ultrasound.

[a]Combination of the data was conducted after assuming the normal distribution of FNPO and OV in the individual studies since all of them included a sample size above 40 and their P5, P50 and P95 suggested only moderate skewness. The standards deviations (SD) of the individual studies were calculated from the means (P50) and P95 values. Means and SD were combined considering the sample sizes of the different studies using the online statistical facilities of the Department of Obstetrics and Gynaecology, Chinese University of Hong Kong (http://department.obg.cuhk.edu.hk/researchsupport/Combine_groups.asp; last accessed August 7, 2013). Combined P5 and P95 values were then calculated from the combined mean and SD.

to them (mainly, transducer frequency ≥8 MHz). This also implies that thresholds will need to be revisited in subsequent years to reflect any further advancement in imaging technology.

The Task Force recognizes that performing follicle counts in a single cross-sectional view of the ovary (FNPS) is the metric in current use to define PCOM. However, normative values for FNPS in the general population are sparse and the utility of this metric compared with FNPO is debatable (Allemand et al., 2006; Lujan et al., 2013). Currently, there is insufficient data to recommend an FNPS threshold to define PCOM.

Lastly, the transabdominal route is not suitable for recording a precise follicle count but in situations when it is the only way to assess ovarian morphology, it allows a reliable assessment of OV (see below).

Ovarian Size to Define PCOM

Proposed thresholds for ovarian enlargement in PCOM

Many studies have shown that increased ovarian size represents an important feature of PCOM (Balen et al., 2003). In fact, ovarian size appears to be increased in the majority of women with PCOS (Carmina et al., 2005) and mean ovarian size is higher in women with PCOS than in normal women matched for age and body weight (Balen et al., 2003; Carmina et al., 2005; Alsamarai et al., 2009).

Less clear is the establishment of an acceptable cut-off for OV between normal and polycystic ovaries (Table II). The Rotterdam

Table II Proposed thresholds for follicle number and ovarian volume in polycystic ovaries.

Author (year)	Proposed threshold	Technical aspects	Statistical analysis	Clinical populations	
				PCOS	**Controls**
Adams *et al.* (1985)	≥10 (2–8 mm)	TA 2D Offline FNPS	Arbitrary	*n* = 55 PCO plus amenorrhoea, hirsutism and/or infertility	–
Yeh *et al.* (1987)	>5 (5–8 mm) in one ovary + >4 in other ovary OV > 7.5 ml	TA 2D Real-time FNPO	Sp 100% Se 82% Max OV in controls	*n* = 74 Hirsutism and/or amenorrhoea plus high androgens or LH:FSH	*n* = 25 Patients with abdominal pain or dysmenorrhoea
Pache *et al.* (1992)	≥12 OV > 8 ml	TV 2D Real-time FNPO	Sp 100 Max OV in Controls	*n* = 52 Infertility and amenorrhoea plus increased LH and/or androgens	*n* = 29 Healthy volunteers
Fox (1999)	≥15 (2–10 mm)	TV 2D Real-time FNPO	Sp 100%	*n* = 29 Hirsutism, amenorrhoea and elevated LH	*n* = 40 Patients with male factor infertility
Atiomo *et al.* (2000)	10 (2–8 mm) OV > 9 ml	Mostly TV 2D Offline FNPS	Arbitrary Sp LO 68% Sp RO 85% Se LO 82% Se RO 72% Sp LO 68% Sp RO 44% Se LO 74% Se RO 69%	*n* = 32 Amenorrhoea plus increased free androgen index or LH:FSH	*n* = 40 Volunteer hospital staff
Jonard *et al.* (2003)	≥12 (2–9 mm)	TV 2D Real-time FNPO	AUC 0.937 Sp 99% Se 75%	*n* = 214 Amenorrhoea or hirsutism plus increased LH/androgens or ovarian area	*n* = 112 Male factor or tubal infertility patients PCO as an exclusion criterion
Jonard *et al.* (2005)	≥12 (2–9 mm) OV > 7 ml	TV 2D Real-time FNPO	AUC 0.956 Sp 97% Se 79% AUC 0.905 Sp 68% Se 91%	*n* = 98 Amenorrhoea plus hirsutism and/or increased androgens	*n* = 57 Male factor or tubal infertility patients
Allemand *et al.* (2006)	≥20 (FNPO) ≥10 (FNPS) OV ≥ 13 ml	TV 3D Offline	AUC 0.987 Sp 100% Se 70% AUC 0.990 Sp 100% Se 90% AUC 0.948 Sp 100% Se 50%	*n* = 10 Amenorrhoea plus hirsutism and/or increased androgens	*n* = 29 Male factor or tubal infertility patients
Chen *et al.* (2008)	≥10 (<10 mm) OV > 6.4 ml	Mostly TV 2D Real-time FNPO	AUC 0.909 Sp 89% Se 85% AUC 0.898 Sp 86% Se 81%	*n* = 432 Amenorrhoea plus hirsutism and/or increased androgens	*n* = 153 Healthy volunteers
Kosus *et al.* (2011)	≥8 (<10 mm) OV > 6.43 ml	TV 2D Real-time FNPO	AUC 0.998 Sp100% Se 95% AUC 0.938 Sp 81% Se 95%	*n* = 251 Hirsutism and/or increased androgens plus amenorrhoea and/or PCO	*n* = 65 Healthy volunteers

Continued

Table II *Continued*

Author (year)	Proposed threshold	Technical aspects	Statistical analysis	Clinical populations	
				PCOS	Controls
Dewailly et al. (2011)	≥19 (<10 mm) OV > 7 ml	TV 2D Real-time FNPO	AUC 0.949 Sp 92% Se 81% AUC 0.923 Sp 89% Se 87%	n = 62 Hyperandrogenism and amenorrhoea	n = 66 Patients with regular menses and no hyperandrogenism Separated by cluster analysis using age, androgens, LH, AMH, OV and FNPO as variables
Lujan et al. (2013)	≥26 (FNPO) ≥9 (FNPS) OV > 10 ml	TV 2D Offline Grid	AUC 0.969 Sp 94% Se 85% κ = 0.71 AUC 0.880 Sp 84% Se 81% κ = 0.72 AUC 0.873 Sp 84% Se 81% κ = 0.82	n = 98 Amenorrhoea plus hirsutism and/or increased androgens	n = 70 Healthy volunteers

AMH, anti-Müllerian hormone; AUC, area under the curve; FNPO, follicle number per ovary; FNPS, follicle number per cross section; FSH, follicle-stimulating hormone; κ, inter-observer agreement when making diagnosis; LH, luteinizing hormone; LO, left ovary; OV, ovarian volume; PCO, polycystic ovaries; RO, right ovary; Se, sensitivity; Sp, specificity; TA, transabdominal; TV, transvaginal.

consensus statement suggested a threshold of 10 ml based on expert opinion (Balen et al., 2003). Since then, several groups of researchers have proposed much lower cut-off values including 6.4 ml (Kosus et al., 2011), 6.7 ml (Chen et al., 2008), 7.0 ml (Jonard et al., 2005; Dewailly et al., 2011) and 7.5 ml (Carmina et al., 2005).

These different OV threshold values might depend on the variable clinical and metabolic characteristics of the populations studied, and particularly on ethnicity, body mass index and insulin levels. In women with PCOS, the mean ovarian size appears to be higher in populations characterized by large prevalence of weight excess, such as those in Canada and the USA (Alsamarai et al., 2009; Lujan et al., 2013), intermediate in European countries (Carmina et al., 2005; Jonard et al., 2005), and lower in East-Asian countries (Chen et al., 2008). In addition, a positive correlation between ovarian size and circulating insulin levels has been demonstrated by several studies (Carmina et al., 2005; Alsamarai et al., 2009). Collectively, there may be reason to support distinct OV thresholds for lean versus overweight or obese populations, but these data are not yet available. However, in a recent description of baseline characteristics of a large cohort (n = 750) of women with PCOS submitted to a randomized, double-blinded clinical trial to determine first-line ovulation induction protocol, BMI had a slight and non-significant effect on OV (Legro et al., 2013).

Ovarian size over the lifespan

Ovarian size also varies with age, reaching a maximum during adolescence (1.3–3.8 years post-menarche), slowly decreasing during adulthood and rapidly shrinking after menopause. The finding of an elevated mean ovarian size during adolescence that decreases with each decade of life has been demonstrated using magnetic resonance imaging (MRI) and ultrasonography. Using MRI, Well et al. (2007) observed a marginal decrease in OV between the second and the fourth decades of life (from 9.5 ± 3.3 ml to 8.5 ± 3.3 ml, approximately a 10% reduction), with ovaries being half their size by the late fifth decade of life (Well et al., 2007). While the absolute size of the ovaries was notably smaller in similar studies using ultrasonography (Pavlik et al., 2000; Garel et al., 2001), the following changes in OV over the lifespan were noted by Pavlik et al. (2000) using ultrasound: age <30 years, 6.6 ± 0.2 ml; 30–39 years, 6.1 ± 0.1 ml; 40–49 years, 4.8 ± 0.0 ml; 50–59 years, 2.6 ± 0.0 ml; 60–69 years, 2.1 ± 0.0 ml; and ≥70 years, 1.8 ± 0.1 ml.

Because relatively small changes appear to occur between the ages of 20 and 39 years, it is unlikely that age-specific thresholds for OV may be truly needed for this population. In contrast, the definition of PCOM in adolescence or in women over the age of 40 years would require careful consideration of natural changes in OV with age.

Measurements of ovarian size by 3D ultrasonography

Three-dimensional ultrasound has been shown to provide an objective tool for quantifying OV. Several studies have measured OV using 3D ultrasound. Mean OV varied between 10.6 and 16.7 ml in women with polycystic ovaries and from 5.2 to 8.7 ml in healthy women of reproductive age (Lam et al., 2007, 2009; Pascual et al., 2008; Battaglia et al., 2012; Jokubkiene et al., 2012). Although two studies comparing 2D and 3D demonstrated a strong correlation in measurements between the two modalities (Pascual et al., 2008; Battaglia et al., 2012), there were significant discrepancies in overall ovarian size between the different studies, suggesting technical and inter-observer variability.

Table III Factors contributing to variations in thresholds for follicle number in polycystic ovaries.

	Inconsistent parameter among studies	Considerations
Clinical populations	Definition of PCOS	Potential to yield heterogeneous cohorts
		PCO as an inclusion criterion is controversial
	Inclusion criteria for controls	Recruitment methods for controls often not specified
		Appropriateness of subfertile women as controls
		PCO as an exclusion criterion is controversial
	Age	Thresholds do not apply to women <18 and >35 years
	Ethnicity	Follicle counts may vary among ethnic populations
Statistical approach	Arbitrary cut-offs	Biased by the interpreter
	Based on 100% specificity	Biased at the expense of test sensitivity
	ROC curve analysis with Youden's Index	Balances test sensitivity and specificity
	95th percentile of control population	Concedes a false-negative rate
		Concedes a false-positive rate
Technical issues	Newer versus older technology	More follicles can be visualized using newer ultrasound technology
	TA versus TV ultrasound	TA approaches are indicated for certain clinical populations
		Visualization is poorer using low-frequency TA approaches, particularly with obesity
	Real-time versus offline counts	Increased duration for post hoc analyses
		Offline methods yield higher counts
		Potential for increased precision in follicle counts made offline
	2D versus 3D follicle counts	Increased cost of 3D equipment
		3D affords shorter scan time for patients
		3D allows for multi-planar and volume-based assessments of follicle counts from stored image files
		3D multi-planar view has highest reliability in follicle counts
		3D methods yield lower follicle counts
		Automated assessment of follicle counts by reconstructed volumes requires further validation

2D, two-dimensional; 3D, three-dimensional ; PCO, polycystic ovaries; PCOS, polycystic ovary syndrome; ROC, receiver operating characteristic; TA, transabdominal; TV, transvaginal.

Recommendations

OV appears as a good surrogate marker of PCOM although, compared with FNPO, OV had lower sensitivity for discriminating between patients with PCOS and controls in all the studies comparing both parameters (Table II). Therefore, this Task Force recommends using OV for the diagnosis of PCOM in instances when the image quality does not allow a reliable estimate of FNPO, especially when the transvaginal route is not feasible. The use of in-house reference normal values is highly recommended but, if unavailable, the existing OV ≥10 ml threshold can be used conservatively.

Other Imaging Variables Used to Define PCOM

There is considerable debate whether the following assessments should be utilized and included in the criteria of PCOM.

Specific assessment of ovarian stroma

By 3D ultrasonography, stromal volume can be measured through calculation and subtraction of total follicular volume from the total OV. For example, Chinese women with PCOS, although having smaller stromal volumes than Caucasian women with PCOS (9.74 versus 10.79 ml), had higher ovarian stromal volume compared with women without PCOS (9.74 versus 4.07 ml) (Chen et al., 2008; Lam et al., 2009).

The ratio of ovarian stroma to total ovarian size may be a good criterion for diagnosis of PCOS, with a cut-off value of 0.32 indicating an association with hyperandrogenaemia (Fulghesu et al., 2007). However, to date there are few studies corroborating the diagnostic potential of this variable. In general, ovarian stromal volume and total ovarian size are well correlated and hence, there may not be any additional value to including stromal size measurements in clinical practice.

Ovarian blood flow in PCOS

Increased OV has been associated not only with enhanced stromal echogenicity but also with increased vascularity. Even though the introduction of 3D ultrasound has allowed better and more objective assessment of ovarian morphology and vascularization (Raine-Fenning et al., 2003, 2004), results pertaining to differences in ovarian blood flow in PCOS have been conflicting.

A prospective study by Adali et al. (2009) showed higher ovarian stromal blood flow and reduced uterine perfusion in patients with PCOS compared with age-matched women without PCOS. Similarly, Battaglia et al. (2012) reported that assessments of ovarian vascularization, as judged by 3D power Doppler, were significantly increased in PCOS patients compared with controls. These findings are consistent with those of other studies (Battaglia et al., 1995; Zaidi et al., 1995; Lam et al., 2009) but disagree with the results of other studies using 2D and 3D Doppler technology that did not observe any differences in vascularization between women with PCOS and controls (Jarvela et al., 2003; Ng et al., 2005; Younis et al., 2011).
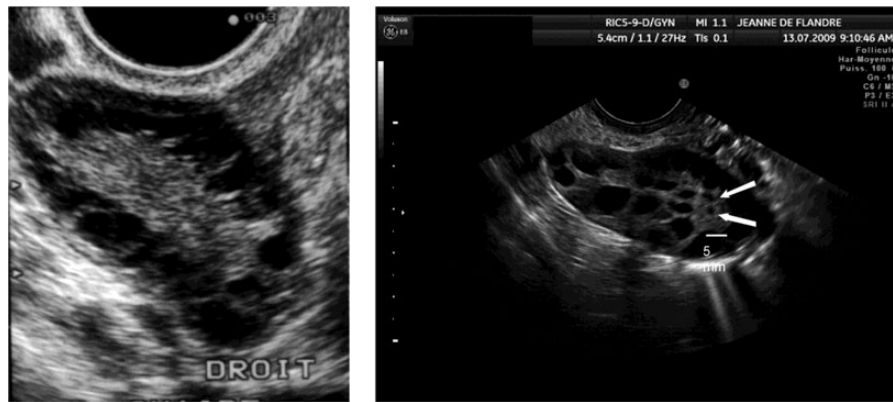
Figure 1 Picture of PCOM obtained with old (2001, left panel) and new (2009, right panel) ultrasound equipment. Small follicles ≤2 mm in diameter (arrows) can be visualized and counted with the new equipment. Reproduced from Dewailly et al. (2011) with permission from Oxford University Press, Copyright 2011.
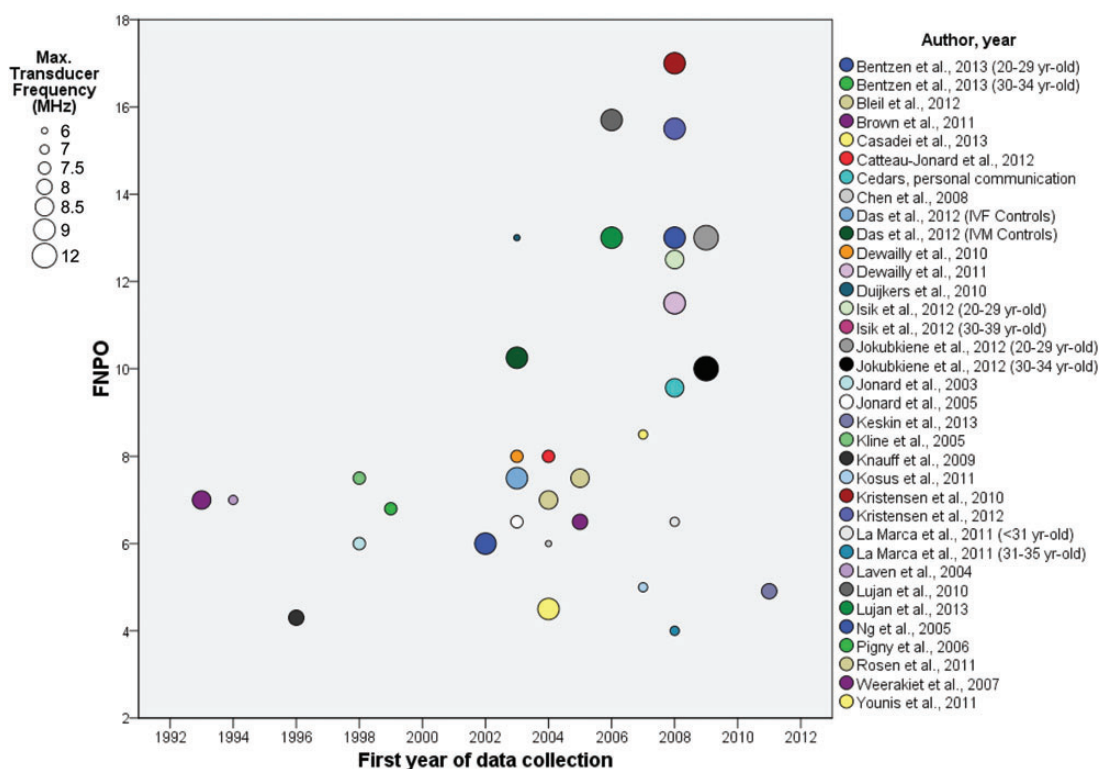
Figure 2 Changes in reported values for the mean or the median (depending on available data) of follicle number per ovary (FNPO) in healthy women with regular menstrual cycles over time. FNPO reported in control populations by various groups are plotted with respect to the first year of data collection. There has been an increase in mean/median values for FNPO over time consistent with the notion that advances in imaging technology, particularly increased transducer frequency (≥8 MHz), allow for improved detection of antral follicles on ultrasonography. Follicle counts were made in the early follicular phase for all but two studies. In total, 30 studies met all the criteria (see Methods section) and were included. Four studies reported follicle counts for separate age groups in control populations and are reported as separate end-points. Colours indicate the first author and year of the manuscript (Jonard et al., 2003, 2005; Laven et al., 2004; van Rooij et al., 2004; Kline et al., 2005; Ng et al., 2005; Pigny et al., 2006; Weerakiet et al., 2007; Chen et al., 2008; Knauff et al., 2009; Dewailly et al., 2010, 2011; Duijkers and Klipping, 2010; Kristensen et al., 2010, 2012; Rosen et al., 2010; Brown et al., 2011; Kosus et al., 2011; La Marca et al., 2011; Younis et al., 2011; Bleil et al., 2012; Catteau-Jonard et al., 2012; Das et al., 2012; Isik et al., 2012; Jokubkiene et al., 2012; Bentzen et al., 2013; Casadei et al., 2013; Keskin et al., 2013; Lujan et al., 2013).

Stromal vascularity may be altered in PCOS through the effects of hyperinsulinaemia on vascular smooth muscle relaxation, increased angiogenesis and dilatation of vessels in ovarian stromal tissue (Steinberg et al., 1994). However, similar to the results concerning stromal vascularization in PCOS, in general, the data supporting a relationship between insulin resistance and ovarian blood flow are inconsistent (Loverro et al., 2001; Ajossa et al., 2002; Chekir et al., 2005). Collectively, the inability to standardize measurements, the use of small sample sizes, differences in study populations (i.e. ethnicity, inconsistent criteria of polycystic ovaries), the use of inappropriate control subjects (e.g. patients with low ovarian reserve) and the lack of adjustment for confounders such as age and body mass index are all factors that that could explain the discrepancies among studies. At present, the lack of uniform data and absence of cut-off values make vascular indices impractical for discriminating between polycystic and normal ovaries.

Increased Serum AMH Concentrations as a Surrogate Marker of PCOM

Rationale for the use of serum AMH concentrations

Increased serum levels of anti-Müllerian hormone (AMH), a peptide produced by the granulosa cells (GC) of ovarian follicles, are the most prominent endocrine abnormality associated with PCOM and, in particular, with ovarian follicle excess (Pigny et al., 2003; Laven et al., 2004). Serum AMH concentrations appear to be increased in PCOS patients because their ovaries exhibit an increased number of AMH-producing pre-antral and small antral follicles (Weenen et al., 2004) and because GC production of AMH is greatly increased (Pellatt et al., 2010). There seems to be a consistent relationship between AMH serum levels and ultrasound estimates of FNPO and OV (Pigny et al., 2003; Laven et al., 2004; Piltonen et al., 2005; Dewailly et al., 2011; Villarroel et al., 2011; Eilertsen et al., 2012; Robin et al., 2012).

The difficult issue of assaying serum AMH concentrations

Unfortunately, universally accepted methods and assays to measure serum concentrations of AMH are still lacking. Up till 2010, about half of all published studies used the Diagnostic Systems Laboratories (DSL) assay whereas the other half used the Immunotech (IOT) assay (Nelson and La Marca, 2011). These two assays utilized two different antibodies and different standards and consequently crude values differed substantially, with the IOT assay yielding higher AMH concentrations in most of series. With the recent consolidation of these two companies by Beckman Coulter (Beckman Coulter, Inc., Chaska, MN, USA), and their sole ownership of the patent to measure mammalian AMH, there is finally a single commercially available assay, the AMH Gen II assay, which will fully replace the DSL and IOT assays (Nelson and La Marca, 2011).

The Gen II assay uses the DSL Gen I antibody with the IOT standards. It is generally believed that higher AMH levels are recorded with the IOT and Gen II assay compared with the DSL assay (Nelson and La Marca, 2011) but the use of conversion factors from one assay to the other is controversial (Rustamov et al., 2012).

Although there are some similarities between the IOT and Gen II assays, as far as the standards are concerned, the latter seems to be affected by proteolysis and most likely by interference with the serum complement in undiluted samples from patients (unpublished data). Further, there is currently no international standard in accordance with the International Federation of Clinical Chemistry. Moreover, there seems to be no eagerness at the moment for national or international bodies to establish such a standard because of the limited use of the assay as well as the fact that there is only one manufacturer at the moment. Therefore, it is difficult to compare the available series.

What are the results so far?

Most of the published reports assessing the role of serum AMH posit that it might play a role in facilitating the diagnosis of PCOS (Cook et al., 2002; Pigny et al., 2003; Lin et al., 2011; Woo et al., 2012) according to the 2003 Rotterdam consensus and 2006 AE-PCOS Society position statements. However, only a limited number of studies have tried to assess specifically whether AMH serum concentrations might be an effective surrogate marker of PCOM to be used in the above-mentioned classifications (Laven et al., 2004; Piltonen et al., 2005; Pigny et al., 2006; Dewailly et al., 2011; Villarroel et al., 2011; Eilertsen et al., 2012).

In adults, in a study by Dewailly et al. (2011), serum AMH concentrations measured using the IOT assay showed an area under the ROC curve of 0.973 for the diagnosis of PCOM, with a threshold value of ≥35 pmol/l showing 92% sensitivity and 97% specificity. In this study, specific thresholds for AMH and FNPO were calculated concomitantly without using pre-determined values for FNPO. In addition, women with supposedly asymptomatic PCOM (see below) were excluded from the control group of regularly menstruating women by cluster analysis. As AMH results with the ROC curve analysis were even better than those obtained for ultrasound assessment of FNPO, these investigators concluded that a serum AMH level of >35 pmol/l might replace the finding of PCOM in the definition of PCOS (Dewailly et al., 2011). In a smaller series, using the same AMH assay, ROC analysis yielded an AMH ≥33 pmol/l threshold (Casadei et al., 2013).

In other studies, the serum AMH level was tested against PCOM as pre-defined by an FNPO ≥12, a threshold that is now likely obsolete (see above). Therefore, it is not surprising that lower sensitivity and/or specificity were obtained with the ROC curve analysis. Eilertsen et al. (2012), using the DSL assay, suggested a 20 pmol/l cut-off value for serum AMH concentrations as a marker of PCOM, with 80% sensitivity and 72% specificity. Again grouping patients according to an FNPO ≥12 threshold for PCOM and pooling the results from DSL and Gen II assays using a conversion factor of 1.4, Homburg et al. (2013) indicated that a serum AMH threshold of 48 pmol/l had an excellent specificity (98%) but only at the expense of a poor sensitivity (60%). In addition to the key issue of using pre-determined threshold values for FNPO, the normative data for serum AMH concentrations are not easy to compare because 'normal' controls might well be recruited from infertility clinics and such controls may not necessarily be representative of the general population.

In adolescents, the diagnostic value of serum AMH concentrations has also been studied since ultrasound is often unreliable to detect PCOM in this population. A study in Chilean adolescents identified a cut-off serum

AMH concentration of 60 pmol/l (with the IOT assay) to diagnose PCOM in regularly menstruating adolescents, with a sensitivity and specificity of 64 and 90% (area under the ROC curve = 0.873, confidence interval 0.782–0.963) (Villarroel et al., 2011). The results were not as good in Australian adolescents with the same assay (area under the ROC curve = 0.67, confidence interval 0.60–0.75) leading the authors to conclude that serum AMH concentrations were a questionable surrogate marker of PCOM in adolescents (Hart et al., 2010).

Recommendations

Given the uncertainty around AMH assays, we decided not to systematically review the currently available data concerning the value of serum AMH in diagnosing PCOM, although another group has done so recently (Iliodromiti et al., 2013). The Task Force recommends against considering an increased serum AMH concentration as surrogate marker of PCOM for clinical practice and research until an accurate AMH assay, which produces reliable and reproducible results, is available in the future.

What Is the Meaning of PCOM?

The answer to this question varies greatly depending on the population under study.

PCOM in hyperandrogenic and/or dysovulatory women

Despite not being included in the diagnostic criteria derived from the consensus conference sponsored in 1990 by the National Institutes of Health (NIH) (Zawadzki and Dunaif, 1992), PCOM is now considered one of the criteria for the diagnosis of PCOS in the two most recent definitions of PCOS, namely the 2003 Rotterdam consensus (The Rotterdam ESHRE/ASRM-sponsored PCOS consensus workshop group, 2004a, b) and the 2006 Androgen Excess & PCOS (AE-PCOS) Society criteria (Azziz et al., 2006, 2009). Of note, in the most recent NIH-sponsored evidence-based methodology workshop on PCOS held in 2012, an independent panel of experts recommended that PCOM continues to be considered as diagnostic criterion for PCOS (Johnson et al., 2012). However, some investigators still contest the inclusion of this feature in diagnostic classifications for PCOS, since it is also observed in apparently normal women (see below).

Available data suggest an inherited basis for the aggregation of PCOM within families with PCOS (Govind et al., 1999). There is also the near-universal finding of PCOM across ethnic and racial groups in women otherwise diagnosed with PCOS. Besides the data from Northern European Caucasian populations, the largest study from China (Zhang et al., 2013) evaluated 719 cases and 685 controls and confirmed that 92% of cases met Rotterdam criteria for FNPO. Diamanti-Kandarakis and Panidis (2007) similarly found that 90% of Greek women presenting with hirsutism and menstrual irregularity had PCOM. Kumarapeli et al. (2008) performed a population-based study in Southeast Asia, identifying probable PCOS cases by a questionnaire sampling menstrual history and symptoms of hyperandrogenism. Only 1% of the defined control population fulfilled diagnostic criteria for PCOS upon clinical evaluation while 17.6% meet ultrasound criteria (Kumarapeli et al., 2008). Of women defined as probable cases, 96.7% met ultrasound criteria defined by Rotterdam criteria (Kumarapeli et al., 2008). An earlier study using the Adams criteria for PCOM (Welt et al., 2006) did identify

some ethnic differences, with African American women having higher follicle counts and OV, but in all ethnicities >90% of women with PCOS showed PCOM on ultrasound examination.

The follicle excess in PCOS is tightly correlated with hyperandrogenism and experimental data on animal models (Vendola et al., 1999) or clinical data about female-to-male transsexuals (Baba et al., 2007) suggest that this is a causal relationship (reviewed in Homburg, 2009).

Recently it has been shown by Principal Component Analysis that the FNPO is one item of the androgen component of PCOS having even better sensitivity than serum androgen measurements (Dewailly et al., 2010). Therefore, the presence of PCOM may be regarded as a sign of hyperandrogenism and the same might apply for elevated serum AMH concentrations (Dewailly et al., 2010).

The follicle excess and high serum AMH level are also intimately linked to the ovulation disorder of PCOS (Catteau-Jonard et al., 2012). Therefore some authors propose a simplified diagnostic procedure for PCOS, i.e. oligo- or anovulation in conjunction with hyperandrogenism. In case one of these criteria was lacking, PCOM and/or a high serum AMH level could be used as a substitute for either oligo- or anovulation or hyperandrogenism, provided other specific disorders have been excluded (Dewailly et al., 2010). This reconciles the above-mentioned two definitions for PCOS (The Rotterdam ESHRE/ASRM-sponsored PCOS consensus workshop group, 2004a, b; Azziz et al., 2006, 2009) that were conflicting about requiring or not hyperandrogenism as a mandatory item.

PCOM in the general population

PCOM is also encountered in the general normal population in as many as 30% of young women (Dewailly et al., 2010; Villarroel et al., 2011). Actually, this figure is even higher if one applies the former FNPO ≥12 threshold when using modern ultrasound equipment, such as in recent series (Table I and Fig. 2). The relevance of this issue has been discussed above.

Taken as a whole, most series showed that PCOM in adult premenopausal healthy women was not related to metabolic variables (see below), although their serum AMH and androgen concentrations were often slightly higher than those of women without PCOM (reviewed in Johnstone et al., 2010).

This does not mean, however, that every clinically normal woman with PCOM has an occult androgen excess disorder. This issue was addressed by Mortensen et al. (2009) who compared 32 asymptomatic volunteers with PCOM to 21 similar volunteers showing normal ovaries on ultrasound. The PCOM subjects showed variable response to GnRH agonist, ACTH and oral glucose testing. Of the women in the PCOM group, 47% showed a GnRH agonist response like that observed in patients with PCOS, and 25% had elevated free testosterone of whom a third had an abnormal GnRH agonist response (Mortensen et al., 2009). It therefore seems as if PCOM formed a distinct but heterogeneous population with respect to ovarian function, ranging from normal (53%) to occult PCOS (25%). However, these conclusions were derived from the results of androgen testing whose specificity and sensitivity were not evaluated according to specific thresholds. To avoid such a subjective bias, cluster analysis was recently used to isolate homogenous subgroups within a control population (Dewailly et al., 2011). Two subgroups were yielded by the analysis, and the clustering was primarily based on the serum AMH level and then FNPO and

OV. The subgroup with the highest values for these parameters represented most likely control patients with PCOM (37% in this study) and was mathematically homogeneous, as was the other subgroup with no marker of PCOM.

The question thus arises as to whether asymptomatic women with PCOM constitute a heterogeneous population in terms of ovarian dysfunction ranging from entirely healthy ovulatory women to women with mild occult PCOS (Mortensen et al., 2009) or, alternatively, constitute a homogenous population representing the milder end of the PCOS spectrum (Ng et al., 2006; Catteau-Jonard et al., 2012). This alternative hypothesis is not supported by the single study available that addresses the follow-up of normal adult women presenting with PCOM. Indeed, PCOM did not predict the future development of PCOS, and even within PCOS patients, not all women meeting PCOM criteria at baseline fulfilled such criteria later in life (Murphy et al., 2006). However, the putative predictive value of PCOM for the future development of PCOS in young women has still to be evaluated, as the mean ages at baseline and follow-up visits were 28 and 38, respectively. The finding that not all women diagnosed with PCOS at initial screening maintained such a diagnosis later in life confirms studies showing that PCOS symptoms, including PCOM, may improve after the age of 30 (Brown et al., 2011). Therefore, the putative prediction of PCOS by PCOM should be investigated in younger populations than the one used for this study.

To our best knowledge, only one longitudinal study about healthy adolescent girls is available so far (Codner et al., 2011). PCOM was observed in 40, 35 and 33.3% of the ultrasonography studies prospectively performed at 2, 3 and 4 years after menarche, respectively, with a poor concordance among girls in the results during the follow-up and no association of PCOM with metabolic variables and androgen concentrations (Codner et al., 2011). One might question the performance of ultrasound in such a setting as only the transabdominal route was used. There is no longitudinal data on the potential surrogate of PCOM, namely increased serum AMH concentrations, which may prove easier to use in such populations. Lastly, little is known about the impact of a significant increase in weight in a previously normal weight woman with PCOM.

In summary, the actual meaning of PCOM in the asymptomatic general population is unknown at present. Because other symptoms of PCOS including menstrual disturbances and cutaneous signs of androgen excess such as acne are very prevalent in the first years after menarche, the presence of PCOM in adolescents must be weighted carefully before making a diagnosis of PCOS based solely on these findings. However, because the possibility exists that isolated PCOM may be a forerunner of ovarian dysfunction in some cases, clinical and possibly ultrasonographic and/or AMH follow-up of these women may be a reasonable approach. Finally, considering the major impact that age exerts on ovarian follicle count and OV, establishing appropriate age-specific cut-off values for these variables is definitely needed in order to improve the clinical usefulness of ultrasound evaluation of PCOM and/or serum AMH determination as criteria for the diagnosis of PCOS.

PCOM in particular situations

Women undergoing treatment with oral contraceptive pills

A small controlled study showed that oral contraceptive pill (OCP) administration may reduce ovarian size and follicle counts in women with PCOS and controls (Somunkiran et al., 2007), whereas serum AMH concentrations, being independent from gonadotrophin influence, do not change during such treatment and might be considered a new PCOS marker in patients who are already on contraceptive treatment (Somunkiran et al., 2007). However, a larger study showed that OCP treatment resulted in no differences in FNPO and OV among women with PCOS, even though serum androgens levels were reduced in the patients who were receiving OCPs and still met criteria for PCOS (Mulders et al., 2005). Giving these conflicting results, ultrasound assessment of PCOM should be conducted preferably in women who have not received recent treatment with OCPs.

Women with functional hypothalamic anovulation and hyperprolactinaemia

Intriguingly, PCOM has been observed in as many as 30–50% of patients with functional hypothalamic anovulation (FHA), a situation where the serum LH and insulin levels are low (reviewed in Sum and Warren, 2009). However, the definition of PCOM was heterogeneous between these different studies and the populations were small in some of them. These data have been recently revisited using cluster analysis (Robin et al., 2012). It turned out that the presence of PCOM in women with FHA could have different meanings. More often, it would replicate what is observed in the general population with a similar prevalence of ~30%. In a minority (10%) of cases however, the presence of PCOM could correspond to women with a pre-existing PCOS that has been 'switched-off' by LH and insulin suppression and that could therefore convert back to PCOS subsequently, as previously suggested by others (Wang and Lobo, 2008).

A high prevalence of 'polycystic' or 'multifollicular' ovaries has been reported in non-PCOS women diagnosed with hyperprolactinaemia (Ardaens et al., 1991). However, the literature is scarce about this issue and has not been revisited with the new criteria for PCOM. A possible explanation is a fortuitous association between hyperprolactinaemia and the 'common' PCOM of the general population. This is different from the association between genuine PCOS and hyperprolactinaemia, where symptoms of PCOS may be 'switched-off' by LH suppression due to hyperprolactinaemia and resume once prolactin levels are corrected.

In these situations, there is a risk of misdiagnosing PCOS if the true origin of amenorrhoea or oligomenorrhoea were not recognized, as these patients exhibit two items of the Rotterdam definition for PCOS. It is important to remember that this definition, as well as those of the NIH and AE-PCOS cited above, cannot be applied if other specific causes of ovulatory dysfunction and/or hyperandrogenism such as hyperprolactinaemia, nonclassic congenital adrenal hyperplasia or androgen-secreting tumours have not been previously ruled out.

Polycystic Ovaries as an Indicator of Ill Health

It is important to establish whether PCOM on its own has any health consequences independent of the other features commonly found in PCOS. The literature is lacking on categorical answers to this question and therefore less direct approaches are needed.

Is there any evidence in healthy women that PCOM confers risk of ill health?

The only incontrovertible circumstance in which PCOM has been shown to be potentially dangerous is when a normal woman with PCOM requires gonadotrophin therapy, for example, as part of *in vitro* fertilization treatment (Jayaprakasan *et al.*, 2012). Under these iatrogenic circumstances, the ovary responds to FSH extremely vigorously and there is a substantial risk of ovarian hyperstimulation syndrome (OHSS) and its attendant side effects. This indicates an extreme sensitivity of the ovary to FSH.

In a clinical study, Hassan and Killick (2003) were unable to show any impact of PCOM on time to pregnancy in women with no other symptoms of PCOS. There have been relatively few studies comparing IVF outcomes in patients with PCOM compared with PCOS. Those who have done case control studies indicate pregnancy rates similar to those in women with normal ovaries or PCOS but an OHSS occurrence similar to that of patients with PCOS (Swanton *et al.*, 2010). There is therefore little evidence to suggest the sole presence of PCOM has any significant risk to subsequent health when other symptoms are absent. While the hormonal profile may be intermediate between normal and PCOS, there is lack of evidence for significant metabolic consequences.

One of the problems in finding PCOM without PCOS is that in some countries, health funds penalize individuals who might be construed to have PCOS on the basis of solitary PCOM. This leads clinicians in these jurisdictions not to do ultrasounds or to avoid using the Rotterdam criteria. Solving this problem is not a medical issue but rather a matter of educating authorities that set the political and social agenda.

Does PCOM in PCOS confer any extra health risk?

A number of studies suggest that the presence of PCOM in the diagnosis of PCOS does not seem to affect clinical or metabolic outcomes when weight and body mass index are taken into consideration (Moran and Teede, 2009). The majority of women with PCOS also have PCOM and a comparison of those with and without PCOM has generally shown similar glucose tolerance, insulin resistance and lipid profiles. One report suggests that PCOM may increase insulin resistance, but this was a very small study (Najmabadi *et al.*, 1997). Cardiovascular risk does not appear to be accentuated by PCOM in PCOS groups (Loucks *et al.*, 2000; Guastella *et al.*, 2010) and, in patients with hyperandrogenism and anovulation, having PCOM was not associated with higher insulin or lipid levels (Guastella *et al.*, 2010). In addition, in normoandrogenic anovulatory patients with PCOM, insulin levels and insulin sensitivity were normal (Guastella *et al.*, 2010).

The majority of studies report less adverse metabolic problems for women with ovulatory PCOS in whom hyperandrogenism and PCOM are combined. However, these women are generally less obese and, when weight is taken into consideration, there appear to be few metabolic differences between ovulatory and non-ovulatory women with PCOM. When non-hyperandrogenaemic women with anovulation and PCOM are compared with controls matched for obesity, there is little evidence for an increased abnormal metabolic profile (Moran and Teede, 2009).

In conclusion, the presence or absence of PCOM does not appear to alter the degree of clinical or metabolic presentation in women with PCOS. PCOM has not been shown to have any significant health consequences in isolation from other symptoms of PCOS. However, there may be some mild biochemical and hormonal features in normal ovulatory non-hyperandrogenic women that might disqualify them as controls for comparative studies, as discussed above and below.

Defining PCOM in 2013: Which Guidelines?

Based on the current concern that PCOM may be over-diagnosed when using the former FNPO ≥ 12 threshold, and until we have newer and more accurate markers for PCOM, the Task Force *recommends* the following guidelines.

Guidelines for clinical practice

On the one hand, assessing PCOM should not be considered mandatory for clinical practice. Assessing PCOM is not really useful if the patient already meets the original 1990 NIH criteria of hyperandrogenism in combination with oligo-anovulation (after excluding specific aetiologies) because PCOM is present in most of these women, as discussed above. Wherein confirmation of PCOM might provide diagnostic confirmation, practitioners should use ultrasound judiciously in light of the negative consequences that a diagnosis of PCOM may have on access to care and insurability of patients in some countries. Nevertheless, clinicians must be aware that anovulation may indicate other ovarian disorders, including larger cysts or tumours, and ovarian ultrasound may prove useful in ruling out such disorders.

On the other hand, in situations of isolated hyperandrogenism or oligo-anovulation, 'mild' PCOS is the most likely aetiology, once other specific diagnoses have been excluded. Establishing the presence of PCOM in ultrasound is needed to confirm the diagnosis of PCOS. Using newer ultrasound technology that affords maximal resolution of ovarian follicles (i.e. transducer frequency ≥ 8 MHz), an FNPO ≥ 25 is diagnostic of PCOM. When such precise ultrasound technology is not available, older ultrasound systems permit the careful measure of OV that may serve as surrogate marker for PCOM, using a ≥ 10 ml threshold. However, this criterion has less sensitivity than FNPO for the diagnosis of PCOM, as discussed above and shown in Table II.

From a pragmatic point of view, it is not strictly useful to make the difference between mild PCOS and either idiopathic hyperandrogenism or idiopathic WHO type 2 anovulation, respectively, provided other diagnoses have been ruled out (see above). Ignoring PCOM and thus PCOS in these situations does not matter very much since therapeutic management will be the same and no specific follow-up is required for those mild cases of PCOS (see above).

PCOM is to be expected in a significant number of normo-ovulatory non-hyperandrogenic women undergoing an infertility work-up (i.e. for tubal or male factor infertility). In these cases the issue is not to define whether or not these women have PCOM, but actually to predict OHSS if ovarian stimulation is considered. No consensual FNPO or AMH predictive thresholds are available at this time (Broer *et al.*, 2011). Therefore, it is the responsibility of each centre to define in-house values beyond which the risk of OHSS is clinically relevant.

Guidelines for research

Research requires an exquisite phenotyping of the populations being studied, and this issue is especially important in control women. This

has been deeply debated by the expert panel during the preparation of this manuscript. All members of the Task Force agreed that, in studies dealing with mild PCOS cases, PCOM must be included as a criterion for PCOS, according to well-defined thresholds for FNPO and OV.

However, there was some disagreement regarding how these thresholds are established, primarily as to whether or not PCOM should be excluded *a priori* from the control populations used to establish such thresholds.

One of the experts argued that PCOM has been shown to be associated with subtle 'PCO-like' abnormalities, mainly, an increased serum AMH level (reviewed in Robin et al., 2012), that might pollute control groups. Therefore, if one requires exclusion in case of hyperandrogenism or oligo-anovulation, it seemed logical to this expert to require also exclusion of controls presenting solely with the third item of the Rotterdam classification (i.e. PCOM) inasmuch as excessive FNPO has been shown to be an index of hyperandrogenism (Dewailly et al., 2010).

Other experts argued that it is not possible to exclude PCOM from controls without knowing the threshold values to make this exclusion, since these threshold values are exactly what we seek from such control population. Some studies have proceeded to such exclusion by using the former FNPO ≥12 threshold (Eilertsen et al., 2012; Homburg et al., 2013). This may have biased the results if newer technologies were used since presumably many controls without PCOM were excluded. Despite not being optimal, using OV as a surrogate for FNPO might be acceptable since this variable does not appear to be affected by the age of the ultrasound equipment (Casadei et al., 2013). There is also the possibility of recognizing the subgroup of normal women with PCOM without using any threshold by performing cluster analysis (Dewailly et al., 2011). However, the adequacy of this approach has to be confirmed by other groups.

The majority of the experts decided that the control populations should be constituted of women with normal menstrual cycles and no evidence of hyperandrogenism, regardless of their ovarian morphology. For the FNPO, it was decided to use the 95th percentiles of such control populations from the most complete recent studies, which included a total of 1127 women aged 18–40 years (Table I). The combined 95th percentile of FNPO of all studies was 23 follicles. Noteworthy, this figure is very close to the FNPO ≥25 threshold showing excellent diagnostic performance as analysed by ROC curve analysis in two recent studies as described above (Dewailly et al., 2011; Lujan et al., 2013).

Therefore, the Task Force recommends using FNPO ≥25 as the most appropriate threshold for the diagnosis of PCOM, as this threshold is consistent with normative data obtained from the general population and has an excellent diagnostic performance as a surrogate marker of PCOM. To be valid, the FNPO ≥25 threshold requires the use of newer technology that affords maximal resolution of ovarian follicles (i.e. transducer frequency ≥8 MHz), but such equipment are usually available for research. Real-time methods should follow the standardization that has been recently proposed for obtaining a follicle count (Broekmans et al., 2010), with the main difference being to continue to consider follicle counts per individual ovary and not the sum of both ovaries. Offline methods, with either 2D or 3D ultrasound, must be applied after completion of a learning curve. With either technique, observer variability must be evaluated and incorporated into the appropriate sections of manuscripts.

Table IV Main recommendations.

(1) The threshold for FNPO defining PCOM should be ≥25 follicles per whole ovary.

   (a) This threshold applies to use of newer imaging technology (essentially transducer frequency ≥8 MHz),

   (b) FNPO is recommended over OV since FNPO has been shown to have greater predictive power for PCOS and less variability among populations aged 18–35 years

   (c) Real-time methods should follow recently proposed standardization. Offline methods, with either 2D or 3D ultrasound, must be applied after completion of a learning curve and standardization.

(2) The threshold for OV should remain at ≥10 ml.

   OV may have a role in instances when image quality does not allow for reliable estimates of FNPO.

(3) The use of the AMH assay as a surrogate to ultrasound is for research purpose only at the present time. Only in-house AMH thresholds for PCOM can be used until there is standardization of the assay techniques.

FNPO, follicle number pre ovary; OV, ovarian volume.

For OV, the Task Force recommends using the former threshold at 10 ml since this is very close to the 95th percentile of the 1021 control women described in Table I, although there was more spread with OV than with FNPO. An OV ≥10 ml threshold should be used with either 2D or 3D ultrasound when normative in-house values are not available. Possibly, a higher threshold should be used in adolescents (Carmina et al., 2010) and lower values should be applied to women >40 years old (Carmina et al., 2012) and some European and East-Asian countries (Jonard et al., 2005; Chen et al., 2008; Kosus et al., 2011). Hence, the Task Force recommends that normative data in populations, according to the ethnicity, age (adolescent, adult, adult aging women) and the body weight (lean or obese women), be established. It must be emphasized however, that OV has been shown consistently to have less diagnostic potential than FNPO to discriminate between PCOS and controls (Table II).

Conclusions

Our main recommendations are summarized in Table IV. We recognize that new technology is currently modifying previous criteria for PCOM and that the former threshold established for FNPO at the Rotterdam conference is no longer valid. Using an FNPO ≥12 threshold leads to the over diagnosis of both PCOM and PCOS, especially if FNPO is determined with newer ultrasound technology. On the contrary, the OV ≥10 ml threshold remains valid, although this value might be lower or higher in some specific populations (see above), and has a lower sensitivity compared with FNPO.

Based on recent studies, the Task Force recommends setting the new threshold for FNPO at ≥25 follicles per ovary and that the threshold for OV remains at ≥10 ml. However, there is need for standardization of follicle counting techniques and for further validation of the threshold defining follicle excess. Moreover, consideration of the relevance of these thresholds in diverse populations in terms of age, race and ethnicity is warranted.

An increased serum AMH concentration generates great expectations as a surrogate marker of PCOM, but we must still wait for the full standardization of currently available assays. At the present time, the AMH assay is sold for research purposes only, and only in-house AMH thresholds for PCOM can be used, provided that each group follows stringent inclusion and exclusion criteria in the selection of the control populations used to establish such thresholds.

Authors' roles

All authors were responsible for the literature search and writing of the section they were allocated to. D.D. was responsible for the overall logistical aspects of this study. M.E.L. was responsible for the record of longitudinal data, building of Fig. 2 and statistics about FNPO. All authors shared their data and read and approved the final paper. H.F.E.-M. helped in the final editing of the manuscript.

Funding

The authors of this study did not receive any funding from any source.

Conflict of interest

None declared.

References

Adali E, Kolusari A, Adali F, Yildizhan R, Kurdoglu M, Sahin HG. Doppler analysis of uterine perfusion and ovarian stromal blood flow in polycystic ovary syndrome. *Int J Gynaecol Obstet* 2009;**105**:154–157.

Adams J, Franks S, Polson DW, Mason HD, Abdulwahid N, Tucker M, Price J, Jacobs HS. Multifollicular ovaries: clinical and endocrine features and response to pulsatile gonadotropin releasing hormone. *Lancet* 1985;**2**:8469–8470.

Adams J, Polson DW, Franks S. Prevalence of polycystic ovaries in women with anovulation and idiopathic hirsutism. *Br Med J (Clin Res Ed)* 1986;**293**:355–359.

Ajossa S, Guerriero S, Paoletti AM, Orru M, Melis GB. Hyperinsulinemia and uterine perfusion in patients with polycystic ovary syndrome. *Ultrasound Obstet Gynecol* 2002;**20**:276–280.

Allemand MC, Tummon IS, Phy JL, Foong SC, Dumesic DA, Session DR. Diagnosis of polycystic ovaries by three-dimensional transvaginal ultrasound. *Fertil Steril* 2006;**85**:214–219.

Alsamarai S, Adams JM, Murphy MK, Post MD, Hayden DL, Hall JE, Welt CK. Criteria for polycystic ovarian morphology in polycystic ovary syndrome as a function of age. *J Clin Endocrinol Metab* 2009;**94**:4961–4970.

Amer SA, Li TC, Bygrave C, Sprigg A, Saravelos H, Cooke ID. An evaluation of the inter-observer and intra-observer variability of the ultrasound diagnosis of polycystic ovaries. *Hum Reprod* 2002;**17**:1616–1622.

Ardaens Y, Robert Y, Lemaitre L, Fossati P, Dewailly D. Polycystic ovarian disease—contribution of vaginal endosonography and reassessment of ultrasonic diagnosis. *Fertil Steril* 1991;**55**:1062–1068.

Atiomo WU, Pearson S, Shaw S, Prentice A, Dubbins P. Ultrasound criteria in the diagnosis of polycystic ovary syndrome (PCOS). *Ultrasound Med Biol* 2000;**26**:977–980.

Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D et al. Grading quality of evidence and strength of recommendations. *BMJ* 2004;**328**:1490.

Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, Janssen OE, Legro RS, Norman RJ, Taylor AE et al. Position statement: criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: an Androgen Excess Society guideline. *J Clin Endocrinol Metab* 2006;**91**:4237–4245.

Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, Janssen OE, Legro RS, Norman RJ, Taylor AE et al. The Androgen Excess and PCOS Society criteria for the polycystic ovary syndrome: the complete task force report. *Fertil Steril* 2009;**91**:456–488.

Baba T, Endo T, Honnma H, Kitajima Y, Hayashi T, Ikeda H, Masumori N, Kamiya H, Moriwaka O, Saito T. Association between polycystic ovary syndrome and female-to-male transsexuality. *Hum Reprod* 2007;**22**:1011–1016.

Balen AH, Laven JS, Tan SL, Dewailly D. Ultrasound assessment of the polycystic ovary: international consensus definitions. *Hum Reprod Update* 2003;**9**:505–514.

Battaglia C, Artini PG, Dambrogio G, Genazzani AD, Genazzani AR. The role of color Doppler imaging in the diagnosis of polycystic-ovary-syndrome. *Am J Obstet Gynecol* 1995;**172**:108–113.

Battaglia C, Battaglia B, Morotti E, Paradisi R, Zanetti I, Meriggiola MC, Venturoli S. Two- and three-dimensional sonographic and color Doppler techniques for diagnosis of polycystic ovary syndrome. The stromal/ovarian volume ratio as a new diagnostic criterion. *J Ultrasound Med* 2012;**31**:1015–1024.

Bentzen JG, Forman JL, Johannsen TH, Pinborg A, Larsen EC, Andersen AN. Ovarian antral follicle subclasses and anti-Müllerian hormone during normal reproductive aging. *J Clin Endocrinol Metab* 2013;**98**:1602–1611.

Bleil ME, Adler NE, Pasch LA, Sternfeld B, Gregorich SE, Rosen MP, Cedars MI. Psychological stress and reproductive aging among pre-menopausal women. *Hum Reprod* 2012;**27**:2720–2728.

Broekmans FJ, de Ziegler D, Howles CM, Gougeon A, Trew G, Olivennes F. The antral follicle count: practical recommendations for better standardization. *Fertil Steril* 2010;**94**:1044–1051.

Broer SL, Dolleman M, Opmeer BC, Fauser BC, Mol BW, Broekmans FJM. AMH and AFC as predictors of excessive response in controlled ovarian hyperstimulation: a meta-analysis. *Hum Reprod Update* 2011;**17**:46–54.

Brown ZA, Louwers YV, Fong SL, Valkenburg O, Birnie E, de Jong FH, Fauser BCJM, Laven JSE. The phenotype of polycystic ovary syndrome ameliorates with aging. *Fertil Steril* 2011;**96**:1259–1265.

Carmina E, Orio F, Palomba S, Longo RA, Lombardi G, Lobo RA. Ovarian size and blood flow in women with polycystic ovary syndrome and their correlations with endocrine parameters. *Fertil Steril* 2005;**84**:413–419.

Carmina E, Oberfield SE, Lobo RA. The diagnosis of polycystic ovary syndrome in adolescents. *Am J Obstet Gynecol* 2010;**203**:201.

Carmina E, Campagna AM, Lobo RA. A 20-year follow-up of young women with polycystic ovary syndrome. *Obstet Gynecol* 2012;**119**:263–269.

Casadei L, Madrigale A, Puca F, Manicuti C, Emidi E, Piccione E, Dewailly D. The role of serum anti-Mullerian hormone (AMH) in the hormonal diagnosis of polycystic ovary syndrome. *Gynecol Endocrinol* 2013;**29**:545–550.

Catteau-Jonard S, Bancquart J, Poncelet E, Lefebvre-Maunoury C, Robin G, Dewailly D. Polycystic ovaries at ultrasound: normal variant or silent polycystic ovary syndrome? *Ultrasound Obstet Gynecol* 2012;**40**:223–229.

Chekir C, Nakatsuka M, Kamada Y, Noguchi S, Sasaki A, Hiramatsu Y. Impaired uterine perfusion associated with metabolic disorders in women with polycystic ovary syndrome. *Acta Obstet Gynecol Scand* 2005;**84**:189–195.

Chen Y, Li L, Chen X, Zhang Q, Wang W, Li Y, Yang D. Ovarian volume and follicle number in the diagnosis of polycystic ovary syndrome in Chinese women. *Ultrasound Obstet Gynecol* 2008;**32**:700–703.

Codner E, Villarroel C, Eyzaguirre FC, Lopez P, Merino PM, Perez-Bravo F, Iniguez G, Cassorla F. Polycystic ovarian morphology in postmenarchal adolescents. *Fertil Steril* 2011;**95**:702–U810.

Cook CL, Siow Y, Brenner AG, Fallat ME. Relationship between serum mullerian-inhibiting substance and other reproductive hormones in untreated women with polycystic ovary syndrome and normal women. *Fertil Steril* 2002;**77**:141–146.

Das M, Shehata F, Son WY, Tulandi T, Holzer H. Ovarian reserve and response to IVF and in vitro maturation treatment following chemotherapy. *Hum Reprod* 2012;**27**:2509–2514.

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**:477–483.

Deb S, Campbell BK, Clewes JS, Raine-Fenning NJ. Quantitative analysis of antral follicle number and size: a comparison of two-dimensional and automated three-dimensional ultrasound techniques. *Ultrasound Obstet Gynecol* 2010;**35**:354–360.

Deb S, Campbell BK, Clewes JS, Pincott-Allen C, Raine-Fenning NJ. Intracycle variation in number of antral follicles stratified by size and in endocrine markers of ovarian reserve in women with normal ovulatory menstrual cycles. *Ultrasound Obstet Gynecol* 2013;**41**:216–222.

Dewailly D, Pigny P, Soudan B, Catteau-Jonard S, Decanter C, Poncelet E, Duhamel A. Reconciling the definitions of polycystic ovary syndrome: the ovarian follicle number and serum anti-Mullerian hormone concentrations aggregate with the markers of hyperandrogenism. *J Clin Endocrinol Metab* 2010;**95**:4399–4405.

Dewailly D, Gronier H, Poncelet E, Robin G, Leroy M, Pigny P, Duhamel A, Catteau-Jonard S. 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**:3123–3129.

Diamanti-Kandarakis E, Panidis D. Unravelling the phenotypic map of polycystic ovary syndrome (PCOS): a prospective study of 634 women with PCOS. *Clin Endocrinol (Oxf)* 2007;**67**:735–742.

Duijkers IJM, Klipping C. Polycystic ovaries, as defined by the 2003 Rotterdam consensus criteria, are found to be very common in young healthy women. *Gynecol Endocrinol* 2010;**26**:152–160.

Eilertsen TB, Vanky E, Carlsen SM. Anti-Mullerian hormone in the diagnosis of polycystic ovary syndrome: can morphologic description be replaced? *Hum Reprod* 2012;**27**:2494–2502.

Fox R. Transvaginal ultrasound appearances of the ovary in normal women and hirsute women with oligomenorrhoea. *Aust N Z J Obstet Gynaecol* 1999;**39**:63–68.

Fox R, Hull M. Ultrasound diagnosis of polycystic ovaries. *Ann N Y Acad Sci* 1993; **687**:217–223.

Fulghesu AM, Angioni S, Frau E, Belosi C, Apa R, Mioni R, Xamin N, Capobianco GP, Dessole S, Fruzzetti F et al. Ultrasound in polycystic ovary syndrome—the measuring of ovarian stroma and relationship with circulating androgens: results of a multicentric study. *Hum Reprod* 2007;**22**:2501–2508.

Garel L, Dubois J, Grignon A, Filiatrault D, Van Vliet G. US of the pediatric female pelvis: a clinical perspective. *Radiographics* 2001;**21**:1393–1407.

Govind A, Obhrai MS, Clayton RN. Polycystic ovaries are inherited as an autosomal dominant trait: Analysis of 29 polycystic ovary syndrome and 10 control families. *J Clin Endocrinol Metab* 1999;**84**:38–43.

Guastella E, Longo RA, Carmina E. Clinical and endocrine characteristics of the main polycystic ovary syndrome phenotypes. *Fertil Steril* 2010;**94**:2197–2201.

Hart R, Doherty DA, Norman RJ, Franks S, Dickinson JE, Hickey M, Sloboda DM. Serum antimullerian hormone (AMH) levels are elevated in adolescent girls with polycystic ovaries and the polycystic ovarian syndrome (PCOS). *Fertil Steril* 2010;**94**: 1118–1121.

Hassan MAM, Killick SR. Ultrasound diagnosis of polycystic ovaries in women who have no symptoms of polycystic ovary syndrome is not associated with subfecundity or subfertility. *Fertil Steril* 2003;**80**:966–975.

Homburg R. Androgen circle of polycystic ovary syndrome. *Hum Reprod* 2009; **24**:1548–1555.

Homburg R, Ray A, Bhide P, Gudi A, Shah A, Timms P, Grayson K. The relationship of serum anti-Mullerian hormone with polycystic ovarian morphology and polycystic ovary syndrome: a prospective cohort study. *Hum Reprod* 2013;**28**:1077–1083.

Hughesdon PE. Morphology and morphogenesis of the Stein-Leventhal ovary and of so-called 'hyperthecosis'. *Obstet Gynecol Surv* 1982;**37**:59–77.

Iliodromiti S, Kelsey TW, Anderson RA, Nelson SM. Can anti-Mullerian hormone predict the diagnosis of polycystic ovary syndrome? A systematic review and meta-analysis of extracted data. *J Clin Endocrinol Metab* 2013. doi: 10.1210/jc.2013-1393.

Isik S, Ozcan HN, Ozuguz U, Tutuncu YA, Berker D, Alimli AG, Akbaba G, Karademir MA, Guler S. Evaluation of ovarian reserve based on hormonal parameters, ovarian volume, and antral follicle count in women with type 2 diabetes mellitus. *J Clin Endocrinol Metab* 2012;**97**:261–269.

Jarvela IY, Sladkevicius P, Tekay AH, Campbell S, Nargund G. Intraobserver and interobserver variability of ovarian volume, gray-scale and color flow indices obtained using transvaginal three-dimensional power Doppler ultrasonography. *Ultrasound Obstet Gynecol* 2003;**21**:277–282.

Jayaprakasan K, Walker KF, Clewes JS, Johnson IR, Raine-Fenning NJ. The interobserver reliability of off-line antral follicle counts made from stored three-dimensional ultrasound data: a comparative study of different measurement techniques. *Ultrasound Obstet Gynecol* 2007;**29**:335–341.

Jayaprakasan K, Chan Y, Islam R, Haoula Z, Hopkisson J, Coomarasamy A, Raine-Fenning N. Prediction of in vitro fertilization outcome at different antral follicle count thresholds in a prospective cohort of 1,012 women. *Fertil Steril* 2012;**98**:657–663.

Johnson TRB, Kaplan LK, Ouyang P, Rizza RA. *Evidence-Based Methodology Workshop on Polycystic Ovary Syndrome*. Bethesda, Maryland: National Institutes of Health, 2012.

Johnstone EB, Rosen MP, Neril R, Trevithick D, Sternfeld B, Murphy R, Addauan-Andersen C, McConnell D, Reijo Pera R, Cedars MI. The polycystic

ovary post-Rotterdam: a common, age-dependent finding in ovulatory women without metabolic significance. *J Clin Endocrinol Metab* 2010;**95**:4965–4972.

Jokubkiene L, Sladkevicius P, Valentin L. Number of antral follicles, ovarian volume, and vascular indices in asymptomatic women 20 to 39 years old as assessed by 3-dimensional sonography: a prospective cross-sectional study. *J Ultrasound Med* 2012;**31**:1635–1649.

Jonard S, Robert Y, Cortet-Rudelli C, Pigny P, Decanter C, Dewailly D. Ultrasound examination of polycystic ovaries: is it worth counting the follicles? *Hum Reprod* 2003;**18**:598–603.

Jonard S, Robert Y, Dewailly D. Revisiting the ovarian volume as a diagnostic criterion for polycystic ovaries. *Hum Reprod* 2005;**20**:2893–2898.

Keskin U, Duru NK, Ercan CM, Dede M, Yenen MC, Ergun A. Changes in the ovarian stromal blood flow in patients treated with laparoscopic salpingostomy for ectopic pregnancy. *J Obstet Gynaecol* 2013;**33**:399–402.

Kline J, Kinney A, Kelly A, Reuss ML, Levin B. Predictors of antral follicle count during the reproductive years. *Hum Reprod* 2005;**20**:2179–2189.

Knauff EA, Eijkemans MJ, Lambalk CB, ten Kate-Booij MJ, Hoek A, Beerendonk CC, Laven JS, Goverde AJ, Broekmans FJ, Themmen AP et al. Anti-Mullerian hormone, inhibin B, and antral follicle count in young women with ovarian failure. *J Clin Endocrinol Metab* 2009;**94**:786–792.

Kosus N, Kosus A, Turhan NO, Kamalak Z. Do threshold values of ovarian volume and follicle number for diagnosing polycystic ovarian syndrome in Turkish women differ from western countries? *Eur J Obstet Gynecol Reprod Biol* 2011; **154**:177–181.

Kristensen SL, Ramlau-Hansen CH, Ernst E, Olsen SF, Bonde JP, Vested A, Toft G. A very large proportion of young Danish women have polycystic ovaries: is a revision of the Rotterdam criteria needed? *Hum Reprod* 2010;**25**:3117–3122.

Kristensen SL, Ramlau-Hansen CH, Andersen CY, Ernst E, Olsen SF, Bonde JP, Vested A, Toft G. The association between circulating levels of antimüllerian hormone and follicle number, androgens, and menstrual cycle characteristics in young women. *Fertil Steril* 2012;**97**:779–785.

Kumarapeli V, Seneviratne RD, Wijeyaratne CN, Yapa RMSC, Dodampahala SH. A simple screening approach for assessing community prevalence and phenotype of polycystic ovary syndrome in a semiurban population in Sri Lanka. *Am J Epidemiol* 2008;**168**:321–328.

Lam PM, Johnson IR, Raine-Fenning NJ. Three-dimensional ultrasound features of the polycystic ovary and the effect of different phenotypic expressions on these parameters. *Hum Reprod* 2007;**22**:3116–3123.

Lam P, Raine-Fenning N, Cheung L, Haines C. Three-dimensional ultrasound features of the polycystic ovary in Chinese women. *Ultrasound Obstet Gynecol* 2009;**34**:196–200.

La Marca A, Spada E, Sighinolfi G, Argento C, Tirelli A, Giulini S, Milani S, Volpe A. Age-specific nomogram for the decline in antral follicle count throughout the reproductive period. *Fertil Steril* 2011;**95**:684–688.

Lamazou F, Arbo E, Salama S, Grynberg M, Frydman R, Fanchin R. Reliability of automated volumetric measurement of multiple growing follicles in controlled ovarian hyperstimulation. *Fertil Steril* 2010;**94**:2172–2176.

Laven JSE, Mulders AGMGJ, Visser JA, Themmen AP, De Jong FH, Fauser BCJM. Anti-Mullerian hormone serum concentrations in normoovulatory and anovulatory women of reproductive age. *J Clin Endocrinol Metab* 2004;**89**: 318–323.

Legro RS, Brzyski RG, Diamond MP, Coutifaris C, Schlaff WD, Alvero R, Casson P, Christman GM, Huang H, Yan Q et al.; for the National Institute of Child Health and Human Development Reproductive Medicine Network. The Pregnancy in Polycystic Ovary Syndrome II study: baseline characteristics and effects of obesity from a multicenter randomized clinical trial. *Fertil Steril* 2013 Oct 21. (Epub ahead of print)

Lin YH, Chiu WC, Wu CH, Tzeng CR, Hsu CS, Hsu MI. Antimullerian hormone and polycystic ovary syndrome. *Fertil Steril* 2011;**96**:230–235.

Loucks TL, Talbott EO, McHugh KP, Keelan M, Berga SL, Guzick DS. Do polycystic-appearing ovaries affect the risk of cardiovascular disease among women with polycystic ovary syndrome? *Fertil Steril* 2000;**74**:547–552.

Loverro G, Vicino M, Lorusso F, Vimercati A, Greco P, Selvaggi L. Polycystic ovary syndrome: relationship between insulin sensitivity, sex hormone levels and ovarian stromal blood flow. *Gynecol Endocrinol* 2001;**15**:142–149.

Lujan ME, Chizen DR, Peppin AK, Kriegler S, Leswick DA, Blotski TG, Pierson RA. Improving inter-observer variability in the evaluation of ultrasonographic features of polycystic ovaries. *Reprod Biol Endocrinol* 2008;**6**:30–41.

Lujan ME, Chizen DR, Peppin AK, Dhir A, Pierson RA. Assessment of ultrasonographic features of polycystic ovaries is associated with modest levels of inter-observer agreement. *J Ovarian Res* 2009;**2**:6.

Lujan ME, Brooks ED, Kepley AL, Chizen DR, Pierson RA, Peppin AK. Grid analysis improves reliability in follicle counts made by ultrasonography in women with polycystic ovary syndrome. *Ultrasound Med Biol* 2010;**36**:712–718.

Lujan ME, Jarrett BY, Brooks ED, Reines JK, Peppin AK, Muhn N, Haider E, Pierson RA, Chizen DR. Updated ultrasound criteria for polycystic ovary syndrome: reliable thresholds for elevated follicle population and ovarian volume. *Hum Reprod* 2013; **28**:1361–1368.

Merce LT, Gomez B, Engels V, Bau S, Bajo JM. Intraobserver and interobserver reproducibility of ovarian volume, antral follicle count, and vascularity indices obtained with transvaginal 3-dimensional ultrasonography, power Doppler angiography, and the virtual organ computer-aided analysis imaging program. *J Ultrasound Med* 2005;**24**:1279–1287.

Moran L, Teede H. Metabolic features of the reproductive phenotypes of polycystic ovary syndrome. *Hum Reprod Update* 2009;**15**:477–488.

Mortensen M, Ehrmann DA, Littlejohn E, Rosenfeld RL. Asymptomatic volunteers with a polycystic ovary are a functionally distinct but heterogeneous population. *J Clin Endocrinol Metab* 2009;**94**:1579–1586.

Mulders AG, ten Kate-Booij M, Pal R, De Kruif M, Nekrui L, Oostra BA, Fauser BC, Laven JS. Influence of oral contraceptive pills on phenotype expression in women with polycystic ovary syndrome. *Reprod Biomed Online* 2005;**11**: 690–696.

Murphy MK, Hall JE, Adams JM, Lee H, Welt CK. Polycystic ovarian morphology in normal women does not predict the development of polycystic ovary syndrome. *J Clin Endocrinol Metab* 2006;**91**:3878–3884.

Najmabadi S, Wilcox JG, Acacio BD, Thornton MH, Kolb BA, Paulson RJ. The significance of polycystic-appearing ovaries versus normal-appearing ovaries in patients with polycystic ovary syndrome. *Fertil Steril* 1997;**67**:631–635.

Nelson SM, La Marca A. The journey from the old to the new AMH assay: how to avoid getting lost in the values. *Reprod Biomed Online* 2011;**23**:411–420.

Ng EHY, Chan CCW, Yeung WSB, Ho PC. Comparison of ovarian stromal blood flow between fertile women with normal ovaries and infertile women with polycystic ovary syndrome. *Hum Reprod* 2005;**20**:1881–1886.

Ng EHY, Chan CCW, Ho PC. Are there differences in ultrasound parameters between Chinese women with polycystic ovaries only and with polycystic ovary syndrome? *Eur J Obstet Gynecol Reprod Biol* 2006;**125**:92–98.

Pache TD, Wladimiroff JW, Hop WC, Fauser BC. How to discriminate between normal and polycystic ovaries: transvaginal US study. *Radiology* 1992;**183**: 421–423.

Pascual MA, Graupera B, Hereter L, Tresserra F, Rodriguez I, Alcazar JL. Assessment of ovarian vascularization in the polycystic ovary by three-dimensional power Doppler ultrasonography. *Gynecol Endocrinol* 2008;**24**:631–636.

Pavlik EJ, DePriest PD, Gallion HH, Ueland FR, Reedy MB, Kryscio RJ, van Nagell JR. Ovarian volume related to age. *Gynecol Oncol* 2000;**77**:410–412.

Pellatt L, Rice S, Mason HD. Anti-Mullerian hormone and polycystic ovary syndrome: a mountain too high? *Reproduction* 2010;**139**:825–833.

Pigny P, Merlen E, Robert Y, Cortet-Rudelli C, Decanter C, Jonard S, Dewailly D. Elevated serum level of anti-mullerian hormone in patients with polycystic ovary syndrome: relationship to the ovarian follicle excess and to the follicular arrest. *J Clin Endocrinol Metab* 2003;**88**:5957–5962.

Pigny P, Jonard S, Robert Y, Dewailly D. Serum anti-Mullerian hormone as a surrogate for antral follicle count for definition of the polycystic ovary syndrome. *J Clin Endocrinol Metab* 2006;**91**:941–945.

Piltonen T, Morin-Papunen L, Koivunen R, Perheentupa A, Ruokonen A, Tapanainen JS. Serum anti-Mullerian hormone levels remain high until late reproductive age and decrease during metformin therapy in women with polycystic ovary syndrome. *Hum Reprod* 2005;**20**:1820–1826.

Raine-Fenning NJ, Campbell BK, Clewes JS, Johnson IR. The interobserver reliability of ovarian volume measurement is improved with three-dimensional ultrasound, but dependent upon technique. *Ultrasound Med Biol* 2003;**29**:1685–1690.

Raine-Fenning NJ, Campbell BK, Clewes JS, Kendall NR, Johnson IR. The interobserver reliability of three-dimensional power Doppler data acquisition within the female pelvis. *Ultrasound Obstet Gynecol* 2004;**23**:501–508.

Raine-Fenning N, Jayaprakasan K, Clewes J, Joergner I, Bonaki SD, Chamberlain S, Devlin L, Priddle H, Johnson I. SonoAVC: a novel method of automatic volume calculation. *Ultrasound Obstet Gynecol* 2008;**31**:691–696.

Robin G, Gallo C, Catteau-Jonard S, Lefebvre-Maunoury C, Pigny P, Duhamel A, Dewailly D. Polycystic ovary-like abnormalities (PCO-L) in women with functional hypothalamic amenorrhea. *J Clin Endocrinol Metab* 2012;**97**:4236–4243.

Rosen MP, Sternfeld B, Schuh-Huerta SM, Reijo Pera RA, McCulloch CE, Cedars MI. Antral follicle count: absence of significant midlife decline. *Fertil Steril* 2010; **94**:2182–2185.

Rustamov O, Smith A, Roberts SA, Yates AP, Fitzgerald C, Krishnan M, Nardo LG, Pemberton PW. Anti-Mullerian hormone: poor assay reproducibility in a large cohort of subjects suggests sample instability. *Hum Reprod* 2012;**27**:3085–3091.

Salama S, Arbo E, Lamazou F, Levaillant JM, Frydman R, Fanchin R. Reproducibility and reliability of automated volumetric measurement of single preovulatory follicles using SonoAVC. *Fertil Steril* 2010;**93**:2069–2073.

Saxton DW, Farquhar CM, Rae T, Beard RW, Anderson MC, Wadsworth J. Accuracy of ultrasound measurements of female pelvic organs. *Br J Obstet Gynaecol* 1990; **97**:695–699.

Scheffer GJ, Broekmans FJM, Bancsi LF, Habbema JDF, Te Velde ER. Quantitative transvaginal two- and three-dimensional sonography of the ovaries: reproducibility of antral follicle counts. *Ultrasound Obstet Gynecol* 2002;**20**:270–275.

Somunkiran A, Yavuz T, Yucel O, Ozdemir I. Anti-Mullerian hormone levels during hormonal contraception in women with polycystic ovary syndrome. *Eur J Obstet Gynecol Reprod Biol* 2007;**134**:196–201.

Stein IF, Leventhal ML. Amenorrhea associated with bilateral polycystic ovaries. *Am J Obstet Gynecol* 1935;**29**:181–191.

Steinberg HO, Brechtel G, Johnson A, Fineberg N, Baron AD. Insulin-mediated skeletal-muscle vasodilation is nitric-oxide dependent—a novel action of insulin to increase nitric-oxide release. *J Clin Invest* 1994;**94**:1172–1179.

Sum M, Warren MP. Hypothalamic amenorrhea in young women with underlying polycystic ovary syndrome. *Fertil Steril* 2009;**92**:2106–2108.

Sun L, Fu QZ. Three-dimensional transrectal ultrasonography in adolescent patients with polycystic ovarian syndrome. *Int J Gynaecol Obstet* 2007;**98**:34–38.

Swanton A, Storey L, McVeigh E, Child T. IVF outcome in women with PCOS, PCO and normal ovarian morphology. *Eur J Obstet Gynecol Reprod Biol* 2010;**149**:68–71.

Swiglo BA, Murad MH, Schunemann HJ, Kunz R, Vigersky RA, Guyatt GH, Montori VM. A case for clarity, consistency, and helpfulness: state-of-the-art clinical practice guidelines in endocrinology using the grading of recommendations, assessment, development, and evaluation system. *J Clin Endocrinol Metab* 2008; **93**:666–673.

Takahashi K, Ozaki T, Okada M, Uchida A, Kitao M. Relationship between ultrasonography and histopathological changes in polycystic ovarian syndrome. *Hum Reprod* 1994;**9**:2255–2258.

The 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* 2004a;**19**:41–47.

The 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* 2004b;**81**:19–25.

van Rooij IA, Tonkelaar I, Broekmans FJ, Looman CW, Scheffer GJ, de Jong FH, Themmen AP, te Velde ER. Anti-mullerian hormone is a promising predictor for the occurrence of the menopausal transition. *Menopause* 2004;**11**:601–606.

Vendola K, Zhou J, Wang J, Bondy CA. Androgens promote insulin-like growth factor-I and insulin-like growth factor-I receptor gene expression in the primate ovary. *Hum Reprod* 1999;**14**:2328–2332.

Villarroel C, Merino PM, Lopez P, Eyzaguirre FC, Van Velzen A, Iniguez G, Codner E. Polycystic ovarian morphology in adolescents with regular menstrual cycles is associated with elevated anti-Mullerian hormone. *Hum Reprod* 2011;**26**: 2861–2868.

Wang JG, Lobo RA. The complex relationship between hypothalamic amenorrhea and polycystic ovary syndrome. *J Clin Endocrinol Metab* 2008;**93**:1394–1397.

Weenen C, Laven JS, Von Bergh AR, Cranfield M, Groome NP, Visser JA, Kramer P, Fauser BC, Themmen AP. Anti-Mullerian hormone expression pattern in the human ovary: potential implications for initial and cyclic follicle recruitment. *Mol Hum Reprod* 2004;**10**:77–83.

Weerakiet S, Lertvikool S, Tingthanatikul Y, Wansumrith S, Leelaphiwat S, Jultanmas R. Ovarian reserve in women with polycystic ovary syndrome who underwent laparoscopic ovarian drilling. *Gynecol Endocrinol* 2007;**23**:455–460.

Well D, Yang H, Houseni M, Iruvuri S, Alzeair S, Sansovini M, Wintering N, Alavi A, Torigian DA. Age-related structural and metabolic changes in the pelvic reproductive end organs. *Semin Nucl Med* 2007;**37**:173–184.

Welt CK, Arason G, Gudmundsson JA, Adams J, Palsdottir H, Gudlaugsdottir G, Ingadottir G, Crowley WF. Defining constant versus variable phenotypic features of women with polycystic ovary syndrome using different ethnic groups and populations. *J Clin Endocrinol Metab* 2006;**91**:4361–4368.

Woo HY, Kim KH, Rhee EJ, Park H, Lee MK. Differences of the association of anti-Mullerian hormone with clinical or biochemical characteristics between women with and without polycystic ovary syndrome. *Endocr J* 2012;**59**:781–790.

Yeh HC, Futterweit W, Thornton JC. Polycystic ovarian disease: US features in 104 patients. *Radiology* 1987;**163**:111–116.

Younis JS, Jadaon JE, Haddad S, Izhaki I, Ben-Ami M. Prospective evaluation of basal stromal Doppler studies in women with good ovarian reserve and infertility undergoing in vitro fertilization-embryo transfer treatment: patients with polycystic ovary syndrome versus ovulatory patients. *Fertil Steril* 2011; **95**:1754–1758.

Zaidi J, Campbell S, Pittrof R, Kyeimensah A, Shaker A, Jacobs HS, Tan SL. Ovarian stromal blood-flow in women with polycystic ovaries—a possible new marker for diagnosis. *Hum Reprod* 1995;**10**:1992–1996.

Zawadzki JA, Dunaif A. Diagnostic criteria for polycystic ovary syndrome: towards a rational approach. In: Dunaif A, Givens JR, Haseltine FP, Merriam GR (eds). *Polycystic Ovary Syndrome*. Boston: Blackwell Scientific, 1992, 377–384.

Zhang H, Guo C, Zhu F, Qu P, Lin W, Xiong J. Clinical characteristics, metabolic features, and phenotype of Chinese women with polycystic ovary syndrome: a large-scale case–control study. *Arch Gynecol Obstet* 2013;**287**: 525–531.