Trends in Endocrinology & Metabolism

CellPress
REVIEWS

Review

Anti-Müllerian Hormone in PCOS: A Review Informing International Guidelines

Helena Teede,[1,*] Marie Misso,[1] Eliza C. Tassone,[1] Didier Dewailly,[2] Ernest Hy Ng,[3] Ricardo Azziz,[4] Robert J. Norman,[5] Marianne Andersen,[6] Stephen Franks,[7] Kathleen Hoeger,[8] Samantha Hutchison,[9] Sharon Oberfield,[10] Duru Shah,[11] Femke Hohmann,[12] Sasha Ottey,[13] Preeti Dabadghao,[14] and Joop S.E. Laven[15]

Polycystic ovary syndrome (PCOS) affects 8–13% of women. The Rotterdam diagnostic criteria include polycystic ovarian morphology (PCOM) on ultrasound, but given recognized challenges, serum anti-Müllerian hormone (AMH) is proposed as an alternative. To inform international PCOS guidelines, a systematic review was completed. Key identified gaps include large international studies in well-defined populations across the lifespan, clustering of AMH with PCOS features, relationships to long-term health outcomes, and improved quality, assay standardization, and sample handling, all needed to determine cut offs. Here we identify research priorities to address these gaps and enhance AMH utility in PCOS. Once issues are addressed, AMH levels could replace more costly and less accessible ultrasound in PCOS diagnosis.

Highlights

This systematic review investigates whether serum anti-Müllerian hormone (AMH) is an effective alternative for the detection of PCOM and/or diagnosis of PCOS.

There is significant heterogeneity in studies conducted in adolescents and adults, with a number of limitations identified.

Studies have lacked well-defined PCOS and control populations that varied across the lifespan, used inconsistent methods for defining cut offs, variably defined PCOM in comparator studies, and had methodological assay and sample handling challenges.

Challenges in Ultrasound PCOM Detection

PCOS is the most common endocrine disorder affecting women of reproductive age, with a reported prevalence of 8–13% [1–5]. The condition is heterogeneous [6] and women may present with reproductive, endocrine, metabolic, and psychosocial symptoms, which vary across their lifespan [7]. The Rotterdam criteria require that women fulfil two of the following three criteria to be diagnosed with PCOS: oligo- or anovulation, clinical and/or biochemical signs of hyperandrogenism, and/or polycystic ovaries on ultrasound [8–10], with the exclusion of other relevant disorders.

Within the diagnostic criteria, PCOM on ultrasonography is defined by either total ovarian volume or follicle number per ovary (FNPO). Original cut offs for PCOM were based on limited evidence [11] and were recently revised in the new international PCOS guidelines, while also highlighting the controversy and challenges with this criterion [1–4]. Determining FNPO is operator and equipment dependent, limiting accuracy and reproducibility. Equipment advances increase sensitivity and in turn FNPO counts [1–4]. Ultrasound involves expensive equipment and trained personnel, leading to increasing costs and impacting accessibility. The ultrasound approach (transabdominal or transvaginal) impacts accuracy, and in some women transvaginal ultrasound is unacceptable or may be perceived as invasive. Multifollicular appearance on ultrasound overlaps with PCOM diagnostic cut offs especially in adolescents, while in older women with PCOS cut-off values might be considerably lower [11]. Recent international PCOS guidelines now recommend against using ultrasound in PCOS diagnosis within 8 years of menarche and call for greater accuracy in PCOS diagnostic criteria worldwide [1–4].

AMH as a Potential Alternative to Ultrasound PCOM Detection

AMH is a polypeptide of the *transforming growth factor beta* (TGFβ) family, solely secreted by granulosa cells of the pre-antral and small antral ovarian follicles [12]. AMH has been shown in

[1]Monash Centre for Health Research and Implementation, School of Public Health and Preventive Medicine, Monash University and Monash Health, Locked Bag 29, Clayton, VIC 3168, Australia
[2]Faculté de Médecine, Université de Lille, Lille, France
[3]Department of Obstetrics and Gynaecology, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong Special Administrative Region, China
[4]Department of Health Policy, Management, and Behavior, School of Public Health, University at Albany, SUNY, Albany, NY 12144, USA
[5]Robinson Research Institute, University of Adelaide and Fertility SA, Adelaide, SA, Australia
[6]Department of Endocrinology, Odense University Hospital, Clinical Institute, University of Southern Denmark, 5000 Odense C, Denmark
[7]Imperial College London, Institute of Reproductive and Developmental Biology, Hammersmith Hospital, London W12 0NN, UK

animal models of PCOS to have a possible causal role in the development of the disorder through *in utero* exposure of the fetus to high AMH levels [13]. In women, AMH inhibits the recruitment of primordial follicles out of the resting oocyte pool and may suppress follicle-stimulating hormone (FSH) action contributing to ovulatory disturbances [14]. Overall, serum AMH levels are significantly higher in women with PCOS than in normal ovulatory women [15,16]. These data have led to the hypothesis that AMH could be a valuable surrogate marker or an alternative to ultrasound FNPO count for the detection of PCOM or in the overall diagnosis of PCOS [14].

Recognized challenges in the use of AMH measurement in PCOS include variations across the lifespan and problems with defining PCOM for comparison. AMH assays may also display differential responses to pre-analytical proteolysis, conformational changes of the AMH dimer, or the presence of interfering substances [17]. Appreciable sample-to-sample variability and substantial discrepancies in between-assay conversion factors suggest assay performance issues. These issues were prioritized and addressed in the recent International evidence-based guideline for the assessment and management of PCOS [1–4]. The aim of this systematic review is to investigate whether AMH is effective for the detection of PCOM and/or diagnosis of PCOS to inform international evidence-based guidelines in PCOS.

Methods

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement [18] and was prepared to inform recommendations in the updated and expanded evidence-based guideline for the assessment and management of PCOS [4].The methodology used for the development of this guideline is aligned with Australia's National Health and Medical Research Council (NHMRC) [19], the European Society of Human Reproduction and Embryology (ESHRE) [20], and the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) methodology [21] and is described in detail in the full guideline [4].

This systematic review addressed the evidence for the following two clinical questions: (i) Is AMH effective to diagnose PCOS?; and (ii) Is AMH effective to detect PCOM?

Systematic Search for Evidence

A systematic search strategy was designed to identify the best available evidence to answer the two clinical questions [22]. The search string comprised terms related to PCOS, PCOM, diagnosis, and AMH and was developed to retrieve articles addressing women with PCOS in all cultural, geographical, and socioeconomic backgrounds and settings. The search strategy was limited to English-language studies in humans and there were no limits on year of publication. A study design filter was not used.

Selection Criteria

The Population of interest, Intervention, Comparison, and Outcome (PICO) framework was used to guide the selection criteria for each clinical question presented in this systematic review and these were developed *a priori* by the multidisciplinary guideline development group [22]. These included reporting of results in the format of threshold, sensitivity, specificity, area under the curve, and precision.

Databases

The following electronic databases were searched on 26 June 2017: Medline (Ovid) – Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R), and Ovid OLDMEDLINE(R) 1950 to Present; EMBASE (Ovid); All EBM (Ovid) – including The Cochrane

[8]Department of Obstetrics and Gynecology, Division of Reproductive Endocrinology, University of Rochester, Rochester, NY 14642, USA
[9]Diabetes and Endocrine Units, Monash Health, Clayton, VIC 3168, Australia
[10]Department of Pediatrics, Division of Endocrinology, Diabetes, and Metabolism, Columbia University Medical Center, New York, NY 10032, USA
[11]Gynaecworld, Center for Women's Health and Fertility, Maharashtra 400036, India
[12]Huisartsenpraktijk Hohmann & De Vet, Rotterdam 3061, The Netherlands
[13]PCOS Challenge, Inc., 931 Monroe Drive, NE Suite A-470 Atlanta, GA 30308, USA
[14]Department of Endocrinology, Sanjay Gandhi Postgraduate Institute of Medical Sciences, Raebareli Road, Lucknow 226014, India
[15]Division of Reproductive Medicine, Department of Obstetrics and Gynecology, Erasmus MC University Medical Center, Rotterdam, The Netherlands

*Correspondence: Helena.Teede@monash.edu (H. Teede).

Database of Systematic Reviews, DARE, CENTRAL, and ACP Journal Club; and PsycInfo (Ovid) and CINAHL.

Evidence Processing

Studies were selected and appraised by one highly experienced reviewer (M.M.) in consultation with colleagues using study selection criteria [22] established *a priori*. The retrieved articles were first reviewed by title and abstract, and then full articles were retrieved for further assessment if the information given suggested that the study met the inclusion criteria.

Assessment of Methodological Quality

The methodological quality (i.e., risk of bias) of each of the included studies was assessed by one reviewer for the adolescent studies (E.B.) and one reviewer for the adult studies (E.C.T.), using a critical appraisal template developed *a priori* [23]. Individual quality items were investigated using a descriptive component approach that assessed attrition bias, reporting bias, selection bias, performance bias, potential confounding, and the appropriateness of the statistical analysis. Any disagreement or uncertainty was resolved by a discussion with a third reviewer (M.M.) and within the team of authors of this manuscript. Using this approach each study was allocated a risk-of-bias rating of low, moderate, or high.

Data Extraction

Data were extracted directly into customized tables for the characteristics of the included studies and results by one reviewer (M.M.). Information was extracted on general study characteristics (lead author, year of publication, study design, country), participants [number, age category (adolescents or adults), body mass index (BMI), AMH, PCOS diagnostic criteria, medication status], and diagnostic accuracy results (threshold, sensitivity, specificity, area under the curve, and precision). Due to the timeline-intensive nature of conducting evidence synthesis for an international guideline, authors were not contacted in instances of missing data or for data conversions.

Data Synthesis

Due to the heterogeneity in diagnostic criteria and/or threshold/cut-off values, meta-analyses (for pooled sensitivity and specificity estimates) have not been performed and thus the study data are presented narratively and in tabular form. True- and false-positive, and true- and false-negative, values for the diagnostic accuracy of AMH for PCOS and PCOM were calculated in Review Manager 5.3 using the sensitivity and specificity data extracted from included studies (M.M. and E.C.T.). AMH data presented as ng/ml were converted to SI units, pmol/l (conversion factor of 7.1429).

Results

A total of 313 potentially relevant studies were identified in the electronic database search, of which 41 duplicates were excluded. The remaining 272 articles were reviewed by title and abstract and 230 were excluded. Forty-two articles were retrieved for full-text screening, of which 29 studies [14,24–51] addressed the diagnostic accuracy of AMH for PCOS and/or PCOM and thus met the inclusion criteria for the clinical questions presented in this review, while 13 full-text articles were excluded (Figure 1). A table of the excluded studies with reasons for their exclusion can be found in section 1.5 of the technical report for the international evidence-based guideline for the assessment and management of polycystic ovary syndrome [22].

One of the 29 studies identified was a systematic review [32] and included nine of the studies identified here. However, it also included studies that did not meet the inclusion criteria for this evidence review and was missing additional studies published more recently that were identified by this review's search; therefore, it was not used in this systematic review.

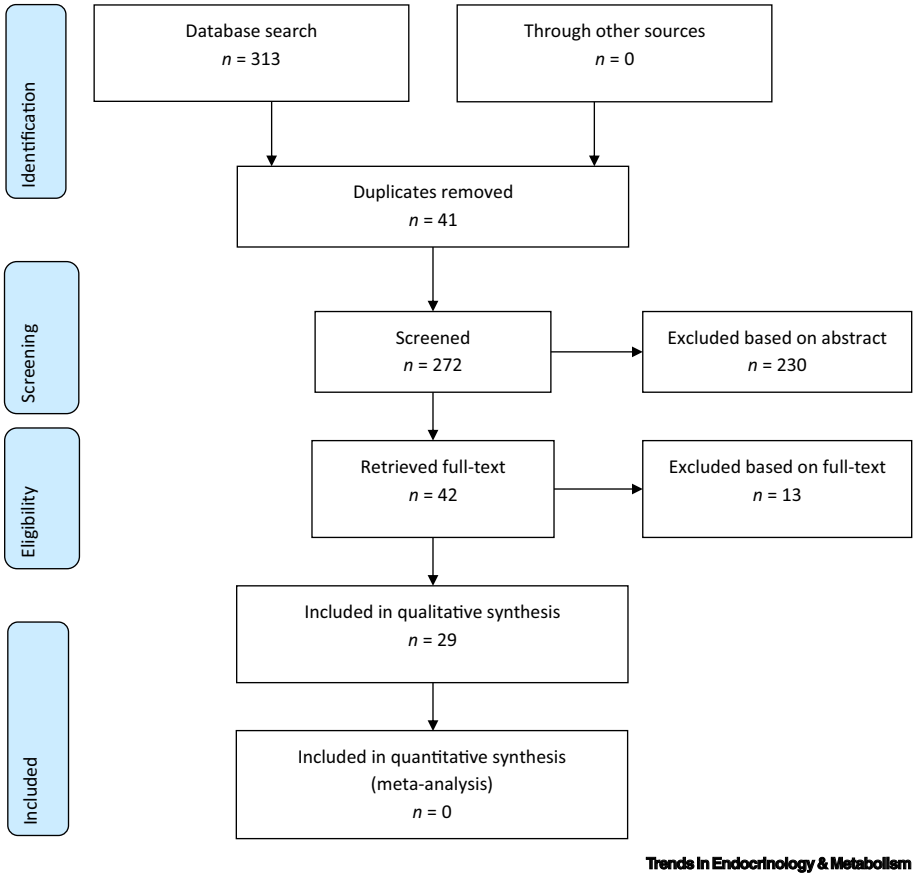

Figure 1. PRISMA Flow Diagram

Characteristics of Included Studies

Table S1 in the supplemental information online includes the key characteristics of the included studies, with four addressing the diagnostic accuracy of AMH for PCOS and PCOM [29,30,36,41] and one addressing PCOM only [46]. Of the 28 studies, six studies included adolescent participants for diagnosis of PCOS [30,34,43,44,46,49] and one of these addressed PCOS and PCOM [30]. The remaining 21 studies [14,24–29,31,35–42,45,47,48,50,51] included adult participants for diagnosis of PCOS, where three of these addressed PCOS and PCOM [29,36,41]; the remaining 18 studies addressed PCOS alone [14,24–28,31,35,37–40,42,45,47,48,50,51]. Of the studies in adolescents, one was in overweight and obese participants [33] and in one study BMI was unclear [49]. Of the studies in adults, one included lean and obese participants [25] and five studies [24,29,35,50,51] included overweight and obese participants.

Participant numbers ranged from 31 to 633 participants for adolescents and from 44 to 606 for adults. The studies were conducted across a range of settings including university departments, outpatient hospital clinics, and laboratories, in countries including Australia, Indonesia, South Korea, Iran, Chile, USA, Turkey, Italy, Taiwan, Croatia, France, Norway, UK, Germany, Denmark, China, and India.

Quality Appraisal of Included Studies

The six studies that included adolescent participants ranged in quality from low to high risk of bias, while the majority of adult studies were at high risk of bias [22]. Reasons for these ratings

include: selection criteria were not explicitly stated; it was unclear whether participants were entered into the study appropriately (randomly or consecutively); case–control design; inclusion of PCOM cases among controls; and inadequacies around the application of index and reference tests – in particular, suboptimal choice about the best compromise between sensitivity and specificity by receiver operating characteristic (ROC) curve analysis. Moderate or high risk of bias was noted in the interpretation of the results.

Diagnostic Accuracy of AMH for PCOS

In adolescents, there were five studies, of which one was found to have a low risk of bias [30], two were of moderate risk of bias [33,34,44], and two were of high risk of bias [43,49], demonstrating areas under the ROC curve of AMH for the diagnosis of PCOS, ranging from 0.5 to 0.88 (Table 1); the threshold cut-off values ranged from 25 to 44 pmol/l.

In adults, there were 21 studies, of which five were found to have a moderate risk of bias [27, 39–41,47] and 16 were of high risk of bias [14,24–26,28,29,31,35–38,42,45,48,50,51], demonstrating areas under the ROC curve of AMH for the diagnosis of PCOS ranging from 0.66 to 0.994 (Table 1); the threshold cut-off values ranged from 10 to 57 pmol/l. Although mean serum AMH levels in adolescent and adult PCOS women were significantly higher than those of non-PCOS participants in all studies, there was significant overlap between the cases and controls. The sensitivity, specificity, and AUC was generally higher in adults than in adolescents, acknowledging that the evidence is of limited quality and that study populations varied widely across studies in terms of recruitment and definitions of both PCOS and control populations.

Diagnostic Accuracy of AMH for PCOM

In adolescents, there was one study of low risk of bias demonstrating an area under the ROC of AMH for the diagnosis of PCOM of 0.87 [46] (Table 2); the threshold cut-off value was 50 pmol/l. In adults, there were four relevant studies, one of which was found to have a low risk of bias [30], one of moderate risk of bias [41], and two of high risk of bias [29,36], demonstrating areas under the ROC of AMH for the diagnosis for PCOM of 0.67 to 0.92 (Table 2). The threshold ranged from 20 to 30 pmol/l. Although serum AMH levels in adolescent and adult PCOM women are significantly higher than those of non-PCOM counterparts in all studies, there is significant overlap between cases and controls.

Identified Gaps in the AMH Literature

This systematic review presents rigorous synthesis of peer-reviewed literature assessing whether AMH is effective for the detection of PCOM and diagnosis of PCOS, in both adolescents and adults, with results informing the international guideline on assessment and management of PCOS. The 28 included studies were rated, with the majority having a moderate or high risk of bias. Heterogeneity was significant with identified challenges including poorly defined study populations, variation across the lifespan, ill-defined approaches to AMH cut offs and challenges with aligning with PCOM, and assay evolution and technical challenges.

The systematic review revealed significant heterogeneity in the accuracy of AMH in reflecting PCOM and in assisting the diagnosis of PCOS. Key contributors to this heterogeneity include the inappropriate selection of participants and the lack of well-defined study populations (those with or without PCOS or features of PCOS in the control populations). It is crucial that participants are entered into studies based on explicit, well-defined, and transparent selection criteria. Study populations need to be generalizable and ideally community recruited, rather than from high-risk subgroups including those presenting with infertility. Comparators or controls need to be very clearly and consistently defined. Entrance to the studies needs to be either random or consecutive and studies need to be adequately powered to detect the specified outcome. The majority of

Table 1. Diagnostic Accuracy of AMH for PCOS[a]

Study ID	Threshold	Diagnostic criteria[b]	PCOS	Non-PCOS	Sensitivity	Specificity	True positive	False positive	True negative	False negative	AUC	Precision
Adolescents												
Hart 2010 [30]	30 pmol/l	Rotterdam	64	149	53.1	69.8	34	45	104	30	0.64	CI = 0.55–0.72 $P = 0.002$
	30 pmol/l	NIH	36	177	52.8	66.1					0.61	CI = 0.49–0.72 $P = 0.048$
Kim 2016 and 2017 [33,34]	44.71 pmol/l	NIH	46	43	67	81					0.788	0.687–0.868 $P < 0.000\ 1$
Sopher 2014 [43]	24.29 pmol/l	NIH	15	16	40	93.8					NR	NR
Tokmak 2015 [44]	100 pmol/l	Rotterdam Youden index	43	47	48.8	77.1					0.579	0.453–0.705 $P = 0.198$
Yetim 2016 [49]	43.57 pmol/l	Rotterdam	53	26	81.1	92.3	43	2	24	10	0.88	CI = 0.80–0.96 $P < 0.001$
Adults												
Carmina 2016 [24]	>33.57 pmol/l	Rotterdam	113	47	79	96	89	2	45	24	0.952	SD = 0.014
	>33.57 pmol/l	A and B	78	47	91	96					0.982	SD = 0.002
	>33.57 pmol/l	C	20	47	50	96					NR	NR
	33.57 pmol/l	D >	15	47	53	96					NR	NR
Casadei 2013 [51]	33 pmol/l	NIH	22	22	95	95					0.970	CI = 0.02–0.92
Cassar 2014 [25]	>30 pmol/l	Rotterdam	43	35	82	79	35	7	28	8	0.829	CI = 0.736–0.923 $P < 0.001$
Chao 2012 [26]	25pmol/l	Rotterdam	31	24	74	79	23	5	19	8	NR	NR
Dewailly 2014 [27]	28 pmol/l	Rotterdam	95	521	84.2	97.5	80	13	508	15	0.948	CI = 0.915–0.982
	28 pmol/l	HA + PCOM	67	521	61.2	97.5					0.894	CI = 0.852–0.936
	28 pmol/l	OA + PCOM	110	521	81.8	97.5					0.938	CI = 0.908–0.969
Dewailly 2011 [28]	35 pmol/l	Rotterdam	62	66	92	97	57	2	64	5	0.973	CI = 0.947–0.998
Eilertsen 2012 [29]	10 pmol/l	Rotterdam	56	206	98.2	94.8	55	11	195	1	0.992	CI = 0.986–0.999
	20 pmol/l	AES	44	218	95.5	97.2					0.994	CI = 0.987–1.000
Homburg 2013 [31]	48 pmol/l	Rotterdam	90	90	60	98.2	54	2	88	36	0.805	NR

Köninger 2014 [35]	25 pmol/l	Rotterdam mild	21	48	71.4	89.6	15	5	43	6	0.80	CI = 0.65–0.91
	25 pmol/l	Rotterdam severe	59	48	84.7	89.6	50	5	43	9	0.88	CI = 0.80–0.95
Lauritsen 2014 [36]	18 pmol/l	Rotterdam	74	373	91.8	98.1	68	7	366	6	0.994	CI = 0.990–0.999
Li 2010 [37]	57.14 pmol/l (8 ng/ml)	Rotterdam	47	40	61.7	70	29	12	28	18	0.664	CI = 0.551–0.778
Li 2012 [38]	28 pmol/l	Rotterdam	131	61	65	62	85	23	38	46	0.68	CI = 0.60–0.76 $P < 0.01$
	30.21 pmol/l	HA+	62	61	82	64					0.82	CI = 0.72–0.92 $P < 0.01$
	26.86 pmol/l	HA-	69	61	64	62					0.66	CI = 0.56–0.75 $P < 0.01$
Pigny 2006 [14]	60 pmol/l	Rotterdam	73	96	67	92	49	8	88	24	0.851	CI = 0.796–0.905
Pigny 2016 [39]	57.28 pmol/l	Rotterdam equivalent	47	48	74.5	91.7	35	4	44	12	0.944	CI = 0.901–0.987
Sahmay 2013 [40]	28.14 pmol/l	Rotterdam	419	151	80	89.8	335	15	136	84	0.916	CI = 0.897–0.935 $P < 0.000\ 1$
Sahmay 2014 [41]	27.14 pmol/l	AES	195	411	80	80.2					0.87	0.84–0.90 $P < 0.001$
	27.14 pmol/l	Rotterdam	228	378	81.6	85.1	186	56	322	42	0.89	0.87–0.92 $P < 0.001$
	27.14 pmol/l	NIH	164	442	80.7	74.7					0.86	0.82–0.89 $P < 0.001$
Saikumar 2013 [42]	23.86 pmol/l	Rotterdam	60	60	98	93	59	4	56	1	0.956	NR
Tremellen 2015 [45]	≥re pmol/l	Rotterdam	43	113	83.7	82.3	36	20	93	7	0.917	NR
Wiweko 2014 [47]	31.79 pmol/l	Rotterdam	71	71	76.1	74.6	54	18	53	17	0.870	CI = 0.81–0.92
Woo 2012 [48]	55.86 pmol/l	Rotterdam	87	53	75.9	86.8	66	7	46	21	0.868	CI = 0.801–0.919
Zadehmodarres 2015 [50]	22.5 pmol/l	Rotterdam	60	57	70.37	77.36	42	13	44	18	NR	NR

[a]Phenotype A, anovulation, hyperandrogenism, and PCO; Phenotype B, ANOV-PCOS, anovulatory with hyperandrogenism and normal ovaries; Phenotype C, OV-PCOS, ovulatory with normal menses, hyperandrogenism, and PCO; Phenotype D, NH-PCOS, anovulatory with normal androgen levels and no symptoms of hyperandrogenism and PCO; PM, PCOS mild, PCO + OA; PS, PCOS severe, all three criteria.
[b]See table of characteristics for definition.

Table 2. Diagnostic Accuracy of AMH for PCOM

Study ID	Threshold	Diagnostic criteria	PCOS	Non-PCOS	Sensitivity	Specificity	True positive	False positive	True negative	False negative	AUC	Precision
Villarroel 2011 [46]	50.25 pmol/l	Rotterdam	25	49	84.0	83.7	21	8	41	4	0.873	CI = 0.782–0.963 $P < 0.000\ 1$
Eilertsen 2012 [29]	20 pmol/l	Rotterdam	113	149	79.6	72.5	90	41	108	23	0.896	CI = 0.855–0.937
Hart 2010 [30]	30 pmol/l	Rotterdam	75	132	54.7	72.7	41	36	96	34	0.67	CI = 0.60–0.75 $P < 0.001$
Lauritsen 2014 [36]	20 pmol/l	Rotterdam	74	373	82.0	84.6	61	57	316	13	0.906	CI = 0.878–0.933
Sahmay 2014 [41]	27.14 pmol/l	Unclear	Unclear	Unclear	83	87					0.92	CI = 0.90–0.93 $P < 0.001$

available studies fail to fulfil these criteria, leading to a moderate-to-high risk of bias and poor reliability. This needs to be addressed before progress can be made in understanding the role of AMH assays in PCOS.

Follicle development varies across the lifespan and is increased in adolescence, falling subsequently until menopause, when oocytes are depleted. There is a need for age-specific cut offs for both PCOM and AMH. Here, the sensitivity, specificity, and area under the ROC curve suggest greater accuracy of AMH in PCOS diagnosis in adults than in adolescents and it may be that the role of AMH in PCOS diagnosis will align with that of PCOM. The new international guidelines now recommend against the use of ultrasound in the diagnosis of PCOS until 8 years post-menarche (Box 1) [1–4]; however, more research is needed to determine age-specific cut offs and acceptable accuracy at given life stages. Given that AMH is also not appropriate for diagnosis in adolescents or adults at present, both hyperandrogenism and ovulatory dysfunction are currently required for diagnosis in adolescents.

Another key challenge with the literature is the significant variably in the way the cut-off values were defined. Traditionally in determining cut-off values in biochemical tests as 'normal' range, a cut-off of the 95th centile is applied to deliver 95% specificity. However, this is not appropriate for defining diagnostic cut offs for a clinical condition. Here, more important considerations include clustering with other clinical features such as hirsutism, hyperandrogenism, and oligo-anovulation or prediction of long-term health outcomes such as fertility. For example, the establishment of gestational diabetes, hypertension, or obesity cut offs was based on long-term health risks, not simply percentiles [52–54]. In the case of AMH, the majority of studies defined the cut offs at the 95th centile, which is not a valid biological cut off. Further research on clustering of AMH with other features of PCOS and the relationship between AMH and long-term health outcomes is now vital.

Other considerations were the significant variability in follicle numbers and development, in PCOM and in AMH across the lifespan. Levels are high in adolescence and overlap considerably with those who do not have other features of PCOS. This makes it difficult to differentiate PCOS from controls on AMH levels [30]. Levels fall in later life, especially after menopause [55]. Age-specific reference ranges are thus vital [56] and it is likely that, aligned with PCOM as a diagnostic feature of PCOS, AMH will be of most use where overlap is least notable, beyond the early post-menarche years.

Box 1. Ultrasound and PCOM Recommendations: International Evidence-Based Guideline [1–4]

Ultrasound should not be used for the diagnosis of PCOS in those with a gynecological age of <8 years (<8 years after menarche), due to the high incidence of multifollicular ovaries in this life stage [clinical consensus recommendation (CCR)].

- The threshold for PCOM should be revised regularly with advancing ultrasound technology, and age-specific cut-off values for PCOM should be defined (CCR).
- The transvaginal ultrasound approach is preferred in the diagnosis of PCOS, if sexually active and if acceptable to the individual being assessed (CCR).
- Using endovaginal ultrasound transducers with a frequency bandwidth that includes 8 MHz, the threshold for PCOM should be on either ovary, a follicle number per ovary of ≥20 and/or an ovarian volume ≥10 ml, ensuring no corpora lutea, cysts, or dominant follicles are present (CCR).
- If using older technology, the threshold for PCOM could be an ovarian volume ≥10 ml on either ovary [clinical practice point (CPP)].
- In patients with irregular menstrual cycles and hyperandrogenism, an ovarian ultrasound is not necessary for PCOS diagnosis; however, ultrasound will identify the complete PCOS phenotype (CPP).
- In transabdominal ultrasound, reporting is best focused on ovarian volume with a threshold of ≥10 ml, given the difficulty of reliably assessing follicle number with this approach (CPP).
- Clear protocols are recommended for reporting follicle number per ovary and ovarian volume on ultrasound. Recommended minimum reporting standards include:
  - Last menstrual period
  - Transducer bandwidth frequency
  - Approach/route assessed
  - Total follicle number per ovary measuring 2–9 mm
  - Three dimensions and volume of each ovary
  - Reporting of endometrial thickness and appearance is preferred – three-layer endometrial assessment may be useful to screen for endometrial pathology
  - Other ovarian and uterine pathology, as well as ovarian cysts, corpus luteum, dominant follicles ≥10 mm (CPP).
- There is a need for training in careful and meticulous follicle counting per ovary, to improve reporting (CPP).

The relationship between AMH with PCOM was also an important consideration (Box 1). Investigators have used the PCOS definition established in 2003 at the Rotterdam conference [10] (i.e., 12 follicles of 2–9 mm diameter per ovary) to define this PCOS diagnostic criterion. This cut off suffers from the same challenges as applying the 95th-centile cut offs to define PCOM and is highly variable by life stage and dependent on advancing ultrasound equipment. Therefore, with the latest ultrasound equipment, the new international guidelines have redefined the PCOM cut offs to a threshold of ≥20 FNPO and have specified that ultrasound-defined PCOM is no longer appropriate in PCOS diagnosis within 8 years post-menarche, given the overlap between PCOS and controls [1–4]. Similar challenges in defining PCOM (cut offs at the 95th centile, changes across the life stage, and technical challenges) mandate further research on clustering of PCOM with other features of PCOS and the relationship between PCOM and long-term health outcomes.

In addition, there are technical issues regarding the assays for serum AMH, leading to further heterogeneity in results. About one-half of the studies were performed using either the Diagnostic Systems Lab (DSL) or Immunotech (IOT) assays, for which concordance in values is problematic. Furthermore, these assays are not marketed anymore. There are very few data with the new automated-platform assays [39]. There is rising awareness on the impact of sample handling, transport, and storage conditions, factors that are under-reported in the literature. There is also a clear need for an international reference standard for AMH and for robust independent evaluation of commercial assays in routine clinical samples with well-defined sample handling and processing protocols [17]. Overall there is an urgent need for international standardization to improve comparability among assays, and the challenge of determining the optimal assay and the issues concerning sample storage and processing need to be addressed before clinical utility can be recommended (Box 2) [1–4].

Box 2. AMH Recommendations: International Evidence-Based Guideline [1–4]

- Serum AMH levels should not yet be used as an alternative for the detection of PCOM or as a single test for the diagnosis of PCOS [evidence-based recommendation (EBR)].
- There is emerging evidence that with improved standardization of assays and established cut-off levels or thresholds based on large-scale validation in populations of different ages and ethnicities, AMH assays will be more accurate in the detection of PCOM (CPP).

Future Steps for AMH in PCOS

- PCOM needs to be consistently defined and follow international guidelines to allow comparison with AMH levels.
- The inclusion of controls with PCOM should be avoided, as mentioned above. This requires a particular statistical approach (cluster analysis). Age-stratified thresholds need to be defined.
- Standardized optimal assays need to be applied.
- AMH is a potential future substitute for detecting PCOM; however, further research is needed including establishing universal threshold for elevated serum AMH level that requires validation in large populations of different ages and ethnicities.

Outstanding Questions

Consistently defined and appropriate study and control populations and biologically relevant cut-off values that reflect clustering of clinical features and are relevant to health outcomes, are life-stage specific, and more clearly define PCOM on ultrasound, with improved accuracy and standardization of assays and handling procedures, are needed in future studies.

How can we improve accuracy and standardisation of assays and handling procedures?

Is there an optimal assay?

What are the biologically relevant AMH diagnostic cut off values that reflect clustering with other clinical features, and that are life-stage specific?

What are the AMH values that accurately predict PCOM on ultrasound?

Do AMH levels have a role in predicting clinical features and guideline treatment of PCOS, outside of diagnosis?

Limitations

A single protocol document for all 40 systematic reviews completed as part of the international PCOS guideline was developed and signed off by all 70 guideline development group expert, consumer, and health professional members. These protocols are publicly available at https://www.monash.edu/__data/assets/pdf_file/0020/1412282/PCOS-Guideline_Technical-report.pdf; however, each individual protocol was not registered. This review was limited to studies published in English, thus putting the review at risk of language bias. Also, we did not contact study authors for missing information or data conversions.

Concluding Remarks

AMH may play a key role in the pathogenesis of PCOS; however, key issues must be addressed before it can be applied clinically to the detection of PCOM or in the diagnosis of PCOS (see Outstanding Questions). These include consistently defined and appropriate study and control populations, biologically relevant cut-off values that reflect clustering of clinical features and are relevant to health outcomes, are life-stage specific, and more clearly define PCOM on ultrasound, and improved accuracy and standardization of assays and handling procedures. With improved standardization of emerging assays and established internationally approved cut-off levels/thresholds based on large-scale validation in defined populations of different ages, AMH may become useful in the clinical detection of PCOM and the diagnosis of PCOS. However, until these issues are addressed, AMH is not clinically applicable and useful in detecting PCOM or diagnosing PCOS and is not recommended outside research in the new international evidence-based guidelines for the assessment and management of PCOS [1–4].

Author Contributions

M.M. with input from all authors designed the search strategy. M.M. ran the database searches, screened articles, selected articles, performed data extraction, performed data conversions, completed the statistical analyses, and contributed to the write up of the manuscript. E.C.T. critically appraised articles and contributed to the write up of the manuscript. H.T. contributed to the write up of the manuscript. All authors assisted in interpretation of the synthesized literature, critically revised the manuscript, and approved the final version for submission.

Acknowledgments

We thank Estifanos Baye for performing the critical appraisals of the adolescent studies included in this review. The guideline and associated systematic reviews were primarily funded by the Australian National Health and Medical Research Council of Australia (NHMRC) supported by a partnership with the European Society of Human Reproduction and Embryology and the

American Society for Reproductive Medicine. Guideline development group members did not receive payment. Travel expenses were covered by the sponsoring organizations.

Disclaimer Statement

H.T. received NHMRC grant funding to the institution to undertake this work. R.A. is a consultant for Medtronics, Ansh Labs, Spruce Biosciences, and Longitude Capital and on the advisory board of Martin PET Imaging. J.S.E.L. has received grants from Ferring and Euroscreen.

Supplemental Information

Supplemental information associated with this article can be found online at https://doi.org/10.1016/j.tem.2019.04.006

References

1. Teede, H.J. *et al.* (2018) Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Hum. Reprod.* 33, 1602–1618
2. Teede, H.J. *et al.* (2018) Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Fertil. Steril.* 110, 364–379
3. Teede, H.J. *et al.* (2018) Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Clin. Endocrinol.* 89, 251–268
4. Centre for Research Excellence in Polycystic Ovary Syndrome (2018) *International Evidence-Based Guidelines for the Assessment and Management of Polycystic Ovary Syndrome*, Monash University
5. Bozdag, G. *et al.* (2016) The prevalence and phenotypic features of polycystic ovary syndrome: a systematic review and meta-analysis. *Hum. Reprod.* 31, 2841–2855
6. Sirmans, S.M. and Pate, K.A. (2013) Epidemiology, diagnosis, and management of polycystic ovary syndrome. *Clin. Epidemiol.* 6, 1–13
7. Teede, H. *et al.* (2010) Polycystic ovary syndrome: a complex condition with psychological, reproductive and metabolic manifestations that impacts on health across the lifespan. *BMC Med.* 8, 41
8. Teede, H.J. *et al.* (2011) Assessment and management of polycystic ovary syndrome: summary of an evidence-based guideline. *Med. J. Aust.* 195, S65–S112
9. Legro, R.S. *et al.* (2013) Diagnosis and treatment of polycystic ovary syndrome: an Endocrine Society Clinical Practice Guideline. *J. Clin. Endocrinol. Metab.* 98, 4565–4592
10. The Rotterdam ESHRE/ASRM Sponsored PCOS Consensus Workshop Group (2004) Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum. Reprod.* 19, 41–47
11. Balen, A.H. *et al.* (2003) Ultrasound assessment of the polycystic ovary: international consensus definitions. *Hum. Reprod. Update* 9, 505–514
12. Durlinger, A.L. *et al.* (2002) Regulation of ovarian function: the role of anti-Müllerian hormone. *Reproduction* 124, 601–609
13. Tata, B. *et al.* (2018) Elevated prenatal anti-Müllerian hormone reprograms the fetus and induces polycystic ovary syndrome in adulthood. *Nat. Med.* 24, 834–846
14. Pigny, P. *et al.* (2006) Serum anti-Müllerian hormone as a surrogate for antral follicle count for definition of the polycystic ovary syndrome. *J. Clin. Endocrinol. Metab.* 91, 941–945
15. Cook, C.L. *et al.* (2002) Relationship between serum Müllerian-inhibiting substance and other reproductive hormones in untreated women with polycystic ovary syndrome and normal women. *Fertil. Steril.* 77, 141–146
16. Seifer, D.B. and MacLaughlin, D.T. (2007) Müllerian inhibiting substance is an ovarian growth factor of emerging clinical significance. *Fertil. Steril.* 88, 539–546
17. Rustamov, O. *et al.* (2014) The measurement of anti-Müllerian hormone: a critical appraisal. *J. Clin. Endocrinol. Metab.* 99, 723–732
18. Moher, D. *et al.* (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 339, b2535
19. National Health and Medical Research Council (2007) *NHMRC Standards and Procedures for Externally Developed Guidelines*, National Health and Medical Research Council
20. Vermeulen, N. *et al.* (2014) *Manual for ESHRE Guideline Development*, European Society for Human Reproduction and Embryology
21. The GRADE Working Group (2009) *GRADE Handbook for Grading Quality of Evidence and Strength of Recommendation* (Version 3.2, updated March 2009), GRADE
22. Anon (2018) *International Evidence-Based guideline for the Assessment and Management of Polycystic Ovary Syndrome. Section 1.5*, Monash University
23. Centre for Clinical Effectiveness (2010) *Critical Appraisal Templates*, Southern Health
24. Carmina, E. *et al.* (2016) AMH measurement versus ovarian ultrasound in the diagnosis of polycystic ovary syndrome in different phenotypes. *Endocr. Pract.* 22, 287–293
25. Cassar, S. *et al.* (2014) Polycystic ovary syndrome and anti-Müllerian hormone: role of insulin resistance, androgens, obesity and gonadotrophins. *Clin. Endocrinol.* 81, 899–906
26. Chao, K.C. *et al.* (2012) Anti-Müllerian hormone serum level as a predictive marker of ovarian function in Taiwanese women. *J. Chin. Med. Assoc.* 75, 70–74
27. Dewailly, D. *et al.* (2014) Using cluster analysis to identify a homogeneous subpopulation of women with polycystic ovarian morphology in a population of non-hyperandrogenic women with regular menstrual cycles. *Hum. Reprod.* 29, 2536–2543
28. Dewailly, D. *et al.* (2011) 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.* 26, 3123–3129
29. Eilertsen, T.B. *et al.* (2012) Anti-Müllerian hormone in the diagnosis of polycystic ovary syndrome: can morphologic description be replaced? *Hum. Reprod.* 27, 2494–2502
30. Hart, R. *et al.* (2010) Serum antiMüllerian hormone (AMH) levels are elevated in adolescent girls with polycystic ovaries and the polycystic ovarian syndrome (PCOS). *Fertil. Steril.* 94, 1118–1121
31. Homburg, R. *et al.* (2013) The relationship of serum anti-Müllerian hormone with polycystic ovarian morphology and polycystic ovary syndrome: a prospective cohort study. *Hum. Reprod.* 28, 1077–1083
32. Iliodromiti, S. *et al.* (2013) Can anti-Müllerian hormone predict the diagnosis of polycystic ovary syndrome? A systematic review and meta-analysis of extracted data. *J. Clin. Endocrinol. Metab.* 98, 3332–3340
33. Kim, J.Y. *et al.* (2017) Anti-Müllerian hormone in obese adolescent girls with polycystic ovary syndrome. *J. Adolesc. Health* 60, 333–339
34. Kim, J.Y. *et al.* (2016) Anti-Müllerian hormone (AMH) in obese adolescent girls with polycystic ovary syndrome (PCOS): cross-sectional and treatment-associated longitudinal changes. 98th Annual Meeting and Expo of the Endocrine Society, ENDO, 2016. *Endocr. Rev.* 37
35. Koninger, A. *et al.* (2014) Anti-Müllerian hormone: an indicator for the severity of polycystic ovarian syndrome. *Arch. Gynecol. Obstet.* 290, 1023–1030

36. Lauritsen, M.P. *et al.* (2014) The prevalence of polycystic ovary syndrome in a normal population according to the Rotterdam criteria versus revised criteria including anti-Müllerian hormone. *Hum. Reprod.* 29, 791–801
37. Li, L. *et al.* (2010) Elevated serum anti-Müllerian hormone in adolescent and young adult Chinese patients with polycystic ovary syndrome. *Wien. Klin. Wochenschr.* 122, 519–524
38. Li, Y. *et al.* (2012) Different diagnostic power of anti-Müllerian hormone in evaluating women with polycystic ovaries with and without hyperandrogenism. *J. Assist. Reprod. Genet.* 29, 1147–1151
39. Pigny, P. *et al.* (2016) Comparative assessment of five serum antiMüllerian hormone assays for the diagnosis of polycystic ovary syndrome. *Fertil. Steril.* 105, 1063–1069.e3
40. Sahmay, S. *et al.* (2013) Elevated serum levels of anti-Müllerian hormone can be introduced as a new diagnostic marker for polycystic ovary syndrome. *Acta Obstet. Gynecol. Scand.* 92, 1369–1374
41. Sahmay, S. *et al.* (2014) Diagnosis of polycystic ovary syndrome: AMH in combination with clinical symptoms. *J. Assist. Rep. Genet.* 31, 213–220
42. Saikumar, P. *et al.* (2013) Anti Müllerian hormone: a potential marker for recruited non growing follicle of ovarian pool in women with polycystic ovarian syndrome. *J. Clin. Diagn. Res.* 7, 1866–1869
43. Sopher, A.B. *et al.* (2014) Anti-Müllerian hormone may be a useful adjunct in the diagnosis of polycystic ovary syndrome in nonobese adolescents. *J. Pediatr. Endocrinol. Metab.* 27, 1175–1179
44. Tokmak, A. *et al.* (2015) Is anti-Müllerian hormone a good diagnostic marker for adolescent and young adult patients with polycystic ovary syndrome? *Turk. J. Obstet. Gynecol.* 12, 199–204
45. Tremellen, K. and Zander-Fox, D. (2015) Serum anti-Müllerian hormone assessment of ovarian reserve and polycystic ovary syndrome status over the reproductive lifespan. *Austr. N. Z. J. Obstet. Gynaecol.* 55, 384–389
46. Villarroel, C. *et al.* (2011) Polycystic ovarian morphology in adolescents with regular menstrual cycles is associated with elevated anti-Müllerian hormone. *Hum. Reprod.* 26, 2861–2868
47. Wiweko, B. *et al.* (2014) Anti-Müllerian hormone as a diagnostic and prognostic tool for PCOS patients. *J. Assist. Reprod. Genet.* 31, 1311–1316
48. Woo, H.Y. *et al.* (2012) Differences of the association of anti-Müllerian hormone with clinical or biochemical characteristics between women with and without polycystic ovary syndrome. *Endocr. J.* 59, 781–790
49. Yetim, A. *et al.* (2016) Anti-Müllerian hormone and inhibin-A, but not inhibin-B or insulin-like peptide-3, may be used as surrogates in the diagnosis of polycystic ovary syndrome in adolescents: preliminary results. *J. Clin. Res. Pediatr. Endocrinol.* 8, 288–297
50. Zadehmodarres, S. *et al.* (2015) Anti-Müllerian hormone level and polycystic ovarian syndrome diagnosis. *Iran. J. Reprod. Med.* 13, 227–230
51. Casadei, L. *et al.* (2013) The role of serum anti-Müllerian hormone (AMH) in the hormonal diagnosis of polycystic ovary syndrome. *Gynecol. Endocrinol.* 29, 545–550
52. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus (2003) Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* 26, S5–S20
53. Flegal, K.M. *et al.* (2012) Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999–2010. *JAMA* 307, 491–497
54. Schwartz, L.M. and Woloshin, S. (1999) Changing disease definitions: implications for disease prevalence. Analysis of the Third National Health and Nutrition Examination Survey, 1988–1994. *Eff. Clin. Pract.* 2, 76–85
55. Ledger, W.L. (2010) Clinical utility of measurement of anti-Müllerian hormone in reproductive endocrinology. *J. Clin. Endocrinol. Metab.* 95, 5144–5154
56. de Vet, A. *et al.* (2002) AntiMüllerian hormone serum levels: a putative marker for ovarian aging. *Fertil. Steril.* 77, 357–362