Title: Anti-müllerian hormone, PCOM and diagnosis of PCOS: A systematic review to inform the international PCOS guideline

Short title: Accuracy of AMH as a marker for PCOM and PCOS

Authors: 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], Joop S.E. Laven[15]

Address:

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, France

3. Department of Obstetrics and Gynaecology, Li Ka Shing Faculty of Medicine, the University of Hong Kong, Hong Kong Special Administrative Region, China

4. Dept. of Health Policy, Management, and Behaviour, School of Public Health, University at Albany, SUNY, Albany, NY, USA 12144

5. Robinson Research Institute, University of Adelaide and Fertility SA, Adelaide, South 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 & Developmental Biology, Hammersmith Hospital London W12 0NN, United Kingdom

8. Department of Obstetrics and Gynecology, Division of Reproductive Endocrinology, University of Rochester, Rochester, NY USA

9. Diabetes and Endocrine Units, Monash Health, Clayton, Victoria, 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 & Fertility, Kwality House, 1st Floor, Kemps Corner, Mumbai- 26, Maharashtra, India

12. Huisartsenpraktijk Hohmann & De Vet, Rotterdam, 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 India 226014.

15. Division of Reproductive Medicine, Department of Obstetrics and Gynecology, Erasmus MC University Medical Center, Rotterdam, Netherlands

*Corresponding author:

Professor Helena Teede

Monash Centre for Health Research and Implementation, School of Public Health and Preventive Medicine, Monash University, Locked Bag 29, Clayton, VIC 3168, Australia.

Telephone: +613 8572 2644

Fax: +613 9594 7550
Email: Helena.Teede@monash.edu

ORCID iD: https://orcid.org/0000-0001-7609-577X

Funding

The guideline was 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 organisations.

Disclosure statement

HT received NHMRC grant funding to the institution to undertake this work. RA is a consultant for Medtronics, Ansh Labs, Spruce Biosciences, and Longitude Capital, and on the advisory board of Martin PET Imaging. JSEL has received grants from Ferring and Euroscreen. MM, ECT, DD, EHN, RJN, MA, SF, KH, SH, SO, DS, FH, SO, PD have nothing to declare.

Abstract

Context:

Polycystic ovary syndrome (PCOS) affects 8-13% of women on Rotterdam diagnostic criteria, which includes polycystic ovarian morphology (PCOM) on ultrasound. This systematic review aims to investigate whether serum Anti-Müllerian Hormone (AMH) is an effective alternative to detection of PCOM and/or the diagnosis of PCOS, to inform international PCOS guidelines.

Evidence acquisition

Electronic databases were searched systematically. Articles were assessed against selection criteria and risk of bias.

Evidence synthesis:

Twenty-nine articles on AMH levels met inclusion criteria, with a moderate to high risk of bias and significant heterogeneity. The studies lacked well-defined PCOS and control populations that varied across the life-span; used inconsistent methods for defining cut offs, variably defined PCOM in comparator studies and had methodological assay and sample handling challenges. Heterogeneity prevented meta-analysis.

Conclusions:

This systematic review reveals key gaps to be overcome before serum AMH can be recommended clinically to detect PCOM or diagnose PCOS. Large scale international collaborative studies in well-defined populations across the life span, exploring how AMH clusters with PCOS features and relates to long term health outcomes is needed to define cut offs. Improved quality and standardization of assays and sample handling are also needed. This work has directly informed international guidelines and sets the scene for research to address clear identified gaps to enhance clinical utility of serum

AMH in PCOS. Once these issues are addressed, AMH levels could replace more costly and less accessible ultrasound in the diagnosis of PCOS.

Key words: Polycystic Ovary Syndrome, review, Anti-Müllerian Hormone, adolescent, adult, diagnostic accuracy

Précis

A systematic review on AMH in PCOS does not support a clinical role and identifies research gaps.

Introduction

Polycystic ovary syndrome (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, polycystic ovarian morphology (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, whilst also highlighting the controversy and challenges with this criteria [1-4]. Determining FNPO is operator and equipment-dependent, limiting accuracy and reproducibility. Equipment advances increase sensitivity and in turn FNPO counts [1-4,12,13]. Ultrasound involves expensive equipment and trained personnel, leading to increasing costs and impacting on accessibility. The ultrasound approach (transabdominal or transvaginal) impacts on accuracy, and in some women transvaginal ultrasound is unacceptable or may be perceived as invasive. Multi-follicular appearance on ultrasound overlaps with PCOM diagnostic cut offs especially in adolescents, whilst 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 called for greater accuracy in PCOS diagnostic criteria worldwide [1-4].

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 [14]. AMH has been shown in animal models of PCOS to have a possible causal role in development of the disorder through in-utero exposure of the fetus to high AMH levels [15]. 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 [16] . Overall, serum AMH levels are significantly higher in women with PCOS compared with normal ovulatory women [17,18]. These data has led to the hypothesis that AMH could be a valuable surrogate marker or an alternative to ultrasound FNPO count for detection of PCOM or in the overall diagnosing of PCOS [16].

Recognised challenges in the use of AMH measurement in PCOS include variations across the life span and problems with defining PCOM for comparison. AMH assays may also display a differential response to pre-analytical proteolysis, conformational changes of the AMH dimer, or the presence of interfering substances [19]. Appreciable sample-to-sample variability and substantial discrepancies in between-assay conversion factors, suggests assay performance issues. These issues were prioritised 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 [20] 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 development of this guideline is aligned with Australia's National Health and Medical Research Council (NHMRC) [21], the European Society of Human Reproduction and Embryology (ESHRE) [22], and the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) methodology [23], and is described in detail in the full guideline [4].

This systematic review addressed the evidence for the following two clinical questions:

1: Is AMH effective to diagnose PCOS?

2: 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 [24]. 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 multi-disciplinary guideline development group [24]. 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 June 26th 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 Database of Systematic Reviews, DARE, CENTRAL and ACP Journal Club; PsycInfo (Ovid) and CINAHL

Evidence processing

Studies were selected and appraised by one highly experienced reviewer (MM) in consultation with colleagues using study selection criteria [24] established *a priori*. The retrieved articles were first

reviewed by title and abstract, and then full articles will be retrieved for further assessment if the information given suggests that the study meets the inclusion criteria.

Assessment of methodological quality

Methodological quality (i.e. risk of bias) of each of the included studies was assessed by one reviewer for the adolescent studies (EB) and one reviewer for the adult studies (ECT), using a critical appraisal template developed *a priori* [25]. Individual quality items were investigated using a descriptive component approach that assessed attrition bias, reporting bias, selection bias, performance bias, potential confounding, and appropriateness of the statistical analysis. Any disagreement or uncertainty was resolved by a discussion with a third reviewer (MM) and within the team of authors of this manuscript. Using this approach each study was allocated a risk of bias rating of either low, moderate, or high.

Data extraction

Data were extracted directly into customized tables for characteristics of included studies and results by one reviewer (MM). Information was extracted on general study characteristics (lead author, year of publication, study design, country), participants (number, age category (adolescents or adults), 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 (MM and ECT). 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 [16,26-53] addressed diagnostic accuracy of AMH for PCOS and/or PCOM and thus met the inclusion criteria for the clinical questions presented in this review, whilst 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 [24].

INSERT FIGURE 1 HERE

One of the 29 studies identified was a systematic review [34] 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.

Characteristics of included studies

Tables 1 and 2 include key characteristics of included studies with four addressing diagnostic accuracy of AMH for PCOS and PCOM [31,32,38,43], and one addressing PCOM only [48]. Of the 28 studies, six studies included adolescent participants for diagnosis of PCOS [32,36,45,46,48,51] and one of these addressed PCOS and PCOM [32]. The remaining 21 studies [16,26-31,33,37-44,47,49,50,52,53] included adult participants for diagnosis of PCOS, where three of these addressed PCOS and PCOM [31,38,43]; the remaining 18 studies addressed PCOS alone [16,26-30,33,37,39-42,44,47,49,50,52,53]. Of the studies in adolescents, one was in overweight and obese participants [35], and in one study BMI was unclear [51]. Of the studies in adults, one included lean and obese participants [27], and five studies [26,31,37,52,53] 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 which included adolescent participants ranged in quality from low to high risk of bias, whilst the majority of adult studies were at high risk of bias (Supplementary Table 5). 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 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 interpretation of the results.

INSERT TABLE 1 HERE

INSERT TABLE 2 HERE

Diagnostic accuracy of AMH for PCOS

In adolescents, there were five studies, of which one was found to have a low risk of bias [32], two were of moderate risk of bias [35,36,46] and two were of high risk of bias [45,51], demonstrating areas under the ROC curve of AMH for the diagnosis of PCOS, ranging from 0.5 to 0.88 (**Table 3**); 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 [29,41-43,49] and 16 were of high risk of bias [16,26-28,30,31,33,37-40,44,47,50,52,53], demonstrating areas under the ROC curve of AMH for the diagnosis of PCOS ranging from 0.66 to 0.994 (**Table 4**); 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.

INSERT TABLE 3 HERE

INSERT TABLE 4 HERE

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 [48] (**Table 5**); 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 [32], one of moderate risk of bias [43] and two of high risk of bias [31,38], demonstrating areas under the ROC of AMH for the diagnosis for PCOM of 0.67 to 0.92 (**Table 6**). 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.

INSERT TABLE 5 HERE

INSERT TABLE 6 HERE

Discussion

This systematic review presents the most up to date, 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 life span, 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 generalisable 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 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 life span 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 suggests 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.

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, establishing gestational diabetes, hypertension, or obesity cut-offs were based on long-term health risks, not simply percentiles [54-56]. 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 very difficult to differentiate PCOS from controls on AMH levels [57]. Levels fall in later life, especially after menopause [58]. Age specific reference ranges are thus vital [59] 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.

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 [60]**,** i.e. 12 follicles of 2-9mm diameter per ovary, to define this PCOS diagnostic criteria. This cut-off suffers from the same challenges of 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 [1-4,12,13]. With 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.

INSERT BOX 1 HERE

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 is very little data with the new automated platform assays [41]. There is rising awareness on the impact of sample handling, transport, and storage conditions, factors which 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 [19]. Overall there is an urgent need for international standardisation in order to improve comparability amongst assays, 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].

INSERT BOX 2 HERE

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 publically 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.

Conclusion

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. 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, more clearly defining PCOM on ultrasound, and improved accuracy and standardisation of assays and handling procedures. With improved standardisation 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].

Acknowledgements

We thank Estifanos Baye for performing the critical appraisals of the adolescent studies included in this review.

Author's roles

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 synthesised literature, critically revised the manuscript and approved the final version for submission.

References

1. Teede HJ, Misso ML, Costello MF, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Human Reproduction (Oxford, England).* 2018.
2. Teede HJ, Misso ML, Costello MF, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Fertility and sterility.* 2018;110(3):364–379.
3. Teede HJ, Misso ML, Costello MF, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Clin Endocrinol (Oxf).* 2018.
4. *International evidence-based guidelines for the assessment and management of polycystic ovary syndrome.* Melbourne: Monash University; 2018.
5. Bozdag G, Mumusoglu S, Zengin D, Karabulut E, Yildiz BO. The prevalence and phenotypic features of polycystic ovary syndrome: a systematic review and meta-analysis. *Human Reproduction.* 2016;31(12):2841-2855.
6. Sirmans SM, Pate KA. Epidemiology, diagnosis, and management of polycystic ovary syndrome. *Clinical epidemiology.* 2013;6:1-13.
7. Teede H, Deeks A, Moran L. Polycystic ovary syndrome: a complex condition with psychological, reproductive and metabolic manifestations that impacts on health across the lifespan. *BMC medicine.* 2010;8:41.
8. Teede HJ, Misso ML, Deeks AA, et al. Assessment and management of polycystic ovary syndrome: summary of an evidence-based guideline. *The Medical journal of Australia.* 2011;195(6):S65-112.
9. Legro RS, Arslanian SA, Ehrmann DA, et al. Diagnosis and Treatment of Polycystic Ovary Syndrome: An Endocrine Society Clinical Practice Guideline. *The Journal of Clinical Endocrinology & Metabolism.* 2013;98(12):4565-4592.
10. 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). *Human Reproduction.* 2004;19(1):41-47.
11. Balen AH, Laven JS, Tan SL, Dewailly D. Ultrasound assessment of the polycystic ovary: international consensus definitions. *Human reproduction update.* 2003;9(6):505-514.
12. Teede HJ, Misso ML, Costello MF, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Human Reproduction (in press).* 2018.
13. Teede HJ, Misso ML, Costello MF, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Human Reproduction (in press).* 2018.
14. Durlinger AL, Visser JA, Themmen AP. Regulation of ovarian function: the role of anti-Mullerian hormone. *Reproduction (Cambridge, England).* 2002;124(5):601-609.
15. Tata B, Mimouni NEH, Barbotin AL, et al. Elevated prenatal anti-Mullerian hormone reprograms the fetus and induces polycystic ovary syndrome in adulthood. *Nature medicine.* 2018.
16. 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(3):941-945.
17. Cook CL, Siow Y, Brenner AG, Fallat ME. Relationship between serum mullerian-inhibiting substance and other reproductive hormones in untreated women with polycystic ovary syndrome and normal women. *Fertility and sterility.* 2002;77(1):141-146.
18. Seifer DB, MacLaughlin DT. Mullerian Inhibiting Substance is an ovarian growth factor of emerging clinical significance. *Fertility and sterility.* 2007;88(3):539-546.

19. Rustamov O, Smith A, Roberts SA, et al. The measurement of anti-Mullerian hormone: a critical appraisal. *J Clin Endocrinol Metab.* 2014;99(3):723-732.
20. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ.* 2009;339.
21. National Health and Medical Research Council. NHMRC standards and procedures for externally developed guidelines. Australia2007.
22. Vermeulen N, D'ANGELO A, de Sutter P, Nelen W. Manual for ESHRE guideline development. European Society for Human Reproduction and Embriology; 2014.
23. The GRADE Working Group. *GRADE handbook for grading quality of evidence and strength of recommendation.* Vol Version 3.2 [updated March 2009]2009.
24. Technical report for: International evidence‐based guideline for the assessment and management of polycystic ovary syndrome. Section 1.5. Melbourne: Monash University; 2018.
25. Centre for Clinical Effectiveness. Critical Appraisal Templates. In: Southern Health, ed. Melbourne, Australia.2010.
26. Carmina E, Campagna AM, Fruzzetti F, Lobo RA. Amh Measurement Versus Ovarian Ultrasound in the Diagnosis of Polycystic Ovary Syndrome in Different Phenotypes. *Endocr Pract.* 2016;22(3):287-293.
27. Cassar S, Teede HJ, Moran LJ, et al. Polycystic ovary syndrome and anti-Mullerian hormone: role of insulin resistance, androgens, obesity and gonadotrophins. *Clin Endocrinol (Oxf).* 2014;81(6):899-906.
28. Chao KC, Ho CH, Shyong WY, et al. Anti-Mullerian hormone serum level as a predictive marker of ovarian function in Taiwanese women. *J Chin Med Assoc.* 2012;75(2):70-74.
29. Dewailly D, Alebic MS, Duhamel A, Stojanovic N. 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. *Human Reproduction.* 2014;29(11):2536-2543.
30. Dewailly D, Gronier H, Poncelet E, et al. Diagnosis of polycystic ovary syndrome (PCOS): revisiting the threshold values of follicle count on ultrasound and of the serum AMH level for the definition of polycystic ovaries. *Human Reproduction.* 2011;26(11):3123-3129.
31. Eilertsen TB, Vanky E, Carlsen SM. Anti-Mullerian hormone in the diagnosis of polycystic ovary syndrome: can morphologic description be replaced? *Human Reproduction.* 2012;27(8):2494-2502.
32. Hart R, Doherty DA, Norman RJ, et al. Serum antimullerian hormone (AMH) levels are elevated in adolescent girls with polycystic ovaries and the polycystic ovarian syndrome (PCOS). *Fertility & Sterility.* 2010;94(3):1118-1121.
33. Homburg R, Ray A, Bhide P, et al. The relationship of serum anti-Mullerian hormone with polycystic ovarian morphology and polycystic ovary syndrome: a prospective cohort study. *Human Reproduction.* 2013;28(4):1077-1083.
34. Iliodromiti S, Kelsey TW, Anderson RA, Nelson SM. Can anti-Mullerian hormone predict the diagnosis of polycystic ovary syndrome? A systematic review and meta-analysis of extracted data. *J Clin Endocrinol Metab.* 2013;98(8):3332-3340.
35. Kim JY, Tfayli H, Michaliszyn SF, Lee S, Nasr A, Arslanian S. Anti-Mullerian hormone in obese adolescent girls with polycystic ovary syndrome. *Journal of Adolescent Health.* 2017;60(3):333-339.
36. Kim JY, Tfayli HM, Michaliszyn SF, Lee S, Nasr A, Arslanian SA. Anti-mullerian hormone (AMH) in obese adolescent girls with polycystic ovary syndrome (PCOS): cross-sectional and treatment-associated longitudinal changes. *Endocrine reviews. Conference: 98th annual meeting and expo of the endocrine society, ENDO.* 2016;37(2 Supplement 1).
37. Koninger A, Koch L, Edimiris P, et al. Anti-Mullerian Hormone: an indicator for the severity of polycystic ovarian syndrome. *Arch Gynecol Obstet.* 2014;290(5):1023-1030.

38. Lauritsen MP, Bentzen JG, Pinborg A, et al. The prevalence of polycystic ovary syndrome in a normal population according to the Rotterdam criteria versus revised criteria including anti-Mullerian hormone. *Human Reproduction.* 2014;29(4):791-801.
39. Li L, Chen X, Mo Y, Chen Y, Wenig M, Yang D. Elevated serum anti-mullerian hormone in adolescent and young adult Chinese patients with polycystic ovary syndrome. *Wien Klin Wochenschr.* 2010;122(17-18):519-524.
40. Li Y, Ma Y, Chen X, et al. Different diagnostic power of anti-Mullerian hormone in evaluating women with polycystic ovaries with and without hyperandrogenism. *J Assist Reprod Genet.* 2012;29(10):1147-1151.
41. Pigny P, Gorisse E, Ghulam A, et al. Comparative assessment of five serum antimullerian hormone assays for the diagnosis of polycystic ovary syndrome. *Fertility & Sterility.* 2016;105(4):1063-1069.e1063.
42. Sahmay S, Atakul N, Aydogan B, Aydin Y, Imamoglu M, Seyisoglu H. Elevated serum levels of anti-Mullerian hormone can be introduced as a new diagnostic marker for polycystic ovary syndrome. *Acta Obstet Gynecol Scand.* 2013;92(12):1369-1374.
43. Sahmay S, Aydin Y, Oncul M, Senturk LM. Diagnosis of Polycystic Ovary Syndrome: AMH in combination with clinical symptoms. *J Assist Reprod Genet.* 2014;31(2):213-220.
44. Saikumar P, Kalai Selvi VS, Prabhu K, Venkatesh P, Krishna P. Anti mullerian hormone: A potential marker for recruited non growing follicle of ovarian pool in women with polycystic ovarian syndrome. *Journal of Clinical and Diagnostic Research.* 2013;7(9):1866-1869.
45. Sopher AB, Grigoriev G, Laura D, et al. Anti-Mullerian hormone may be a useful adjunct in the diagnosis of polycystic ovary syndrome in nonobese adolescents. *J Pediatr Endocrinol.* 2014;27(11-12):1175-1179.
46. Tokmak A, Timur H, Aksoy RT, Cinar M, Yilmaz N. Is anti-mullerian hormone a good diagnostic marker for adolescent and young adult patients with polycystic ovary syndrome? Anti-mullerian hormon polikistik over sendromlu adolesan ve genc eriskinlerde iyi bir tanisal belirtec midir? *Turk Jinekoloji ve Obstetrik Dernegi Dergisi.* 2015;12(4):199-204.
47. Tremellen K, Zander-Fox D. Serum anti-Mullerian hormone assessment of ovarian reserve and polycystic ovary syndrome status over the reproductive lifespan. *Aust N Z J Obstet Gynaecol.* 2015;55(4):384-389.
48. Villarroel C, Merino PM, Lopez P, et al. Polycystic ovarian morphology in adolescents with regular menstrual cycles is associated with elevated anti-Mullerian hormone. *Human Reproduction.* 2011;26(10):2861-2868.
49. Wiweko B, Maidarti M, Priangga MD, et al. Anti-mullerian hormone as a diagnostic and prognostic tool for PCOS patients. *J Assist Reprod Genet.* 2014;31(10):1311-1316.
50. Woo HY, Kim KH, Rhee EJ, Park H, Lee MK. Differences of the association of anti-Mullerian hormone with clinical or biochemical characteristics between women with and without polycystic ovary syndrome. *Endocr J.* 2012;59(9):781-790.
51. Yetim A, Yetim C, Bas F, et al. Anti-Mullerian 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.* 2016;8(3):288-297.
52. Zadehmodarres S, Heidar Z, Razzaghi Z, Ebrahimi L, Soltanzadeh K, Abed F. Anti-mullerian hormon level and polycystic ovarian syndrome diagnosis. *Iranian Journal of Reproductive Medicine.* 2015;13(4):227-230.
53. Casadei L, Madrigale A, Puca F, et al. The role of serum anti-Mullerian hormone (AMH) in the hormonal diagnosis of polycystic ovary syndrome. *Gynecol Endocrinol.* 2013;29(6):545-550.
54. Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes care.* 2003;26 Suppl 1:S5-20.
55. Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. *Jama.* 2012;307(5):491-497.

525 56. Schwartz LM, Woloshin S. Changing disease definitions: implications for disease prevalence.
526 Analysis of the Third National Health and Nutrition Examination Survey, 1988-1994. *Effective*
527 *clinical practice : ECP.* 1999;2(2):76-85.
528 57. Hart R, Doherty DA, Norman RJ, et al. Serum antimullerian hormone (AMH) levels are
529 elevated in adolescent girls with polycystic ovaries and the polycystic ovarian syndrome
530 (PCOS). *Fertility and sterility.* 2010;94(3):1118-1121.
531 58. Ledger WL. Clinical utility of measurement of anti-mullerian hormone in reproductive
532 endocrinology. *J Clin Endocrinol Metab.* 2010;95(12):5144-5154.
533 59. de Vet A, Laven JS, de Jong FH, Themmen AP, Fauser BC. Antimullerian hormone serum
534 levels: a putative marker for ovarian aging. *Fertility and sterility.* 2002;77(2):357-362.
535 60. Revised 2003 consensus on diagnostic criteria and long‐term health risks related to
536 polycystic ovary syndrome (PCOS). *Human Reproduction.* 2004;19(1):41-47.

537

538

539

540

541

542

543

544

545

546

547

548

549

550

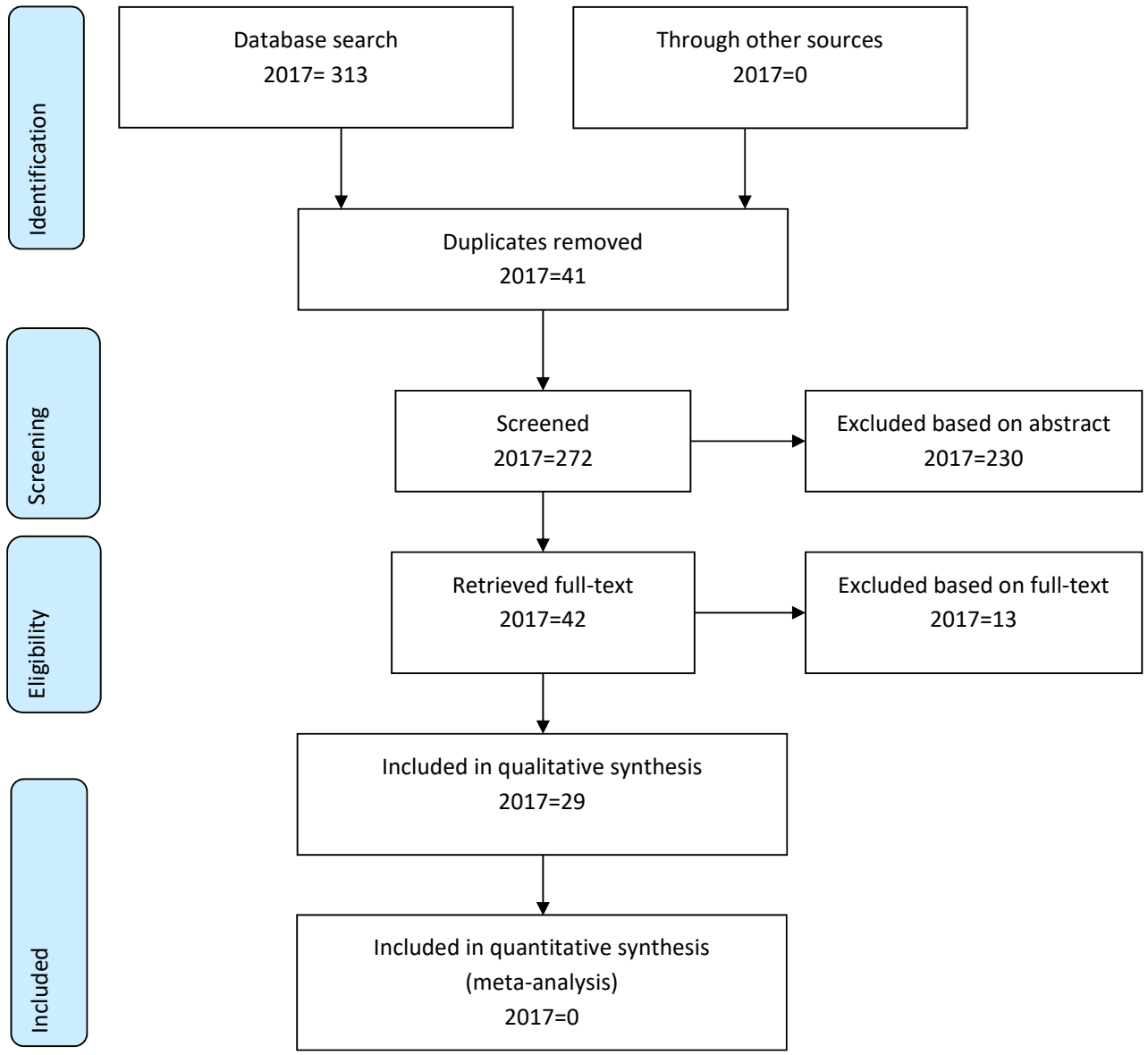

Figure 1: PRISMA flow diagram

571 **Table 1: Key characteristics of included studies – adolescents**

Study ID	Design	ROB	Setting	N	Adults/ Adolescents	BMI	AMH (pmol/L)	Diagnostic criteria	Medication status
Hart 2010 +PCOM [32]	Prospective cohort	Low	Australia	P: 175 C: 458	Adolescents	Rotterdam p=0.001 P: 24.43±5.12 C: 22.07±2.94 NIH p<.001 P: 25.83±5.64 C: 22.15±3.09 PCOM p=0.046 P: 23.5±4.60 C: 22.3±3.03	Unclear medians	Rotterdam NIH	NR
Kim 2016 [36] & 2017 [35]	Prospective cohort	Moderate	USA	P: 46 C: 43	Adolescents	P: 37.7±1.1 C: 33.1±1.1 p=0.003	P: 59.5 C: 30.7	NIH	Excluded
Sopher 2014 [45]	Prospective cohort	High	USA	P: 15 C: 16	Adolescents	P: 0.45 ± 0.79 C: 0.19 ± 0.60 z-score, NS	P: 31.4±24.3 C: 17.1±9.3 p<0.05^	NIH	Not in prior 3m
Tokmak 2015 [46]	Prospective cohort	Moderate	Turkey	P: 43 C: 47	Young adults ~18	P: 22.9±4.7 C: 21.8±2.8 p=0.081	P: 72.1±49.3 C: 67.1±39.3 p=0.198	Rotterdam	Excluded
Villarroel 2011 PCOM only [48]	Prospective cohort	Low	Chile	P: 25 C: 49	Adolescents	PCOM: 22.5±0.5 C: 22.7±0.4 SEM	PCOM: 72.5±6.1 C: 33.4±2.6 p<0.0001	Rotterdam	Excluded
Yetim 2016 [51]	Prospective cohort	High	Turkey	P: 53 C: 26	Adolescents	Unclear	P: 78.7 (11.9–361.4) C: 29 (6.6-85.4) p<0.001	Rotterdam	Excluded

572 ROB, risk of bias; BMI, body mass index; AMH, anti-müllerian hormone; PCOM, polycystic ovarian morphology; P, PCOS; C, control; NIH, National Institute of
573 Health PCOS diagnostic criteria; NR, not reported

574 **Table 2: Key characteristics of included studies – adults**

Study ID	Design	ROB	Setting	N	Adults/ Adolescents	BMI	AMH	Diagnostic criteria	Medication status
Carmina 2016 [26]	Retrospective cohort	High	Italy	P: 113 C: 47	Adults	P: 27.6±6 C: 27±4	P: 65.7±33.6 C: 20.7±5.7 p<.01	Rotterdam	Not in prior 3m
Casadei 2013 [53]	Prospective cohort	High	Italy	P: 22 C: 22	Adults	P: 27.4±5.9 C: 21.9±3.1 P<0.001	P: 69.3±32 C: 18.3±9.3 p<0.001	NIH	NR
Cassar 2014 [27]	Prospective cohort	High	Australia	P: 43 C: 35	Adults	P: 22LN 21OW C: 19LN 16OW	LN P: 64.7±29.8 LN C: 29.8±30.5 OW P: 54.4±30.2 OW C: 17.8±33.7 P<0.05 across all	Rotterdam	Not in prior 3m
Chao 2012 [28]	Prospective cohort	High	Taiwan	P: 31 C: 24	Adults	NR	NR	Rotterdam	Not in prior 2m
Dewailly 2014 [29]	Retrospective cohort	Moderate	Croatia	P: 95 C: 521	Adults	P: 27 (20–40)* C: 23 (19–31)	P: 50.9 (10.2-129.8) C: 12.2 (2.2–26.4)	Rotterdam HA+OA	Excluded
Dewailly 2011 [30]	Retrospective cohort	High	France	P:62 C1A:66	Adults	P: 28.0 (18.7–41.7) C1A: 24.0 (18.7-37.6)	P: 81.2 (25.4–256.2) C1A: 21.0 (10.0–35.0) p=0.0001	Rotterdam	NR
Eilertsen 2012 +PCOM [31]	Retrospective cohort, case-control	High	Norway	PR: 56 CR: 206 PA: 44 CA: 218	Adults	PR: 27.8±5.7 CR: 26.6±5.0 p=0.14 PA: 28.5±5.7 CA: 26.6±5.0 p=0.02	PR: 44.8+27.5 CR: 19.7+16.8 PA: 42.7+26.7 CA: 21.5+19.3 P<0.001 for both	Rotterdam AES	N=40 had hormonal contraceptives
Homburg 2013 [33]	Prospective cohort	High	UK	P: 90 C: 90	Adults	P: 24.9±2.4 C: 24.8±2.6	P: 77.6±61.0 C: 23.6±15.0 P<0.001	Rotterdam	NR

Köninger 2014 [37]	Prospective cohort	High	Germany	PM: 21 PS: 59 C: 48	Adults	PM: 26.7±7.0 PS: 29.1±7.4 C: 24.3±4.4	PM: 36.4±27.9 PS: 63.6±52.1 C: 15.0±8.6	Rotterdam	Not in prior 3m
Lauritsen 2014 + PCOM [38]	Prospective cohort	High	Denmark	P: 74 C: 373	Adults	P: 24.2±4.2 C: 22.9±3.4 p=0.02	P: 35.6(22.2-62.9) C: 17.8 (9.8-29.4) p<0.001	Rotterdam	Excluded
Li 2010 [39]	Prospective cohort	High	China	P: 47 C: 40	Both 17-25	P: 21.25±4.29 C: 20.04±1.83 p=0.083	P: 70.4±35.2 C: 50.9±21.6 p=0.002	Rotterdam	Not in prior 3m
Li 2012 [40]	Prospective cohort	High	China	PHA+: 62 PHA-: 69 C: 61	Adults	PHA+:20.1±5.76 PHA-: 23.35±5.22 C: 20.52±1.58 P<0.05 for all	PHA+: 60.1±32.6 PHA-: 41.5±27.5 C: 26.7±16.1 P<0.05 for all	Rotterdam	NR
Pigny 2006 [16]	Prospective cohort	High	France	P: 73 C: 96	Adults	P: 26.0 (19–39) C: 23.4 (18.2-31.8) mean 5-95 p<0.01	P: 81.6 (26.3-214) C: 33.5 (8.3– 68.1) mean 5-95, P<0.001	Rotterdam	Not in prior 3m
Pigny 2016 [41] Compares assays	Retrospective cohort	Moderate	France	P: 47 C: 48	Adults	NR	NR	Rotterdam equivalent	NR
Sahmay 2013 [42]	Prospective cohort	Moderate	Turkey	P: 419 C: 151	Adults	P: 25.43±4.6** C: 25.4±4.4 NS	P: 52.4±28.9 C: 16.0±12.1 p<0.001	Rotterdam	Not in prior 6m
Sahmay 2014 + PCOM [43]	Prospective cohort	Moderate	Turkey	AES P: 195 C: 411 ROT P: 228 C: 378 NIH P: 164 C:442	Adults	AES P: 25.7±4.6 C: 26.2±7.5 Rotterdam P: 25.5±4.6 C: 26.3±6.4 NIH P: 25.9±4.6 C: 26.1±6.9	AES p<0.05 for all P: 62.9±47.9 C: 18.6±16.4 Rotterdam P: 61.4±46.4 C: 16.4±11.4 NIH P: 65.7±50.7 C: 20.0±18.6	Rotterdam, AES, NIH	Not in prior 6m

Saikumar 2013 [44]	Prospective cohort	High	India	P: 60 C: 60	Adults Infertile	P: 27.5±2.65 C: 25.0±3.3 NS	P: 31.3±16.0 C: 16.2±5.8 p<0.001	Rotterdam	NR
Tremellen 2015 [47]	Retrospective cohort	High	Australia	P: 43 C: 113	Adults	NR	P: 51 (40.5–74.7) C: 14.2 (5.9–29.5) p<0.0001	Rotterdam	NR
Wiweko 2014 [49]	Prospective case-control	Moderate	Indonesia	P: 71 C: 71	Adults	P: 25.86 (17.85–39.14) C: 25.00±5.77 p=0.072	P: 67.9±36.5 C: 25.2±13.9 p<0.001	Rotterdam	Not in prior 3m
Woo 2012 [50]	Prospective cohort	High	South Korea	P: 87 C: 53	Adults	P: 21.3±3.40 C: 20.1±1.88 p=0.007	P: 82.7±45.1 C: 38.4±21.4 p<0.001	Rotterdam	Not in prior 6m
Zadehmodarres 2015 [52]	Prospective cohort	High	Iran	P: 60 C: 57	Adults	P: 29.02±6.53 C: 28.76±3.41 p=0.389	P: 51.0±46.6 C: 23.9±24.6 p=0.001	Rotterdam	No OCP in prior 1m No OI in prior 6m

Data are presented as mean±standard deviation or median (interquartile range); *medians with 5th–95th percentiles; ROB, risk of bias; BMI, body mass index; AMH, anti-müllerian hormone; P, PCOS; C, controls; NIH, National Institute of Health PCOS diagnostic criteria; NR, not reported; LN, lean; OW, overweight; AES, Androgen Excess Society PCOS diagnostic criteria; PM, PCOS mild, PCO+OA; PCOS and oligoanovulation, PS, PCOS severe, all three criteria. PHA+, PCOS with hyperandrogenism; PHA–, PCOS with normal androgen levels and no clinical hyperandrogenism. ^p value adjusted for menstrual age

Table 3: Diagnostic accuracy of AMH for PCOS- studies in adolescents

Study ID	Threshold	Diagnostic criteria*	PCOS	Non-PCOS	Sensitivity	Specificity	True positive	False positive	True negative	False negative	AUC	Precision

Hart 2010	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 & 2017	44.71 pmol/L	NIH	46	43	67	81					0.788	0.687-0.868 p<.0001
Sopher 2014	24.29 pmol/L	NIH	15	16	40	93.8					NR	NR
Tokmak 2015	100 pmol/L	Rotterdam Youden index	43	47	48.8	77.1					0.579	0.453-0.705 p=0.198
Yetim 2016	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

583

584 **Table 4: Diagnostic accuracy of AMH for PCOS- studies in adults**

Study ID	Threshold	Diagnostic criteria*	PCOS	Non-PCOS	Sensitivity	Specificity	True positive	False positive	True negative	False negative	AUC	Precision
Carmina 2016	>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	33pmol/L	NIH	22	22	95	95					0.970	CI=0.02–0.92
Cassar 2014	>30 pmol/L	Rotterdam	43	35	82	79	35	7	28	8	0.829	CI=0.736–0.923 P <0.001
Chao 2012	25pmol/L	Rotterdam	31	24	74	79	23	5	19	8	NR	NR
Dewailly 2014	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	35 pmol/L	Rotterdam	62	66	92	97	57	2	64	5	0.973	CI=0.947–0.998
Eilertsen 2012	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	48 pmol/L	Rotterdam	90	90	60	98.2	54	2	88	36	0.805	NR
Köninger 2014	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	18 pmol/L	Rotterdam	74	373	91.8	98.1	68	7	366	6	0.994	CI=0.990–0.999
Li 2010	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	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	60 pmol/L	Rotterdam	73	96	67	92	49	8	88	24	0.851	CI=0.796–0.905
Pigny 2016	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	28.14 pmol/L	Rotterdam	419	151	80	89.8	335	15	136	84	0.916	CI=0.897–0.935 p < 0.0001
Sahmay 2014	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	23.86 pmol/L	Rotterdam	60	60	98	93	59	4	56	1	0.956	NR
Tremellen 2015	≥36 pmol/L	Rotterdam	43	113	83.7	82.3	36	20	93	7	0.917	NR
Wiweko 2014	31.79 pmol/L	Rotterdam	71	71	76.1	74.6	54	18	53	17	0.870	CI=0.81–0.92
Woo 2012	55.86 pmol/L	Rotterdam	87	53	75.9	86.8	66	7	46	21	0.868	CI=0.801-0.919
Zadehmodarres 2015	22.5 pmol/L	Rotterdam	60	57	70.37	77.36	42	13	44	18	NR	NR

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

588

589 **Adolescents**

590 **Table 5: Diagnostic accuracy of AMH for PCOM- study in adolescents**

Study ID	Threshold	Diagnostic criteria	PCOS	Non-PCOS	Sensitivity	Specificity	True positive	False positive	True negative	False negative	AUC	Precision
Villarroel 2011	50.25 pmol/l	Rotterdam	25	49	84.0	83.7	21	8	41	4	0.873	CI=0.782–0.963 p<0.0001

591

592

593 **Adults**

594 **Table 6: Diagnostic accuracy of AMH for PCOM- study in adults**

Study ID	Threshold	Diagnostic criteria	PCOS	Non-PCOS	Sensitivity	Specificity	True positive	False positive	True negative	False negative	AUC	Precision
Eilertsen 2012	20 pmol/L	Rotterdam	113	149	79.6	72.5	90	41	108	23	0.896	CI=0.855-0.937
Hart 2010	30 pmol/L	Rotterdam	75	132	54.7	72.7	41	36	96	34	0.67	CI=0.60–0.75 p<.001
Lauritsen 2014	20 pmol/L	Rotterdam	74	373	82.0	84.6	61	57	316	13	0.906	CI=0.878–0.933

Sahmay 2014	27.14 pmol/L	Unclear	Unclear	Unclear	83	87					0.92	CI=0.90–0.93 p<0.001

595

596

Box 1. Ultrasound and PCOM recommendations in the International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome [1-4]

Ultrasound should not be used for the diagnosis of PCOS in those with a gynaecological age of < 8 years (< 8 years after menarche), due to the high incidence of multi-follicular ovaries in this life stage **(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 8MHz, the threshold for PCOM should be on either ovary, a follicle number per ovary of ≥20 and/or an ovarian volume ≥ 10ml, ensuring no corpora lutea, cysts or dominant follicles are present **(CCR)**
- If using older technology, the threshold for PCOM could be an ovarian volume ≥ 10ml on either ovary **(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 ≥ 10ml, 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 2-9mm
    - Three dimensions and volume of each ovary
    - Reporting of endometrial thickness and appearance is preferred – 3-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 ≥ equal 10mm **(CPP)**
- There is a need for training in careful and meticulous follicle counting per ovary, to improve reporting **(CPP)**

597 CCR= clinical consensus recommendations; CPP= clinical practice point

Box 2. AMH recommendations in the International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome[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 **(EBR)**
- There is emerging evidence that with improved standardisation 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 previously mentioned. 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.

598 EBR= evidence-based recommendation; CPP= clinical practice point

599