National Institute for Health and Care Excellence

Polycystic ovary syndrome

Adaptation report 2 – Prevalence, screening and management of psychological features and models of care

NICE guideline [NGXX]

July 2026

Draft for Consultation

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN: [XXX]

Contents

2 Prevalence, screening and management of psychological features and models of care

The adaptation reports were produced using the reviews from the International Guideline (IG). Any further details can be found in the technical report from the IG, including results of the analyses and full study references.

General principles

Recommendations from the International evidence-based guideline for PCOS*

Practice points:

- Psychological features are common and important component of PCOS that all health professionals should be aware of.
- Funding bodies should recognise that PCOS is highly prevalent, has significantly higher psychological disorders which should be prioritised and funded accordingly.

The IG included general principles practice points at the beginning of their section on prevalence, screening and management of psychological features and models of care, which were not evidence-based. Practice points are those which arose from the IG committee discussion of the evidence and EBR and CR recommendations.

NICE economic evidence

No health economic evidence on the general principles of care for PCOS was identified in the health economic literature search conducted by NICE.

NICE recommendations

The relevant recommendations for this section are Rec 1.2.1 and 1.2.2.

The committee's discussion

Clinical

The committee made two new recommendations relating to overall care for people with PCOS. This was not an area that was specifically covered by the IG and as such there is no evidence base presented in the IG for this area. The committee felt it was important to ensure continuity of care between services and to highlight the importance of patient preference when making decisions about treatment options. As such two new NICE recommendations were made to provide guidance on creating a patient focussed service.

Health economic

The committee noted that the recommendations made for this section of the guideline are reflective of best clinical practice. The committee also acknowledged that these recommendations are overarching for the entirety of the guideline and therefore concluded that overall cost-effectiveness is discussed iteratively throughout the guideline for each respective review question. The committee also noted that no significant resource impact would be associated with the recommendations made for this section of the guideline in relation to general principles of care.

2.1 Quality of life

Review questions 2.1: In women with PCOS, what is the prevalence and severity of reduced QoL?

2.1.1 Recommendations from the International evidence-based guideline for PCOS*

Evidence based recommendation:

2.1.1 Healthcare professionals and women should recognise the adverse impact of PCOS and/or PCOS features on quality of life in adults.

Practice point:

2.1.2 Women with PCOS should be asked about their perception of PCOS related symptoms, impact on quality of life, key concerns and priorities for management.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist: In women with PCOS, what is the prevalence and severity of reduced QoL?

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(PICO was clear and detailed).*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(detailed PICO available which is suitable for the review)*.
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(criteria is detailed and explains the use of accuracy data where methods could be compared, which is the*

Section	Question	Answer
		practice point of the question that mentions tools/methods to assess QoL. Study quality exclusions detailed and appropriate, but they state any study lower than a cross-sectional would be an exclusion without naming the types of studies it would exclude, and the review includes case-control studies. There were two questions covered by the review Q2.2 is a narrative review on which QoL domain is most affected).
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(restrictions appear appropriate but are not clearly reported. This was an updated review, and a new search was done but does not state dates, but in the search strategy most databases searched from 2017)*.
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(English language filter used, no others noted)*.
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(The PICO has a lot of detail but there are some areas it could be more clearly reported)*.
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Yes *(Sources were bibliographic databases (Medline; Embase; PsycInfo; CINAHL; Cochrane "All EBM") and references brought forward from prior guideline. Registries are less useful for this type of question - non-interventional studies do not usually have to be pre-registered. Embase is a useful source for conference abstracts)*.
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	No *(Other than including refs from previous guideline no non-database sources listed in PRISMA flow chart (p1302))*.

Section	Question	Answer
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(There was a reasonable mix of search terms for Medline but there is an error in the CINAHL strategy with random full stops on line 10 that affect retrieval. No translation of MeSH terms for other databases which should have been done. For example, PsycInfo has narrower terms beneath "Quality of Life" (e.g. "Health-Related Quality of Life") which would have been missed. The numbers are also very low given that backdating with a simplified version of the Embase strategy without all the same terms 720 results were retrieved which is more than was retrieved for all databases combined in the IG)*.
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(English language limit is appropriate).*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(2 reviewers plus evidence team for consultation - appropriate for reducing error in study selection).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	Unclear *(There were problems with the search strategy and the impact of this is unclear).*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(minimised by having 2 reviewers for study selection along with evidence team for discussion where needed)*.
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(detailed study table with study characteristics table and structured extraction form were present allowing interpretation of results)*.
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Probably yes *(all study results appear to have been included, 36 studies in SR, 27 were suitable for GRADE/evidence profiles.*

Section	Question	Answer
		Assume this is because some were included in just the narrative review).
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(quality appraisal for each included study was completed which included a risk of bias check list and GRADE conducted)*.
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Probably yes *(minimised by 2 reviewers appraising studies alongside evidence team)*.
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Risk of bias was appraised and GRADE conducted on results. Some of the evidence is presented narratively)*.
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(36 studies included in SR, 27 in GRADE profiles/tables. Detailed meta-analysis/descriptive analysis summary lists how many studies were included in each meta-analysis for each QoL domain. Appears that studies are accounted for in analysis)*.
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(no information was given regarding whether the planned analysis of the results was undertaken. Adolescents, ethnicity and phenotype were subgroups stated in the PICO. Adolescent tools are analysed separately in comparison 2)*.
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Unclear *(meta-analysis and descriptive analysis was used appropriately for this review question, study designs were appropriate for analysis, except for case control studies were included whereas we would not normally include where there are sufficient cross-sectional studies. Median results were meta-analysed, whereas we would not usually include medians*

Section	Question	Answer
		because this data is likely to be skewed).
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Probably yes *(heterogeneity high across all domains on forest plots. They discuss heterogeneity, it is downgraded in GRADE where relevant and all forest plots used random effects for all outcomes regardless of heterogeneity and SMD)*.
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Yes *(funnel plots present but lack of descriptive summary)*.
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Unclear *(biases downgraded in extraction table but not a lot of detail for each domain, just overall risk of bias. GRADE is conducted on evidence, but studies were observational and were mainly assessed to be low risk of bias whereas we would have started them as low quality in GRADE for risk of bias, but they seem to start as high quality. The GRADE ratings overall do not match the number of downgraded domains in GRADE).*
Synthesis and findings	Concerns regarding the synthesis and findings	Unclear *(The evidence has been assessed for risk of bias and downgraded in GRADE, however it is not downgraded as much as NICE would downgrade. Heterogeneity is acknowledged but is not resolved).*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(The PICO has a lot of detail but there are some areas it could be more clearly reported).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	Unclear *(there were problems with the search and the impact of this is unclear),*

Section	Question	Answer
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Risk of bias was appraised and GRADE conducted on results. Some of the evidence is presented narratively).*
Judging risk of bias	Concerns regarding the synthesis and findings	*Unclear* *(The evidence has been assessed for risk of bias and downgraded in GRADE, however it is not downgraded as much as NICE would downgrade. Heterogeneity is acknowledged but is not resolved).*
Overall review ratings	Overall risk of bias	Moderate *(there were problems with the search terminology so it is unclear how this would affect the number of studies identified and there are differences from NICE in the downgrading of GRADE).*
Overall review ratings	Applicability as a source of data	Fully applicable

Evidence to recommendations justification: There were lower quality of life scores in the PCOS group for all tools (SF-36, WHOQOL-BREF and PCOSQ) measuring quality of life, however there was a very high level of heterogeneity across all outcomes which was not resolved by subgroup or sensitivity analyses. Two studies did not use norms-based scoring, but removal of these studies did not change the heterogeneity. Three studies reported medians and IQRs, and this was used as the mean and standard deviation within the analysis. The review does not state what would be a clinically meaningful difference between the PCOS and non-PCOS groups, but the overall statistical trend is towards lower quality of life scores for those with PCOS. The strength of the evidence was low to very low for GRADE. Limited findings from adolescent studies (Paediatric Quality of life inventory) did not show a

difference between those with and without PCOS. The panel discussion from the IG acknowledged that the certainty of evidence is low and is based on observational data, but the evidence is consistent across disease specific and generic tools. Therefore, the EBR “Professionals and women should recognise the adverse impact of PCOS and/or PCOS features on quality of life in adults” is not too prescriptive, and therefore an acceptable recommendation given the evidence.

IG economic evidence

No health economic evidence was identified in the IG for review question 2.1 on quality of life.

2.1.2 NICE economic evidence

Included studies

A single health economic search was performed by NICE to identify published economic evaluations of relevance to all review questions in this guideline. See the literature search strategy in Appendix A.

No economic studies were identified which were applicable to this review question (see economic study selection flow chart in Appendix B).

Excluded studies

No economic studies were reviewed at full text and excluded from this review.

Economic model

No original health economic modelling was conducted for review question 2.1 on quality of life. This review question was not concerned with addressing decisions between competing alternatives and therefore does relate to resource optimisation.

2.1.3 NICE recommendations

The relevant recommendation for this section is Rec 1.1.1.

2.1.4 The committee's discussion and interpretation of the evidence

Clinical

The committee acknowledged the lay members' experiences of how PCOS had been detrimental to their quality of life and psychological wellbeing. They felt that earlier knowledge of their condition would have helped them understand their experiences more fully. A recurring theme was the need for compassion and for health care practitioners to recognise and understand the impact of PCOS. This recommendation was made to encourage practitioners to have meaningful conversations with people about their PCOS and to recognise how its effects can be wide-ranging, can change across the life course, and be cumulative. Ongoing discussion of the priorities and concerns of people with PCOS is therefore essential.

The IG included an EBR for healthcare professionals to recognise the adverse impact of PCOS on quality of life, along with a practice point encouraging them to ask about people's perceptions of their PCOS-related symptoms, the impact on their quality of life, and their key concerns and priorities for management. The NICE recommendation adapted this into standard NICE style and made more active by suggesting a kind, sensitive and age-appropriate approach to be taken in line with NICE's guidelines on patient experience; asking for permission before discussing sensitive subjects was an additional point; and asking about their experiences echoed the practice point of asking about their perception of symptoms, quality of life impact and concerns and priorities. Many lay members described how their quality of life had been affected as early as menarche. In addition, the term *psychological wellbeing* was added to the recommendation.

Health economic

No health economic evidence was included in the IG or in NICE's health economic literature review for review question 2.1 on quality of life.

The committee contextualised the one EBR in the IG (2.1.1) and the one practice point recommendation (2.1.2). The committee noted that these contextualised recommendations were concerned with asking people about their experiences of PCOS and how it affects their quality of life and psychological wellbeing. It was acknowledged that this would be conducted at existing healthcare appointments and therefore the cost of implementing this would likely be minimal and take the form of additional healthcare professional's time. The committee emphasised the importance of this recommendation, and the positive impact it would have for people with PCOS, due to the number of people with PCOS who feel misunderstood or inadequately cared for in current practice. The committee acknowledged the potential for additional costs in terms of additional staff time but concluded that these recommendations would likely be implemented at existing healthcare appointments. The committee also acknowledged that these recommendations align to the principles covered in NICE's existing guidelines (such as NICE's patient experience guideline). The committee were therefore confident that these recommendations are cost-effective.

The committee also discussed the benefits associated with these recommendations, noting the potential for improved mental health by reducing the stigma of mental health for people with PCOS and simultaneously raising awareness of the challenges people with PCOS may face.

Overall, the committee concluded that it was unlikely any significant resource impact would be associated with their recommendation.

2.2 Depression and anxiety

Review question 2.2: In women with PCOS, what is the prevalence and severity of depression and anxiety?

2.2.1 Recommendations from the International evidence-based guideline for PCOS*

Evidence-based recommendations:

2.2.1 Healthcare professionals should be aware of the high prevalence of moderate to severe depressive symptoms and depression in adults and adolescents with PCOS and should screen for depression in all adults and adolescents with PCOS, using regionally validated screening tools.

2.2.2 Healthcare professionals should be aware of the high prevalence of moderate to severe anxiety symptoms and anxiety disorders in adults and should screen for anxiety in all adults with PCOS, using regionally validated screening tools.

Consensus recommendation:

2.2.3 If moderate or severe depressive or anxiety symptoms are detected, practitioners should further assess, refer appropriately or offer treatment.

Practice point:

2.2.4 Severity of symptoms and clinical diagnosis of depression or anxiety should guide management. The optimal interval for anxiety and depression screening is not known. A pragmatic approach could include screening at diagnosis with repeat screening based on clinical judgement, risk factors, comorbidities and life events, including the perinatal period. Screening for mental health disorders comprises assessment of risk factors, symptoms, and risk of self-harm and suicidal intent.

*Copyright Monash University Melbourne Australia 2023. Reproduced with permission from Monash University.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(clear and detailed PICO).*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(detailed PICO available which is suitable for the review)*.
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(eligibility criteria clearly described in PICO).*
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(restrictions on date are present but not well described, any observational studies included)*.
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(English language and full publication limits are appropriate for review question).*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(criteria adequately detailed in PICO).*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Yes *(4 named databases used, CINAHL not used as not used in Cooney et al which they are updating. Bibliographic databases and references brought forward from previous review. Trial registries are less important for prevalence-type questions).*

Section	Question	Answer
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably yes *(12 additional studies were included from systematic reviews identified in the search but no other mention of manual searching methods. General methods states that they checked with experts for any relevant studies).*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Yes *(full search strategy is detailed in the review, the search terms and dates appear appropriate for the review question. Some Embase terms are not main ones but seem to map over effectively).*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(not explicitly stated but appears to have been appropriate. English language limit discouraged by ROBIS but unlikely to induce bias).*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(2 reviewers for title and abstract review, suitable for review).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	Low *(No searching beyond databases but not this is not overly problematic for a prevalence question).*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(2 reviewers for study selection and review).*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(detailed study table available for interpretation of results).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(evidence summary describes studies included in meta-analysis and which could not be included with reasons).*
Data collection and	Was risk of bias (or methodological quality)	Yes *(detailed risk of bias assessment*

Section	Question	Answer
study appraisal	formally assessed using appropriate criteria?	*table for each study but no full study extraction).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Yes *(2 reviewers completed this).*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(risk of bias provided and they provided details of which studies were able to be meta-analysed).*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(47 studies included in GRADE tables, 43 for depression, 27 for anxiety, evidence summary details that a lot of studies included in the review were not suitable for meta-analysis).*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(no information was given regarding the planned analysis of the results in the PICO but they sub-grouped by adults and adolescents and by QoL tool and studies that used clinical interviews and sensitivity analyses conducted for 'trait' or 'state' scores only, but heterogeneity was not resolved).*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(meta-analysis was used appropriately for this review question; study designs were appropriate for analysis except we would not include case control studies when enough cross-sectional studies were available).*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Probably yes *(significant heterogeneity between studies, I^2 73% p<0.001, the authors appear to have taken this into account in their summary. Random effects model was used, and they downgraded in GRADE for inconsistency).*

Section	Question	Answer
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Probably yes *(funnel plots displayed for each outcome but lack of summary).*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Probably yes *(there is no structured extraction form, but GRADE is conducted and reason for downgrading for risk of bias is presented. However, they downgrade once for risk of bias because evidence was from observational studies and most being of moderate or high risk of bias. Also, there are more than one domain downgraded but overall GRADE rating is higher than we would rate).*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(the synthesis was appropriate, biases assessed and downgraded in GRADE, inconsistency addressed but not resolved).*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(criteria adequately detailed in PICO).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	Low *(no searching beyond databases but not a huge worry for a prevalence question).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(risk of bias provided and they provided details of which studies were able to be meta-analysed).*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(the synthesis was appropriate, biases assessed and downgraded in GRADE, inconsistency addressed but not resolved).*
Overall review ratings	Overall risk of bias	Low

Section	Question	Answer
		(eligibility criteria, identification of studies, risk of bias and synthesis and findings were all appropriate).
Overall review ratings	Applicability as a source of data	Fully applicable

Evidence to recommendations justification: Many studies were included in the evidence synthesis and the IG stated that the evidence pointed toward statistically significant high prevalence of depression in those with PCOS compared to those who did not. Prevalence of depression in adults had an OR of 2.63 (95% CI 2.12 to 3.28) from 41 studies (n=53,254) and anxiety an OR of 2.89 (95% CI 2.27 to 3.68) from 23 studies (n=14,519); adolescents an OR of 2.26 (95% CI 1.36 to 3.76) for depression from 6 studies (n=1098) and anxiety an OR of 0.92 (95% CI 0.11 to 7.96) from 3 studies (n=455). However, it should be noted most studies had a high or moderate risk of bias, with overall GRADE rating low or very low, except for adolescents' prevalence of depression with a GRADE rating of moderate. There was high heterogeneity for adult studies, which was not resolved by subgroup analysis of studies that used a clinical interview. The studies were conducted in various countries with only one study from the UK (Barry 2011), which only included depression score data, but did follow the trend of higher prevalence in those with PCOS.

Depression scores in adults showed the most notable difference between those with and without PCOS with an overall SMD of 0.76 (95% CI 0.58 to 0.94). Adolescents' depression scores showed an SMD of 0.41 (95% CI 0.13 to 0.70). Anxiety scores SMD 0.58 (95% CI 0.41 to 0.76) in adults and SMD 0.23 (95% CI -0.19 to 0.64). There was a lot of variation in method of measurement, however they did analyse the anxiety and depression scores by method used. This found significant subgroup differences between tools used for depression (BDI tool showed a high mean difference of 4.91 (95% CI 4.06 to 5.76) or SMD of 1.02 (95% CI 0.74 to 1.31) with high heterogeneity.

Anxiety scores showed a significant subgroup difference, but mean differences were not as profound.

Therefore, the evidence justifies the assertation that healthcare professionals should be aware of the high prevalence of moderate to severe depressive symptoms, depression and anxiety in adults and adolescents with PCOS and the high prevalence of moderate to severe anxiety symptoms and anxiety disorders in adults.

IG economic evidence

No health economic evidence was identified in the IG for review question 2.1 on depression and anxiety.

The IG noted in their discussion of the evidence that the guideline developers working on this review question had different opinions on the resource requirements to implement their recommendations associated with this review question. However, no further information was provided in the IG detailing these differences of opinion.

It was also noted that equitable access to appropriately trained health professionals for mental health varies considerably. It was also acknowledged that workforce training may need to be considered to implement their recommendations – noting the possible variations in the presentation of mental health disorders and the need to conduct screening in a culturally sensitive manner.

2.2.2 NICE economic evidence

Included studies

A single health economic search was performed by NICE to identify published economic evaluations of relevance to all review questions in this guideline. See the literature search strategy in Appendix A.

No economic studies were identified which were applicable to this review question (see economic study selection flow chart in Appendix B).

Excluded studies

No economic studies were reviewed at full text and excluded from this review.

Economic model

No original health economic modelling was conducted for review question 2.2 on depression and anxiety as the committee concluded that any recommendations with the potential for associated costs, or resource impact, for the NHS would likely already be covered by NICE's existing guidelines on anxiety and depression.

2.2.3 NICE recommendations

The relevant recommendations for this section are Rec 1.14.1 and 1.14.2.

2.2.4 The committee's discussion and interpretation of the evidence

Clinical

The committee agreed with the IG on the high prevalence of depression and anxiety in those with PCOS and so contextualised the recommendation relating to this. However, the identification, assessment and management of depression and anxiety is covered by NICE guidelines and so a cross-reference to these guidelines was made.

Health economic

No health economic evidence was identified in the IG or in NICE's health economic literature search for review question 2.2 on depression and anxiety.

The committee contextualised the two EBRs from the IG into one recommendation. The committee noted that as this recommendation was concerned with making people aware of the prevalence of depression and anxiety for people with PCOS, no resource implications are associated with this recommendation.

The committee also cross-referred to NICE's existing guidelines on depression (NG222 and NG134) and anxiety (CG113), which provides information on how both depression and anxiety should be identified, assessed and managed.

2.3 Psychosexual function

Review question 2.3: In women with PCOS what is the prevalence and severity of psychosexual dysfunction?

2.3.1 Recommendations from the International evidence-based guideline for PCOS*

Consensus recommendations:

2.3.1 Healthcare professionals could consider the multiple factors that can influence psychosexual function in PCOS including higher weight, hirsutism, mood disorders, infertility and PCOS medications.

2.3.2 Permission to discuss psychosexual function should be sought noting that the diagnosis of psychosexual dysfunction requires both low psychosexual function, combined with related distress.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Criteria well described, studies appear to meet this).*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(detailed PICO available which is suitable for the review).*
Study eligibility criteria	Were eligibility criteria unambiguous?	Probably yes *(eligibility criteria clearly described in PICO although the comparison was those without PCOS and also those under 14 years of age. It is not clear why under 14 is specified).*

Section	Question	Answer
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Probably yes *(restrictions on date are present but are not well described, updated review but no date limits. No limits on study quality, any original study included e.g. case-control, RCT, cross-sectional, whereas we would prioritise higher quality studies).*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(English language limit is appropriate).*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(it is not clear why under 14 years of age is important but otherwise appropriately detailed).*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(no trial registry records but not as relevant for prevalence-type questions as interventions. Conference abstracts are excluded - presumably at the sifting stage. Other than that, there was a reasonable selection of bibliographic databases).*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	No *(no information is given about additional methods such as manual searches. Google Scholar is the only alternative source listed in PRISMA flow chart and I would consider that similar to a database)*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Yes *(full search strategy is detailed in the review, the search terms and dates appear appropriate for the review question. Search strategies are reasonable and accurately translated).*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(not explicitly stated but appears to have been appropriate. Exclusion criteria*

Section	Question	Answer
		discounts theses, conferences non-English language papers. These limits are not incorporated in the search but would be equivalent to NICE standard practice).
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(2 reviewers for study selection and appraisal).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	Low *(limited searching beyond databases is the only real concern but for a prevalence-type question (compared to effectiveness/intervention-type) this is not a great concern).*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(2 reviewers likely reduced errors in this area).*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(study table with study characteristics present allowing interpretation of results)*.
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(2 studies excluded from initial 29 includes due to outliers, 27 included in meta-analyses).*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(risk of bias assessments have been completed and appear appropriate, provide details in structured extraction form for individual studies and GRADE conducted).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Probably yes *(reasonable to have 2 reviewers for this).*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(sufficient study characteristics and risk of bias assessment and GRADE conducted).*

Section	Question	Answer
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(2 studies excluded for outliers, remaining 27 used for meta-analyses, details still provided for studies).*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(no information was given regarding the planned analysis of the results except comparing those over and under 14 years. However, sensitivity and subgroup analyses were conducted).*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Probably yes *(meta-analyses were conducted appropriately for this review question, study designs were appropriate for analysis. Mean differences meta-analysed, no SMD. We would not have meta-analysed the different scales without SMD but they do just look at the method of the majority alone (FSFI) and subgroup).*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Probably yes *(serious inconsistency noted in results table for outcomes of total sexual function, sexual desire, sexual arousal, lubrication, orgasm, satisfaction and pain. All other outcomes inconsistency is not applicable. Random effects analysis used but still high heterogeneity. Subgroup analyses by fertility status and BMI conducted and showed a difference between subgroups for fertility for total sexual function and fertility and BMI for total satisfaction and all were downgraded in GRADE for serious inconsistency).*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Probably yes *(funnel plots available for sexual function with results clustered, however due to study/intervention types a written explanation would have been useful. Sensitivity analysis for FSFI method of measurement only, which did not resolve the heterogeneity).*
Synthesis and findings	Were biases in primary studies minimal or	Probably yes *(no obvious discussion about this but*

Section	Question	Answer
	addressed in the synthesis?	*appears reasonable and GRADE is conducted).*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(high heterogeneity but it is addressed).*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(Not clear why under 14 years of age is important but otherwise appropriately detailed).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	Low *(Limited searching beyond databases is the only real concern but for a prevalence-type question (compared to effectiveness/intervention-type) this is not a great concern).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(sufficient study characteristics and risk of bias assessment and GRADE conducted).*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(high heterogeneity but it is addressed).*
Overall review ratings	Overall risk of bias	Low *(low for all concerns although there was a lot of heterogeneity in the analyses, and some effect sizes were very small and SMD was not used when there were different scales used).*
Overall review ratings	Applicability as a source of data	Partially applicable

Evidence to recommendations justification: The IG states that pooled analysis showed that there were no differences in sexual desire (MD -0.22 (95% CI -0.47 to 0.03) (16 studies; n=2498) or pain (-0.27 (95% CI -0.57 to 0.03) (13 studies n=2003) between women with and without PCOS, but women with PCOS had lower total sexual function, sexual arousal, lubrication, orgasm and satisfaction. Total sexual function had an MD of -2.42 (95% CI -

3.26 to -1.58) (17 studies n=2143); Sexual arousal had MD -0.36 (95% CI -0.59 to -0.13) in 14 studies (n=2177); lubrication MD -0.47 (95% CI -0.75 to -0.20) in 14 studies (n=2136), orgasm MD -0.35 (95% CI -0.52 to -0.17) (15 studies n=2207) and satisfaction MD -1.48 (95% CI -2.21 to -0.75) in 19 studies n=2994). However, this was based on statistical significance and clinical difference was not discussed in the IG. The mean differences for sexual arousal, lubrication and orgasm were not showing really large differences between the groups, but the trend was towards lower values for the PCOS group. Visual Analogue Score (VAS) in 2-4 studies showed no difference for importance of sexual thoughts and fantasies, importance of sexual satisfaction and pain during intercourse, however these were not pooled. Lower values were found in sex life satisfaction, body hair impact, social contact impact and sexual attractiveness. There was a high level of heterogeneity and the IG notes that the subgroup numbers of the current meta-analysis were too small to result in clinically relevant data. The IG concludes that it was a small but statistically significant reduction of sexual function in all assessed domains: desire, arousal, lubrication, orgasm and pain but that the relationship between PCOS and/or varying characteristics of PCOS and psychosexual dysfunction remains unclear. Therefore, this justifies the lack of any EBRs, but CRs were made regarding health care practitioners to consider the multiple factors that can influence psychosexual function in PCOS, including higher weight, hirsutism, mood disorders, infertility and PCOS medications. Furthermore permission to discuss psychosexual function should be sought noting that the diagnosis of psychosexual dysfunction requires both low psychosexual function, combined with related distress.

IG clinical evidence

No health economic evidence was identified in the IG for review question 2.3 on psychosexual function.

2.3.2 NICE economic evidence

Included studies

A single health economic search was performed by NICE to identify published economic evaluations of relevance to all review questions in this guideline. See the literature search strategy in Appendix A.

No economic studies were identified which were applicable to this review question (see economic study selection flow chart in Appendix B)

Excluded studies

No economic studies were reviewed at full text and excluded from this review.

Economic model

No original health economic modelling was conducted for review question 2.3 on psychosexual function as this review was concerned with the prevalence and severity of psychosexual function and therefore not appropriate for health economic analysis.

2.3.3 NICE recommendations

The relevant recommendation for this section is Rec 1.1.1.

2.3.4 The committee's discussion and interpretation of the

Clinical

The committee agreed with the IG CRs and adapted them into the standard NICE recommendation style and combined both into one recommendation. The recommendation was changed to involve discussion, rather than just considering, the impact of various factors and psychosexual function.

Health economic

No health economic evidence was identified in the IG or in NICE's health economic literature review for review question 2.3 on psychosexual function.

The committee noted that the two contextualised recommendations were of an informative nature and therefore no resource implications are associated with these recommendations.

2.4 Body image

Review question 2.4: In women with PCOS, what is the prevalence and severity of body image distress?

2.4.1 Recommendations from the International evidence-based guideline for PCOS*

Evidence-based recommendation:

2.4.1 Healthcare professionals should be aware that features of PCOS can have a negative impact on body image.

*Copyright Monash University Melbourne Australia 2023. Reproduced with permission from Monash University.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(criteria well described, studies appear to meet this).*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(detailed PICO available which is suitable for the review, with only validated questionnaires from studies to assess body image distress).*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(eligibility criteria clearly described in PICO).*
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(There are no date restrictions as this is a new review and they state in PICO no limits on publication date).*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information	Yes *(study types selected appear appropriate for review question).*

Section	Question	Answer
	appropriate (e.g. publication status or format, language, availability of data)?	
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(PICO is detailed but could have more details about date limit).*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Yes *(appropriate number (7) and range of databases used. No searches for unpublished studies/registry records but arguably less important for prevalence-type questions. Range of databases for published material is fine).*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Yes *(a limited number of additional references (n=3) were identified through reference list searching. Protocol has option of expert contact if no other evidence found but it does not look like this was necessary).*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(limited free-text terms on body image. MeSH terms not translated for Embase/PsycInfo so key index terms like body dissatisfaction/ (Embase) would have been missed. It is not clear why both Medline and PubMed were included in the search, with very different numbers of results).*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(similar limits to NICE standard exclusions (e.g. English language, letters etc). Presumably applied during the sift rather than the search as the limits do not appear in the strategies)*.
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(2 reviewers in consultation with the evidence team for study selection and appraisal).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(there were problems with the search strategies, likely to miss relevant papers*

Section	Question	Answer
		and limited other search methods that might otherwise compensate for that).
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(2 reviewers in consultation with the evidence team for study selection and appraisal).*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(study table with study characteristics and structured study extraction form present allowing interpretation of results).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(9 studies included and in PRISMA, 3 studies identified from a reference list search were excluded)*.
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(risk of bias assessments have been completed in structured extraction form and appear appropriate).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Probably yes *(2 reviewers in consultation with evidence team likely reduced errors in this area).*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(adequate study characteristics presented and risk of bias assessment conducted using appropriate structured criteria).*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(all 9 studies found were included in the results).*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(no information was given regarding the planned analysis of the results except subgroups in PICO were adolescents, ethnicity, phenotype, pregnancy).*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(Meta-analysis was conducted appropriately for this review question, study designs were appropriate for analysis as all were case-control studies).*

Section	Question	Answer
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Probably yes *(there were a couple of outcomes that had high I^2 but too few studies to explore. They used random effects for one where there was high inconsistency but not for another. No serious inconsistency was noted in GRADE table).*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Probably yes *(funnel plots available however very few results for each outcome, too few for funnel plots).*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Probably yes *(All studies were moderate mainly due to non-comparable populations of cases and controls and this is downgraded in GRADE. GRADE ratings all started at moderate because observational so low for all outcomes).*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(all studies included, synthesis was appropriate although there was inconsistency).*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(PICO is detailed but could have more details about date limit).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(there were problems with the search strategies, likely to miss relevant papers and limited other search methods that might otherwise compensate for that).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(adequate study characteristics presented and risk of bias assessment conducted using appropriate structured criteria).*
Judging risk of bias	Concerns regarding the synthesis and findings	Low

Section	Question	Answer
		(all studies included, synthesis was appropriate although there was inconsistency).
Overall review ratings	Overall risk of bias	Low *(no major concerns across the review except for concerns regarding search strategies).*
Overall review ratings	Applicability as a source of data	Fully applicable

Evidence to recommendations justification: The multidimensional body-self relations questionnaire (MBSRQ-AS) (a self-reported assessment of body image) appearance evaluation subdomain score in women with and without PCOS showed the most difference between groups with an MD -0.78 (95% CI -0.90 to -0.65) showing higher dissatisfaction in the PCOS group, with no heterogeneity from three studies. MBSRQ-AS overweight preoccupation showed a lower dissatisfaction score in the control group for two studies MD 0.60 (95% CI 0.46 to 0.74) with similar results however this is reported as worse in the PCOS group. Otherwise, the outcomes were all worse for PCOS except for body image concerns – the body esteem scale for adolescents and adults (BESAA) weight, BESAA attribution and body image scale. All studies were case-control studies with serious risk of bias with one to 3 studies per outcome. Therefore, this is not a lot of strong evidence on which to base an EBR. The EBR stated “health professionals and women should be aware that features of PCOS can have a negative impact on body image” [Strong recommendation for the option], which is not really prescriptive so is appropriate. They also state that a research priority is required to determine clinically meaningful differences in body image scores in PCOS and more research of body image in the adolescent population.

IG economic evidence

No health economic evidence was identified in the IG for review question 2.4 on body image.

The IG noted that screening may have resource implications in terms of impact on length of consultation and noted that intervention may require referral to other health practitioners. It was acknowledged that implementation and access to resources may vary depending on local circumstances and therefore a change in clinical practice may be required.

2.4.2 NICE economic evidence

Included studies

A single health economic search was performed by NICE to identify published economic evaluations of relevance to all review questions in this guideline. See the literature search strategy in Appendix A.

No economic studies were identified which were applicable to this review question (see economic study selection flow chart in Appendix B)

Excluded studies

No economic studies were reviewed at full text and excluded from this review.

Economic model

No original health economic modelling was conducted for review question 2.4 on body image as this review was concerned with the prevalence and severity of body image distress and therefore not appropriate for health economic analysis.

2.4.3 NICE recommendations

The relevant recommendation for this section is Rec 1.1.1.

2.4.4 The committee's discussion and interpretation of the evidence

Clinical

The IG made one EBR for review question 2.4 on body image. The committee agreed with the IG recommendation that features of PCOS can have a negative effect on body image and adapted it into the standard NICE recommendation style. The NICE recommendation was made more active regarding discussion of the impact of PCOS on body image, rather than just being aware of it.

Health economic

No health economic evidence was identified in the IG or in NICE's health economic literature review for review question 2.4 on body image.

The guideline committee contextualised the one-evidence based recommendation made in the IG for review question 2.4. The committee noted that this recommendation is of an informative nature and therefore no resource implications are associated with recommendation.

2.5 Eating disorders and disordered eating

Review question 2.5: In women with PCOS what is the prevalence and severity of disordered eating?

2.5.1 Recommendations from the International evidence-based guideline for PCOS*

Evidence based recommendations:

2.5.1 Eating disorders and disordered eating should be considered in PCOS, regardless of weight, especially in the context of weight management and lifestyle interventions (see sections 2.4 and 3.6)

Practice point:

2.5.2 If disordered eating or eating disorders are suspected, appropriately qualified practitioners should further assess via a full diagnostic interview. If an eating disorder or disordered eating is detected, appropriate management and support should be offered.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Criteria well described, studies appear to meet this).*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Probably yes *(detailed PICO available which is suitable for the review. However, we do not usually include development studies, only validation, where there is a lot of evidence).*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(eligibility criteria for a risk tool review are relevant and clearly described in PICO)*.
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Probably yes *(Updated review but no date limits. No limits on study quality, any original study e.g. case-control, RCT, cross-sectional, whereas we would prioritise higher quality studies study types selected appear appropriate for review question. Also, they include development and validation studies, whereas we would prioritise validation studies).*

Section	Question	Answer
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Probably yes *(English language limit is appropriate).*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(PICO is very detailed, however may include studies that we would not unless we did not have a lot of studies).*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Yes *(appropriate number (4) and range of databases used. Medline; Cochrane; PsycInfo; Embase. No trial registries but not so relevant for prevalence-type questions).*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	No *(no information is given about additional methods such as manual searches. References from prior review brought forward for this update but other than that no additional methods listed).*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Probably no *(MeSH terms not properly translated for Embase (e.g. exp "Feeding and Eating Disorders"/ should be exp "Eating Disorders"/) but appear to map onto related Emtree terms OK. Ovid default treats multi-word lines as directional, so Medline line 2 (eating disorder*.tw.) would pick up mention of eating disorders but not disordered eating, the phrase used in the question. MeSH should compensate for this to some degree but overall impact is hard to gauge).*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(not explicitly stated but appears to have been appropriate. No restrictions beyond English language*).

Section	Question	Answer
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(2 reviewers for study selection and appraisal).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	Unclear *(There are errors in the search strategies. Unlikely to be hugely impactful but cannot say definitively).*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(2 reviewers for study selection and appraisal).*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(study table with study characteristics present allowing interpretation of results).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(14 studies included in PRISMA and review, meta-analysis and GRADE tables).*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(risk of bias assessments have been completed in a structured extraction form and appear appropriate, GRADE conducted)*.
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Probably yes *(2 reviewers likely reduced errors in this area)*.
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(2 reviewers, study characteristics present, risk of bias conducted in appropriate form and GRADE conducted. All studies accounted for).*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(all 14 studies found were included in the results).*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	Unclear *(no information was given regarding the planned analysis of the results*

Section	Question	Answer
		except for adolescent, ethnicity and phenotype was included as subgroups in PICO. Not possible to subgroup as it would leave one study for adolescents. Rotterdam criteria was included as a sensitivity analysis but not set out in protocol).
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(Meta-analysis was used appropriately for this review question, study designs were appropriate for analysis).*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Unclear *(not too much heterogeneity except for disordered eating, results are presented for sensitivity analyses of studies that used Rotterdam criteria for diagnosis which did not have heterogeneity, however all outcomes are rated as no serious inconsistency, whereas disordered eating studies should have been downgraded. Random effects used for all meta-analyses).*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Probably yes *(funnel plots available for the various outcomes).*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Probably yes *(no obvious discussion about this but appears reasonable and was included in GRADE tables).*
Synthesis and findings	Concerns regarding the synthesis and findings	*Unclear* *(Grading for inconsistency in GRADE for disordered eating was unclear. Sensitivity analysis included but was not a pre-defined analysis).*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(the PICO is very detailed, however may include studies that we would*

Section	Question	Answer
		not unless we did not have a lot of studies).
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	Unclear *(There are errors in the search strategies. Unlikely to be hugely impactful but cannot say definitively).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(2 reviewers, study characteristics present, risk of bias conducted in appropriate form and GRADE conducted. All studies accounted for).*
Judging risk of bias	Concerns regarding the synthesis and findings	*Unclear* *(The grading for inconsistency in GRADE for disordered eating was unclear. Sensitivity analysis included but was not a pre-defined analysis).*
Overall review ratings	Overall risk of bias	Unclear *(Unclear areas where errors in the search and grading for inconsistency in GRADE and pre-defined analysis. Otherwise, concerns were low).*
Overall review ratings	Applicability as a source of data	Fully applicable

Evidence to recommendations justification: The IG states that there was an increased odds of any eating disorder in women with PCOS (OR 1.53 (95% CI 1.29 to 1.82) in 10 studies and in the sub-analysis of only studies where PCOS diagnosis was confirmed by Rotterdam criteria (OR 2.88 (95% CI 1.55 to 5.34) in 6 studies. There was an increased odds of binge eating disorder in women with PCOS (2.09 (95% CI 0.18 to 3.72) in 5 studies and sub-analysis of Rotterdam criteria OR 2.70 (95% CI 1.47 to 4.97) in 4 studies. Where prevalence of disordered eating is defined by a score above the cut-off of a validated eating disorder questionnaire there was no increased odds for adult women or for the sub-analysis confirmed by Rotterdam criteria (OR 1.77

(95% CI 0.63 to 4.91) in 6 studies and OR 2.21 (95% CI 0.87 to 5.59) in 4 studies; no increased odds of bulimia nervosa in women with PCOS (OR 1.34 (95% CI 1.17 to 1.54) in 6 studies and in the sub-analysis confirmed by Rotterdam criteria OR 1.56 (95% CI 0.56 to 4.36) in 4 studies. There were no increased odds of anorexia nervosa in women with PCOS OR 0.94 (95% CI 0.69 to 1.28) in 5 studies. Subgroup analysis was not conducted because only one study used Rotterdam criteria. The evidence ranged from moderate to very low in GRADE. All results favoured PCOS but were only significant for increased odds of eating disorder or binge eating disorder. They state that eating disorders can only be diagnosed via clinical intervention and only one small study (n=49) used this and found prevalence of any disorder was higher compared to controls. The EBR states that eating disorders and disordered eating should be considered in PCOS, regardless of BMI and especially in the context of weight management and lifestyle interventions, which is a conditional (weak recommendation for the option) in the IG. This would be a weak recommendation in NICE terms also. The research priorities in the IG include prevalence of eating disorder/disordered eating in women with PCOS, including different subgroups, using a structured clinical interview and considering all types of eating disorders.

IG economic evidence

No health economic evidence was identified in the IG for review question 2.5 on eating disorders and disordered eating.

The IG noted that the costs of clinical interviews for diagnosis would be considerable, given the time taken to conduct interviews for people with suspected disordered eating and the clinical expertise required of the interviewer. In addition, it was noted that if disordered eating interventions are required, the cost of implementing these interventions would be considerable. The IG did however note that there is evidence that eating disorders/disordered eating treatment is cost-effective in the general

population, although no further information was provided to underpin this statement. The IG also acknowledged that treatment costs can be reduced via use of a stepped care approach.

It was noted that some healthcare settings may not have the resources or capability to screen, assess, diagnose and treat disordered eating.

2.5.2 NICE economic evidence

Included studies

A single health economic search was performed by NICE to identify published economic evaluations of relevance to all review questions in this guideline. See the literature search strategy in Appendix A.

No economic studies were identified which were applicable to this review question (see economic study selection flow chart in Appendix B).

Excluded studies

No economic studies were reviewed at full text and excluded from this review.

Economic model

No original health economic modelling was conducted for review question 2.5 on eating disorders and disordered eating as this review was concerned with the prevalence of disordered eating and therefore not appropriate for health economic analysis. In addition, the committee were aware of NICE's existing guideline on eating disorders (NG69) and concluded that treatment and identification of eating disorders would not significantly differ for people with PCOS compared to the general population. The committee therefore concluded that the recommendations made in NG69 would apply to this PCOS guideline whereby the cost-effectiveness of the recommendations produced in NG69 has already been assessed.

2.5.3 NICE recommendations

The relevant recommendation for this section is Rec 1.15.1.

2.5.4 The committee's discussion and interpretation of the evidence

Clinical

The IG made one EBR (2.5.1) and one practice point recommendation (2.5.2) for this review question. The committee agreed with the IG EBR, but thought that the existing NICE guideline on eating disorders covered this area adequately. It provides information on identifying, assessing and managing eating disorders. The committee, therefore, cross-referred to NICE's guideline on eating disorders. The committee also made a separate recommendation in agreement with part of the IG's practice point recommendation, not to exclude the possibility of an eating disorder or disordered eating in a person with PCOS solely based on their weight. The committee thought this was important to highlight as a specific PCOS-related recommendation.

Health economic

No health economic evidence was identified in the IG or in NICE's health economic literature review for review question 2.5 on eating disorders and disordered eating.

The guideline committee cross-referred to NICE's existing guideline on eating disorders (NG69) noting that this provides information on identification, assessment and treatment for eating disorders.

The committee also contextualised a section of the one practice point recommendation the IG made which detailed that possibility of an eating disorder should not be excluded based on weight alone. The committee noted

that because this recommendation is of an informative nature, no resource implications are associated with this recommendation.

2.6 Information resources, models of care, cultural and linguistic considerations

Review question 2.6: the IG was split into four reviews:

- 2.6.1 Information needs
- 2.6.2 Models of care
- 2.6.3 Support to manage PCOS
- 2.6.4 Patient care

The clinical information pertaining to these reviews is detailed according to each topic, as indicated in the bullet points above. The overarching health economic evidence relating review question 2.6 is detailed in section 2.6.4 on patient care.

2.6 (1) Information needs

Review question 2.6.1: What are the information, resource and education needs of women, adolescents, culturally and linguistically diverse groups and healthcare providers regarding PCOS?

2.6.1 Recommendations from the International evidence-based guideline for PCOS*

Evidence-based recommendations:

2.6.1.1 Tailored information, education and resources that are high-quality, culturally appropriate and inclusive should be provided to all with PCOS.

2.6.1.2 Information, education and resources are a high priority for patients with PCOS and should be provided in a respectful and empathic manner.

Consensus recommendation:

2.6.1.3 Entities responsible for health professional education should ensure that information and education on PCOS is systemically embedded at all levels of health professional training to address knowledge gaps.

Practice points:

2.6.1.4 The diversity of the population should be considered when adapting practice paradigms. Healthcare professional opportunities should be optimised at all stages of graduate and postgraduate training, continuing professional development and in practice support resources.

2.6.1.5 Women should be counselled on the risk of misinformation and guided to evidence-based resources.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist (2.6.1)

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Criteria well described, studies appear to meet this)*.
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(detailed PICO available for qualitative and quantitative review which was suitable for the review).*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(eligibility criteria clearly described in PICO).*
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(it is a new review so no date limits).*

Section	Question	Answer
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Probably yes *(study types selected appear appropriate for review question except they state that they would include development and validation studies whereas we would prioritise validation studies).*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(appropriate limits on studies although development studies are included which we would exclude if we had enough validation studies).*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably no *(appropriate number (5) and range of databases used. Sources are all bibliographic databases. Reasonable selection within that. No attempt to identify unpublished reports. Abstracts and protocols specifically excluded).*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	No *(no information is given about additional methods such as manual searches Nothing bar database searches reported in PRISMA chart).*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Probably no *(full search strategy is detailed in the review, the search terms and dates appear appropriate for the review question note the reviewers did the search and appraisal for 2.6.1 and 2.6.4 together. Structure of the search is: PCOS AND (health workers or patients) AND info* *OR* *PCOS AND health workers AND perspectives* *Including terms for workers and patients may exclude papers - the scope of the question is basically everyone involved so the population would normally just be PCOS rather than, for example, PCOS*

Section	Question	Answer
		and "patients". More significantly - no translation of MeSH terms for Embase. Some terms also are not translated for PsycInfo e.g. patient preference/ is a null term (preferences/ and/or client attitudes are the alternative).
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably no *(It is not explicitly stated but appears to have been appropriate. Exclusion (in sifting criteria rather than the search) of comparative intervention studies is questionable.* *English language limit is applied in line with NICE practice. The date limit of 1990+ is applied but not justified. It would be better to have some rationale)*.
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(2 reviewers for study selection and appraisal).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(There are an accumulation of small concerns with this question - no aspect of the search is entirely satisfactory. Unusual basic structure; issues with translation; unjustified exclusion criteria etc).*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(2 reviewers for study selection and appraisal).*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(study table with study characteristics and structured individual study extraction form present allowing interpretation of results).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(41 studies included in review that were in PRISMA).*
Data collection and	Was risk of bias (or methodological quality)	Yes *(study quality appraisal available detailing bias assessments etc however*

Section	Question	Answer
study appraisal	formally assessed using appropriate criteria?	*questions seem to relate to case-control studies when most of the studies are cross-sectional and given nature of the question cross-sectional is more appropriate. There is a quality appraisal for all qualitative studies. There were no GRADE tables because of the broad nature of the clinical question and the inclusion of studies that of heterogeneous designs, populations or aims, therefore evidence presented narratively in part 2).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Probably yes *(2 reviewers likely reduced errors in this area).*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(2 reviewers, all studies are accounted for, risk of bias assessed and structured extraction form used however may be more appropriate for case-control studies).*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(all 41 studies found were included in the results, 25 studies for studies related to women with PCOS and 16 for studies related to healthcare professionals).*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(no information was given regarding the planned analysis of the results and results were presented narratively).*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Not applicable *(no meta-analysis was performed therefore a proper synthesis of evidence was not done, a narrative was written in part 2).*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Not applicable *(no meta-analysis performed).*
Synthesis and findings	Were the findings robust, e.g. as demonstrated	Not applicable *(no meta-analysis performed).*

Section	Question	Answer
	through funnel plot or sensitivity analyses?	
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Not applicable *(no meta-analysis performed).*
Synthesis and findings	Concerns regarding the synthesis and findings	High *(results presented narratively).*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(appropriate limits on studies although development studies are included which we would exclude if we had enough validation studies).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Accumulation of small concerns with this question - no aspect of the search is entirely satisfactory. Unusual basic structure; issues with translation; unjustified exclusion criteria etc).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(2 reviewers, all studies are accounted for, risk of bias assessed and structured extraction form used however may be more appropriate for case-control studies).*
Judging risk of bias	Concerns regarding the synthesis and findings	High *(results presented narratively).*
Overall review ratings	Overall risk of bias	N/A *(results were not synthesised so cannot comment on overall risk of bias).*
Overall review ratings	Applicability as a source of data	Partially applicable

Evidence to recommendations justification: The IG states that due to the broad nature of the clinical question and the inclusion of studies of

heterogeneous designs, populations and/or aims, it was not feasible to generate GRADE evidence profile tables. There is not an evidence synthesis for a qualitative/mixed methods review except findings are presented in Part 2 of the question. However, they did make EBRs regarding provision of tailored information, education and resources for people with PCOS.

2.6 (2) Models of care

Review question 2.6.2: What are the characteristics of available models of care implemented in PCOS clinic or service?

2.6.1 Recommendations from the International evidence-based guideline for PCOS*

Consensus recommendation:

2.6.2.1 Models of care should prioritise equitable access to evidence-based primary care with pathways for escalation to integrated specialist and multidisciplinary services as required.

Practice point:

2.6.2.2 Strategies to deliver optimal models of care could include health professional education, care pathways, virtual care, broader health professional engagement (e.g. nurse practitioners) and coordination tools.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist (2.6.2)

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Criteria well described, studies appear to meet this).*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(The question asked what are the characteristics of available models of care? and includes any primary study or guidelines or systematic review with outcomes from comparison of a model of care compared to none, usual care or others, and excludes specific interventions, so is appropriate).*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(eligibility criteria clearly described in PICO).*
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(No restrictions on date as it is a new review but is not stated in PICO).*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(English language limit is appropriate, study types selected are appropriate for review question).*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(Limits are appropriate).*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Yes *(appropriate number (5) and range of databases used. Question is after descriptive reports. Usual ROBIS requirement for unpublished material is less relevant in this context. Existing list of sources (Medline, Embase, Cochrane, PsycInfo, CINAHL + expert input) seems entirely appropriate).*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably yes *(no information is given about additional methods such as manual searches. All included studies reported*

Section	Question	Answer
		to be identified "by key contacts" but actual methods not described).
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Probably no *(full search strategy is detailed in the review. Basic search is appropriately structured but no information on some translations (PsycInfo) and no translation of MeSH terms for Embase apparent. Some CINAHL terms are missing spacing (e.g. MH "Patient CarePlans+" rather than MH "Patient Care Plans+") - may be a transcription error or may be an actual error that would miss relevant material)*.
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(not explicitly stated but appears to have been appropriate. English language limit used - fine by NICE methods manual. Publication type limits are pragmatic and appropriate).*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(two reviewers for title and abstract with evidence team used when needed for study appraisal).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	Unclear *(Unclear how included studies were sourced - none came from the search. Search strategy translations are not 100% accurate).*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(two reviewers should have minimised error when appraising studies, along with input from evidence team).*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(study table with study characteristics present allowing interpretation of results).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(5 studies included in review, 0 reports from key contacts included - all excluded for incorrect study design, 2*

Section	Question	Answer
		studies covered service evaluation, 4 studies covered patient evaluation).
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Probably yes *(all 5 studies have detailed study characteristics and quality appraisals completed including commentary on bias in structured extraction form. Risk of bias was conducted on individual studies but there was not a comparison group for any study but scored low or moderate risk of bias).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Yes *(this should have been minimised by having 2 reviewers plus the evidence team).*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(2 reviewers, adequate structured extraction form, all studies identified are included).*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(descriptive analysis of all 5 studies was included)*.
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(no information was given regarding the planned analysis of the results).*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Not applicable *(no meta-analysis, only descriptive analysis).*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Not applicable *(no meta-analysis, only descriptive analysis).*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Not applicable *(no meta-analysis, only descriptive analysis).*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Not applicable *(no meta-analysis, only descriptive analysis and no GRADE conducted).*

Section	Question	Answer
Synthesis and findings	Concerns regarding the synthesis and findings	*High* *(No meta-analysis, only descriptive analysis).*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(the limits are appropriate).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	Unclear *(Unclear how included studies were sourced - none came from the search. The search strategy translations are not 100% accurate).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(2 reviewers, adequate structured extraction form, all studies identified are included).*
Judging risk of bias	Concerns regarding the synthesis and findings	*High* *(no meta-analysis, only descriptive analysis)*
Overall review ratings	Overall risk of bias	Moderate *(no meta-analysis or GRADE conducted due to heterogeneity of designs, populations and/or aims).*
Overall review ratings	Applicability as a source of data	Partially applicable *(Models are from different countries so may have limited relevance to the UK).*

Evidence to recommendations justification: The evidence was unable to be meta-analysed as the models of care were from different countries, had different aims or different populations. Therefore, given the lack of evidence of sufficient quality an EBR was not made. The CR is very strong ‘should prioritise equitable access to evidence-based care with pathways for escalation to integrated specialist and multidisciplinary services as required’ given that it is not evidence-based.

2.6 (3) Support to manage PCOS

Review question 2.6.3: How can we best support women to navigate the impact of PCOS on family and interpersonal relationships?

2.6.1 Recommendations from the International evidence-based guideline for PCOS*

Consensus recommendation:

2.6.3.1 Public health actors should consider increasing societal awareness and education on PCOS to reduce stigma and marginalisation.

Practice point:

2.6.3.2 Culturally appropriate resources and education on PCOS across the life span for families of those with the condition, should be considered.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist (2.6.3)

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	No *(review was a level 4 narrative review, as such no search strategy or screening was completed).*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	No *(review was a level 4 narrative review, as such no search strategy or screening was completed).*
Study eligibility criteria	Were eligibility criteria unambiguous?	No *(review was a level 4 narrative review, as such no search*

Section	Question	Answer
		strategy or screening was completed).
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	No *(review was a level 4 narrative review, as such no search strategy or screening was completed).*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	No *(review was a level 4 narrative review, as such no search strategy or screening was completed).*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	High *(a narrative review was chosen rather than searching for studies).*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	No information *(narrative review- no study selection process).*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	No information *(narrative review- no study selection process).*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No information *(narrative review- no study selection process).*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	No information *(narrative review- no study selection process).*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	No information *(narrative review- no study selection process).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	Unclear *(No original searches carried out for this question. Unclear how cited studies were collected).*

Section	Question	Answer
Data collection and study appraisal	Were efforts made to minimise error in data collection?	No *(narrative review - no risk of bias assessment).*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	No *(narrative review - no risk of bias assessment).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	No *(narrative review - no risk of bias assessment).*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	No *(narrative review - no risk of bias assessment).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	No *(narrative review - no risk of bias assessment).*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	N/A *(narrative review conducted).*
Synthesis and findings	Did the synthesis include all studies that it should?	No information *(narrative review- no studies to assess).*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(narrative review- no studies to assess).*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	No information *(narrative review- no studies to assess).*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	No information *(narrative review- no studies to assess).*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	No information *(narrative review- no studies to assess).*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	No information *(narrative review- no studies to assess).*
Synthesis and findings	Concerns regarding the synthesis and findings	N/A

Section	Question	Answer
		(narrative review conducted).
Judging risk of bias	Concerns regarding specification of study eligibility	High *(narrative review conducted).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	N/A *(narrative review conducted).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	N/A *(narrative review conducted).*
Judging risk of bias	Concerns regarding the synthesis and findings	N/A *(narrative review conducted).*
Overall review ratings	Overall risk of bias	High *(narrative review leading to consensus recommendation, included a "grade direction and strength" assessment of the recommendation which was assigned weak).*
Overall review ratings	Applicability as a source of data	Fully applicable

Evidence to recommendations justification: No evidence was included in the review, and a narrative review was decided upon prior to starting the review, hence there is no EBRs.

2.6 (4) Patient care

Review question 2.6.4: What are the key challenges for those with PCOS when interacting with healthcare professionals about polycystic ovary syndrome and related features?

2.6.1 Recommendations from the International evidence-based guideline for PCOS*

Evidence-based recommendations:

2.6.4.1 Healthcare professionals should employ shared decision making and support patient agency or ability to take independent actions to manage their health and care.

2.6.4.2 The importance of being knowledgeable about PCOS, of applying evidence-based practices when sharing news on diagnosis, treatment and health implications, and of ascertaining and focusing on patient priorities, should be recognised.

Consensus recommendation:

2.6.4.3 Healthcare system leaders should enable system wide changes to support health professional training, knowledge and practice in sharing news optimally, shared decision making and patient agency, including ensuring adequate consultation time and accessible resources.

Practice points:

2.6.4.4 Evidence-based strategies for shared decision making and for sharing news (such as the SPIKES framework) are readily available and should be used to inform PCOS care.

2.6.4.5 All healthcare professionals partnering with women with PCOS should be knowledgeable in sharing news, in shared decision making and in supporting patient self-management.

2.6.4.6 Evidence-based strategies and resources can be used to support patient activation, which refers to modifiable knowledge, skills, ability, confidence and willingness to self-manage one’s own health and care.

1 **IG clinical evidence**

2 **Critical appraisal - ROBIS systematic review checklist (2.6.4)**

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Criteria well described, studies appear to meet this).*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(detailed PICO available which is suitable for the review note- this review was carried out with 2.6.1. Qualitative and quantitative studies of satisfaction surveys looking at people with PCOS' interactions with HCPs in healthcare settings matches the question).*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(eligibility criteria clearly described in PICO).*
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(New systematic review so no date restrictions is present but not well described in PICO).*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(study types selected are appropriate for review question – qualitative and quantitative surveys).*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(PICO is adequately detailed).*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(appropriate number (5) and range of databases used. Search strategy is the same as for Q2.6.1 but omission of unpublished reports is arguably less of an issue for this question. List of databases is reasonable. No additional sources beyond bibliographic databases).*

Section	Question	Answer
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	No *(no additional search methods reported in PRISMA flow chart).*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Probably no *(full search strategy is detailed in the review. Various small errors e.g. no translation of MeSH terms for Embase. Some terms also aren't translated for PsycInfo e.g. patient preference/ is a null term (preferences/ and/or client attitudes are the alternative).*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(not explicitly stated but appears to have been appropriate. English language limit was applied. The date limit of 1990+ was also applied but not justified. It would be better to have some rationale)*.
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(two reviewers assessed studies at title and abstract level with full text retrieved if a decision could not be made, evidence team was also consulted when needed).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	Unclear *(Various small errors - unclear how these interact and whether they'd induce a genuine bias as opposed to missing studies at random).*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(2 reviewers plus evidence team).*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Probably yes *(study table with study characteristics present allowing interpretation of results, however, would have been useful to have individual a full study extraction).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Probably yes *(28 studies included for 2.6.4 in PRISMA and included study list, only 27 appear to be included in the study quality appraisal table, however this reference is quoted in*

Section	Question	Answer
		the summary of findings table as contributing to the review missing reference in next column if needed).
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(risk of bias assessment has been completed for each study and is presented in tables).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Yes *(2 reviewers plus evidence team).*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(study collection adequate and quality appraisal conducted and presented).*
Synthesis and findings	Did the synthesis include all studies that it should?	Probably yes *(results from all studies appear to be included in the summary of findings table however only 27/28 studies have a quality appraisal. There only seems to be n=26 in the quality appraisal table and references 7 and 18 not mentioned in the evidence synthesis. Not clear why these two were not included although in write-up it says one was a SR and one was a comparative study of people with and without PCOS).*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(no information was given regarding the planned analysis of the results. Subgroups mentioned in the PICO was adolescents, ethnicity and phenotype but was not analysed).*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Probably yes *(no meta-analysis completed, descriptive analysis used which seems appropriate given study types and outcomes Most studies included were qualitative, or a couple of mixed methods studies and one cross-sectional study. It says that most studies required participants to have received a diagnosis of PCOS, however the PICOS states that if studies*

Section	Question	Answer
		do not have a diagnosis of PCOS they would be excluded).
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Yes *(Quality review but coherence is assessed by GRADE-CerQual).*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	N/A *(Not available as qualitative review)*.
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Probably yes *(GRADE-CerQual used but says due to broad nature of the clinical question and the inclusion of studies of heterogeneous designs, populations and/or aims, it was not feasible to generate GRADE evidence profile tables for this question, but individual studies ROB is assessed).*
Synthesis and findings	Concerns regarding the synthesis and findings	Unclear *(unsure on impact of the risk of bias assessments).*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(PICO is adequately detailed).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	Unclear *(Various small errors - unclear how these interact and whether they'd induce a genuine bias as opposed to missing studies at random).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(study collection adequate and quality appraisal conducted and presented).*
Judging risk of bias	Concerns regarding the synthesis and findings	Unclear *(unsure on impact of the risk of bias assessments).*
Overall review ratings	Overall risk of bias	Low *(low and unclear concerns for some domains such as unsure on impact of the*

Section	Question	Answer
		risk of bias assessments but themes from studies seem appropriate).
Overall review ratings	Applicability as a source of data	Fully applicable

Evidence to recommendations justification: Themes included in 28 studies were found in the qualitative review and included: 1. Interactions were challenging when bad news was shared in a way that did not safeguard wellbeing (high confidence); 2. Interactions were challenging when they did not provide opportunities to facilitate shared decision-making about outcome that matter to people (high confidence); 3. Interactions were challenging when healthcare professionals did not support patient agency (their ability to take independent actions to manage their health and care (moderate confidence). GRADE was not conducted due to the broad nature of the clinical question and inclusion of heterogenous designs, populations and/or aims. The EBRs were justified by the findings of the IG evidence review, however these areas are covered by NICE guidelines on patient experience and shared-decision making.

IG economic evidence

No health economic evidence was identified in the IG for review question 2.6 on information resources, models of care, relationships and interactions.

2.6.2 NICE economic evidence

Included studies

A single health economic search was performed by NICE to identify published economic evaluations of relevance to all review questions in this guideline. See the literature search strategy in Appendix A.

No economic studies were identified which were applicable to this review question (see economic study selection flow chart in Appendix B).

Excluded studies

No economic studies were reviewed at full text and excluded from this review.

Economic model

No original health economic modelling was conducted for review question 2.6 as the committee concluded that the recommendations made in the section of the IG were covered by existing NICE guidance on patient experience and shared decision-making.

2.6.3 NICE recommendations

No recommendations were made. Cross-reference was made to the patient experience and shared decision-making NICE guidelines.

2.6.4 The committee's discussion and interpretation of the evidence

Clinical

The committee agreed that these issues would be covered comprehensively by the patient experience and shared decision-making NICE guidelines and so cross-referred to these.

Health economic

No health economic evidence was identified in the IG or in NICE's health economic literature search for review question 2.6 on information needs.

No recommendations were contextualised from the IG for this section of the guideline as the committee cross-referred to NICE's existing guidelines on

patient experience and shared decision-making. No health economic or resource implications are therefore associated with cross-referral.

2.7 Psychological therapy

Review question 2.7: Is psychological therapy effective for management and support of depression and/or anxiety, disordered eating, body image distress, self-esteem, feminine identity or psychosexual dysfunction in women with PCOS?

2.7.1 Recommendations from the International evidence-based guideline for PCOS*

Consensus recommendations:

2.7.1 Women with PCOS diagnosed with depression, anxiety, and/or eating disorders should be offered psychological therapy guided by regional general population guidelines and the preference of the woman with PCOS.

2.7.2 Women with PCOS with disordered eating, body image distress, low self-esteem, problems with feminine identity, or psychosexual dysfunction should be offered evidence-based treatments (e.g. cognitive behaviour therapy) where appropriate.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	No *(review was a level 4 narrative*

Section	Question	Answer
		review, as such no search strategy or screening was completed).
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	No *(review was a level 4 narrative review, as such no search strategy or screening was completed).*
Study eligibility criteria	Were eligibility criteria unambiguous?	No *(review was a level 4 narrative review, as such no search strategy or screening was completed).*
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	No *(review was a level 4 narrative review, as such no search strategy or screening was completed).*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	No *(review was a level 4 narrative review, as such no search strategy or screening was completed).*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	High *(narrative review conducted).*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	No information *(No original searches carried out: "This question was allocated as a narrative review. Hence, no search or screening was undertaken, and recommendations will be consensus based").*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	No information *(No original searches carried out: "This question was allocated as a narrative review. Hence, no search or screening was undertaken, and recommendations will be consensus based").*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No information *(No original searches carried out: "This question was allocated as a narrative review. Hence, no search or screening was undertaken, and*

Section	Question	Answer
		recommendations will be consensus based").
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	No information *(No original searches carried out: "This question was allocated as a narrative review. Hence, no search or screening was undertaken, and recommendations will be consensus based").*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	No information *(No original searches carried out: "This question was allocated as a narrative review. Hence, no search or screening was undertaken, and recommendations will be consensus based").*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	Unclear *(Narrative review conducted. No original searches carried out for this question. Unclear how cited studies were sourced).*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	No *(narrative review - no risk of bias assessment).*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	No *(narrative review - no risk of bias assessment).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	No *(narrative review - no risk of bias assessment).*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	No *(narrative review - no risk of bias assessment).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	No *(narrative review - no risk of bias assessment).*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	N/A *(narrative review conducted).*

Section	Question	Answer
Synthesis and findings	Did the synthesis include all studies that it should?	No information *(narrative review- no studies to assess).*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(narrative review- no studies to assess).*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	No information *(narrative review- no studies to assess).*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	No information *(narrative review- no studies to assess).*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	No information *(narrative review- no studies to assess).*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	No information *(narrative review- no studies to assess).*
Synthesis and findings	Concerns regarding the synthesis and findings	N/A *(narrative review conducted).*
Judging risk of bias	Concerns regarding specification of study eligibility	High *(narrative review conducted).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	Unclear *(narrative review conducted).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	N/A *(narrative review conducted).*
Judging risk of bias	Concerns regarding the synthesis and findings	N/A *(narrative review conducted).*
Overall review ratings	Overall risk of bias	High *(narrative review leading to 2 consensus recommendations, included a "grade direction and strength" assessment on of the recommendations which was*

Section	Question	Answer
		assigned strong recommendation for the option).
Overall review ratings	Applicability as a source of data	Fully applicable

Evidence to recommendations justification: No evidence was searched for as a narrative review was decided on prior to searching for any evidence. The CR is very strong considering there is no supporting evidence.

IG economic evidence

No health economic evidence was identified in the IG for review question 2.7 on psychological therapy.

The IG noted that evidence-based psychological treatments tend to be well accepted and that the availability of psychological therapy has been improved with telehealth and online interventions. It was also noted that the costs of providing psychological interventions are reduced using stepped care interventions, and that online programs can be effective for some, particularly when used with a guide or coach. The IG also stated that the implementation of evidence-based programs, within a stepped care framework, improves cost-effectiveness, however no further evidence was provided to underpin this statement.

The IG stated that longer treatments tend to be more effective and noted that the time demands for healthcare professionals associated with providing psychological treatments can be reduced by the use of stepped-care approaches and telehealth delivered interventions. The IG also noted that evidence suggests that psychological services are poorly integrated into models of care for PCOS.

2.7.2 NICE economic evidence

Included studies

A single health economic search was performed by NICE to identify published economic evaluations of relevance to all review questions in this guideline. See the literature search strategy in Appendix A.

No economic studies were identified which were applicable to this review question (see economic study selection flow chart in Appendix B)

Excluded studies

No economic studies were reviewed at full text and excluded from this review.

Economic model

No original health economic modelling was conducted for review question 2.7 on psychological therapy as the committee agreed that psychological therapy would be covered in NICE's existing guidelines on anxiety and depression.

2.7.3 NICE recommendations

The relevant recommendations for this section are Rec 1.14.2 and 1.14.3.

2.7.4 The committee's discussion and interpretation of the evidence

Clinical

No evidence was provided in the IG for psychological therapy for those diagnosed with depression, anxiety and/or eating disorders so a recommendation specifically for PCOS was not possible. Furthermore, no evidence was reported for evidence-based treatments for those with PCOS and disordered eating, body image distress, low self-esteem, problems with feminine identity, or psychosexual dysfunction. The committee acknowledged

the lack of evidence in the IG and felt the NICE guidelines could be cross-referred to for the management of depression and anxiety.

The committee, therefore, cross-referred to NICE's existing guidelines on depression (NG222 and NG134) and anxiety (CG113) for this section of the guideline. The committee noted that these guidelines provide information on how depression and anxiety should be identified, assessed and managed.

Health economic

No health economic evidence was identified in the IG or in NICE's health economic literature review for review question 2.7 on psychological therapy.

The committee acknowledged the lack of clinical and cost-effectiveness evidence on psychological therapy in the IG. The committee also noted that the criteria for accessing psychological therapy would likely be the same for those with PCOS compared to the general population. As the committee cross-referred to NICE's existing guidelines, no health economic, or resource implications, are associated with cross-referral.

2.8 Antidepressant and anxiolytic treatment

Review question 2.8: Are anti-depressants and anxiolytics effective for management and support of depression and/or anxiety or disordered eating in women with PCOS?

2.8.1 Recommendations from the International evidence-based guideline for PCOS*

Consensus recommendation:

2.8.1 Psychological therapy could be considered first-line management, and antidepressant medications considered in adults where mental health

disorders are clearly documented and persistent, or if suicidal symptoms are present, based on general population guidelines.

Practice point:

2.8.2 Lifestyle intervention and other therapies (e.g. COCP, metformin, laser hair removal) that target PCOS features should be considered, given their potential to improve psychological symptoms.

Where pharmacological treatment for anxiety and depression is offered in PCOS, healthcare professionals should apply caution:

- to avoid inappropriate treatment with antidepressants or anxiolytics
- to limit use of agents that exacerbate PCOS symptoms, including weight gain.

Healthcare professionals should be aware that not managing anxiety and depression may impact adherence to PCOS treatment/management.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Criteria well described, studies appear to meet this).*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(detailed PICO available which is suitable for the review).*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(eligibility criteria clearly described in PICO).*

Section	Question	Answer
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(this is an updated review, but no date limits provided in PICO).*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(Limit of English language is appropriate, study types selected, and outcomes are appropriate for review question).*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(limits are appropriate and criteria mostly described in detail).*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	No *(appropriate number (5) and range of databases used. No sources used beyond bibliographic databases so no chance of retrieving unpublished RCTs).*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	No *(no information is given about additional methods such as manual searches. PRISMA chart implies includes from prior review would have been brought forward had it not been an empty review. Beyond that, no additional search methods used).*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(full search strategy is detailed in the review. RCT filter is the McMaster sensitive one: MEDLINE | Health Information Research Unit should be 99% sensitivity but the term random.mp (Medline line 29, p1953) is not truncated, and so the filter misses over a quarter of references tagged with randomized controlled trial.pt, which is 93% of the RCTs in Medline. Also, similar error in PsycInfo, no proper Embase translation, line combinations missing spacing in CINAHL (S24, p1953) etc).*

Section	Question	Answer
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(not explicitly stated but appears to have been appropriate. English language used - not recommended under ROBIS but pragmatic and standard NICE practice. Study types appropriate to the question).*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(two reviewers assessed studies at title and abstract level with full text retrieved if a decision could not be made).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Search filters and other elements of the search have been run in a way that would demonstrably miss a significant proportion of relevant RCTs).*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(2 reviewers plus evidence team).*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(study table with study characteristics and full study extraction present allowing interpretation of results).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(1 study met inclusion criteria for review).*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(risk of bias included in study characteristics table and structured individual study extraction form).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Yes *(2 reviewers plus evidence team).*
Data collection and	Concerns regarding methods used to collect data and appraise studies	Low

Section	Question	Answer
study appraisal		*(no concerns as studies seem appropriately included and appraised, however only one study found).*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(only one included study so clearly has been included).*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(no information was given regarding the planned analysis of the results).*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	N/A *(not applicable as only one study).*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	N/A *(not applicable as only one study).*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	N/A *(not applicable as only one study).*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Unclear *(Risk of bias was conducted for the individual study, however it was rated moderate risk of bias and says it was downgraded for high risk of bias due to lack or randomisation or blinding but in the structured answers it says it was randomised and had blinding).*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(only one study so meta-analysis not possible and risk of bias seems contradictory).*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(limits are appropriate and criteria mostly described in detail).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Search filters and other elements of the search have been run in a way*

Section	Question	Answer
		that would demonstrably miss a significant proportion of relevant RCTs).
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(no concerns as studies seem appropriately included and appraised, however only one study found).*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(there was only one study so meta-analysis was not possible, and risk of bias seems contradictory).*
Overall review ratings	Overall risk of bias	High *(Search filters and other elements of the search have been run in a way that would demonstrably miss a significant proportion of relevant RCTs).*
Overall review ratings	Applicability as a source of data	Fully applicable

Evidence to recommendations justification: Only one study was identified, therefore no EBRs were made. The CR is very strong considering it is not based on any evidence.

IG economic evidence

No health economic evidence was identified in the IG for review question 2.8 on antidepressant and anxiolytic treatment.

The IG noted that healthcare professionals need to be adequately trained in the management of common mental health disorders. It was also noted that in many countries it is not usual practice to screen adolescents or adults with PCOS for depression and/or anxiety symptoms and therefore doing so may identify affected patients who would otherwise be missed. The IG acknowledged that screening may have resource implications such as an

impact on the length of consultations in the form of additional time required to complete an appropriate care plan, however this can be reduced using screening tools. The IG also noted that if depression and/or anxiety symptoms are detected, intervention may require referral to other health practitioners.

The IG stated that a pragmatic approach to adopting their recommendations may be to screen all women and adolescents at the time of PCOS diagnosis and where appropriate, at the time of their regular physical health checks for PCOS. The need to consider an individual's risk factor for anxiety and depression, to inform if additional screening is warranted was also noted.

NICE economic evidence

Included studies

A single health economic search was performed by NICE to identify published economic evaluations of relevance to all review questions in this guideline. See the literature search strategy in Appendix A.

No economic studies were identified which were applicable to this review question (see economic study selection flow chart in Appendix B).

Excluded studies

No economic studies were reviewed at full text and excluded from this review.

Economic model

No original health economic modelling was conducted for review question 2.8 on antidepressant and anxiolytic treatment as the committee agreed that prescription of these would be covered in NICE's existing guidelines on anxiety and depression.

2.8.2 NICE recommendations

No recommendations were made.

2.8.3 The committee's discussion and interpretation of the evidence

Clinical

No EBRs were included in the IG, but a recommendation was made regarding psychological therapy as first-line management, and antidepressant medications considered in adults where mental health disorders are clearly documented and persistent, or if suicidal symptoms are present. The committee acknowledged the lack of evidence for people with PCOS in the IG and as such felt no new recommendations could be made here. .

The committee, therefore, cross-referred to NICE's existing guidelines on depression (NG222 and NG134) and anxiety (CG113) for this section of the guideline. The committee noted that these guidelines provide information on how depression and anxiety should be managed, including information on when psychological therapy, antidepressants and anxiolytic should be prescribed.

Health economic

No health economic evidence was identified in the IG or in NICE's health economic literature review for review question 2.8 on antidepressant and anxiolytic treatment.

The committee acknowledged that treatment for anxiety and depression would be the same for those with PCOS compared to the general population. The committee also noted that the lack of clinical and cost-effectiveness evidence on antidepressant and anxiolytic treatment in the IG. As the

committee cross-referred to NICE's existing guidelines, no health economic, or resource implications, are associated with cross-referral.

Appendix A Health economic literature review search strategy

The searches for the cost effectiveness evidence were run on 5 December 2024 and re-run on 25 March 2026. The following databases were searched: Medline (Ovid), Embase (Ovid; Econlit (Ovid) and the International HTA Database. Limits were applied to remove study types. The validated NICE cost utility filter was used on MEDLINE and Embase. English language limits were applied, and the search was run for evidence published since 2009.

A NICE Senior Information Specialist (SIS) conducted the searches. The MEDLINE strategy was quality assured by another NICE SIS. All translated search strategies were peer reviewed to ensure their accuracy. Both procedures were adapted from the 2015 PRESS Guideline Statement.

The Medline strategy is presented below

1 Polycystic Ovary Syndrome/

2 ((polycystic or poly cystic) adj4 ovar*).tw.

3 pco*.tw.

4 ((degenerat* or sclerocystic) adj4 ovar*).tw.

5 stein leventhal.tw.

6 Anovulation/

7 anovulat*.tw.

8 (oligo ovulat* or oligoovulat*).tw.

9 ((hyperandrogen* or hyper androgen*) adj4 ovar*).tw.

10 or/1-9 55812

11 Economics/

12 Value of life/

13 exp "Costs and Cost Analysis"/

14 exp Economics, Hospital/

15 exp Economics, Medical/

16 Economics, Nursing/

17 Economics, Pharmaceutical/

18 exp "Fees and Charges"/

19 exp Budgets/

20 budget*.ti,ab.

21 cost*.ti.

22 (economic* or pharmaco?economic*).ti.

23 (price* or pricing*).ti,ab.

24 (cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.

25 (financ* or fee or fees).ti,ab.

26 (value adj2 (money or monetary)).ti,ab.

27 or/11-26

1 28 10 and 27

2 29 letter.pt. or letter/

3 30 note.pt.

4 31 editorial.pt.

5 32 case report/ or case study/

6 33 (letter or comment*).ti.

7 34 or/29-33

8 35 randomized controlled trial/ or random*.ti,ab.

9 36 34 not 35

10 37 animals/

11 38 exp Animals, Laboratory/

12 39 exp Animal Experimentation/

13 40 exp Models, Animal/

14 41 exp Rodentia/

15 42 (rat or rats or mouse or mice or rodent*).ti.

16 43 or/37-42

17 44 43 not humans/

18 45 36 or 44

19 46 28 not 45

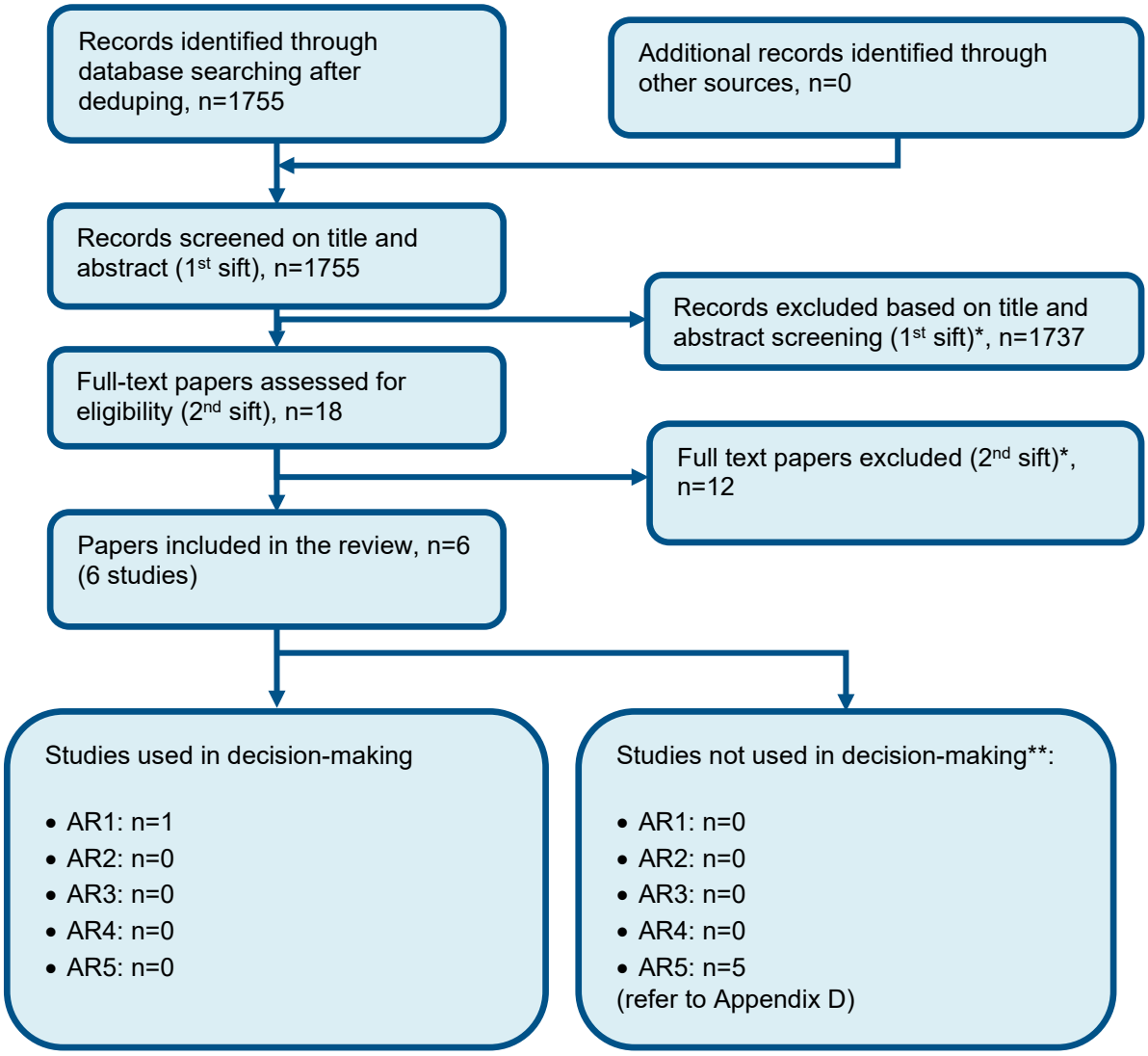
1 47 limit 46 to english language/

2 48 limit 47 to ed=20090101-20241

3 49 limit 47 to dt=20090101-20241205

4 50 48 or 49

Appendix B Health economic PRISMA diagram



* Not an economic evaluation, non-relevant population, intervention, comparison, design, setting or perspective; non-English language, not a full paper

**please refer to Review strategy described in the Economic review protocol in Methods document (Appendix B)

Appendix C Economic evidence tables

None

Appendix D Excluded health economic studies

None

Appendix E Health economic model

No original economic modelling was undertaken for the review questions in section two of the IG.

Appendix F References

None

Copyright for reproduced material

https://doi.org/10.26180/24003834.v1

https://doi.org/10.26180/23625288.v1

Suggested citation: Helena Teede et al.

International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome 2023. Monash University.

https://doi.org/10.26180/24003834.v1

Melbourne, Australia 2023

Monash University ISBN: 978-0-6458209-0-4