human reproduction **ORIGINAL ARTICLE** *Reproductive endocrinology*

Anti-Mullerian hormone in the diagnosis of polycystic ovary syndrome: can morphologic description be replaced?

Tina B. Eilertsen[1,2,*], Eszter Vanky[2,3], and Sven M. Carlsen[4,5]

[1]Department of Obstetrics and Gynaecology, Hospital of Namsos, Nord-Troendelag Hospital Trust, Namsos, Norway [2]Department of Laboratory Medicine, Children's and Women's Health, Norwegian University of Science and Technology, Trondheim, Norway [3]Department of Obstetrics and Gynaecology, University Hospital of Trondheim, Trondheim, Norway [4]Unit for Applied Clinical Research, Department of Cancer Research and Molecular Medicine, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway [5]Department of Endocrinology, St. Olavs Hospital, University Hospital of Trondheim, Trondheim, Norway

*Correspondence address. Tel: +47 47015313; Fax: +4774215732; E-mail: tina.eilertsen@hnt.no

STUDY QUESTION: Can anti-Müllerian hormone (AMH) level replace the morphologic description in the diagnosis of polycystic ovary syndrome (PCOS) and what is the relationship between AMH and different diagnostic criteria of PCOS?

SUMMARY ANSWER: AMH may be a good substitute for polycystic ovarian morphology (PCOM) in diagnosing PCOS.

WHAT IS KNOWN ALREADY: AMH has been suggested as an alternative to antral follicle count (AFC) in diagnosing PCOS. Cut-off values for AMH studied so far show an acceptable specificity but a rather poor sensitivity, leaving up to one-third of PCOS women undiagnosed.

STUDY DESIGN, SIZE, DURATION: We used data from a cross-sectional, case–control study on women with prior preterm birth and their controls, i.e. women with prior full-term birth. Among 262 women, 56 met the Rotterdam criteria (PCOS-R) and 44 the Androgen Excess-PCOS Society (PCOS-AES) criteria of PCOS.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Fasting blood samples were collected, a transvaginal ultrasound investigation and a clinical examination were performed. PCOS-R and PCOS-AES were re-diagnosed by replacing PCOM with AMH. Main outcome measures were the prevalence of PCOS, PCOM, hirsutism, oligoamenorrhoea and serum levels of AMH and androgens.

MAIN RESULTS AND THE ROLE OF CHANCE: When replacing PCOM with AMH, the specificity and sensitivity for identifying PCOS were 97.1 and 94.6% according to the PCOS-R criteria and 97.2 and 95.5% according to the PCOS-AES criteria, respectively, at an AMH cut-off value of 20 pmol/l.

LIMITATIONS, REASONS FOR CAUTION: The results need to be confirmed when international standards and methods for AMH measurements are established.

WIDER IMPLICATIONS OF THE FINDINGS: AMH may be a good substitute for PCOM in diagnosing PCOS.

STUDY FUNDING/COMPETING INTEREST(S): This study was financed by the Cooperative of Central Norway Regional Health Authority and Norwegian University of Science and Technology. The authors have no interests to disclose.

TRIAL REGISTRATION NUMBER: This study is registered at www.clinicaltrials.gov as NCT01355536.

Key words: AMH / PCOS / AFC / hirsutism / cut-off / sensitivity / specificity

Introduction

During the last decades it has become evident that polycystic ovary syndrome (PCOS) is more than just hirsutism and infertility (Dunaif and Thomas, 2001). Most PCOS women are insulin resistant (Robinson et al., 1993; Chang et al., 2005). PCOS women have increased risk of hypertension, altered glucose metabolism and probably increased life time risk of cardiovascular diseases (Moran et al., 2010). They

have a higher prevalence of miscarriage, pregnancy-induced hypertension, pre-eclampsia, gestational diabetes mellitus, complicated deliveries and preterm births (Boomsma et al., 2006; Roos et al., 2011, 2011). The Androgen Excess-PCOS Society (AES) have published guidelines for follow-up on PCOS women, where the aim is intervention to prevent long-term consequences (Wild et al., 2010).

Diagnosing PCOS is a challenge with changing criteria and different definitions. While the National Institutes of Health (NIH) criteria from 1990 are strict and include only anovulation and hyperandrogenism (HA; Dunaif and Thomas, 2001), the Rotterdam consensus criteria of 2003 are broader and include polycystic ovarian morphology (PCOM) as a criterion (Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2004). More recently, the AES have proposed that HA should be mandatory for the diagnosis (Azziz et al., 2006). Obtaining good data on ovarian morphology demand a time- and resource-consuming ultrasound (US) examination by a specialist with appropriate skills. There has been a lack of standardization in assessing PCOM, leading to differences in results (Broekmans et al., 2010). PCOM changes through the menstrual cycle and oral contraceptive use, making a standardized, reliable assessment of PCOM even more difficult (Somunkiran et al., 2007). In addition, US imaging changes over time due to the rapidly improving quality of US equipment (Kristensen et al., 2010).

PCOS is not a rare condition. The prevalence among women of fertile age is 6–10% using the NIH criteria (Knochenhauer et al., 1998; Azziz et al., 2004) and 14–17% using the broader Rotterdam criteria (March et al., 2010; Tehrani et al., 2011). There are reasons to believe that many PCOS women in the general population are undiagnosed (March et al., 2010; Eilertsen et al., 2012). In addition, obesity worsens the metabolic and endocrine profile in PCOS, and the obesity epidemic may lead to increased prevalence (Hoeger, 2006).

This underlines the importance of identifying women with PCOS due to the need for follow-up on short- and long-term health risks. A standardized diagnostic tool with minimal inter-observer variation is needed.

Anti-Mullerian hormone (AMH), also known as Mullerian inhibiting substance is produced by the granulosa cells of small antral follicles (De Meyts et al., 1999; Durlinger et al., 2002; Weenen et al., 2004). AMH has an inhibiting role in the ovary, contributing to follicular arrest (Pellatt et al., 2010). AMH levels in women are low until the age of 8, rise rapidly until puberty and decline steadily from the age of 25 until menopause, when AMH production ceases. Recently, extensive normative data of AMH levels in women were published (Hagen et al., 2010).

There is a good correlation between AMH and antral follicle count (AFC; de Vet et al., 2002; van Rooji et al., 2002; Fanchin et al., 2003; Laven et al., 2004). Accordingly, AMH has been proposed as a marker of PCOS and as a substitute for AFC in the PCOS diagnosis (Pigny et al., 2003; Piltonen et al., 2005; Broekmans et al., 2008). AMH also correlates with the other criteria of PCOS: oligoamenorrhoea (OA) and HA (van Rooji et al., 2002; Fanchin et al., 2003; Laven et al., 2004; Eldar-Geva et al., 2005; Piltonen et al., 2005; Carlsen et al., 2009; Nardo et al., 2009; Lin et al., 2011). Cut-off values of AMH have been proposed but have varying sensitivity and specificity (Pigny et al., 2006; Lin et al., 2011). Recently, results with high sensitivity and specificity for AMH as a PCOS marker was published.

However, in this study women with PCOM only were excluded from the control group (Dewailly et al., 2011).

It still remains to be shown whether AMH is a useful tool in diagnosing PCOS in adults, and to determine the appropriate cut-off levels that show the highest sensitivity and specificity in a non-selected adult population.

The aim of the present study was to evaluate how AMH actually performs when used instead of PCOM when diagnosing PCOS. Thus, we hypothesized that diagnosing PCOS using the criteria such as OA, HA and/or AMH (an AMH-based PCOS diagnosis) would identify the original PCOS diagnosis with high sensitivity and specificity.

Materials and Methods

Study design

In the present study we used data from a former cross-sectional, case–control study on women with prior preterm birth. Details according to the study design are recently published (Eilertsen et al., 2012). Half of the women had given preterm birth, while the other half had given birth at term. In all, 262 participants were enrolled and data collected between October 2006 and April 2008. In the prior published study, 21 women were excluded because of prior twin deliveries but are included in this study. The mean time interval between the prior birth and the data collection was 5 years and 1 month. The participating women were recruited from a general population in a well-defined geographic area.

In addition to an obstetric history of preterm or term births as required to participate in the original study, the only inclusion criteria were having Namsos Hospital as their local hospital during pregnancy. Exclusion criteria were (i) lack of communication skills in Norwegian or English, (ii) on-going breastfeeding or (iii) pregnancy. Women who had moved from the region at the time of the study were also excluded.

Identical investigational procedures were performed on all participants. All the evaluations, the US examination included, were carried out by the same investigator (T.B.E.), excluding intraobserver variations. We aimed to schedule the inclusion visit early in the menstrual cycle (Days 1–5).

The study was approved by the Regional Committee for Research Ethics in Health Region IV in Norway. Informed consent was signed by all women before inclusion in the study. The Helsinki Declaration was followed throughout the study.

Measurements

Fasting venous blood samples were taken. Blood samples were centrifuged at room temperature within 30 min. Serum was stored at −70°C until analysis.

The participant's medical history data were recorded. Blood pressure was recorded using the automatic device Boso-Medicus (Bosch & Sohn, Jungingen, Germany). A medical examination including height, body weight, hip and waist measurements, was performed. Weight was recorded using electronic scales (Seca alpha model 770; Seca, Hamburg, Germany). Hirsutism was scored by the main investigator (T.B.E.) using the Ferriman–Gallwey (FG) score (Ferriman and Gallwey, 1961). The length of the participant's menstrual cycles was recorded.

Gynaecological examination including a transvaginal US examination was performed with the US equipment General Electric Logiq Book XP with vaginal probe 7, 5 MHz (General Electric Medical Systems, Solingen, Germany). The size of the ovaries was measured in three dimensions and the total numbers of follicles 2–9 mm in diameter were counted. The ovarian volume was calculated by the formula: height × width × depth × 0.5.

All women were diagnosed both according to the Rotterdam 2003 criteria (PCOS-R; Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2004) and the Androgen Excess-PCOS Society criteria (PCOS-AES; Azziz *et al.*, 2006). According to the Rotterdam 2003 consensus, two out of three criteria have to be met: PCOM, OA or HA. To meet the AES criteria, HA is mandatory in addition to PCOM and/or OA. Oligoanovulation was regarded to be present if the participants reported OA.

PCOM is defined as ≥ 12 follicles of $2-9$ mm in diameter and/or ovarian volume ≥ 10 ml in at least one ovary. OA is defined as the length of the menstrual cycle >35 days or <10 periods per year. HA may be biochemical and/or clinical. Biochemical HA is defined as testosterone >2.5 nmol/l, free testosterone index (FTI) ≥ 0.6 and/or androstenedione ≥ 10.0 nmol/l, while clinical HA is defined as the FG score ≥ 8. The same definition of HA, OA and PCOM is used when setting the PCOS-R and the PCOS-AES diagnosis.

To construct receiver operating characteristic (ROC) curves with AMH as the diagnostic criteria instead of PCOM, we created new variables by replacing PCOM with AMH at different values. Several thresholds of AMH were used, then combined with HA and/or OA as appropriate to obtain a substitute for the PCOS-R or PCOS-AES diagnoses, respectively. These AMH-based PCOS-R and AMH-based PCOS-AES variables are used in the assays and the original PCOS-R or PCOS-AES are set as actual state or the gold standard when constructing ROC curves. AMH-based PCOS-R diagnosis is set if two out of three criteria are met: OA, HA and/or AMH above different levels. AMH-based PCOS-AES diagnosis is set if HA is present together with OA and/or AMH above different levels.

We also constructed ROC curves with the true AMH values in the assay with AFC and PCOM as actual states, respectively.

This was done to compare the diagnostic power of AMH when used in combination with the other diagnostic criteria and when used alone.

Hormonal immunoassays

For testosterone and androstenedione analyses, we used organic solvent extraction (dichloromethane for testosterone and ethyl ether for androstenedione) prior to quantification. Androstenedione was measured by a competitive immunoassay based on antibody-coated tubes, the Coat-A-Count® Direct Androstenedione procedure (Siemens Medical Solutions Diagnostics, Los Angeles, CA, USA), using reagents and calibrators supplied by the manufacturer. AMH were measured by an enzymatically amplified two-site immunoassay, the ACTIVE® MIS/AMH enzyme-linked immunosorbent assay (ELISA), using reagents and calibrators supplied by the manufacturer (Diagnostic Systems Laboratories, Inc, Webster, TX, USA). Sex hormone-binding globulin (SHBG) and testosterone were measured by enzyme immunoassay for the quantitative determination in serum, using reagents and calibrators supplied by the same manufacturer (DRG Instruments GmbH, Marburg, Germany). All measurements were performed in singles. Intra- and inter-assay coefficients of variation were 5.6 and 2.2% for androstenedione, 4.2 and 7.7% for AMH, 6.6 and 5.5% for SHBG and 9.5 and 14.0% for testosterone. FTI was calculated as testosterone/SHBG $\times$ 10.

All analyses were performed on kits from the same batch.

Statistics

All statistical analyses were performed with SPSS version 16.0 for Windows (SPSS Inc., USA). Groups were compared with *t*-tests for independent samples, Pearson's χ^2 test or Fischer's exact test used for 2×2 tables as appropriate. Values are given as means and standard deviations (SD) or absolute numbers and percentages in parenthesis.

Relationships between AMH and other variables were evaluated by Pearson's correlation coefficient (PCC). *P*-values <0.05 were considered significant. No adjustment for multiple testing was performed.

ROC curves were constructed to evaluate the capacity of AMH to identify PCOS. Sensitivity against $1 -$ specificity was plotted at each threshold level and the area under the curve (AUC) was computed. The AUC represents the probability of identifying PCOS and a value of 0.5 means that the test is no better than chance. ROC curves were constructed using AMH-based PCOS-R and AMH-based PCOS-AES (see section Measurements) with the original PCOS-R and PCOS-AES diagnoses as actual states, respectively. ROC curves were also constructed using AFC and PCOM as actual states.

Results

Study groups and characteristics

Among the 262 participants, 56 (21.4%) met the PCOS-R criteria and 44 (16.8%) met the PCOS-AES criteria (Table I). All women who met the PCOS-AES criteria also met the PCOS-R criteria. There is a 100% overlap between the groups, with the same women diagnosed by

Table I Demographic data according to PCOS status[a].

Variables[b]	PCOS-R (*n* = 56)	Non-PCOS-R (*n* = 206)	*P*-value	PCOS-AES (*n* = 44)	Non-PCOS-AES (*n* = 218)	*P*-value
Age (years)	33.3 ± 5.5	35.3 ± 5.0	0.01	33.5 ± 5.8	35.2 ± 5.0	0.05
Parity (no.)	2.5 ± 1.0	2.6 ± 1.0	0.33	2.5 ± 1.1	2.6 ± 1.0	0.48
BMI (kg/m^2)	27.8 ± 5.7	26.6 ± 5.0	0.14	28.5 ± 5.7	26.6 ± 5.0	0.02
Height (cm)	166.7 ± 5.2	165.8 ± 6.0	0.35	166.1 ± 5.0	166.0 ± 6.0	0.96
Weight (kg)	77.0 ± 15.0	73.3 ± 14.5	0.10	78.2 ± 14.6	73.3 ± 14.6	0.04
Waist−hip ratio	0.83 ± 0.07	0.81 ± 0.08	0.12	0.83 ± 0.08	0.81 ± 0.07	0.12
SBP (mmHg)	122 ± 13	121 ± 12	0.52	124 ± 13	121 ± 12	0.18
DPB (mmHg)	73 ± 11	73 ± 9	0.91	73 ± 11	73 ± 9	0.99
Smoking (no.)	12 (21)	43 (21)	0.93	12 (27)	43 (20)	0.26

S/DBP, systolic/diastolic blood pressure.

[a]All women are diagnosed both by the PCOS-R and the PCOS-AES criteria: PCOS according to the Rotterdam criteria (PCOS-R) and the Androgen Excess-PCOS Society criteria (PCOS-AES). All women who meet the PCOS-AES criteria also meet the PCOS-R criteria, accordingly there is an overlap between the groups.

[b]Data are presented as the mean $\pm$ SD or as numbers (percentage).

Table II AFC, ovarian volume and AMH levels according to the diagnostic criteria of PCOS and PCOS diagnosis.

		Total n = 262	AFC[a] (no.)	P-value	Ovarian volume[a] (ml)	P-value	AMH (pmol/l)	P-value
PCOS criteria								
Hirsutism	Yes	50	14.6 ± 8.2		8.6 ± 4.2		34.5 ± 26.2	
	No	212	9.5 ± 5.7	P < 0.001	6.4 ± 3.1	P < 0.001	22.8 ± 20.5	P = 0.004
Biochemical HA	Yes	18	15.9 ± 10.1		8.3 ± 4.0		38.3 ± 24.0	
	No	244	10.1 ± 6.0	P < 0.001	6.7 ± 3.4	NS	24.1 ± 21.7	P = 0.024
Total HA	Yes	60	14.8 ± 8.6		8.6 ± 4.2		36.0 ± 26.2	
	No	202	9.2 ± 5.1	P < 0.001	6.3 ± 3.0	P < 0.001	21.8 ± 19.6	P < 0.001
Oligoamenorrhoea	Yes	32	16.9 ± 8.2		10.3 ± 3.5		42.1 ± 28.8	
	No	230	9.5 ± 5.7	P < 0.001	6.3 ± 3.1	P < 0.001	22.7 ± 20.0	P < 0.001
PCOM	Yes	113	15.9 ± 6.2		9.2 ± 3.3		39.0 ± 24.7	
	No	149	6.3 ± 2.5	P < 0.001	4.8 ± 1.9	P < 0.001	14.5 ± 11.8	P < 0.001
PCOS diagnosis								
PCOS-R	Yes	56	17.8 ± 7.4		10.0 ± 3.7		44.8 ± 27.5	
	No	206	8.4 ± 4.5	P < 0.001	5.8 ± 2.7	P < 0.001	19.7 ± 16.8	P < 0.001
PCOS-AES	Yes	44	17.9 ± 7.8		10.2 ± 3.9		42.7 ± 26. 7	
	No	218	8.9 ± 5.1	P < 0.001	6.1 ± 2.9	P < 0.001	21.5 ± 19.3	P < 0.001

Data are presented as the mean ± SD.
[a]AFC and ovarian volume are the mean of both ovaries.

different criteria. PCOS-AES women had higher BMI and weight than non-PCOS women. Both PCOS-R and PCOS-AES women were younger than non-PCOS women. There were no differences between either of the two PCOS groups and the non-PCOS women with respect to parity, height, waist−hip ratio, blood pressure or smoking (Table I).

Table II shows the AFC count, ovarian volume and AMH levels according to PCOS-R, PCOS-AES and the different diagnostic criteria. The mean AMH level in PCOS-R women is 44.8 pmol/l versus 19.7 pmol/l in non-PCOS-R women. The mean AMH level in PCOS-AES women is 42.7 pmol/l versus 21.5 pmol/l in non-PCOS-AES women.

Nine out of 56 (16.1%) PCOS women became pregnant following assisted reproductive techniques (ARTs) versus 11 out of 206 (5.3%) in the control group (P = 0.007).

AMH; sensitivity, specificity and ROC curves

Several levels of AMH were analyzed to test its ability to identify PCOS-R, PCOS-AES, AFC ≥12 and PCOM. The ROC curves obtained when PCOM is replaced by AMH at different levels are shown in Fig. 1a and b. Figure 1a shows the AMH-based PCOS variable identifying PCOS-R, with an AUC equal to 0.992. Figure 1b shows the AMH-based PCOS variable identifying PCOS-AES with AUC equal to 0.994. The AUC for AMH identifying AFC ≥12 is 0.891 (Fig. 1c) and for AMH identifying PCOM is 0.896 (Fig. 1d). The sensitivity and specificity for some of the analyzed cut-off levels are listed in Table III.

Figure 2 shows AMH levels in women with zero or one PCOS criterion and the different PCOS phenotypes. It also illustrates the groups at risk of getting wrongly diagnosed or failing to be diagnosed with PCOS when PCOM is replaced by AMH. Here exemplified with an AMH cut-off value set at 20 pmol/l. The figure also illustrates that women with no PCOS criterion, PCOM only, the HA + OA

phenotype and the HA + OA + PCOM phenotype are unaffected if AMH are used instead of PCOM.
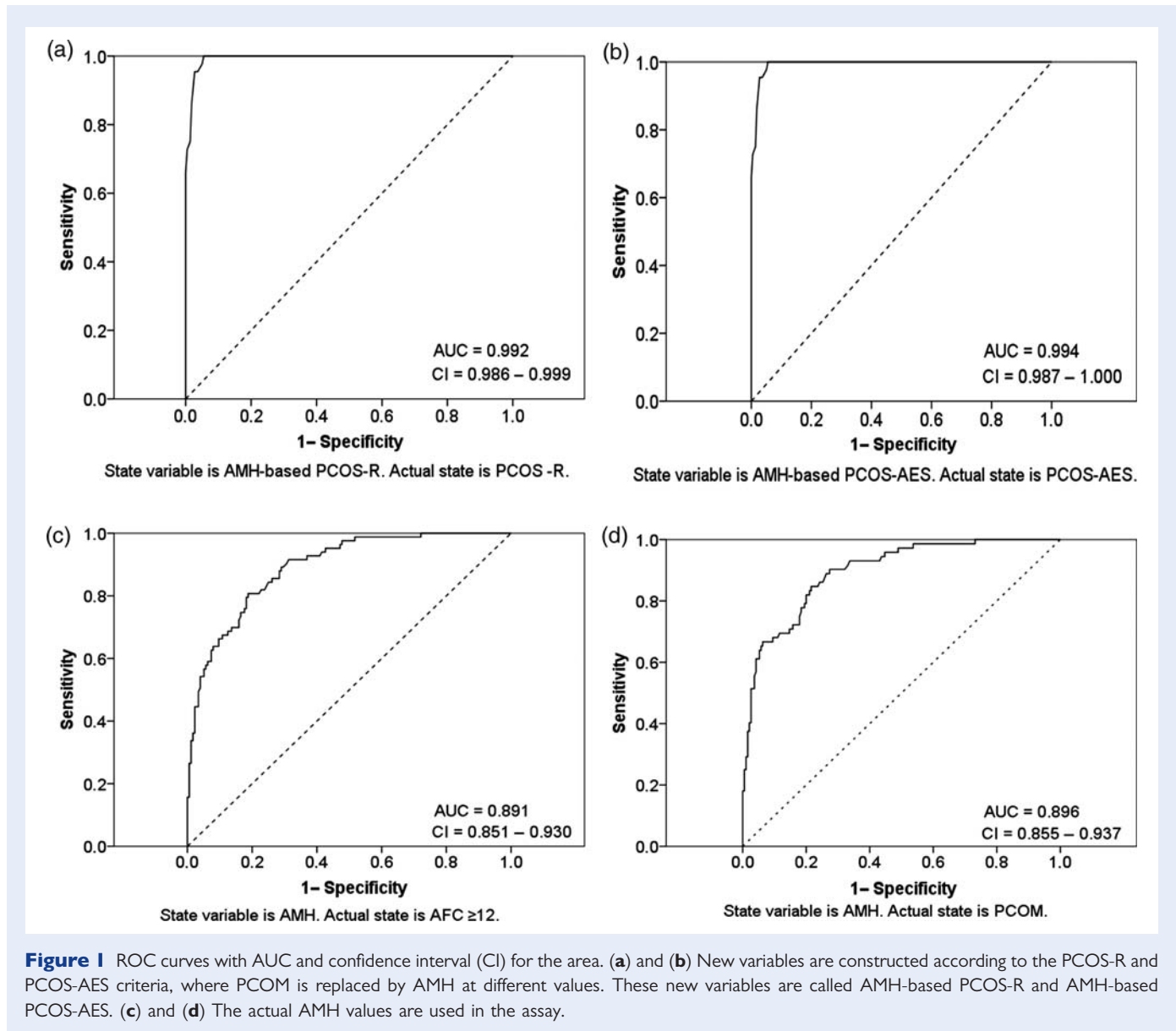
Hormonal contraceptive use

A total of 40 women used hormonal contraception, four within the PCOS groups (two according to PCOS-AES and four according to PCOS-R). The ovarian volume was lower among the contraceptive users, but the AFC or AMH levels were not different. There was also a trend toward less HA in the contraceptive-user group, with androstenedione and SHBG levels being significantly lower (Table IV). ROC curve analyses done on the participants when hormonal contraceptive users were excluded did not change the results noteworthy; the AMH-based PCOS-R assay gave a sensitivity of 94.2 and a specificity of 96.5, while the AMH-based PCOS-AES assay gave a sensitivity of 95.2 and a specificity of 96.7 at a AMH cut-off level of 20 pmol/l.

AMH correlations

There is a strong positive correlation between AMH and AFC, with a PCC of 0.788. There is a weak but significant positive correlation between both AMH and AFC with FTI (PCC = 0.22 and 0.24), testosterone (PCC = 0.34 and 0.28), androstenedione (PCC = 0.35 and 0.33), FG score (PCC = 0.22 and 0.35) and menstrual cycle length (PCC = 0.25 and 0.33) and a negative correlation with age (PCC = 0.25 and 0.33).

Discussion

The main finding of this study is that the power of AMH to diagnose PCOS increases substantially when combined with the other diagnostic criteria of PCOS. If used in this way, AMH can be substituted for PCOM and is equally good in diagnosing PCOS-R and PCOS-AES.

Figure 1 ROC curves with AUC and confidence interval (CI) for the area. (**a**) and (**b**) New variables are constructed according to the PCOS-R and PCOS-AES criteria, where PCOM is replaced by AMH at different values. These new variables are called AMH-based PCOS-R and AMH-based PCOS-AES. (**c**) and (**d**) The actual AMH values are used in the assay.

According to our results, AMH is not suitable as a single screening tool for PCOS, independent of the other diagnostic criteria, or to identify PCOM or AFC ≥ 12 with both acceptable sensitivity and specificity. However, when replacing PCOM with AMH in women with PCOS-R or PCOS-AES, the accuracy is remarkable. In the present study an AMH cut-off value of 20 pmol/l shows a high specificity and sensitivity, and high AUC values indicate that this may be a very accurate diagnostic procedure.

As shown in Figure 2, groups with no PCOS criterion, the PCOM-only group, the HA + OA phenotype and HA + OA + PCOM phenotype remain unaffected when PCOM are replaced by AMH. The HA + OA and the HA + OA + PCOM phenotypes still have criteria sufficient to remain in the PCOS group even with low AMH concentrations. Likewise, women with no PCOS criterion or PCOM only will not be diagnosed with PCOS even with AMH concentrations higher than the cut-off level, because they do not have other criteria sufficient for the diagnosis. It is just the HA-only group, the

OA-only group, the HA + PCOM phenotype and the OA + PCOM phenotype that might be diagnosed wrongly or fail to be diagnosed. This contributes substantially to the high specificity and sensitivity in this model and is probably the main explanation. These results seem obvious and the findings should not be a surprise. However, to our knowledge, they have not previously been shown or demonstrated in a simple model.

Two groups are at risk of being wrongly diagnosed with PCOS if AMH levels are >20 pmol/l: the HA-only and the OA-only groups. As shown in Fig. 2, in the present sample most women from both these groups had AMH levels <20 pmol/l, maintaining a high specificity at this AMH cut-off value. In addition, two phenotypes, the HA + PCOM group and the OA + PCOM group, will wrongly be excluded if AMH levels are below the cut-off value; the latter only if PCOS-R criteria are used. Figure 2 shows that women in these two groups mostly (25/27 and 11/12, respectively) have AMH levels >20 pmol/l in this sample, maintaining a high sensitivity as well. The results are equally

Table III Sensitivity and specificity of AMH-based PCOS-R and AMH-based PCOS-AES at identifying cases diagnosed by PCOS-R and PCOS-AES.

| Assay[a] (total, n = 262)	Actual state: PCOS-R		Assay[a]	Actual state: PCOS-AES	
	Sensitivity (%) [Yes (n = 56)]	Specificity (%) [No (n = 206)]		Sensitivity (%) [Yes (n = 44)]	Specificity (%) [No (n = 218)]
AMH-based PCOS-R[b] (pmol/l)			AMH-based PCOS-AES[c] (pmol/l)		
10	98.2	94.8	10	97.7	95.0
20	94.6	97.1	20	95.5	97.2
25	85.7	98.1	25	86.4	98.2
30	75.0	98.5	30	75.0	98.6
40	69.6	99.5	40	72.7	99.5
Assay[d]	**Actual state: AFC ≥12**		**Assay[d]**	**Actual state: PCOM**	
	Sensitivity (%) [Yes (n = 83)]	Specificity (%) [No (n = 176)]		Sensitivity (%) [Yes (n = 113)]	Specificity (%) [No (n = 149)]
AMH pmol/l			AMH pmol/l		
10	98.8	39.8	10	92.9	44.3
20	91.6	69.8	20	79.6	72.5
30	67.5	87.5	30	54.0	88.6
40	54.2	95.5	40	42.5	96.6

Sensitivity and specificity of AMH at identifying AFC ≥12 and PCOM. PCOS-R, PCOS diagnosis according to the Rotterdam criteria; PCOS-AES, PCOS diagnosis according to the Androgen Excess-PCOS Society criteria (AFC >12 and/or volume of one or both ovaries ≥10 ml).
[a]PCOM is replaced by the AMH concentrations shown resulting in AMH-based PCOS diagnoses.
[b]AMH-based PCOS-R diagnosis is set if two out of three of the following criteria are met: oligomenorrhea, HA, AMH above the concentration listed.
[c]AMH-based PCOS-AES diagnosis is set if HA is present together with oligoamenorrhoea and/or AMH above the concentration listed.
[d]The true AMH values are used in the assay.
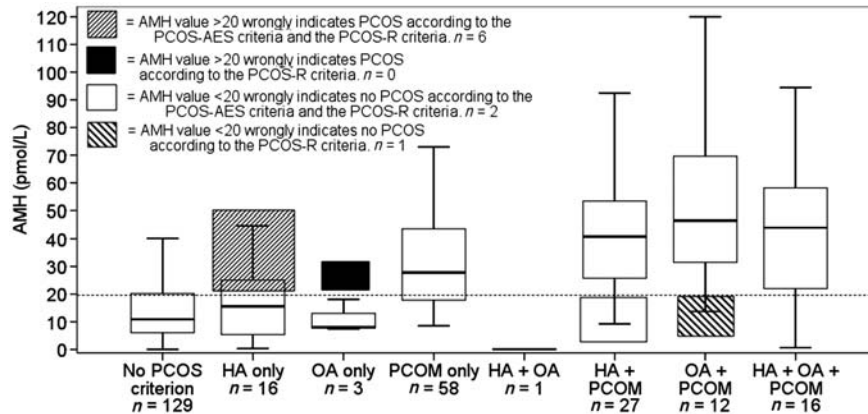
Figure 2 A box plot of AMH levels in 262 women with zero or one PCOS criterion and different PCOS phenotypes. The shaded, white and black boxes illustrate where AMH at a cut-off level of 20 pmol/l would fail to perform in diagnosing PCOS-R and PCOS-AES if AMH was used instead of PCOM. The boxes show the 25th and the 75th percentile with a line representing the mean. The whiskers show the 5th and 95th percentile. HA, hyperandrogenism; OA, oligoamenorrhoea; PCOM, polycystic ovarian morphology; PCOS-R, PCOS according to the Rotterdam criteria; PCOS-AES, PCOS according to the Androgen Excess-PCOS society criteria.

good regardless of which diagnostic criteria used: PCOS-R or PCOS-AES. However, the number of participants in each subgroup is small, and this has to be considered when interpreting the results.

The correlation of AMH to AFC is well known and strong. In addition, AMH correlates weakly but significant to OA and biochemical HA (van Rooji et al., 2002; Fanchin et al., 2003; Laven et al., 2004; Eldar-Geva et al., 2005; Piltonen et al., 2005; Carlsen et al., 2009; Nardo et al., 2009; Lin et al., 2011). This is confirmed in the present study. Dewailly et al. (2010, 2011) report that AMH not only reflect the AFC, but also to a great extent HA. In our material,

Table IV Hormonal contraceptive use.

	Hormonal contraceptive use (n = 40)	No hormonal contraceptive use (n = 222)	P-value
Ovarian volume (ml)	5.2 ± 2.8	7.2 ± 3.5	0.01
AFC	10.2 ± 6.3	10.5 ± 6.6	0.81
AMH (pmol/l)	23.1 ± 22.6	25.4 ± 22.0	0.55
FG score	3.5 ± 3.7	4.6 ± 4.1	0.08
Testosterone (nmol/l)	1.01 ± 0.45	1.16 ± 0.46	0.06
Androstenedione (nmol/l)	4.26 ± 2.12	5.11 ± 2.34	0.03
SHBG (nmol/l)	90.3 ± 85. 8	60. 3 ± 32.8	0.04
FTI	0.21 ± 0.19	0.25 ± 0.17	0.28
PCOS-R (no.)	4 (10.0)	52 (23.4)	0.06
PCOS-AES (no.)	2 (5.0)	42 (18.9)	0.03

Data are presented as the mean ± SD or as numbers (percentage).

we also found a weak but significant correlation between AMH and clinical HA. The correlation of AMH to all PCOS criteria, not only AFC, contributes as well to the good performance of AMH replacing PCOM in PCOS diagnostics.

Our data suggest that AMH cut-off values can be set lower than previously suggested (Pigny et al., 2006; Dewailly et al., 2011), with higher specificity and sensitivity. Given the challenges in obtaining good US imaging of ovarian morphology, the substantial inter-observer variation and the ever-evolving US technology, it seems that AMH should be discussed as the standard criterion in PCOS diagnostics, and not only as a substitute when PCOM data are unavailable.

These findings open the possibility of a simpler way to diagnose PCOS-R and PCOS-AES. According to the high prevalence of the syndrome, the high level of undiagnosed women and the long-term health risk these women face, this seems necessary. If there were an easy way to obtain HA as well, diagnosis could be obtained by a simple question about menstrual regularity and a blood sample. This seems unlikely because the determination of the degree of hirsutism demands a thorough clinical examination and exploration of the woman's hair removing practice. Hirsutism contributes to a great extent to the number of hyperandrogenic women in this study, and is important for the PCOS diagnosis. Hirsutism is probably a symptom of long-term androgen exposure and once manifest, it does not disappear. Therefore, hirsutism still must be considered an important measure of HA. The overall prevalence of hirsutism in our material is 19.1% and is mainly explained by the high prevalence of PCOS. The prevalence of hirsutism in an unbiased, healthy population of premenopausal women is 12.2% (Sanchon et al., 2012) and the prevalence of hirsutism in PCOS women is reported up to 70% (Fauser et al., 2012).

The mean AMH level in non-PCOS women or controls differs between studies. This affects the results and makes studies difficult to compare. The mean AMH level in non-PCOS-R women in this study is 19.7 pmol/l. This is in agreement with the reported normative data (Hagen et al., 2010). Control groups from infertility clinics, recruiting non-PCOS women may be biased and the AMH levels may not reflect a general population.

There are several different kits for AMH analyses on the market dominated by Diagnostic Systems Lab (DSL) and Immunotec (IOT) assays. These two companies have recently merged and marketed a new commercially available assay, the AMH Gen II assay (Nelson and La, 2011). The new assay provides new analytical methods and reference levels. Accordingly, research results with previous and new AMH kits will be difficult to compare and interpret. Nevertheless, the main message of this study is the high sensitivity and specificity reported, which will be possible to replicate, if true. The consensus on AMH cut-off levels will be difficult to reach until an AMH standard assay has been established.

The strength of the study is that the population is fairly non-selected and comes from a well-defined geographic area even if the single center design may be considered a weakness. The women were invited based on prior birth history and gives the study a resemblance of retrospective design. However, the blood samples, the PCOS diagnoses and the clinical examination were collected, set and performed at inclusion as a cross-sectional case–control study. All participants included gave birth at some time in their past. This may be used as an argument against this being a fairly general population. On the other hand, most women in their mid-30s will belong to such a population.

The prevalence of PCOS is high in the present population, probably due to the findings of the original study with increased prevalence of PCOS among women with previous preterm birth (Eilertsen et al., 2012). However, the prevalence of PCOS in the control group is 14.2%. This is in accordance with the reported PCOS prevalence in general populations (March et al., 2010; Tehrani et al., 2011). Another possible weakness of the study is that some women using hormonal contraceptives are included in the study. This may on the other hand underestimate the prevalence of PCOS. When we excluded contraceptive users from our analyses, the results were essentially unchanged.

A possible selection bias is that all women included had a history of childbirth. Infertility is a major aspect of PCOS. PCOS women who become pregnant may have a more favorable form of the syndrome and this may be reflected in the results of the hormone analyses, including AMH. On the other hand, significantly more women with former infertility and pregnancy following ART are seen in the PCOS group than in the control group. This may balance this bias to some extent. In addition, we have shown previously that most

PCOS women who want to become pregnant, achieve pregnancy, with or without assistance (Vanky *et al.*, 2009).

In the present study PCOS women are significantly younger than non-PCOS women. PCOS features of the ovaries tend to be less prominent with age (Duijkers and Klipping, 2010) and may be the explanation of this. It has to be kept in mind that the AMH level also declines with age (Hagen *et al.*, 2010), reflecting the same changes. The use of PCOM versus the use of AMH when setting the PCOS diagnoses will probably not be affected by age.

In conclusion, PCOM can be replaced by AMH when diagnosing PCOS, both according to the PCOS-R criteria and the PCOS-AES criteria. Sensitivity and specificity is high even at low AMH levels. Future studies should use universally accepted methods for AMH measurements and international standards should be established. If a high sensitivity and specificity is confirmed by others, AMH may replace US examination of the ovaries in PCOS diagnosis.

Authors' roles

T.B.E., E.V. and S.M.C. contributed equally to the conception and design of the study, analyzes and interpretation of data. T.B.E. collected data and drafted the manuscript. E.V. and S.M.C. revised the manuscript and approved the final version.

Funding

Conflict of interest

None declared.

References

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, amanti-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.

Boomsma CM, Eijkemans MJ, Hughes EG, Visser GH, Fauser BC, Macklon NS. A meta-analysis of pregnancy outcomes in women with polycystic ovary syndrome. *Hum Reprod Update* 2006;**12**:673–683.

Broekmans FJ, Visser JA, Laven JS, Broer SL, Themmen AP, Fauser BC. Anti-Mullerian hormone and ovarian dysfunction. *Trends Endocrinol Metab* 2008;**19**:340–347.

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

Carlsen SM, Vanky E, Fleming R. Anti-Mullerian hormone concentrations in androgen-suppressed women with polycystic ovary syndrome. *Hum Reprod* 2009;**24**:1732–1738.

Chang WY, Knochenhauer ES, Bartolucci AA, Azziz R. Phenotypic spectrum of polycystic ovary syndrome: clinical and biochemical characterization of the three major clinical subgroups. *Fertil Steril* 2005; **83**:1717–1723.

De Meyts ER, Jorgensen N, Graem N, Muller J, Cate RL, Skakkebaek NE. Expression of anti-Mullerian hormone during normal and pathological gonadal development: association with differentiation of Sertoli and granulosa cells. *J Clin Endocrinol Metab* 1999;**84**:3836–3844.

de Vet A, Laven JS, de Jong FH, Themmen AP, Fauser BC. Antimullerian hormone serum levels: a putative marker for ovarian aging. *Fertil Steril* 2002;**77**:357–362.

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.

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* 2010;**26**:152–160.

Dunaif A, Thomas A. Current concepts in the polycystic ovary syndrome. *Annu Rev Med* 2001;**52**:401–419.

Durlinger AL, Visser JA, Themmen AP. Regulation of ovarian function: the role of anti-Mullerian hormone. *Reproduction* 2002;**124**:601–609.

Eilertsen T, Vanky E, Carlsen S. Increased prevalence of diabetes and polycystic ovary syndrome in women with a history of preterm birth: a case–control study. *BJOG* 2012;**119**:266–275.

Eldar-Geva T, Margalioth EJ, Gal M, Ben-Chetrit A, Algur N, Zylber-Haran E, Brooks B, Huerta M, Spitz IM. Serum anti-Mullerian hormone levels during controlled ovarian hyperstimulation in women with polycystic ovaries with and without hyperandrogenism. *Hum Reprod* 2005;**20**:1814–1819.

Fanchin R, Schonauer LM, Righini C, Guibourdenche J, Frydman R, Taieb J. Serum anti-Mullerian hormone is more strongly related to ovarian follicular status than serum inhibin B, estradiol, FSH and LH on day 3. *Hum Reprod* 2003;**18**:323–327.

Fauser BC, Tarlatzis BC, Rebar RW, Legro RS, Balen AH, Lobo R, Carmina E, Chang J, Yildiz BO, Laven JS *et al.* Consensus on women's health aspects of polycystic ovary syndrome (PCOS): the Amsterdam ESHRE/ASRM-Sponsored 3rd PCOS Consensus Workshop Group. *Fertil Steril* 2012;**97**:28–38.

Ferriman D, Gallwey J. Clinical assessment of body hair growth in women. *J Clin Endocrinol Metab* 1961;**21**:1440–1447.

Hagen CP, Aksglaede L, Sorensen K, Main KM, Boas M, Cleemann L, Holm K, Gravholt CH, Andersson AM, Pedersen AT *et al.* Serum levels of anti-Mullerian hormone as a marker of ovarian function in 926 healthy females from birth to adulthood and in 172 Turner syndrome patients. *J Clin Endocrinol Metab* 2010;**95**:5003–5010.

Hoeger KM. Role of lifestyle modification in the management of polycystic ovary syndrome. *Best Pract Res Clin Endocrinol Metab* 2006;**20**:293–310.

Knochenhauer ES, Key TJ, Kahsar-Miller M, Waggoner W, Boots LR, Azziz R. Prevalence of the polycystic ovary syndrome in unselected black and white women of the southeastern United States: a prospective study. *J Clin Endocrinol Metab* 1998;**83**:3078–3082.

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.

Laven JS, Mulders AG, Visser JA, Themmen AP, de Jong FH, Fauser BC. Anti-Mullerian hormone serum concentrations in normoovulatory and

anovulatory women of reproductive age. *J Clin Endocrinol Metab* 2004; **89**:318–323.

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

March WA, Moore VM, Willson KJ, Phillips DI, Norman RJ, Davies MJ. The prevalence of polycystic ovary syndrome in a community sample assessed under contrasting diagnostic criteria. *Hum Reprod* 2010;**25**:544–551.

Moran LJ, Misso ML, Wild RA, Norman RJ. Impaired glucose tolerance, type 2 diabetes and metabolic syndrome in polycystic ovary syndrome: a systematic review and meta-analysis. *Hum Reprod Update* 2010;**16**:347–363.

Nardo LG, Yates AP, Roberts SA, Pemberton P, Laing I. The relationships between AMH, androgens, insulin resistance and basal ovarian follicular status in non-obese subfertile women with and without polycystic ovary syndrome. *Hum Reprod* 2009;**24**:2917–2923.

Nelson SM, La MA. 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.

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.

Robinson S, Kiddy D, Gelding SV, Willis D, Niththyananthan R, Bush A, Johnston DG, Franks S. The relationship of insulin insensitivity to menstrual pattern in women with hyperandrogenism and polycystic ovaries. *Clin Endocrinol (Oxf)* 1993;**39**:351–355.

Roos N, Kieler H, Sahlin L, Ekman-Ordeberg G, Falconer H, Stephansson O. Risk of adverse pregnancy outcomes in women with polycystic ovary syndrome: population based cohort study. *BMJ* 2011; **343**:d6309.

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.

Sanchon R, Gambineri A, Alpanes M, Martinez-Garcia MA, Pasquali R, Escobar-Morreale HF. Prevalence of functional disorders of androgen excess in unselected premenopausal women: a study in blood donors. *Hum Reprod* 2012;**27**:1209–1216.

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.

Tehrani FR, Simbar M, Tohidi M, Hosseinpanah F, Azizi F. The prevalence of polycystic ovary syndrome in a community sample of Iranian population: Iranian PCOS prevalence study. *Reprod Biol Endocrinol* 2011;**9**:39.

van Rooij I, Broekmans FJ, te Velde ER, Fauser BC, Bancsi LF, de Jong FH, Themmen AP. Serum anti-Mullerian hormone levels: a novel measure of ovarian reserve. *Hum Reprod* 2002;**17**:3065–3071.

Vanky E, Backe B, Carlsen SM. Sex ratio and pregnancy complications according to mode of conception in women with polycystic ovary syndrome. *Acta Obstet Gynecol Scand* 2009;**88**:1261–1266.

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.

Wild RA, Carmina E, Amanti-Kandarakis E, Dokras A, Escobar-Morreale HF, Futterweit W, Lobo R, Norman RJ, Talbott E, Dumesic DA. Assessment of cardiovascular risk and prevention of cardiovascular disease in women with the polycystic ovary syndrome: a consensus statement by the Androgen Excess and Polycystic Ovary Syndrome (AE-PCOS) Society. *J Clin Endocrinol Metab* 2010; **95**:2038–2049.