Harmonising research outcomes for polycystic ovary syndrome: An international multi-stakeholder core outcome set

Bassel H. Al Wattar[1,2], Helena Teede[3,4], Rhonda Garad[3,4]; Steve Franks[5], Adam Balen[6], Priya Bhide[1,7], Terhi Piltonen[8], Daniela Romualdi[9,10], Joop Laven[11], Mala Thondan[12], Aurora Bueno-Cavanillas[13,14,15], Ngawai Moss[1], Caroline Andrews[16], Rachel Hawkes[16], Ben W. Mol[17], Khalid S. Khan[1], Shakila Thangaratinam[1]

[1]Barts Research Centre for Women's Health (BARC), Women's Health Research Unit, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

[2]Warwick Medical School, University of Warwick, Coventry, UK

[3]National Health and Medical Research Council Centre for Research Excellence in PCOS, Monash Centre for Health Research and Implementation, Monash University Melbourne, Vic., Australia

[4]Endocrine and Diabetes Units, Monash Health, Melbourne, Vic., Australia

[5]Imperial College School of Medicine, Institute of Reproductive and Developmental Biology, Hammersmith Hospital, London, UK

[6]Leeds Teaching Hospitals, Leeds UK

[7]Homerton University Hospital NHS Foundation Trust, London, UK

[8]Department of Obstetrics and Gynecology, PEDEGO Research Unit, Medical Research Centre, Oulu University Hospital, University of Oulu, Oulu, Finland

[9]Department of Woman and Child Health and Public Health, Woman Health Area, Fondazione Policlinico Universitario A.Gemelli, Rome, Italy

[10]Department of Woman and Child Health, Azienda Ospedaliera Card.Panico, Tricase, Italy

[11]Div Reproductive Endocrinology and Infertility, Dept of Obstetrics and Gynecology, Erasmus University Medical Centre, Rotterdam, The Netherlands.

[12] Harp Family Medical Centre, Melbourne, VIC

[13] Department of Preventive Medicine and Public Health, University of Granada, Granada, Spain.

[14] Consortium for Biomedical Research in Epidemiology & Public Health (CIBER Epidemiología y Salud Pública-CIBERESP), Madrid, Spain.

[15] Inst Invest Biosanitaria Ibs GRANADA, Granada, Spain.

[16]Verity – The PCOS Self Help Group, Surry, UK

[17]Department of Obstetrics and Gynaecology, University of Monash, Melbourne, Australia.

Corresponding author:

Bassel H.Al Wattar - Women's Health Research Unit, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London UK

b.wattar@qmul.ac.uk

Running title: Core outcomes for polycystic ovary syndrome

Abstract

STUDY QUESTION: What are the key core outcomes to be reported in studies on polycystic ovary syndrome (PCOS)?

SUMMARY ANSWER: We identified three generic and 30 specific core outcomes in six domains (**AUTHOR:** the main text and Fig. 1 state seven domains. Please would you recheck or clarify?): metabolic (eight), reproductive (7) (**AUTHOR:** correct, as below?), pregnancy (10), oncological (one), psychological (one), and long-term outcomes (one).

WHAT IS KNOWN ALREADY: Research reporting PCOS is heterogeneous with high variation in outcome selection, definition and quality.

STUDY DESIGN, SIZE, DURATION: Evidence synthesis and a modified Delphi method with e-surveys were used as well as a consultation meeting.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Overall, 71 health professionals and 123 lay consumers (women with lived experience of PCOS and members of advocacy and peer support groups) (**AUTHOR:** it may be helpful for the reader to briefly describe who was included the lay consumer group. Thank you.) from 17 high-, middle- and low-income countries were involved in this analysis.

MAIN RESULTS AND THE ROLE OF CHANCE: The final core outcome set included three generic outcomes (BMI, quality of life, treatment satisfaction) that are applicable to all studies on women with PCOS and 30 specific outcomes that were categorised into six domains (**AUTHOR:** the main text and Fig. 1 state seven domains. Please would you recheck or clarify?): eight metabolic outcomes (waist circumference, type 2 diabetes, insulin resistance, impaired glucose tolerance, hypertension, coronary heart disease, lipids profile, venous thromboembolic disease); seven reproductive outcomes [viable pregnancy (confirmed by ultrasound including singleton, twins, and higher multiples), clinical and biochemical

hyperandrogenism, menstrual regularity, reproductive hormonal profile, chronic anovulation, ovulation stimulation success including the number of stimulated follicles≥12mm, incidence and severity of ovarian hyperstimulation syndrome]; 10 pregnancy outcomes (live birth, miscarriage, stillbirth, neonatal mortality, gestational weight gain, gestational diabetes, pre-term birth, hypertensive disease in pregnancy, baby birth weight, major congenital abnormalities); three psychological outcomes (depression, anxiety, eating disorders); one oncological (abnormal endometrial proliferation including atypical endometrial hyperplasia and endometrial cancer); and one outcome in the long-term domain (long-term offspring metabolic and developmental outcomes).

LIMITATIONS, REASONS FOR CAUTION: We involved lay consumers in all stages of study through e-surveys but not through focus groups, thereby limiting our understanding of their choices. We did not address the variations in the definitions and measurement tools for some of the core outcomes.

WIDER IMPLICATIONS OF THE FINDINGS: Implementing this core outcome set in future studies on women with PCOS will improve the quality of reporting and aid evidence synthesis.

STUDY FUNDING/COMPETING INTEREST(S): Evidence synthesis was funded through the Australian government, National Health and Medical Research Council (NHMRC) Centre for Research Excellence in PCOS and HT is funded through an NHMRC fellowship. BHA is funded through an NIHR lectureship. All authors have no competing interest to declare.

Keywords: Polycystic ovary syndrome, stakeholder, Delphi, core outcome, reporting.

Introduction

Polycystic ovary syndrome (PCOS) is the commonest chronic endocrine condition, affecting 8-13% of women of reproductive age (Bozdag *et al.*, 2016). With a variety of metabolic, reproductive and psychological features, PCOS predisposes women to adverse health outcomes such as diabetes, metabolic syndrome, depression and subfertility (Azziz *et al.*, 2016; Teede *et al.*, 2010). Care for women with PCOS remains fragmented across various health professionals, including primary care physicians, gynaecologists, endocrinologists, fertility specialists, specialist nurses, dieticians and allied health professionals, often leading to delayed diagnosis and inconsistent clinical management internationally (Teede *et al.*, 2010). This problem permeates into clinical research on PCOS with poor collaboration across health disciplines and inadequate prioritisation of key clinical outcomes as well as scarce engagement of lay consumers (Tay, Moran, *et al.*, 2018). Selective and heterogeneous outcome reporting is common practice, often hindering meaningful evidence synthesis, increasing research wastage and limiting impact (Khan and O'Donovan, 2014). Consequently, the translation and implementation of evidence in clinical guidelines on PCOS remains limited despite an increasing number of clinical trials (Tay, Joham, *et al.*, 2018).

The use of condition-specific standardised sets of core outcomes as a minimum for reporting across future studies is recommended, to minimise variations in outcome reporting (Williamson *et al.*, 2012). Several core outcomes sets have been successfully developed in an attempt to standardise reporting and improve research quality (Tugwell *et al.*, 2007). We aim to identify those core outcomes to be minimally reported in clinical studies on PCOS using a modified Delphi method involving an international panel of stakeholders.

Materials and Methods

We developed a core outcome set for PCOS research using a prospectively registered protocol available online (Wattar *et al.*, 2018) and reported our findings in line with current recommendations (Kirkham *et al.*, 2016). The study had a dedicated Core Management Group (CMG) responsible for the study design and overall conduct (BHA, HT, RG, and ST) with oversight from the Guideline Development Group (GDG) of the 2018 international evidence-based guideline on the diagnosis and management of PCOS (Teede *et al.*, 2018). Members of both groups took part in the survey anonymously.

Identification of outcomes

We identified a longlist of all relevant outcomes reported in clinical trials on PCOS using 40 systematic reviews conducted by the GDG during the development of the international guideline (Teede *et al.*, 2018). We initially categorised outcomes on this longlist into four main domains: metabolic, reproductive, pregnancy and long-term outcomes. To facilitate the Delphi voting process, we combined outcomes of similar clinical and physiological background under one label e.g. High-Density Lipoprotein, Low-Density Lipoprotein, and Triglycerides were combined under lipids profile. The final longlist was piloted among the CMG members before the start of the Delphi process for its face validity and ease of use; any disagreement was resolved by consensus. We generated lay definitions for all outcomes on the longlist using the University of Michigan simplification guide to medical terms to facilitate the participation of lay consumers in the Delphi process (University of Michigan, n.d.).

Health professionals

We included representatives of each of the following health professional stakeholder groups: endocrinologists, general obstetricians and gynaecologists, fertility specialists, academics,

specialist nurses and midwives, primary care physicians, and allied health specialists. We created a list of candidates per stakeholder group using the contacts of the CMG and the GDG members and leveraged the wider membership of the Androgen Excess and Polycystic ovary syndrome society (AE-PCOS) to expand our pool of international stakeholders (Androgen Excess and PCOS Society, n.d.). We sought stakeholder representation from specific countries to ensure a balanced representation of both developed and developing countries from all five continents.

Modified Delphi method

We asked health professionals to complete a two-round Delphi process using a custom-designed electronic survey on Google Forms. In each round, participants were asked to score each of the outcomes on the longlist using a ten-point Likert scale anchored between zero (labelled 'not important') and 10 (labelled 'very important'). Participants were able to suggest any additional outcomes at the end of the 1st Delphi round; all outcomes identified were incorporated and voted on in the 2nd Delphi round.

At the end of the 1st round, we provided participants with individualised feedback comprising their individual score, the mean score of the whole group of health professionals, and the mean score of the lay consumers' group for each outcome. Feedback was provided using individualised emails with an embedded custom-designed Google form prompting participants to consult those scores before providing their new scores for the 2nd round. The feedback design was aimed to promote reflection and reach consensus among participants by the end of the 2nd Delphi round. Non-responders received three reminders with a personalised message before being excluded from the 2nd round.

We used the following pre-specified consensus criteria: outcomes were included (core) if they had a score of ≥7 by more than 70% of participants and a score of ≤4 by less than 15% of participants. Outcomes were excluded (not core) if they received a score of ≥7 by less than 15% of participants and a score of ≤4 by more than 70% of participants. Outcomes with any other score combinations were considered equivocal and were discussed at the final consultation meeting. Both rounds were moderated by the same researchers (BHA and RG).

Patient and public involvement

We sought input from a lay consumers group on both the study design and the Delphi process. Participants in the lay group were identified as women with lived experience of PCOS with an established diagnosis, or if they cared for their family members such as partners, or individuals with PCOS life-experiences such as leaders of advocacy and peer support groups. We leveraged links to established charities and lay support groups including Verity-PCOS UK and PCOS Challenge to engage their membership and promote participation in our study. Candidates were sent electronic invitations via emails and social media platforms, which included a brief summary of the study objectives, the consensus convergence process and the lay definitions of included outcomes. Participants were asked to score each of the outcomes on the longlist using a 10 points Likert scale anchored between zero (labelled 'not important') and 10 (labelled 'very important'). They were also asked to provide any additional outcomes of relevance to women with PCOS.

Consultation meeting

We held a final consultation meeting involving the CMG and representatives from both the health professionals and lay consumers stakeholder groups. The meeting consisted of group discussions followed by two voting rounds using the same criteria to reach consensus. The

objectives of the meeting were to discuss all equivocal outcomes that did not reach consensus in the Delphi process, to agree and finalise the core outcomes list, and to devise a dissemination and implementation plan of the final core outcome set.

Data analysis

We collected data and Delphi scores using live online password-protected Google forms. Each participant was issued a unique identifier to avoid duplicate entries in the Delphi process. We collected basic demographics on the participants to ensure adequate representations across countries and disciplines. We reported using ranking orders, percentages and natural frequencies. All statistical analyses were conducted using Microsoft Excel 2013 (Microsoft Corp., Redmond, WA, USA).

Results

Participants and longlist of outcomes

In total, 71 health professionals (16 endocrinologists, 14 fertility specialists, two general obstetricians and gynaecologists, 21 academics active in PCOS research, five paediatricians, five specialist nurses and midwives, two primary care physicians, one occupational therapist, one psychologist, one pharmacist, and three dieticians) and 123 lay consumers from 17 countries (Australia, Belgium, Canada, Chile, China, Czech Republic, Estonia, France, India, Italy, Netherlands, South Africa, Spain, Sri Lanka, Sweden, UK, and USA) participated in the Delphi process. (Fig. 1) In the 2nd Delphi round, we received responses from 52 health professionals achieving a 74% response rate.

Initially, 60 outcomes were included in the longlist: 16 metabolic, 17 reproductive, 16 pregnancy, and 11 long-term outcomes. (Table I) Five additional outcomes were suggested by participants at the end of the 1st round and were included in the 2nd round; two outcomes by lay consumers (body image and treatment satisfaction) and three outcomes by health professionals (skin disorders, hepatic and visceral fat, adiponectin levels). At the time of conception of this longlist, we received the findings of the COMMIT core outcome set which identified all core outcomes for reporting on infertility treatment in women's health (Duffy and Farquhar, 2017). We included the following outcomes in our longlist and Delphi process to seek stakeholders' input on their relevance to PCOS research: viable pregnancy confirmed by ultrasound including singleton pregnancy, twin pregnancy, and higher multiples; pregnancy loss including miscarriage and stillbirth; live birth; gestational age at delivery; birthweight; neonatal mortality; and major congenital abnormalities (**AUTHOR:** please would you check that the punctuation is correct for this list? Thank you.). Three outcomes were judged as not particularly relevant to PCOS by the CMG and were not included in the Delphi process: termination of pregnancy, ectopic pregnancy, and time to pregnancy leading to live birth.

Delphi survey

After the 2nd round of the Delphi process, 40 out of 65 outcomes (62%) were identified as important for inclusion in the final core outcome set (Table I). Seven outcomes (7/65, 11%) were considered to be of low importance (endometriosis, adnexal adhesions, sexually transmitted disease, nipple discharge, induction of labour, cervical cancer, and ovarian cancer). All remaining outcomes (18/65, 28%) were equivocal with no clear consensus.

There was clear consensus for twenty-nine outcomes being considered important by both health professionals and lay consumers through all stages of the Delphi. (Table I) Eleven outcomes were identified as important by lay consumers but were not prioritised by health professionals by the end of the 2nd Delphi round (markers of cardiovascular disease, cerebrovascular disease, dysmenorrhea, thyroid function tests, major congenital abnormalities, endometriosis, adnexal adhesions, breast cancer, cervical cancer, ovarian cancer, and ovarian cysts). In contrast, three outcomes were considered to be important by health professionals but not by lay consumers (waist circumference, ovarian hyperstimulation syndrome, and baby birthweight).

Lay consumers' input led to a significant shift in health professionals' opinion, prioritising four outcomes as important by the end of the second Delphi round (coronary heart disease, reproductive hormonal profile, long-term offspring metabolic and development outcomes, and suicide attempts). Of the five additional outcomes added to the 2nd Delphi round, two were considered to be important towards the core outcomes set (skin disorders and treatment satisfaction).

Consultation meeting

Thirteen stakeholders participated in the final consultation meeting: two endocrinologists, four fertility specialists, two primary care physicians, two gynaecologists, and three lay consumers. The meeting panel acknowledged that given the varied clinical presentation of PCOS, it would be impractical to report on all the identified core outcomes in this set in each individual study. Therefore, the panel advocated dividing the final core set into generic outcomes (BMI, quality of life, and treatment satisfaction) to be reported in all future studies

and six specific additional outcome domains (metabolic, reproductive, pregnancy, psychological, oncological, and long-term outcomes) to be considered for reporting depending on the study's design, population characteristics and primary research focus.

Within the metabolic outcomes domain, the panel noted the high variability in measuring and reporting on waist/hip ratio in practice, thus the panel advocated its exclusion from the core set while keeping waist circumference. The panel felt that waist circumference was more relevant to studies investigating metabolic and cardiovascular outcomes in women with PCOS, in contrast to BMI which has correlation in all outcome domains, thus it was kept as a generic outcome. The panel also advocated the exclusion of metabolic syndrome from the core set while maintaining the reporting on its contributing components: type 2 diabetes, hypertension and lipid profile. The panel highlighted that measuring insulin resistance is only recommended in research settings and noted the difficulty of measuring it in clinical practice. They advocated the use of clamp studies, where possible, in mechanistic, experimental and laboratory-based research while substituting with simpler measures, such as oral glucose challenge test area under the curve, in larger-scale clinical studies.

Obstructive sleep apnoea, snoring, and daytime sleepness where voted as equivocal outcomes by both groups in the Delphi process. The panel acknowledged the increased prevalence of obstructive sleep apnoea in women with PCOS and its association with adverse health outcomes. However, those outcomes were not considered critical enough to be included as core.

Venous thromboembolic disease was considered a core outcome given its higher incidence in women with PCOS and the severity of associated morbidity (Okoroh *et al.*, 2015). The panel acknowledged that other adverse events, such as treatment side effects and allergic reaction

(Domecq *et al.*, 2013), could be of critical importance for reporting in clinical trials as per the principals of Good Clinical Practice in clinical research (Guideline, 2002), but none were specifically highlighted as core in this set.

In the reproductive outcomes domain, the panel considered subfertility to be a complementary outcome to live birth and viable pregnancy with high variation in its reporting and follow up periods. Therefore, subfertility was excluded in favour of keeping viable pregnancy, pregnancy loss and live birth as core. The panel deemed heavy menstrual bleeding to be less relevant to women with PCOS in contrast to menstrual regularity, thus the former was voted out of the final core set. Elements of hyperandrogenism (biochemical and clinical e.g hirsutism) were considered equally important and investigators are encouraged to report on both where possible using standardised tools, as highlighted by the 2018 evidence-based guidelines (Teede *et al.*, 2018).

All outcomes adopted from the Core Outcome Measures for Infertility Trials (COMMIT) core set (Duffy and Farquhar, 2017) were voted as core in our Delphi process. To avoid confusion, the panel considered all outcomes in the COMMIT set to be relevant to PCOS fertility studies, thus investigators evaluating reproductive outcomes in women with PCOS are encouraged to consider both sets for reporting on core reproductive outcomes as a minimum.

In the pregnancy outcomes domain, the panel acknowledged the higher risk of both pre-eclampsia and pregnancy-induced hypertension in women with PCOS and advocated the reporting on the full spectrum of hypertensive disease in pregnancy as per established definitions (The National Institute for Health and Care Excellence, 2019). The lay consumers on the panel expressed the importance of breastfeeding in mothers with PCOS to improve

both maternal and offspring outcomes. However, the panel consensus was not to include breastfeeding as a core outcome, as the relationship to PCOS was unclear, but rather to highlight its importance as an outcome favoured by lay consumers.

Both the health professionals and the lay consumers advocated the inclusion of offspring long-term metabolic and developmental outcomes in the core set. The panel acknowledged the evidence suggesting a link between fetal *in utero* exposure in mothers with PCOS and future adverse offspring metabolic and developmental outcomes such as obesity, metabolic syndrome, insulin resistance, and autism (Bell *et al.*, 2018; Kosidou *et al.*, 2016; Sir-Petermann *et al.*, 2009; Wilde *et al.*, 2018). However, the panel was also unable to recommend a set follow up period for the offspring of mothers with PCOS, nor suggest standardised measurement tools for reporting in this group. Given the difficulties associated with reporting on these outcomes, the panel acknowledged they would only be suitable for specific types of clinical studies with planned long-term follow-up. Further work is required to evaluate the prevalence and association of those metabolic and developmental outcomes in the offspring of mothers with PCOS to then prioritise core outcomes of importance for future studies.

Two oncology related outcomes were prioritised by the Delphi process: endometrial hyperplasia and endometrial cancer. Given the high association between both outcomes and the common pathophysiology, the panel advocated combining them into one core outcome reporting on abnormal endometrial proliferation in women with PCOS.

Four psychological outcomes were prioritised by the Delphi process, all highly emphasised by lay consumers (anxiety, depression, eating disorders, and suicidal attempts). The panel

acknowledged the lack of a standardised definition and measurement tools to report on suicidal attempts in the context of randomised trials and therefore excluded it from the final core set, keeping the three remaining psychological outcomes.

Discussion

Summary of findings

In this study, we report on the development of the first core outcome set for harmonising PCOS research worldwide, to our knowledge. The final core set included 33 outcomes categorised in seven clinical practice domains (**AUTHOR:** please would you recheck: six or seven domains?). We leveraged extensive evidence syntheses on PCOS (40 systematic reviews) from the International PCOS guideline to capture the full range of outcomes and engaged a wide multidisciplinary stakeholder panel from high-, middle-, and low-income countries in a Delphi and workshop process. Lay consumer input had a pivotal role in the development of this core set, exemplified by focus on specific outcome domains such as mental health.

Strength and limitations

We used a robust methodology to identify outcomes relevant to PCOS research and to reach consensus among stakeholders. We registered our study prospectively and used predefined consensus criteria to identify outcomes of core importance. Stakeholders participated anonymously in the Delphi process to maintain their autonomy and avoid overt influence of particular individuals or stakeholder groups on the final score (Okoli and Pawlowski, 2004). We ensured sufficient representation of all relevant stakeholder groups from high-, middle-

and low-income countries and collaborated with leading professional charities and lay consumer support groups to expand our pool of participants. We employed a special survey for lay consumers using lay terminology to promote their effective participation in the Delphi process. We held a final consultation meeting and engaged a panel of all participating stakeholder groups promoting an interactive forum to agree on equivocal outcomes, and to discuss the practical implementation of the final core set.

Our findings are limited by the 26% attrition rate in the 2nd Delphi round, which could have influenced the final list of prioritised outcomes. This, however, is not uncommon in Delphi methodology (Dos Santos *et al.*, 2018; Al Wattar *et al.*, 2017). We were unable to hold focus groups or structured interviews with lay consumers, which may have limited our understanding of their choices on key outcomes. Still, we engaged a large number of lay consumers from many countries and ensured adequate representation in the final consultation meeting. To ensure feasibility, we combined some outcomes under one label (e.g. lipid profile); including all individual outcomes in the Delphi process might have changed the final set.

Implications for future research

The diverse clinical features of PCOS demand studies of different design and focus to address the current research need. To aid the implementation of this core set in practice, we divided outcomes into different outcome domains to cover the varied pathophysiology of PCOS. Investigators are encouraged to adapt their primary reporting according to the clinical focus of their study and their established research question, aiming to cover all relevant core outcomes in this set. For example, studies evaluating fertility treatments in a non-pregnant

PCOS population might not be able to report on the core outcomes within the oncology domain, but should aim to report on all generic core outcomes in addition to those in the reproductive domain, while justifying the lack of reporting on any remaining outcome domains. We also encourage investigators to consult all additional core sets that might apply to studies on women with PCOS within the CoRe Outcomes in Women's and Newborn health (CROWN) and the Core Outcome Measures in Effectiveness Trials (COMET) initiatives' databases, given the diverse nature of PCOS. Thus, in the same previous example, researchers evaluating fertility treatments in women with PCOS are encouraged to report on the generic and reproductive outcomes in both this HARP (HARmonising research outcomes for Polycystic ovary syndrome) (**AUTHOR:** can HARP be defined here?) and the COMMIT fertility core outcome sets (Duffy and Farquhar, 2017).

The voice of lay consumers was strong in the development of this core outcome set and led to a significant change in the convergence of consensus among participating stakeholders. This was more evident for mental health, offspring, and pregnancy outcomes. Traditionally, those outcomes have been poorly reported on in the literature (Teede *et al.*, 2018) and we hope that implementing this core set would help to raise their profile, ultimately increasing research impact on women's health and the whole society. A major challenge to adopting all the views of lay consumers was related to the lack of clear definitions and standardised measurement tools for some outcomes especially in the case of long-term offspring follow-up.

We aimed to generate a list of recommended measurement tools to report on the identified core outcome set following on from the recommendations of the international guideline (Teede *et al.*, 2018), however, some outcomes such as insulin resistance lacks unanimity.

Further research work is required to harmonise reporting on these outcomes in PCOS studies with input from all involved stakeholders including lay consumers. However, several core outcomes lacked an internationally standardised measurement tool, such as insulin resistance. We plan to investigate this further to develop, harmonise and standardise relevant missing measurement tools to facilitate the implementation of this core set.

Conclusion

Researchers are encouraged to adopt this core set of 33 outcomes in future studies on women with PCOS to standardise reporting and enable impactful evidence synthesis.

Acknowledgements

The authors acknowledge the support of the Australian Centre for Research Excellence in PCOS who led the guideline development and evidence synthesis and the AE-PCOS society and the Verity UK charity.

Authors' roles

BHA drafted the protocol and the 1st manuscript, moderated the Delphi process and analysed the data; RG helped to moderate the Delphi process and edited the final manuscript; HT led the evidence synthesis process and identified and contacted health professional stakeholders and oversaw the study design and conduct, ST oversaw the study design and conduct and edited the final manuscript; all remaining co-authors helped in data curation and edited the final manuscript.

Funding

Evidence synthesis was funded through the Australian government, National Health and Medical Research Council (NHMRC) Centre for Research Excellence in PCOS and HT is funded through an NHMRC fellowship. BHA is funded through an NIHR lectureship.

Conflict of interest

None

References

Androgen Excess and PCOS Society. Available at: http://www.ae-society.org/.

Azziz R, Carmina E, Chen Z, Dunaif A, Laven JSE, Legro RS, Lizneva D, Natterson-Horowtiz B, Teede HJ, Yildiz BO. Polycystic ovary syndrome. *Nat Rev Dis Prim* 2016;**2**:16057.

Bell GA, Sundaram R, Mumford SL, Park H, Mills J, Bell EM, Broadney M, Yeung EH. Maternal polycystic ovarian syndrome and early offspring development. *Hum Reprod* 2018;**33**:1307–1315.

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. *Hum Reprod* 2016;**31**:2841–2855.

Domecq JP, Prutsky G, Mullan RJ, Sundaresh V, Wang AT, Erwin PJ, Welt C, Ehrmann D, Montori VM, Murad MH. Adverse effects of the common treatments for polycystic ovary syndrome: A systematic review and meta-analysis. *J Clin Endocrinol Metab* 2013;**98**:4646–4654.

Duffy JMN, Farquhar C. COMMIT: Core Outcome Measures for Infertility Trials. 2017. Available at: http://www.comet-initiative.org/studies/details/1023.

Guideline ICHHT. *E6: Note for Guidance on Good Clinical Practice*. PMP/ICH/135/95). European Medicines Agency, 2002.

Khan K, O'Donovan P. The CROWN Initiative: journal editors invite researchers to develop core outcomes in women's health. *BMC Pregnancy Childbirth* 2014;**14**:199.

Kirkham JJ, Gorst S, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, Moher D, Schmitt J, Tugwell P. Core outcome set–STAndards for reporting: the COS-STAR

455 statement. *PLoS Med* 2016;**13**:e1002148.

456 Kosidou K, Dalman C, Widman L, Arver S, Lee BK, Magnusson C, Gardner RM. Maternal
457 polycystic ovary syndrome and the risk of autism spectrum disorders in the offspring: a
458 population-based nationwide study in Sweden. *Mol Psychiatry* 2016;**21**:1441.

459 Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design
460 considerations and applications. *Inf Manag* 2004;**42**:15–29.

461 Okoroh EM, Boulet SL, George MG, Hooper WC. Assessing the intersection of
462 cardiovascular disease, venous thromboembolism, and polycystic ovary syndrome.
463 *Thromb Res* 2015;**136**:1165–1168.

464 Dos Santos F, Drymiotou S, Antequera Martin A, Mol BW, Gale C, Devane D, van't Hooft J,
465 Johnson MJ, Hogg M, Thangaratinam S. Development of a core outcome set for trials
466 on induction of labour: an international multistakeholder Delphi study. *BJOG An Int J*
467 *Obstet Gynaecol* 2018;**125**:1673–1680.

468 Sir-Petermann T, Codner E, Pérez V, Echiburú B, Maliqueo M, Ladron de Guevara A,
469 Preisler J, Crisosto N, Sánchez F, Cassorla F. Metabolic and reproductive features
470 before and during puberty in daughters of women with polycystic ovary syndrome. *J*
471 *Clin Endocrinol Metab* 2009;**94**:1923–1930.

472 Tay CT, Joham AE, Hiam DS, Gadalla MA, Pundir J, Thangaratinam S, Teede HJ, Moran
473 LJ. Pharmacological and surgical treatment of nonreproductive outcomes in polycystic
474 ovary syndrome: An overview of systematic reviews. *Clin Endocrinol (Oxf)*
475 2018;**89**:535–553.

476 Tay CT, Moran LJ, Wijeyaratne CN, Redman LM, Norman RJ, Teede HJ, Joham AE.
477 Integrated Model of Care for Polycystic Ovary Syndrome. *Semin Reprod Med*

2018;**36**:86–94.

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 Med* 2010;**8**:41.

Teede H, Misso M, Costello M, Dokras A, Laven J, Moran L, Piltonen T, Norman R. International evidence-based guideline for the assessment and management of polycystic ovary syndrome. 2018. Available at: monash.edu/medicine/sphpm/mchri/pcos.

The National Institute for Health and Care Excellence. Hypertension in pregnancy: diagnosis and management. 2019. Available at: https://www.nice.org.uk/guidance/ng133.

Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L. OMERACT: an international initiative to improve outcome measurement in rheumatology. *Trials* 2007;**8**:38.

University of Michigan. Simplification Guide to Medical Terms. Available at: https://medicine.umich.edu/medschool/research/office-research/institutional-review-boards/guidance/simplification-guide-medical-terms.

Al Wattar BH, Tamilselvan K, Khan R, Kelso A, Sinha A, Pirie AM, McCorry D, Khan KS, Thangaratinam S. Development of a core outcome set for epilepsy in pregnancy (E-CORE): a national multi-stakeholder modified Delphi consensus study. *BJOG An Int J Obstet Gynaecol* 2017;**124**:661–667.

Wattar BHA, Khan K, Teede H, Thangaratinam S. Developing a Core outcome set for women with polycystic ovary syndrome. 2018. Available at: http://www.comet-initiative.org/studies/details/1115.

Wilde MA de, Eising JB, Gunning MN, Koster MPH, Evelein AM V, Dalmeijer GW, Uiterwaal CSPM, Eijkemans MJC, Ent CK van der, Meijboom FJ. Cardiovascular and

metabolic health of 74 children from women previously diagnosed with polycystic ovary syndrome in comparison with a population-based reference cohort. *Reprod Sci* 2018;**25**:1492–1500.

Williamson P, Altman D, Blazeby J, Clarke M, Gargon E. Driving up the quality and relevance of research through the use of agreed core outcomes. 2012.

Figure legend

Figure 1 Flow chart of the modified Delphi method to develop a core outcome set for polycystic ovary syndrome.

PCOS: polycystic ovary syndrome