Human Reproduction, Vol.26, No.11 pp. 3123–3129, 2011
Advanced Access publication on September 16, 2011 doi:10.1093/humrep/der297

human reproduction

ORIGINAL ARTICLE *Reproductive endocrinology*

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

D. Dewailly[1,*], H. Gronier[1], E. Poncelet[2], G. Robin[1], M. Leroy[1], P. Pigny[3], A. Duhamel[4], and S. Catteau-Jonard[1]

[1]Service de Gynécologie Endocrinienne et de Médecine de la Reproduction, Hôpital Jeanne de Flandre, C.H.R.U., 59037 Lille, France [2]Service de Radiologie, Hôpital Jeanne de Flandre, Lille, France [3]Laboratoire de Biochimie et Hormonologie, Lille, France [4]Unité de Biostatistiques, Université Lille Nord de France, Lille EA2694, France

*Correspondence address. Tel: +33-3-20-44-63-09; Fax: +33-3-20-44-64-07; E-mail: didier.dewailly@chru-lille.fr

Submitted on February 4, 2011; resubmitted on April 29, 2011; accepted on May 9, 2011

BACKGROUND: Polycystic ovarian morphology (PCOM) at ultrasound is currently used in the diagnosis of polycystic ovary syndrome (PCOS). We hypothesized that the previously proposed threshold value of 12 as an excessive number of follicles per ovary (FN) is no longer appropriate because of current technological developments. In this study, we have revisited the thresholds for FN and for the serum Anti-Müllerian hormone (AMH) level (a possible surrogate for FN) for the definition of PCOM.

METHODS: Clinical, hormonal and ultrasound data were consecutively recorded in 240 patients referred to our department between 2008 and 2010 for exploration of hyperandrogenism (HA), menstrual disorders and/or infertility.

RESULTS: According to only their symptoms, patients were grouped as: non-PCOS without HA and with ovulatory cycles (group 1, $n = 105$), presumption of PCOS with only HA or only oligo-anovulation (group 2, $n = 73$) and PCOS with HA and oligo-anovulation (group 3, $n = 62$). By cluster analysis using androgens, LH, FSH, AMH, FN and ovarian volume, group 1 appeared to be constituted of two homogeneous clusters, most likely a non-PCOM non-PCOS subgroup ($n = 66$) and a PCOM, non-PCOS (i.e. asymptomatic) subgroup ($n = 39$). Receiver operating characteristic curve analysis was applied to distinguish the non-PCOM non-PCO members of group 1 and to group 3. For FN and serum AMH respectively, the areas under the curve were 0.949 and 0.973 and the best compromise between sensitivity (81 and 92%) and specificity (92 and 97%) was obtained with a threshold values of 19 follicles and 35 pmol/l (5 ng/ml).

CONCLUSIONS: For the definition of PCOM, the former threshold of >12 for FN is no longer valid. A serum AMH >35 pmol/l (or >5 ng/ml) appears to be more sensitive and specific than a FN >19 and should be therefore included in the current diagnostic classifications for PCOS.

Key words: polycystic ovary syndrome / anti-Müllerian hormone / ovarian follicle / ultrasonography / diagnosis

Introduction

Polycystic ovary syndrome (PCOS) is a common endocrine disorder, affecting up to 10% of women of reproductive age (Norman *et al.*, 2007). Its prevalence varies according to the definition used and to the reference population (Diamanti-Kandarakis *et al.*, 1999; Asuncion *et al.*, 2000; Azziz *et al.*, 2004).

The cardinal features of PCOS are hyperandrogenism (HA) and oligo-anovulation. The metabolic abnormalities often associated with this syndrome (obesity, insulin resistance, hyperinsulinemia and dyslipidemia) are not included in the definition of the syndrome because it is still unclear whether they are intrinsic to the disease or not (Moran and Teede, 2009). The current diagnostic classifications (The Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group, 2004; Azziz *et al.*, 2006) use HA, oligo-anovulation and polycystic ovarian morphology (PCOM) at ultrasound. Whether HA is a necessary criterion remains controversial. By allowing the diagnosis of PCOS with only two items out of the three (HA, oligo-anovulation and

PCOM), the so-called Rotterdam classification includes patients without overt HA (Dewailly et al., 2006), which is still a disputed issue (Azziz, 2006; Franks, 2006).

We recently reported that this is only an apparent controversy, because in fact the presence of PCOM (defined largely by an excess of follicles <10 mm at ovarian ultrasound (U/S) turned out, after principal component analysis, to be itself a sign of HA (Dewailly et al., 2010). The same was observed with a high serum level of Anti-Müllerian hormone (AMH) (Dewailly et al., 2010), a peptide produced by the granulosa cells (GC) of follicles, that is highly correlated to the excess of number of follicles (FN) in patients with PCOS (Pigny et al., 2003, 2006; Laven et al., 2004). We have therefore proposed a simplified classification for the diagnosis of PCOS: oligo-anovulation and HA should first be required. When one of these criteria is not present, FN or AMH could be used as a surrogate for HA or oligo-anovulation (Dewailly et al., 2010).

In practice, this classification can be used only if we have specific thresholds of AMH and FN, beyond which these two parameters can be considered as markers of PCOM. Threshold values have been proposed previously (Balen et al., 2003; Pigny et al., 2006) but their validity is questionable in the view of current technological developments. For instance, the FN threshold proposed in 2003 (Jonard et al., 2003) (i.e. 12 follicles per ovary) leads to a major but artificial increase in the prevalence of PCOM in normal populations (especially in women aged <30 years) when using new U/S equipment (Duijkers and Klipping, 2009; Johnstone et al., 2010; Kristensen et al., 2010). This lead some authors to conclude recently that PCOM has no pathological significance (Johnstone et al., 2010), while others recommended revisiting the threshold for FN (Kristensen et al., 2010).

There is indeed an urgent need to revisit these markers, but setting thresholds to define PCOM is particularly difficult. This is mainly due to the high incidence of asymptomatic women with PCOM on U/S, whose prevalence in the literature varies from ≈10 to ≈30% depending on the equipment and the definitions used (reviewed in Johnstone et al., 2010). In many studies, these women were included in the control groups, limiting the power of statistical procedures, such as receiver operating characteristic (ROC) analysis (Zweig and Campbell, 1993), used to study the sensitivity (Se) and specificity (Spe) of markers of PCOM. In this study, we have tried to circumvent this difficulty by a preliminary step using cluster analysis (Hartigan, 1985). Our hypothesis was that asymptomatic patients with PCOM could constitute a mathematically homogenous group among the control patients. If this were true, it would be possible to isolate and then to exclude those women from the control group. This would allow subsequent definition of new diagnostic thresholds for FN and serum AMH that could be used with acceptable Se and Spe for the detection of PCOM.

Materials and Methods

Patients

Data were obtained from a database including clinical, hormonal and U/S features that were consecutively recorded between 2008 and 2010 from patients referred to our department. These patients were referred for exploration of HA, menstrual disorders and/or infertility due to male factor and/or tubal abnormality. Women with unexplained infertility or endometriosis were excluded. Clinical, hormonal and U/S examinations were performed in the early follicular phase, between Day 2 and 5 of the menstrual cycle. In patients with menstrual disorders, the last menstrual period was either spontaneous or induced by the administration of dydrogesterone (10 mg/day for 7 days). This study was approved by the Institutional Review Board of the University Hospital of Lille. All patients gave their informed consent before inclusion in this study.

Exclusion criteria were the following: age <18 or more than 35 years, suspicion of low ovarian reserve (FSH > 12 IU/l), hyperprolactinemia (serum prolactin >20 ng/ml on two subsequent determinations) or non-classic 21-hydroxylase deficiency [basal 17-hydroxyprogesterone (17-OHP) >5 ng/ml and/or post-adrenocorticotrophic hormone-stimulated value >12 ng/ml]. Ovarian or adrenal tumours were excluded on the basis of serum total testosterone (TT) or dehydroepiandrosterone sulphate (DHEA-S) levels lower than 1.5 ng/ml or 15 μmol/l, respectively. Any patient with criteria for hypothalamic amenorrhea was also excluded. Furthermore, any patient with at least one follicle with a diameter >9 mm at U/S or a serum estradiol (E_2) level above 80 pg/ml was excluded from the study.

Investigations

During the medical examination, patients were specifically asked about their menstrual history. Oligomenorrhoea was defined as an average cycle length of more than 35 days and included women with frank amenorrhea. Clinical HA was defined by the presence of hirsutism (modified Ferriman–Gallwey score over 6) and/or acne located in more than two areas. Hyperandrogenemia was defined as a serum TT level >0.5 ng/ml and/or a serum androstenedione level (A) >2.02 ng/ml, as previously reported (Dewailly et al., 2006).

Prolactin, LH, FSH, E2, OHP, DHEA-S, A and TT levels were measured by immunoassays as described previously (Dewailly et al., 2006). Serum AMH levels were assessed using the second-generation enzyme immunoassay AMH-EIA (ref A16507) provided by Beckman Coulter Immunotech (Villepinte, France), as described previously (Catteau-Jonard et al., 2007).

For every patient, U/S examination was performed with a Voluson E8 Expert (General Electric Systems, VELIZY, France) with a 5–9 MHz transvaginal transducer. U/S measurements were taken in real-time, according to as standardized protocol. The highest possible magnification was used to examine the ovaries. After determination of the longest medial axis of the ovary, the length and thickness were measured and the ovarian volume (OV) was calculated as described previously (Jonard et al., 2003). For each ovary, the total number of all visible follicles smaller than 10 mm in diameter was counted by slow and continuous scanning of the entire ovary, from one margin to the other in longitudinal cross-section. Every operator was asked to count any follicle that can now be detected with the new equipment (Fig. 1) without using any lower cut-off value. For the OV and the FN, the data used for statistical analysis were the mean of recorded values for the left and right ovaries. We excluded from the analysis patients in whom transvaginal ultrasonography was not possible (due to virginity or patient refusal) and those with a history of ovarian surgery.

Statistical methods

All statistical analyses were performed using SPSS and SAS software. Statistical significance between mean values was attributed to two-tailed $P < 0.05$. The results are expressed as median with 5th and 95 percentiles. Comparisons between groups were performed using both Kruskal–Wallis test and a non-parametric analysis of variance (ANOVA) after rank transformation using the methodology suggested by Conover and Iman (1981). The Bonferroni correction was applied for multiple comparisons in *post hoc* tests. Significant relationships between serum AMH, FN

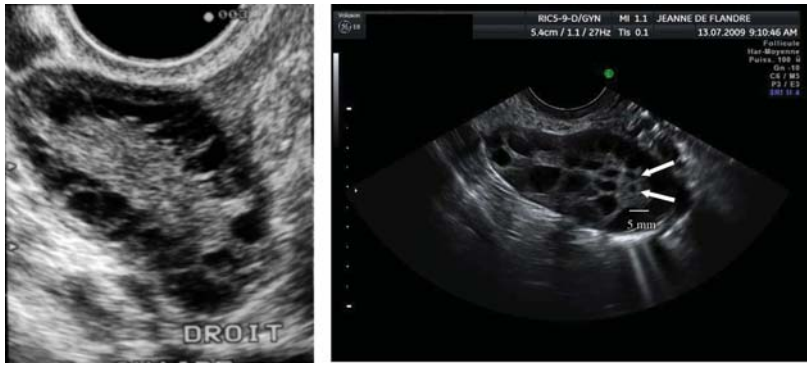

Figure 1 Picture of a PCOM with old (2001) (left) and new (2009) (right) equipment. Small follicles ≤2 mm (arrows) can be visualized and counted with the new equipment.

Table I Main clinical, hormonal and U/S data in the four subgroups of patients.

	Group 1A (n = 66)	Group 1B (n = 39)	Group 2 (n = 73)	Group 3 (n = 62)	P by Kruskal–Wallis test
Age (years)	30.0 (21.9–34.6)[c,d]	28.3 (19.9–33.8)	28.7 (21.4–32.8)[a]	27.6 (20.1–34.0)[a]	0.008
BMI (kg/m^2)	24.0 (18.7–37.6)	24.0 (18.0–39.0)	26.0 (18.0–38.7)	28.0 (18.7–41.7)	0.106
WC (cm)	78.0 (65.8–105.2)	80.0 (64.0–116.0)	80.0 (63.0–112.0)	89.0 (67.1–130.9)	0.065
FSH (IU/l)	5.3 (3.6–8.7)[d]	5.0 (3.4–9.0)	5.0 (3.3–7.2)	4.7 (3.1–6.6)[a]	0.014
LH (IU/l)	3.2 (1.5–6.3)[d,c]	3.8 (2.2–9.0)[d]	4.0 (1.8–11.4)[a]	5.5 (2.3–13.7)[a,b,c]	0.0001
AMH (pmol/l)	21.0 (10.0–35.0)[b,c,d]	47.2 (36.6–65.8)[a,d]	52.3 (18.0–103.4)[a]	81.2 (25.4–256.2)[a,b,c]	0.0001
TT (ng/ml)	0.17 (0.05–0.39)[c,d]	0.21 (0.05–0.42)[c,d]	0.29 (0.10–0.64)[a,b,d]	0.49 (0.22–0.94)[a,b,c]	0.0001
A (ng/ml)	1.23 (0.66–1.90)[c,d]	1.41 (0.77–1.97)[c,d]	1.67 (0.75–2.95)[a,b,d]	2.50 (1.44–4.56)[a,b,c]	0.0001
Follicle number	11.5 (6.2–21.8)[b,c,d]	19.0 (11.3–28.6)[a,d]	21.0 (10.8–41.0)[a,d]	30.5 (15.0–58.7)[a,b,c]	0.0001
Ovarian volume (ml)	5.0 (2.3–9.2)[b,c,d]	6.7 (3.9–12.4)[a,d]	7.1 (3.6–14.9)[a,d]	10.1 (4.6–17.0)[a,b,c]	0.0001

Values are expressed as median with 5–95th percentiles in parentheses.
Non-parametric ANOVA with *post hoc* Bonferonni correction.
[a]Significantly different from group 1A.
[b]Significantly different from group 1B.
[c]Significantly different from group 2.
[d]Significantly different from group 3.

and age were evaluated by the non-parametric Spearman correlation coefficient.

ROC curves were constructed to examine the diagnostic test performance, i.e. the ability to discriminate between groups (Zweig and Campbell, 1993). Se (y-axis) against [1-Spe (x-axis)] was plotted at each threshold level, and the area under the curve (AUC) was computed by the non-parametric Wilcoxon test. The AUC represents the probability of correctly identifying controls and patients with PCOS. A value of 0.5 means that the result is no better than chance.

In order to test our hypothesis that asymptomatic patients with PCOM constitute a mathematically homogenous group among the control patients, we analysed the homogeneity of the control group using a cluster analysis. The variables that are considered as markers of PCOM according to the available literature (reviewed in Johnstone *et al.*, 2010), i.e TT, A, LH, AMH, OV and FN were included in the analysis. The age of patients was also included since it may confound some variables. Cluster analysis is a statistical multivariate classification procedure used to classify patients in different groups or clusters according to different profiles (Hartigan, 1985). These clusters are not defined *a priori* and are such that individuals in a given cluster are close to each other in the sense of a similar measure and individuals in different clusters tend to be dissimilar. The cluster analysis was based on the k-means method. In this method, the similarity between individuals is measured using the usual euclidian distance. The homogeneity of clusters was assessed by the squared correlation ratio (R^2) which is the ratio of the between-cluster variation and the total variation computed from all the variables. The graphical representation of the R^2 values against the number of clusters was used to choose the most appropriate number of cluster. In addition, the R^2 of each variable was computed in order to determine the most important variables in the identification of clusters.

Results

According to their symptoms, the 240 patients included in this study were divided into three groups: group 1 (n = 105) including women

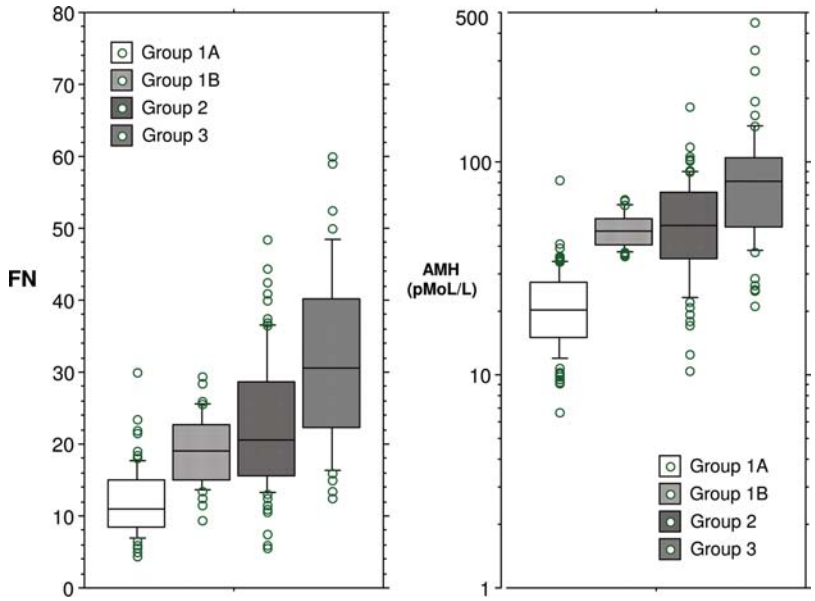

Figure 2 Box-and-whisker plots showing the values of the follicle count (left) and serum AMH level (right, logarithmic scale) in the four subgroups of patients (see text). Horizontal small bars represent the 5–95th percentile range, and the boxes indicate the 25–75th percentile range. The horizontal line in each box corresponds to the median.

without HA (clinical or biological) and with regular menses (non-PCOS group), group 2 ($n = 73$) including women with only HA or only oligo-anovulation (presumption of PCOS) and group 3 ($n = 62$) including women with HA and oligo-anovulation, i.e. patients with genuine PCOS as defined by the current classifications (The Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group, 2004; Azziz *et al*., 2006). U/S data were not used in this classification.

Group 1 was subjected to cluster analysis using age, TT, A, LH, AMH, OV and FN as classifying variables. Since the models with three or four clusters were not accompanied by a significant increase of total R^2 by comparison with the model with two clusters, we considered that two subgroups (1A and 1B) were generated by the analysis, comprising 66 and 39 women, respectively. The critical parameters for this classification were primarily the serum AMH level and then the FN and OV, with R^2 values of 0.72, 0.31 and 0.22, respectively. All other variables, including age, had R^2 values <0.05.

As shown in Table I and Fig. 2, group 1B differed from group 1A by a significantly higher mean rank of AMH, FN and OV, while mean ranks of TT, A, LH and FSH were similar. Group 1B differed from group 2 by lower mean ranks of TT and A, while mean ranks of AMH, FN, OV, FSH and LH were similar. Compared with group 3, group 1B had significantly lower mean ranks of LH, TT, A, AMH, FN and OV (Table I).

Considering that group 1B represented asymptomatic women with PCOM, it was excluded from the data and we then performed a ROC curve analysis using a population gathering groups 1A and 3 (non-PCOM non-PCOS controls and PCOS women, respectively). The AUCs were 0.949 [% confidence interval (CI): 0.915–0.982], 0.923 (% CI: 0.874–0.973) and 0.973 (% CI: 0.947–0.998) for FN, OV and AMH, respectively. The best compromise between Se and Spe was obtained with threshold values of FN at 19, of OV at 7 ml and of AMH at 35 pmol/l (Table II). With these thresholds, the prevalence of elevated AMH, FN and OV was 40, 26.5 and 23% in the initial group 1 (non-PCOS, $n = 105$), respectively, and 78, 59 and 54% in group 2 (presumption of PCOS, $n = 73$), respectively. These figures reflect the prevalences of PCOM in group 1, and of PCOS (using the Rotterdam classification) in group 2, both of which were higher with AMH than with FN or OV.

Table II Results of ROC analysis using a population consisting of groups 1A and 3 (non-PCOM non-PCOS women and those with PCOS, respectively).

	AUC (95% CI)	Threshold	Sensitivity (%)	Specificity (%)
Follicle number	0.949 (0.915–0.982)	17	87	83
		19	**81**	**92**
		21	78	92
Ovarian volume	0.923 (0.874–0.973)	**7 ml**	**87**	**89**
		8 ml	75	92
		9 ml	63	95
		10 ml	50	99.5
Serum AMH	0.973 (0.947–0.998)	30 pmol/l	92	82
		35 pmol/l	**92**	**97**
		40 pmol/l	85	100

AUC, area under the ROC curve; CI, confidence interval.
Bold values indicate best compromise between sensitivity and specificity.

As shown in Fig. 3, a highly significant correlation was observed between the values of AMH and FN in the entire study population

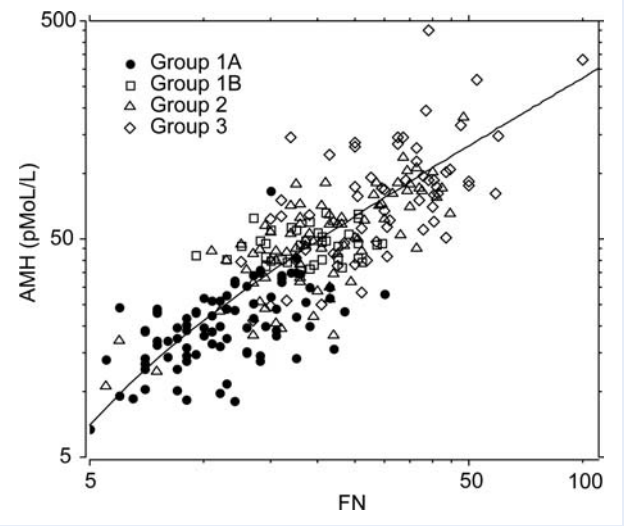

Figure 3 Correlation between FN and serum AMH in the entire population ($n = 240$). See the text for the definition of subgroups.

($r = 0.771$, $P < 0.0001$). Both FN and serum AMH were negatively correlated to age, with modest, although significant, R values (-0.235 and -0.207 respectively, $P < 0.0001$ for both). When the earlier ROC analysis was restricted to the non-PCOM non-PCOS patients who were aged <30, the AUC for AMH was 0.979, with the best compromise at 37 pmol/l (Spe = 100%, Se = 95%) and the AUC for FN was 0.960, with the best compromise at 20 (Spe = 94%, Se = 84%).

Discussion

So far, no single marker other than FN at U/S has been able to separate accurately women with normal ovaries from those with PCOM among a group of asymptomatic women. Consequently, it might seem impossible to exclude the former from a control group without using a given threshold for FN. However, as a whole, women having PCOM differ from those with normal ovaries by having slightly higher mean serum androgens (Adams *et al.*, 2004; Mortensen *et al.*, 2009) or AMH levels (Johnstone *et al.*, 2010), although this was not significant in all studies (reviewed in Johnstone *et al.*, 2010).

By analysing the profiles of the control patients according to the markers of PCOS without using any pre-determined threshold, our cluster analysis has been able to isolate a homogenous subgroup of women within our initial control group. We can confidently assume that these asymptomatic women represented a population with PCOM, since the most relevant variables used by the analysis were first the serum AMH and then the FN and OV. Indeed, an increased serum AMH level has been recently reported in normal women with PCOM (Johnstone *et al.*, 2010). We show here that this increase was to the same degree as in women with HA or oligo-anovulation, i.e. with most probably a moderate form of PCOS. Therefore, in disagreement with others (Johnstone *et al.*, 2010), we think that such data support the hypothesis that PCOM in normal women are not a morphological variant of normal ovaries but rather represent a functional entity that may be considered as a silent form of PCOS (Franks *et al.*, 2008) for which serum AMH could be the best marker.

Once women with presumed PCOM were excluded from the initial non-PCOS group, our ROC analysis confirmed our hypothesis that the FN threshold retained at the Rotterdam conference in 2003 (The Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group, 2004) is now obsolete. Indeed, with this analysis, the best compromise between Se (81%) and Spe (92%) was obtained with a threshold of 19 follicles per ovary, whereas we had previously proposed a threshold of 12 with older equipment (Jonard *et al.*, 2003). This adjustment agrees with the one proposed more recently by Allemand *et al.* (2006) which was 20 per ovary, using 3D U/S. It must be stressed however that only 10 patients with PCOS and 29 controls were included in that study. Conversely, our data differ from those of Lam *et al.* (2007), also with 3D U/S, who reported that their control subjects had a similar FN range to that reported previously with older equipment. However, before using 3D U/S screening, they performed 2D screening with a modern machine and they eliminated from the control group the women who had the former Rotterdam criteria of PCO (i.e. FN > 12).

In our opinion, this important change in the threshold for FN is not attributable to the 3D technique. Indeed, it was reported that follicular counting gave identical results with the 2D and 3D techniques (Jayaprakasan *et al.*, 2007). Likewise, we have not observed any difference in our experience (data not shown), which is why we present results with the 2D technique that is more easily available in practice for follicular counting. In agreement with others (Johnstone *et al.*, 2010), our opinion is that the significant increase in the threshold is due to the improvement of the resolving power of U/S images with new appliances, with 2D or 3D as well. Indeed, it is now possible to detect follicular images <2 mm diameter, which was not previously the case (Fig. 1). These small images are probably not artefacts, as evidenced in this study by the excellent correlation between the FN counted with U/S and the serum AMH levels, with a correlation coefficient being stronger than that obtained when using FN with older U/S machines (Pigny *et al.*, 2003; Laven *et al.*, 2004). This is in keeping with the immuno-cyto-chemistry (ICC) data in humans showing that AMH is maximally expressed in GC from follicles measuring 1–4 mm (Weenen *et al.*, 2004).

Our study indicated that serum AMH level is a more reliable marker of PCOM than FN. First, cluster analysis showed that serum AMH appeared as the strongest parameter to isolate women with PCOM within the non-PCOS group. Second, in the population gathering non-PCOM non-PCOS women and those with PCOS, the area under the ROC curve (AUC) was better with serum AMH than with FN or OV, and this analysis yielded higher rates of Spe and Se (97 and 92%, respectively). These figures are also much better than in our previous study about AMH in 2006 (92 and 67%, respectively) (Pigny *et al.*, 2006). This is explained by a much greater AUC in the present study (0.973 versus 0.851 respectively). In addition, the threshold proposed here is lower than that reported previously (35 versus 60 pmol/l, respectively) (Pigny *et al.*, 2006). Rather than technical issues (the assay procedure did not change), the main reason for these discrepancies is probably the difference in the selection of the non-PCOS reference group. In 2006, using the U/S equipment and criteria of the time (Pigny *et al.*, 2006), we probably failed excluding all asymptomatic women with PCOM from this group. Indeed, in the present study, U/S was less sensitive than AMH in detecting PCOM in the non-PCOS group and in patients with mild PCOS as

Table III Adaptation of the previous classifications for the diagnosis of PCOS, proposing an excessive FN of >19 or serum AMH concentration >35 pmol/l or >5 ng/ml as a surrogate when either oligo-anovulation or HA is missing.

Oligo -anovulation	Clinical and/or biological HA	FN > 19 and/or serum AMH[a]> 35 pmol/l (5 ng/ml)	Diagnosis
+	+	(+/−)[b]	PCOS
+	−	+	PCOS
−	+	+	PCOS
−	−	+	Normal woman with PCOM[c]
+	−	−	Idiopathic anovulation
−	+	−	Idiopathic hyperandrogenism

As with the previous classifications, other causes of oligo-anovulation and/or HA must be excluded before applying this classification.
[a]To be used preferentially.
[b]Not necessary for the diagnosis.
[c]Consider the risk for OHSS.

well. Therefore, our results in 2006 were presumably spoiled by a greater overlap between the non-PCOM non-PCOS group and the PCOS population. Conversely, our results with OV were similar to those in our previous report (Jonard *et al*., 2005), with an optimal threshold at 7 ml and similar Spe but with better Se (89 versus 91% and 87 versus 67%, respectively). This indicates that the higher Se of the new U/S equipments has much less impacted this measure than the follicle counting. However, OV still has a lower Se and Spe than the FN and especially the AMH assay.

Knowing the negative effect of age on the FN and serum AMH values, that we confirm here, although it was weak, some authors have advocated adapting the thresholds to the patients' age (Duijkers and Klipping, 2009; Johnstone *et al*., 2010). Besides the inconvenience of having to validate and then to memorize several thresholds, we think that this adaptation is not crucial. Indeed, when our ROC analysis was restricted to non-PCOM non-PCOS women that were aged <30, the data for both FN and serum AMH were very close to that obtained in the whole population. Conversely, it must be emphasized that the criteria proposed here should not be applied to women aged more than 35. Lastly, we must recognize that our control group included exclusively patients referred to our clinic. It might therefore be not fully representative of the general population. Further studies are required to validate our data in other settings.

In conclusion, for the definition of PCOM and PCOS, serum AMH appears as a sensitive and specific parameter that would probably be easier to reproduce from one to another centre than the follicle count, as the latter is highly dependent on the evolving quality of the machines and/or the operator skill. When a single commercially available AMH assay (Kumar *et al*., 2010) is universally used, it will become urgent that an international consortium validates the threshold for AMH proposed here to suggest the presence of PCOM, i.e. 35 pmol/l (or 5 ng/ml). Since we have previously reported that the serum AMH level is closely related to both the markers of HA and ovulatory disturbance in PCOS (Dewailly *et al*., 2010), the use of this parameter in the current classifications for PCOS (The Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group, 2004; Azziz *et al*., 2006) would allow simplification of the diagnosis of PCOS, as proposed in Table III.

Authors' roles

D.D. participated in the study design, execution, analysis, manuscript drafting and critical discussion. H.G. participated in the study execution, analysis, manuscript drafting and critical discussion. E.P. participated in the study design, execution and critical discussion. G.R. participated in the study execution and critical discussion. M.L. participated in the study execution and critical discussion. P.P. participated in the study execution and critical discussion. A.D. participated in the study design, execution, analysis and critical discussion. S.C.-J. participated in the study design, execution, analysis, manuscript drafting and critical discussion.

Acknowledgements

We thank Mrs Lydie Lombardo and Sylvie Vanoverschelde for their help in collecting the blood samples and caring for patients. We also thank the staff of the Laboratoire de Biochimie et Hormonologie (Centre de Biologie Pathologie) and the Service de Radiologie (Hôpital Jeanne de Flandre, Centre Hospitalier Régional Universitaire de Lille) for their excellent technical help. We thank Ms Mary-Jane Guerry for her help in editing this manuscript in English.

References

Adams JM, Taylor AE, Crowley WF Jr, Hall JE. Polycystic ovarian morphology with regular ovulatory cycles: insights into the pathophysiology of polycystic ovarian syndrome. *J Clin Endocrinol Metab* 2004;**89**:4343–4350.

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.

Asuncion M, Calvo RM, San Millan JL, Sancho J, Avila S, Escobar-Morreale HF. A Prospective study of the prevalence of the polycystic ovary syndrome in unselected Caucasian women from Spain. *J Clin Endocrinol Metab* 2000;**85**:2434–2438.

Azziz R. Controversy in clinical endocrinology: diagnosis of polycystic ovarian syndrome: the Rotterdam criteria are premature. *J Clin Endocrinol Metab* 2006;**91**:781–785.

Azziz R, Woods KS, Reyna R, Key TJ, Knochenhauer ES, Yildiz BO. The prevalence and features of the polycystic ovary syndrome in an unselected population. *J Clin Endocrinol Metab* 2004;**89**:2745–2749.

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

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

Catteau-Jonard S, Pigny P, Reyss AC, Decanter C, Poncelet E, Dewailly D. Changes in serum anti-Müllerian hormone level during low-dose recombinant follicular-stimulating hormone therapy for anovulation in polycystic ovary syndrome. *J Clin Endocrinol Metab* 2007;**92**:4138–4143.

Conover WJ, Iman RL. Rank transformations as a bridge between parametric and nonparametric statistics. Am Stat 1981;**35**:124–129.

Dewailly D, Catteau-Jonard S, Reyss AC, Leroy M, Pigny P. Oligoanovulation with polycystic ovaries but not overt hyperandrogenism. *J Clin Endocrinol Metab* 2006;**91**:3922–3927.

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-Müllerian hormone concentrations aggregate with the markers of hyperandrogenism. *J Clin Endocrinol Metab* 2010;**95**:4399–4405.

Diamanti-Kandarakis E, Kouli CR, Bergiele AT, Filandra FA, Tsianateli TC, Spina GG, Zapanti ED, Bartzis MI. A survey of the polycystic ovary syndrome in the Greek Island of Lesbos : hormonal and metabolic profile. *J Clin Endocrinol Metab* 1999;**84**:4006–4011.

Duijkers IJ, Klipping C. Polycystic ovaries, as defined by the 2003 Rotterdam consensus criteria, are found to be very common in young healthy women. *Gynecol Endocrinol* 2009;**26**:152–160.

Franks S. Controversy in clinical endocrinology: diagnosis of polycystic ovarian syndrome: in defense of the Rotterdam criteria. *J Clin Endocrinol Metab* 2006;**91**:786–789.

Franks S, Webber LJ, Goh M, Valentine A, White DM, Conway GS, Wiltshire S, McCarthy MI. Ovarian morphology is a marker of heritable biochemical traits in sisters with polycystic ovaries. *J Clin Endocrinol Metab* 2008;**93**:3396–3402.

Hartigan JA. Statistical theory in clustering. *J Classif* 1985;**2**:63–76.

Jayaprakasan K, Hilwah N, Kendall NR, Hopkisson JF, Campbell BK, Johnson IR, Raine-Fenning NJ. Does 3D ultrasound offer any advantage in the pretreatment assessment of ovarian reserve and prediction of outcome after assisted reproduction treatment?. *Human Reproduction* 2007;**22**:1932–1941.

Johnstone EB, Rosen MP, Neril R, Trevithick D, Sternfeld B, Murphy R, Addauan-Andersen C, McConnell D, Pera RR, 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.

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.

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.

Kumar A, Kalra B, Patel A, McDavid L, Roudebush WE. Development of a second generation anti-Müllerian hormone (AMH) ELISA. *J Immunol Methods* 2010;**362**:51–59.

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. *Human Reproduction* 2007; **22**:3116–3123.

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

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, Rosenfield RL. Asymptomatic volunteers with a polycystic ovary are a functionally distinct but heterogeneous population. *J Clin Endocrinol Metab* 2009;**94**:1579–1586.

Norman RJ, Dewailly D, Legro RS, Hickey TE. Polycystic ovary syndrome. *Lancet* 2007;**370**:685–697.

Pigny P, Merlen E, Robert Y, Cortet-Rudelli C, Decanter C, Jonard S, Dewailly D. Elevated serum level of anti-Müllerian 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-Müllerian hormone as a surrogate for antral follicle count for definition of the polycystic ovary syndrome. *J Clin Endocrinol Metab* 2006;**91**:941–945.

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* 2004;**19**:41–47.

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

Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* 1993; **39**:561–577.