National Institute for Health and Care Excellence

Polycystic ovary syndrome

Adaptation report 3 – Lifestyle management

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 All NICE guidance is subject to regular review and may be updated or withdrawn.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]

1

Contents

3 Lifestyle management

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.

3.1 Effectiveness of lifestyle interventions

Review question 3.1: In women with PCOS, are lifestyle interventions (compared to minimal or nothing) effective for improving anthropometric, metabolic, reproductive, fertility, quality of life, and emotional wellbeing outcomes?

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

Evidence-based recommendation:

3.1.1 Lifestyle intervention (exercise alone or multicomponent diet combined with exercise and behavioural strategies) should be recommended for all women with PCOS, for improving metabolic health including central adiposity and lipid profile.

Consensus recommendation:

3.1.2 Healthy lifestyle behaviours encompassing healthy eating and/or physical activity should be recommended in all those with PCOS to optimise general health, quality of life, body composition and weight management (maintaining weight, preventing weight gain and/or modest weight loss).

Practice points:

3.1.3 Health professionals should be aware that lifestyle management is a core focus in PCOS management.

3.1.4 Lifestyle management goals and priorities should be co-developed in partnership with women with PCOS, and value women's individualised preferences.

3.1.5 There are benefits to a healthy lifestyle even in the absence of weight loss.

3.1.6 In those with higher weight, weight management can be associated with significant clinical improvements and the following key points need to be considered including:

- a lifelong focus on prevention of further weight gain
- if the goal is to achieve weight loss, a tailored energy deficit could be prescribed for women, considering individual energy requirements, body weight and physical activity levels.
- the value in improvement in central adiposity (e.g. waist circumference, waist-hip ratio) or metabolic health
- the need for ongoing assessment and support.

3.1.7 Healthcare professionals should be aware of weight stigma when discussing lifestyle management with women with PCOS. [see 3.6]

3.1.8 Healthy lifestyle and optimal weight management, in the context of structured, intensive and ongoing clinical support, appears equally effective in PCOS as in the general population.

3.1.9 In those who are not overweight, in the adolescent and at key life points, the focus should be on healthy lifestyle and the prevention of excess weight gain.

3.1.10 Insulin resistance is considered as a pathophysiological factor in PCOS. However, clinically available insulin assays are of limited clinical relevance and should not be used in routine care (refer to recommendation 1.9.12).

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 *(PICOS are clearly defined. PRISMA suggests that study selection was carried out based on the PICO).*

Section	Question	Answer
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Probably yes *(PICOS is clearly defined and appropriate for the review question. QoL is part of question and is reported in summary of studies but is excluded in the protocol).*
Study eligibility criteria	Were eligibility criteria unambiguous?	Probably yes. *(More detail could have been included for the comparison, as it only states usual care).*
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(Limits are appropriate and clearly defined. This is an update of a review, limits for searching the data are applied after the publication of the previous review).*
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 *(Limits are English language and dated to update search are appropriate).*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All signalling questions were answered "yes" or "probably yes", so no potential concerns. The "probably yes" was related to whether the eligibility criteria was appropriate for the review question, as the overall PICOs were defined but QoL is in the question and results, yet it is stated in the PICO that this outcome and patient satisfaction is excluded. More detail could have been given for the comparator, as it just stated 'usual care'. The limits of English language and date were appropriate).*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published	Probably no *(5 databases searched – Medline, Psycinfo, EMBASE, Cinahl and CENTRAL. However, there was no consideration of sources beyond bibliographic*

Section	Question	Answer
	and unpublished reports?	*databases. This would exclude any unpublished reports).*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably yes. *(This work is an attempt to update a prior (2019) Cochrane review (https://doi.org/10.1002/14651858.CD007506.pub4). All included studies from the Cochrane review have been considered for the current update but the update searches have not included any methods to identify studies beyond database searching. The main methods section says that relevant reviews/meta-analysis studies identified by the search strategy were also searched for identification of additional studies, but this is not specified in the review).*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(The English spelling of the controlled vocabulary term randomized controlled trial.pt is used, which is the most important term for identifying randomised controlled trials but it will not work unless the American spelling (''randomized', not 'randomised' is used, therefore this approach may miss RCTs. MesH subject headings have not been translated for other databases, so just searching on text words. There are some typos within the search strategy such as "self elicacy").*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(The search date was from the last search date of the Cochrane review, which was used, but as applied (by publication date rather than by database entry date) there is a small risk of missing papers due to the lag time between references being published and entered into a bibliographic database. Limited to publications in English which is appropriate).*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(Studies were selected and appraised by 5 reviewers).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(A few databases were searched and there were multiple reviewers to minimise error in selection of*

Section	Question	Answer
		studies. However, there were serious concerns about the search terminology, including not using both US/UK terminology for randomised controlled trials and typos in the search strategy).
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(Data was collected and appraised by 5 reviewers alongside the 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 *(Detailed study characteristics tables available, appears appropriate for interpretation of results).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Probably *yes* *(The details of the outcomes extracted from each study is clearly given).*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Probably yes *(Risk of bias was assessed by study and GRADE was used, however the downgrading approach was different for overall grading, for example an outcome is downgraded for risk of bias and indirectness, but it is rated as moderate quality rather than low quality overall. No footnotes for explanation).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Probably yes *(Does not clearly state but looks like appraisal was carried out by multiple reviewers).*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(All the signalling questions were answered with "yes" or "probably yes" as the data was collected and appraised by multiple reviewers; detailed study characteristics tables were provided; and the details of the outcomes extracted from each study is clearly given. Risk of bias was conducted and tables provided. GRADE ratings were provided, however GRADE was rated differently from NICE with downgrading done only once where there were serious ratings for more than 2 domains. There were no GRADE footnotes to see the reasons for downgrading. There was no risk of bias conducted for the Cochrane review).*

Section	Question	Answer
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(PRIMSA shows that 7 new studies and 11 from previous guideline were included in the analysis. 4 from the original Cochrane were excluded because their PCOS definition was unclear and so did not meet their updated PICO).*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	Probably yes *(No pre-defined analyses were reported. However, the methods section appears to be rigorous in detail and all analyses mentioned are addressed in 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?	Unclear *(Meta-analyses were performed based on lifestyle interventions mentioned in the PICO. The types of interventions were varied but there was little heterogeneity found. All lifestyle interventions were combined compared to minimal treatment, however it should be noted that the PICO said it was compared to usual care, without defining what usual care was, but in the analyses used the term minimal treatment. RCTs were included only and outcomes synthesised based on the reported outcomes.* *They state that where medians and interquartile ranges were reported instead of means and SD, they used medians in place of means and the formula of SD= (third quartile-first quartile)/1.35 to calculate the SD. It may be possible to meta-analyse means and medians if the distribution is symmetrical, but medians are often used because the distribution is skewed so may not be appropriate).*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Probably yes *(Only two outcomes in the meta-analyses appeared to have high I^2 values (figure 8 and 9). They state in the evidence summary on page 1981 that a random effects model was used and they are appropriately downgraded in GRADE).*
Synthesis and findings	Were the findings robust, e.g. as demonstrated	No information

Section	Question	Answer
	through funnel plot or sensitivity analyses?	*(There was no information about sensitivity or subgroup analyses).*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Yes *(Risk of bias was reported in the extractions and downgraded for in GRADE).*
Synthesis and findings	Concerns regarding the synthesis and findings	Unclear *(The signalling responses were mainly "yes" or "probably yes" however there was one "unclear". PRISMA flowchart was included and all studies were accounted for. No pre-defined analyses were reported, however the methods section appears rigorous, and all analyses mentioned are addressed in the results. Meta-analyses were performed appropriately to the PICO, with all lifestyle interventions combined. Meta-analysis was only conducted where there were 3 studies or more. Only two outcomes had high I^2 value (figure 8 and 9). They state in the text that a random effects model was used, and it was downgraded appropriately in GRADE. However, they did include medians and calculate IQRs in the meta-analyses, which may not be appropriate if the data is skewed).*
Judging risk of bias	Concerns regarding specification of study eligibility criteria	Low *(All signalling questions were answered "yes" or "probably yes", so no potential concerns. The "probably yes" was related to whether the eligibility criteria was appropriate for the review question, as the overall PICOs were defined but QoL is in the question and results, yet it is stated in the PICO that this outcome and patient satisfaction is excluded. More detail could have been given for the comparator, as it just said, 'usual care'. The limits of English language and date were appropriate).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(A few databases were searched and there were multiple reviewers to minimise error in selection of studies. However, there were serious concerns about the search terminology, including not using both US/UK terminology for randomised controlled trials and typos in the search strategy).*

Section	Question	Answer
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(All the signalling questions were answered with "yes" or "probably yes" as the data was collected and appraised by multiple reviewers; detailed study characteristics tables were provided; and the details of the outcomes extracted from each study is clearly given. Risk of bias was conducted and tables provided. GRADE ratings were provided, however GRADE was rated differently from NICE with downgrading done only once where there were serious ratings for more than 2 domains. There were no GRADE footnotes to see the reasons for downgrading. There was no risk of bias conducted for the Cochrane review).*
Judging risk of bias	Concerns regarding the synthesis and findings	Unclear *(The signalling responses were mainly "yes" or "probably yes" however there was one "unclear". PRISMA flowchart was included and all studies were accounted for. No pre-defined analyses were reported, however the methods section appears rigorous, and all analyses mentioned are addressed in the results. Meta-analyses were performed appropriately to the PICO, with all lifestyle interventions combined. Meta-analysis was only conducted where there were 3 studies or more. Only two outcomes had high I^2 value (figure 8 and 9). They state in the text that a random effects model was used, and it was downgraded appropriately in GRADE. However, they did include medians and calculate IQRs in the meta-analyses, which may not be appropriate if the data is skewed).*
Overall review ratings	Overall risk of bias	Low *(Two domains were rated low, one was rated high for concerns regarding methods used to identify and/or select studies, because of issues with the search terminology. One domain was unclear which was due to the inclusion of medians. The relevance of identified studies to the review's research question was appropriately considered and the reviewers provided results for both those outcomes that were statistically significant and not).*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: The IG used the GRADE framework to describe how they went from the evidence to the recommendations. The overall GRADE rating of the certainty of the evidence for 3.1 was rated as very low. However, the recommendations are 'strong' recommendations, based on the GRADE framework. They justified this as there were relatively few participants (n=634) across the 18 studies, so the recommendations were not only informed by the evidence but also based on additional guidelines for the general population.

The recommendation specifically states that lifestyle interventions are recommended for 'improving metabolic health including central adiposity and lipid profile.' The only outcomes showing statistical significance were waist circumference; waist: hip ratio; Ferriman-Gallwey score; fasting insulin; total cholesterol and low-density lipoprotein cholesterol (LDL-C). There were no differences for body weight; body mass index (BMI), sex hormone-binding globulin (SHBG), Total testosterone, free androgen index (FAI); glucose regulation (fasting); glucose regulation (2-hour postprandial glucose levels); high-density lipoprotein cholesterol (HDL-C) and triglyceride (TG) levels. The recommendations are justified based on general population data which support healthy lifestyle interventions and there are few side effects, but access, cost and equity may be concerns. However, there is some ambiguity to the recommendation as lifestyle interventions are not defined and whether they are relevant to all women with PCOS. The review mixed studies of various weight status so it is not clear whether the findings are only relevant to those who are overweight or obese. The implementation considerations were that substantial resources are likely needed to implement the comprehensive lifestyle modifications. Research priorities in the IG suggest that further high quality, powered randomised controlled trials and pragmatic implementation trials are needed to look at different aspects of lifestyle interventions, such as different ways the interventions can be delivered.

Where no evidence was identified and a CR was made the IG recommends healthy lifestyle behaviours in everyone with PCOS. This would not involve additional NHS resources and is in line with other NICE guideline recommendations.

IG economic evidence

No health economic evidence was identified in the IG for review question 3.1 on the effectiveness of lifestyle interventions.

The IG noted that many interventions reported in the clinical review are resource intensive and would require substantial resources to implement. The IG also made a statement on cost-effectiveness noting that these costs are balanced by health outcome improvements. However, no additional rationale was provided to underpin this statement.

The IG also acknowledged that implementation may be variable due to the availability of local resources and that healthcare funders may not prioritise lifestyle management interventions. The IG also concluded that their recommendations would likely improve equity.

3.1.2 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 3.1.

The committee agreed that other sections of the IG were of a higher priority for original health economic modelling, noting that most review questions in Chapter 3 of the IG are covered by existing NICE guidelines.

Unit costs

Table 1: Unit costs for lifestyle interventions

Resource	Unit costs (per hour, including qualification costs)
Behavioural interventions	
Entry level counsellor	£43 - £44
Counsellor & clinical psychology trainee	£55 - £57
Specialist counsellor & clinical psychologist	£66 - £68
Principal clinical psychologist	£74 - £77
Consultant counsellor	£86 - £89
Consultant clinical psychologist	£103 - £106
Dietary interventions	
Dietician (Band 5)	£43 - £44
Dietician (Band 6)	£55 - £57

Source: PSSRU 2024 (accessed 2 November 2025). Hospital-based staff cost and community-based staff cost respectively presented.

3.1.3 NICE recommendations

The relevant recommendations for this section are Rec 1.1.5, 1.1.6 and 1.13.1.

3.1.4 The committee's discussion and interpretation of the evidence

Clinical evidence

The committee's contextualisation of the IG recommendations on lifestyle interventions was informed primarily by a comparison of the IG's recommendations with recommendations in NICE's guideline on the prevention and management of obesity and overweight (NG246), which, with the exception of pregnant women, applies to all women with PCOS, irrespective of their weight and BMI threshold. The committee noted that there was significant overlap between the IG and NG246, with the NICE guideline also providing additional UK-specific detail therefore the committee decided to cross-refer to the existing NICE guideline for guidance on

lifestyle advice for all women with PCOS, regardless of BMI, and for guidance on lifestyle interventions for women with PCOS who are overweight or obese.

The committee agreed that lifestyle advice was important for all people with PCOS, regardless of BMI and so recommended advice on healthy lifestyle behaviours in line with NG246. The committee did not think the IG provided sufficient justification for lifestyle interventions in non-obese and non-overweight women with PCOS. There was a lack of studies in the IG for lifestyle interventions in these populations, with most of the evidence coming from studies where either the participants were overweight or obese. The committee noted further that the IG itself had acknowledged this gap by identifying effectiveness in the non-overweight group as a research priority The committee therefore made a recommendation for lifestyle interventions to be provided in line with NG246 for those who are obese or overweight.

Health economic

No health economic evidence was provided in the IG to support their recommendation for providing lifestyle interventions for all people with PCOS. However, the IG did note that substantial resources would likely be needed to implement this recommendation. In the absence of any included health economic evidence identified from our literature search, unit costs were presented to the committee to aid their consideration of cost-effectiveness.

The committee acknowledged there is significant overlap with NICE's guideline NG246 and the IG recommendations for those with PCOS who are overweight or obese, noting that lifestyle interventions are recommended for this cohort in NG246. The committee also discussed that there was a lack of clinical and health economic evidence on lifestyle interventions for people with PCOS who have a healthy BMI. The possibility of additional health economic work to determine the cost-effectiveness of lifestyle interventions for those people with PCOS and a healthy BMI was discussed, but due to lack of available clinical evidence it was decided that any HE

analysis would not help inform an assessment of the cost-effectiveness of recommendations. The committee therefore concluded that due the lack of clinical evidence for people with healthy BMI's, recommendations on lifestyle interventions should be confined to those who are overweight or obese. The committee also acknowledged that the resource impact of recommending lifestyle interventions for women with PCOS who are not obese or overweight would likely be significant. The committee, however, emphasised that advice on healthy diet, physical activity and healthy behaviours should be provided to all people with PCOS, irrespective of their BMI, and made recommendations to reflect this.

3.2 Behavioural interventions

Review question 3.2: In women with PCOS, are behavioural interventions in addition to diet and/or exercise (compared to diet and/or exercise alone) effective for improving anthropometric, metabolic, reproductive, fertility, QoL and emotional wellbeing outcomes?

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

Consensus recommendation:

3.2.1 Lifestyle interventions could include behavioural strategies such as goal-setting, self-monitoring, problem-solving, assertiveness training, reinforcing changes and relapse prevention, to optimise weight management, healthy lifestyle and emotional wellbeing in women with PCOS.

Practice point:

3.2.2 Behavioural support could include: goal setting, problem solving, self-monitoring and reviewing, or SMART goals (Specific, Measurable, Achievable, Realistic and Timely).

3.2.3 Comprehensive healthy behavioural or cognitive behavioural interventions could be considered to increase support, engagement, retention, adherence and maintenance of healthy lifestyle and improve health outcomes in women with PCOS.

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 *(PICOS were clearly stated)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(The question was clear and the PICOS matched this)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(There was no ambiguity in eligibility criteria)*
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 *(Yes although no studies were found so they could have included non-randomised studies. There was no date or sample size limits and multiple outcomes 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 *(Only English language as limit and this was appropriate)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All scored yes and one scored probably yes (as they could have looked at non-randomised studies), so there are no concerns for eligibility criteria)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably no *(Embase, Medline, Web of Science, Cochrane Central Register of controlled trials and google scholar used. However, limited consideration for unpublished reports – database sources included conference sources (via Web of Science/Embase) and trial registries (via CENTRAL) but no subsequent attempt to chase up related publications to any non-journal publications found)*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably yes *(No additional search methods mentioned except in the general methods section of the technical report says that 'GDG members were consulted, bibliographies of*

Section	Question	Answer
		relevant reviews/meta-analysis studies identified by the search strategy were also searched for identification of additional studies', but this is not stated in the review)
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 *(Searches are reasonably structured but terminology for interventions is not as extensive as it could be – they seem quite skewed to finding behavioural interventions when they are described in those terms, but would potentially overlook relevant studies if they are described in other ways (e.g. 'counselling' or 'mindfulness interventions').*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Yes *(a date limit is applied which is appropriate).*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(Two reviewers sifted the studies based on the title and abstracts alongside the reviewing team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(A few databases were searched and there were multiple reviewers to minimise error in selection of studies. However, there were issues around extensiveness of search terminology and limited searching beyond published material in standard databases)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(Two reviewers selected and appraised the studies)*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	No information *(No studies identified)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	No information *(No studies were identified)*

Section	Question	Answer
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	No information *(No studies were identified)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	No information *(No studies identified)*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(No studies identified)*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(No studies were included but PRISMA and excluded list match)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(No studies identified)*
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 *(No studies identified)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	No information *(No studies identified)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	No information *(No studies identified.)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	No information *(No studies identified)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(No studies were identified)*

Section	Question	Answer
Judging risk of bias	Concerns regarding specification of study eligibility criteria	Low *(Most of the studies were answered as 'yes' however non-randomised studies could have been included, however in their findings they discuss some randomised controlled trials which did not meet eligibility criteria in the background (see below) so this may be the reason)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(A few databases were searched and there were multiple reviewers to minimise error in selection of studies. However, there were issues around extensiveness of search terminology and limited searching beyond published material in standard databases)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Not applicable (No studies were included).
Judging risk of bias	Concerns regarding the synthesis and findings	Not applicable (No studies were included).
Overall review ratings	Overall risk of bias	There were no studies included in the review and the eligibility criteria and identification and selection of studies was acceptable. There were studies not included that were detailed but this was not included in the recommendations. There was no evidence-based recommendation made, but a consensus recommendation and practice points suggested behavioural strategies may be useful in women with PCOS.
Overall review ratings	Applicability as a source of data	No studies were identified.

IG evidence to recommendations justification: no evidence was found for the systematic review so there were no EBRs, only CR and practice points. The CR

suggested types of behavioural strategies that could be used in women with PCOS such as goal-setting, self-monitoring, problem-solving, assertiveness training, reinforcing changes and relapse prevention.

IG economic evidence

No health economic evidence was identified in the IG for review question 3.2 on behavioural interventions.

The IG noted that overall, there remains a paucity of evidence to support the implementation of additional behavioural therapy on top of lifestyle management for people with PCOS. The IG also noted that additional healthcare professional training and time may limit the acceptability of implementing the recommendations made in the IG and the additional cost of implementing these recommendations may impact the feasibility to adopt their recommendations.

3.2.2 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 model was developed for review question 3.2 as the committee agreed that other areas of the guideline were of a higher priority for original health economic analysis. The committee made this decision primarily since they were aware of NICE's existing guideline on overweight and obesity

management (NG246) and therefore suspected that any recommendations made for this section of the guideline would likely be covered by this existing guideline.

3.2.3 NICE recommendations

No recommendations were made.

3.2.4 The committee's discussion and interpretation of the evidence

Clinical evidence

The committee's contextualisation of the IG recommendations on behavioural strategies was informed primarily by comparison of the IG's recommendations with recommendations in NICE's guideline on the prevention and management of obesity and overweight (NG246), which, with the exception of pregnant women, already applies to all women with PCOS. The committee noted that there was significant overlap between the two and substantial additional information in NG246 compared with section 3.2 of the IG. The committee also highlighted that behavioural strategies may be classified as a lifestyle intervention and were therefore interventions included in IG's included evidence base in evidence review 3.1. For this reason, and because the committee's recommendations on lifestyle interventions themselves cross-refer to NG246, the committee considered guidance on behavioural strategies to be adequately covered by this cross-reference.

Health economic

No health economic evidence was included in the IG to support their recommendations on behavioural interventions. In addition, no health economic evidence was identified in NICE's health economic literature search.

As the recommendations in NICE's PCOS guideline cross-refer to existing NICE guideline recommendations in NG246, no significant resource impact is associated with these recommendations.

3.3 Diet interventions

Review question 3.3: In women with PCOS, are diet interventions (compared to different diets) effective for improving anthropometric, metabolic, fertility, and emotional wellbeing outcomes?

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

Evidence-based recommendation:

3.3.1 Health professionals and women could consider that, there is a lack of evidence supporting any one type of diet composition over another for anthropometric, metabolic, hormonal, reproductive or psychological outcomes.

Consensus recommendation:

3.3.2 Any diet composition consistent with population guidelines for healthy eating has health benefits, and within this, health professionals should advise sustainable healthy eating tailored to an individual's preferences and goals.

Practice point:

3.3.3 Tailoring of dietary changes to food preferences, allowing for a flexible, individual and co-developed approach to achieving nutritional goals and avoiding unduly restrictive and nutritionally unbalanced diets, are important, as per general population recommendations.

3.3.4 Barriers and facilitators to optimise engagement and adherence to dietary change should be discussed, including psychological factors, physical limitations, socioeconomic and sociocultural factors, as well as personal motivators for change. The value of broader family engagement should be considered. Referral to suitably trained allied healthcare professionals needs to be considered when women with PCOS need support with optimising their diet.

*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 *(PICOS is clearly defined. Search dates are clearly stated. This is an update of a previous guideline review)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(The question was about diet interventions compared to different diets to improve anthropometric, metabolic, fertility, and emotional wellbeing outcomes in women with PCOS. The PICO included women diagnosed with PICOS by the NIH 1990, Rotterdam 2003 or AE-PCOS 2006 criteria. The intervention included all types of dietary compositions and compared to all types of dietary compositions that differed from the intervention diet. They included anthropometric, fertility, non-fertility, metabolic, quality of life and emotional wellbeing outcomes. They also included the surrogate measures of insulin resistance. They included SRs or RCTs, which was appropriate for the intervention question)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(They clearly state which population is included and excluded, interventions that are included and excluded. They exclude those usual diet or original patient diet if it has not defined its nutritional composition. It has stated different outcomes under each category and that fat mass measured using skin-fold tests is excluded. Non-RCTs are excluded)*
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g.	Yes *(The search was date restricted in line with update of the previous guideline. No limits on sample size and there were many*

Section	Question	Answer
	date, sample size, study quality, outcomes measured)?	*outcomes included. The study type was limited to systematic reviews and RCTs for the outcomes. This is appropriate given the number of studies found)*
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 *(There were no restrictions on publication or format and availability of data. Limited to English language)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All were scored yes so there was no concerns with the study eligibility criteria. There were few limits except English language. The PICO was clearly defined)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably no *(Searches through Medline OVID, All EMB, PsycINFO, EMBASE and CINHAL. The database searches are okay for published material but no specific searches for unpublished material)*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably yes *(They evaluated the evidence from the previous guideline for inclusion. The methods section of the technical report says that GDG members were consulted, bibliographies of relevant reviews/meta-analysis studies identified by the search strategy were also searched for identification of additional studies but this is not specified in the review except mention of an additional trial identified via a bibliography (p2176) but no details of any systematic effort to find new studies via methods other than database searching)*
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 *(Medline search (p2170) is okay. Date limit is by publication year rather than database entry date, so some scope for papers to be missed because of the lag between publication and inclusion in a database. Reporting of search strategies for other databases is extremely unclear (e.g. is*

Section	Question	Answer
		search at the bottom of page 2 for Medline of Embase), but it looks like subject (MeSH) terms from Medline have not been properly translated for other databases).
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(The date restriction was appropriate as it was an update of the 2018 guideline however as it was applied (by publication date rather than by database entry date) there is a small risk of missing papers due to the lag time between references being published and entered into a bibliographic database; no restrictions on publication format. Limited to English language which is appropriate).*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(Two reviewers sifted the studies based on the title and abstracts alongside the reviewing team).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	Unclear *(There were some domains rated yes or probably yes, such as two reviewers to reduce error in study selection and any restrictions by date or language were acceptable, also some bibliographic searches were apparent. However, there was some unclear reporting for searches in databases beyond Medline).*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(Studies were selected and appraised by 6 reviewers alongside the evidence team. Structured data extraction forms were used).*
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 characteristics table summarises all important information).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Probably yes *(Many studies could not be meta-analysed due to differences but were descriptively presented in tables).*

Section	Question	Answer
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(Risk of bias assessment for each included study is given in the evidence summary).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Probably yes *(Evidence summary mentions reasons for high risk of bias).*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Most scores were yes or probably yes as studies appraised by multiple reviewers, study characteristics summarise all important information. All study results were collected for use in the descriptive analysis, where relevant to the systematic review but not all were meta-analysed due to heterogeneity of types of diet etc. Risk of bias is provided in the evidence summary using appropriate criteria and they detail reasons for high risk of bias)*.
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(PRISMA suggests that 5 out of 12 included RCTs were from the updated search and rest were from previous guideline)*.
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(No details of predefined analyses)*.
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 done for 5 of the included studies comparing a hypocaloric high protein/low carbohydrate diet (40% carbohydrate, 30% protein, 30% fat) vs a control diet (55% carbohydrate, 15% protein, 30% fat). Meta analysis could not be conducted for the rest of the studies due to heterogeneity in the dietary composition of intervention and control and variability in the energy restriction)*.
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Probably yes *(There was heterogeneity for fasting insulin, HOMA-IR and LDL-C, they used the random effects model (on all meta-analyses not just those with high*

Section	Question	Answer
		heterogeneity) but there were too few studies to investigate with subgroup analyses as they had minimum 3. They did not downgrade for inconsistency but included a note (see at bottom of ROBIS)).
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Probably yes *(Evidence summary gives funnel plots for meta-analysis. Sensitivity analyses not possible to be done).*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Yes *(Risk of bias was assessed for all primary studies using an extraction table and GRADE included risk of bias across outcomes and was downgraded according to the primary studies).*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(Mostly scored yes or probably yes, with no information for pre-defined analyses but the protocol was detailed. All meta-analyses used random effects model but there were too few studies to investigate with subgroup analysis).*
Judging risk of bias	Concerns regarding specification of study eligibility criteria	Unclear *(All were scored yes so there were no concerns with the study eligibility criteria. There were few limits except English language. The PICO was clearly defined).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	Unclear *(There were some domains rated yes or probably yes, such as two reviewers to reduce error in study selection and any restrictions by date or language were acceptable, also some bibliographic searches were apparent. However, there was some unclear reporting for searches in databases beyond Medline).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Most scores were yes or probably yes as studies appraised by multiple reviewers, study characteristics summarise all important information. All study results*

Section	Question	Answer
		were collected for use in the descriptive analysis, where relevant to the systematic review but not all were meta-analysed due to heterogeneity of types of diet etc. Risk of bias is provided in the evidence summary using appropriate criteria and they detail reasons for high risk of bias).
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(Mostly scored yes or probably yes, with no information for pre-defined analyses but the protocol was detailed. All meta-analyses used random effects model but there were too few studies to investigate with subgroup analysis).*
Overall review ratings	Overall risk of bias	*All of the domains were low concern except for identifying and selection of studies. Studies that were meta-analysed showed no difference except for HDL-C and this is reflected in the EBR which states that there is a lack of evidence supporting one type of diet composition over another. They explained why they did not meta-analyse for some studies due to dietary composition of intervention and control arms and variability in energy restriction prescribed so these were summarised descriptively. GRADE was not conducted on these studies however they state that these findings suggest that diets with a range of macronutrient compositions could be recommended for women with PCOS and is in line with the general population. The consensus recommendation states that any diet composition consistent with population guidelines for healthy eating has health benefits and therefore should be tailored to individuals*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: 12 RCTs were found and included in this review. The EBR states that there is a lack of evidence supporting one diet over another. The authors were unable to meta-analyse results because of

heterogeneity in the dietary composition of intervention arms, variability in the dietary composition of control arms and variability in the energy restriction prescribed. Some of the studies were small and from a single country (Iran), which included a traditional diet which may not be applicable to other countries. A meta-analysis was conducted on three studies comparing high protein/low carbohydrate diet versus standard diets. Benefit was shown for a high protein diet in reducing HDL-C (high density lipoprotein C). No other statistically significant benefits were found for the other outcomes. The overall findings of the studies on dietary interventions do not support any one diet over another which justifies the evidence based recommendation in the IG.

The CR recommends any diet composition recommended in population guidelines for healthy eating may be beneficial advice and should be tailored to an individual's preferences and goals, which is in line with NICE obesity guidelines.

IG economic evidence

No health economic evidence was included in the IG relating to review question 3.3 on diet interventions.

The IG noted that the potential costs associated with providing dietary interventions could include the development and distribution of educational resources for people with PCOS, education of health professionals, and tailored dietary advice delivered by health professionals which could require longer consultation times.

The IG stated that their recommendations may increase healthcare costs because of increased consultation times and referral to specialist healthcare professionals, however noted that the long-term benefits of dietary and lifestyle modification may reduce the health and economic burden of PCOS. It was also noted that the recommendations made in IG are unlikely to significantly change usual care for most health practitioners, but people with PCOS may want more specific dietary advice to assist them with losing weight or preventing weight gain.

In terms of implementation, it was noted that education on evidence-based dietary recommendations is required for both health professionals and people with PCOS, especially regarding expectations of treatment and misinformation.

3.3.2 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 model was developed for review question 3.3 from the IG as the committee agreed that other areas of the guideline were of a higher priority for original health economic analysis. The committee made this decision primarily since they were aware of NICE's existing guideline on overweight and obesity management (NG246) and therefore suspected that any recommendations made for this section of the guideline would likely be covered by this existing guideline.

3.3.3 NICE recommendations

The relevant recommendation for this section is Rec 1.13.2.

3.3.4 The committee's discussion and interpretation of the evidence

Clinical evidence

The committee decided to contextualise the one EBR from the IG on diet interventions (3.3.1). The committee concluded that the information conveyed in this

recommendation, namely that there is no evidence to support one type of diet composition over another would be useful for healthcare professionals to be aware of when responding to questions from women with PCOS on diet.

The committee decided not to contextualise the other CRs (3.3.2) and practice point recommendations (3.3.3 & 3.3.4) on diet interventions. The committee concluded that the contextualisation of the EBR (3.3.1), whilst also cross-referring to NICE's existing guideline on overweight and obesity (NG246), encapsulated everything they wanted to say on dietary interventions for people with PCOS.

Health economic

No health economic evidence was included in the IG for review question 3.3, and no health economic evidence that was applicable to this review question was identified in the health economic literature search conducted by NICE. The committee therefore considered the cost-effectiveness of the IG recommendations in relation to an NHS context – discussing the costs and health outcomes qualitatively.

The guideline committee decided to contextualise the one EBR from the IG which noted the lack of evidence to support one type of healthy diet composition over another. As this contextualised recommendation provides information to healthcare practitioners and people with PCOS – informing them of the evidence base which indicates that no specific dietary intervention is recommended over another – the only potential cost implication of this recommendation could be the additional staff time required to inform people about a healthy diet. As these costs implications are likely to minimal, no significant resource impact is anticipated from this recommendation.

The guideline committee also cross-referred to NICE's guideline NG246 on overweight and obesity management – noting that there is information within this guideline on healthy eating for the general population, as well as for those who are overweight or obese. No cost implications or significant resource impact will be associated with this cost-referral to existing NICE guidance.

3.4 Exercise interventions

Review question 3.4: In women with PCOS, are exercise interventions (compared to different exercises) effective for improving anthropometric, metabolic, reproductive, fertility, quality of life and emotional wellbeing outcomes?

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

Evidence-based recommendation:

3.4.1 Health professionals and women could consider that, there is no evidence to support any one type and intensity of exercise being better than another for anthropometric, metabolic, hormonal, reproductive or psychological outcomes.

Consensus recommendation:

3.4.2 Any physical activity consistent with population guidelines will have health benefits and within this, health professionals should advise any sustainable physical activity based on individual preferences and goals.

3.4.3 Health professionals should encourage and advise the following in concordance with general population physical activity guidelines:

- All adults should undertake physical activity as doing some physical activity is better than none.
- Adults should limit the amount of time spent being sedentary (e.g. sitting, screen time) as replacing sedentary time with physical activity of any intensity (including light intensity) provides health benefits.

For the prevention of weight gain and maintenance of health, adults (18-64 years) should aim for a minimum of 150 to 300 minutes of moderate intensity activities or 75 to 150 minutes per week of vigorous intensity aerobic activity or an equivalent combination of both spread throughout the week, plus muscle strengthening activities (e.g. resistance/flexibility) on two non-consecutive days per week.

For promotion of greater health benefits including modest weight-loss and prevention of weight-regain, adults (18-64 years) should aim for a minimum of 250 min/week of moderate intensity activities or 150 min/week of vigorous intensities or an equivalent combination of both, plus muscle strengthening activities (e.g. resistance/flexibility), ideally on two non-consecutive days per week.

Adolescents should aim for at least 60 minutes of moderate- to vigorous-intensity physical activity per day including activities that strengthen muscle and bone, at least three times per week.

Practice point:

3.4.4 Physical activity is any bodily movement produced by skeletal muscles that requires energy expenditure. It includes leisure time physical activity, transportation (e.g. walking or cycling), occupational (i.e. work), household chores, playing games, sports or planned exercise, or activities in the context of daily, family and community activities.

3.4.5 Aerobic activity is best performed in bouts of at least 10 minutes duration, aiming to achieve at least 30 minutes daily on most days.

3.4.6 Barriers and facilitators to optimise engagement and adherence to physical activity should be discussed, including psychological factors (e.g. body image concerns, fear of injury, fear of failure, mental health), personal safety concerns, environmental factors, physical limitations, socioeconomic factors, sociocultural factors, as well as personal motivators for change. The value of broader family engagement should be considered. Referral to suitably trained allied health professionals needs to be considered for optimising physical activity in women with PCOS.

3.4.7 Self-monitoring including with fitness tracking devices and technologies for step count and exercise intensity, could be used as an adjunct to support and promote active lifestyles and minimise sedentary behaviours.

*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 *(The PICOS is clearly detailed and appropriate)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(The question is about exercise interventions compared to different exercises for various outcomes, which the PICOS matches)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(The eligibility criteria were well-defined in the PICOS)*
Study eligibility criteria	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(The limits by date were appropriate as they take into consideration the original search dates and an RCT filter is included which is appropriate for the question)*
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 *(The limit by English language was appropriate)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All study eligibility criteria was given a 'yes' score, so this was given low concerns overall)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably no *(This is an update of their original review so original searches were updated. 5 databases were searched - Medline, PsychINFO, EMBASE, All EMB and CINAHL. This is okay for published reports but is insufficient for unpublished reports –*

Section	Question	Answer
		would want to see trial registries included for these)
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably yes *(Only 2 records identified from sources other than databases – from the dates of the included studies it can be assumed these were brought forward from the prior guideline review, though this is not explicitly stated. Otherwise, no additional methods of findings studies reported, however the general methods section of the technical report says that GDG members were consulted, bibliographies of relevant reviews/meta-analysis studies identified by the search strategy were also searched for identification of additional studies, but this is not specified in the review).*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(lacks terms for guidelines and HTAs. Random.mp is not truncated so would miss mentions of randomised or randomized trials. It looks like the McMaster Sensitive RCT "Hedge" (filter) has been attempted to replicate but has failed on this term. It is likely very important given that the sensitive filter does not contain the term randomized controlled trial.pt, which is the most important one for picking up RCTs.* *Search looks to have been run verbatim across multiple databases without modification of subject heading from Medline).*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(The search date can be seen in the search strategy as 1st January 2017 to 31st December 2022, as this was an update of a previous review, however the date limit is applied by publication date rather than the (preferable) database entry date, so a chance of missing studies through indexing lag. Restricted on English language and excluded MEDLINE records)*

Section	Question	Answer
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(Studies were selected and appraised by 2 reviewers alongside the reviewing team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Fundamental errors in search syntax in the RCT filter (line 50, p2311) and no translation of Medline search terms to make them appropriate for Embase and other databases. Suspect that this would have made a big difference to retrieval in Embase and other sources but cannot be totally certain because total number of results is not reported for each database. There is a lack of search terms to pick up guidelines or HTAs despite these being included in the protocol)*
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 *(There were enough details in the summary study characteristics tables)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(All 5 eligible studies were included in the synthesis)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Probably yes *(Individual ROB assessment of each eligible study is reported in the evidence summary)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Probably yes *(There was risk of bias assessment conducted, they say that 2 reviewers appraised but not clear if this is critical appraisal of risk of bias or for appraising relevance to the PICO details)*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(The results are likely to include all relevant details and minimising bias as most questions were answered probably yes or yes)*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(All 5 eligible studies have been included in*

Section	Question	Answer
		the synthesis. Three comparisons were reported, meta-analysis was conducted for one of the comparisons (HIIT vs MICT) which was reported by 3 of the included studies. Descriptive analysis was done for the remaining two comparisons due to being a single study)
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	Probably yes *(There was a protocol but unclear if published. However, the methods section appears rigorous, protocol detailed and all analyses mentioned are addressed in 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?	Probably yes *(For HIIT vs MICT all the outcomes were presented in the same forest plot and SMD used. Results were provided separately for each outcome in a table and in GRADE using mean difference. Studies were meta-analysed even if two studies, which differs to other reviews in the IG. Forest plots not provided for other comparisons)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Unclear *(Serious inconsistency for fasting blood glucose; fasting insulin; HOMA-IR; LDL-C, testosterone. They are downgraded in GRADE for inconsistency, but GRADE table says fixed effects model used. There were too few studies to do subgroup analyses)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	No information *(There were too few studies per outcome for sensitivity analysis or funnel plots)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Probably yes *(The studies were assessed for risk of bias and bias when synthesised included in GRADE rating but then most say unclear risk of bias without more details in footnotes)*
Synthesis and findings	Concerns regarding the synthesis and findings	Unclear *(There were a few areas where it was unclear because details were not reported, particularly around the forest plots where everything was in one forest plot but then*

Section	Question	Answer
		reported in tables/GRADE as the appropriate analysis. Fixed effects models used although downgraded for inconsistency)
Judging risk of bias	Concerns regarding specification of study eligibility criteria	Low *(All study eligibility criteria was given a 'yes' score, so this is low concerns overall)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(A range of databases were searched for published studies and the search limits are adequate and minimisation of bias with 2 reviewers however there were fundamental errors in search syntax in the RCT filter (line 50, p2311) and no translation of Medline search terms to make them appropriate for Embase and other databases. Suspect that this would have made a big difference to retrieval in Embase and other sources but can't be 100% sure because total number of results isn't reported for each database).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(The results are likely to include all relevant details and minimising bias as most questions were answered probably yes or yes)*
Judging risk of bias	Concerns regarding the synthesis and findings	Unclear *(There were a few areas where it was unclear because details were not reported, particularly around the forest plots where everything was in one forest plot but then reported in tables/GRADE as the appropriate analysis. Fixed effects models used although downgraded for inconsistency)*
Overall review ratings	Overall risk of bias	*Yes, all areas were low concerns, except for identification of studies which was high and synthesis and findings was unclear. All GRADE ratings were low to very low, and the meta-analysis/descriptive analysis summary discussed inconsistency there. The recommendations say there is no*

Section	Question	Answer
		evidence to support any one type and intensity of exercise
Overall review ratings	Applicability as a source of data	*Fully applicable*

IG evidence to recommendations justification: there were only 5 RCTs and no differences were found between the different types of exercise, therefore the EBR is justified where it states that there is no evidence to support any one type and intensity of exercise for the included outcomes. This review was specifically looking at comparing different types of exercises in a PCOS population but they state that in general physical activity and structured exercise has been shown to be beneficial in general populations. They made a CR in line with this to recommend 'any physical activity consistent with population guidelines will have health benefits and within this, health professionals should advise any sustainable physical activity based on individual preferences and goals.' Both the EBR and first CR were rated conditional (weak) recommendation for the option. The second CR suggests that health professionals should provide advice in line with general population physical activity guidelines, with strong recommendation for the option, which is in line with the NICE obesity guideline.

IG economic evidence

No health economic evidence was included in the IG relating to review question 3.4 on exercise interventions.

The IG noted that it is important to consider the cost and resource requirements of exercise interventions at a health system level, noting that many low-cost options are available for individuals seeking exercise. Although the IG did not explicitly state these, examples of such could include, walking, running and free exercise classes accessed on the internet. The IG also acknowledged that irrespective of the exercise programme or physical activity recommendation, implementation of their

recommendations may result in increased referral to health professionals, thus potentially increasing health professional time demands.

The IG also noted that engagement of healthcare practitioners and healthcare funders is required for successful implementation of their recommendations. The IG concluded that it was likely feasible for health professionals to give general population physical activity advice but did acknowledge the associated time demands on healthcare professionals if not already implemented in current practice.

The potential drop-out or non-adherence to physical activity was acknowledged, noting that feasibility and sustainability of exercise interventions may vary depending on socioeconomic, sociocultural, physical factors.

The IG also noted that for individuals self-funding their exercise, financial barriers may be an issue.

3.4.2 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 model was developed for review question 3.4 from the IG as the committee agreed that other areas of the guideline were of a higher priority for original health economic analysis. The committee made this decision primarily

since they were aware of NICE's existing guideline on overweight and obesity management (NG246) and therefore suspected that any recommendations made for this section of the guideline would likely be covered by this existing guideline.

3.4.3 NICE recommendations

The relevant recommendation for this section is Rec 1.13.2.

3.4.4 The committee's discussion and interpretation of the evidence

Clinical evidence

The committee decided to contextualise the one EBR from the IG (3.4.1) as they concluded that the information conveyed in this recommendation, namely that there is a lack of evidence to support any one type and intensity of exercise over another (for anthropometric, metabolic, hormonal, reproductive or psychological outcomes) would be useful for healthcare professionals and people with PCOS to be aware of. The committee also cross-referenced existing NICE guidance (NG246), specifically cross-referring to health lifestyle advice and physical activity.

The committee decided not to contextualise the other two CRs on exercise interventions (recommendation 3.4.2 & 3.4.3) and practice point recommendations (3.4.4 to 3.4.7) associated with evidence review 3.4 in the IG as they concluded that the contextualisation of the EBR (3.4.1), in conjunction with cross-referring to NG246, encapsulated everything they wanted to say on exercise interventions for people with PCOS.

Health economic

No health economic evidence was identified in the IG or in NICE's health economic literature search for review question 3.4.

The NICE guideline committee decided to contextualise the one EBR from the IG (3.4.1). This recommendation detailed that there is no evidence to support one type of exercise over another for people with PCOS. The committee noted that because

this recommendation provides information on the current evidence base – and therefore no specific exercise intervention is being recommended – the only potential cost implications associated with this recommendation could be additional staff time required to inform people about the benefits of exercising. The committee noted that informing people of the benefits of exercise is reflective of best clinical practice and therefore any resource implications associated with this recommendation are likely to be small especially given the fact that this information will likely be provided at existing contacts with healthcare professionals. For this section of the guideline, the committee also cross-referred to NICE's existing recommendations on physical activity in NG246.

The committee concluded that as the contextualised recommendation from the IG is of an informative nature, and therefore the cost implications for implementing this are likely to minimal, no significant resource impact is anticipated for this review question.

3.5 Factors affecting weight gain in PCOS

Review question 3.5: Why are women with PCOS at increased risk of weight gain?

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

Evidence based recommendation:

3.5.1 Health professionals and women with PCOS could consider that there is no clear evidence of physiological or behavioural lifestyle differences, related to weight, in women with PCOS compared to women without.

Practice point:

3.5.2 Whilst the specific mechanisms are unclear, it is recognised that many women with PCOS will have underlying mechanisms that drive greater longitudinal weight gain and higher BMI which may:

- underpin greater challenges with weight management.
- highlight the importance of lifelong healthy lifestyle strategies and prevention of excess weight gain.
- assist women with PCOS and health professionals in forming realistic, tailored lifestyle goals.

IG clinical evidence

There were two reviews conducted on 1) extrinsic factors and 2) intrinsic factors potentially related to challenges with weight management.

Critical appraisal - ROBIS systematic review checklist - extrinsic factors

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Probably yes *(PICO is clearly defined. Few details for exclusion criteria)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Probably yes *(There were lots of different types of study design included, case-control, cross-sectional, cohort study and RCTs were all included however the question is about risk factors and may have been better answered by a cohort study with confounders addressed)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Probably yes *(There was no interventions specified in the protocol, but randomised controlled trials were included in addition to non-randomised)*
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 date limits were applied, which is appropriate as it is a new review)*
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 *(Inclusion was limited appropriately to English language)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(Most of the scoring was yes or probably yes. There were few limits except English language. The PICO was defined but a lot of different study designs were included, whereas it may have been better limited or prioritised)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Yes *(4 databases were searched - Medline (Ovid), Web of Science, Scopus and CINAHL. Two searches done for extrinsic and intrinsic – different databases for each. For the extrinsic factors there is no Embase but CINAHL; Scopus and Web of Science to compensate. There is no trial*

Section	Question	Answer
		registry sources, but this is an association question rather than an intervention one, so they are less relevant in this context)
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably yes *(No results from non-database sources included in PRISMA flow chart or mentioned elsewhere). However, the general methods state that relevant reviews/meta-analysis studies identified by the search strategy were also searched for identification of additional studies, but this is not specified in the review).*
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 *(Separate searches for intrinsic and extrinsic factors pre-suppose that the factors identified in the protocol (diet, exercise, hormone levels etc) are determinants of weight gain. There could be others left out (GC to answer). Another approach could be to look for all association studies on weight gain in women with PCOS without pre-specifying possible mechanisms, or to take a hybrid of the two approaches. The search strategy for the extrinsic factors (p2368) looks broadly internally consistent but lack of line references on search strategy makes a proper appraisal impossible. No details provided as to how this has been adapted for other databases either).*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(No date limits and publication type limits are appropriate).*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(2 reviewers sifted the studies based on the title and abstracts alongside the reviewing team)*.
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	*Unclear* *(Insufficient information to conclude whether searches have been run correctly. Search strategy pre-supposes that factors defined a priori are the determinants of excess weight gain in women with PCOS).*

Section	Question	Answer
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(2 reviewers selected and appraised studies alongside the 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 characteristics table appears to be sufficiently detailed).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Unclear *(It appears that through initial searching 64 eligible studies were identified. However, PRISMA indicated that total 70 were included out of which 65 were included for evidence synthesis).*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Probably no *(Table for quality appraisal of included studies shows individual studies risk of bias, rather than full extraction tables like other reviews. The same criterion is used for case-control, cross-sectional and RCT studies. Confounding is only mentioned as whether groups are similar at baseline).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Probably yes *(Says studies were selected and appraised by 2 reviewers using study selection and appraisal criteria (PICO above) established a priori. Unclear whether they were critically appraising the studies).*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Unclear *(Most scoring was probably yes or yes, but probably no for appraisal of risk of bias and there was an inconsistency in the PRISMA and how may studies included and why some studies were not meta-analysed. It was unclear if 2 reviewers critically appraised the studies).*
Synthesis and findings	Did the synthesis include all studies that it should?	Unclear *(Some studies were not included in the synthesis, such as 40 with total energy intake were included in review but only 32 were suitable for meta-analysis).*

Section	Question	Answer
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	Probably yes *(Protocol detailed but no information on if protocol published).*
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 *(Studies were grouped based on the outcomes. There is only one comparison - PCOS vs control. Only observational studies were meta-analysed)*.
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Probably yes *(There was a lot of heterogeneity across all outcomes (I^2 over 90% for most), random effects model is used. GRADE is downgraded twice for inconsistency).*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Yes *(Funnel plots are given where data was meta-analysed).*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Unclear There were no extractions but there is a table showing risk of bias judgements but there should be extractions for all studies. The GRADE table reports many outcomes as being of serious rather than very serious risk of bias, which is unusual given that a lot of case-control studies were included and even a couple studies that did not have a study type recorded and one had no demographic information, dietary and physical tools reported but was rated low risk of bias.
Synthesis and findings	Concerns regarding the synthesis and findings	Unclear *(Most scored yes or probably yes but. It is not clear how some studies were graded as low risk of bias and why some studies results were not included in the meta-analysis)*
Judging risk of bias	Concerns regarding specification of study eligibility criteria	Low *(Most of the scoring was yes or probably yes. There were few limits except English language. The PICO was clearly defined. A lot of different study designs were included but no distinction made between them).*

Section	Question	Answer
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	*Unclear* *(Insufficient information to conclude whether searches have been run correctly. Search strategy pre-supposes that factors defined a priori are the determinants of excess weight gain in women with PCOS).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Unclear *(Most scoring was probably yes or yes, but probably no for appraisal of risk of bias and there was an inconsistency in the PRISMA and how may studies included and why some studies were not meta-analysed. It was unclear if 2 reviewers critically appraised the studies)*.
Judging risk of bias	Concerns regarding the synthesis and findings	Unclear *(Most scored yes or probably yes but. It is not clear how some studies were graded as low risk of bias and why some studies results were not included in the meta-analysis).*
Overall review ratings	Overall risk of bias	*Yes, most of the concerns were unclear however the overall GRADE rating for most of the evidence was very low and the recommendation was that there was no conclusive evidence*
Overall review ratings	Applicability as a source of data	Fully applicable

Critical Appraisal - ROBIS systematic review checklist - intrinsic factors

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Probably yes *(PICO is clearly defined. Few details for exclusion criteria).*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Probably yes *(There were lots of different types of study design included, case-control, cross-sectional, cohort study and RCTs were all included however the question is about risk factors and may have been better answered by a cohort study with confounders addressed)*.

Section	Question	Answer
Study eligibility criteria	Were eligibility criteria unambiguous?	Probably yes *(There was no interventions specified in the protocol, but randomised controlled trials were included in addition to non-randomised)*.
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 date limits were applied. It was a new review).*
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 *(Inclusion was limited appropriately to English language).*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(Most of the scoring was yes or probably yes. There were few limits except English language. The PICO was clearly defined).*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Yes *(5 databases were searched - Medline (Ovid), Embase PsycINFO, AMED and CINAHL. Two searches done for extrinsic and intrinsic – different databases for each. There is no trial registry sources, but this is an association question rather than an intervention so they are less relevant in this context)*.
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably yes *(No results from non-database sources included in PRISMA flow chart or mentioned elsewhere in review except for the general methods section of the technical report says that GDG members were consulted, bibliographies of relevant reviews/meta-analysis studies identified by the search strategy were also searched for identification of additional studies, but this is not specified in the review).*
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 *(Separate searches for intrinsic and extrinsic factors pre-suppose that the factors identified in the protocol (diet,*

Section	Question	Answer
		exercise, hormone levels etc) are determinants of weight gain. There could be others left out (GC to answer). Another approach could be to look for all association studies on weight gain in women with PCOS without pre-specifying possible mechanisms, or to take a hybrid of the two approaches. The search strategy for the intrinsic factors (p2432) has line references but it is not clear what database this is for, or how it was translated for other sources).
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(No date limits and publication type limits are appropriate).*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(2 reviewers sifted the studies based on the title and abstracts alongside the reviewing team).*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	*Unclear* *(Insufficient information to conclude whether searches have been run correctly. Search strategy pre-supposes that factors defined a priori are the determinants of excess weight gain in women with PCOS. GC to state whether appropriate).*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(2 reviewers selected and appraised studies alongside the 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 characteristics table appears to be sufficiently detailed).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Probably yes *(PRISMA indicated that total 31 'original papers' were included out of which 30 were included for evidence synthesis).*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Probably no. *(Table for quality appraisal of included studies shows individual studies risk of bias, rather than full extraction tables like other reviews. The same criterion is used for case-control, cross-sectional and RCT*

Section	Question	Answer
		studies. Confounding is only mentioned as whether groups are similar at baseline).
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Probably yes *(Says studies were selected and appraised by 2 reviewers using study selection and appraisal criteria (PICO above) established a priori. Unclear whether they were critically appraising the studies).*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Unclear *(Most scoring was 'probably yes' or 'yes', except for a 'probably no'. There was an inconsistency in the PRISMA and how may studies included and it was unclear if 2 reviewers critically appraised the studies. There were little details of confounding and risk of bias appraisal was similar for all study types).*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(included all 30 studies in the synthesis based on the outcomes reported).*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	Probably yes *(Protocol detailed but no information on if protocol published).*
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 *(Studies were grouped based on the outcomes. There is only one comparison - PCOS vs control. Only one outcome was meta-analysed and only 5 of the 9 studies were included, they state why the other 4 were no included, one being that they reported outcomes in median (IQR), which differs from what was done in the extrinsic factor review where median was converted to mean).*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Probably yes *(Only one outcome was meta-analysed. The meta-analysis showed a lot of heterogeneity with I^2 value (98%). A random effects model is used. It seems like the CIs of all studies overlapped except one study, which looks like an outlier. GRADE was downgraded accordingly).*

Section	Question	Answer
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Yes *(Funnel plots are given where data was meta-analysed).*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Unclear *(There were no extractions but there is a table showing risk of bias judgements but there should be extractions for all studies, as in other reviews).*
Synthesis and findings	Concerns regarding the synthesis and findings	Unclear *(Most scored yes or probably yes but only one outcome is meta-analysed, and it has high I^2, but this has been downgraded in GRADE accordingly and random effects used. There was an unclear score as it is unclear how biases are addressed in synthesis)*
Overall review ratings	Overall risk of bias	Low
Judging risk of bias	Concerns regarding specification of study eligibility criteria	*(Most of the scoring was yes or probably yes. There were few limits except English language. The PICO was clearly defined)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	*Unclear* *(Insufficient information to conclude whether searches have been run correctly. Search strategy pre-supposes that factors defined a priori are the determinants of excess weight gain in women with PCOS. GC to state whether appropriate).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Unclear *(Most scoring was probably yes or yes, except for a 'probably no'. There was an inconsistency in the PRISMA and how may studies included and it was unclear if 2 reviewers critically appraised the studies. There were little details of confounding and risk of bias appraisal was similar for all study types).*
Judging risk of bias	Concerns regarding the synthesis and findings	Unclear *(Most scored yes or probably yes but only one outcome is meta-analysed, and it has high I^2, but this has been downgraded in GRADE accordingly and random effects*

Section	Question	Answer
		used. There was a 'probably no' as biases do not seem to be addressed in synthesis).
Overall review ratings	Overall risk of bias	*Yes, the main concern was regarding the appraising of risk of bias, particularly confounding factors and for different study types, however all the outcomes were downgraded in GRADE to 'very low overall' and the recommendation was that there is no clear evidence*
Overall review ratings	Applicability as a source of data	*Fully applicable*

IG evidence to recommendations justification: Women with PCOS have reported higher rates of weight gain and excess weight compared to women who do not have PCOS. The EBR found no clear evidence of physiological or behavioural lifestyle differences, related to weight in women with PCOS compared to those without. The review was split into extrinsic factors and intrinsic factors. Extrinsic factors were those that were potentially related to challenges with weight management (i.e lifestyle factors) such as dietary intake, and physical activity; intrinsic factors were those that were potentially related to challenges with weight management (ie energy homeostasis) such as energy intake and energy expenditure. Sixty-five studies were found for the review on extrinsic factors and 30 for the review on intrinsic factors. There was evidence that women with PCOS have differences in lifestyle behaviours (fat intake, physical activity) and some difference in appetite suppressing gut hormones, but this was of low quality. This justifies the EBRs made.

IG economic evidence

No health economic evidence was included in the IG relating to review question 3.5 on increased risk of weight gain.

The IG noted that their recommendations may require longer consultation times or increased referrals to allied health, resulting in additional costs, time and resource requirements. It was also noted that in some settings referral to community-based

lifestyle management programs may be possible, however no further information was provided.

The IG acknowledged that dietary intake and physical activity can differ depending on socio-economic status and therefore access to appropriate healthcare services with sufficient time for assessment and consideration of barriers to weight management should be tailored to reflect these differences.

3.5.2 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 model was developed for review question 3.5 from the IG as the committee agreed that other areas of the guideline were of a higher priority for original health economic analysis. The committee also acknowledged that this review question is concerned with the clinical features of PCOS and therefore health economic modelling would have limited benefit for this review question.

3.5.3 NICE recommendations

The relevant recommendation for this section is Rec 1.1.6.

3.5.4 The committee's discussion and interpretation of the evidence

Clinical evidence

The IG made two recommendations associated with review question 3.5 on factors affecting weight gain in PCOS (3.5.1 & 3.5.2). There was no overlap between section 3.5 of the IG and any existing NICE guidance. Therefore, when contextualising the IG recommendations, the committee considered the ROBIS assessments, the evidence underpinning the recommendations and the content of the recommendations from the IG. After consideration of the above, the committee decided to contextualise the practice point recommendation 3.5.2. This recommendation highlights: the challenges people with PCOS can experience with weight management, the importance of a healthy lifestyle, and the need for goals to be realistic and tailored.

Although recommendation 3.5.1 from the IG is an EBR, the committee concluded that they did not want to contextualise this recommendation. The committee noted that the IG's review question asked the question as to why women with PCOS are at increased risk of weight gain. This review question therefore assumed that women with PCOS have an increased risk of weight gain, however the IG did not present evidence of this. The committee also noted that the IG's review showed a lack of consistent evidence and noted the IG concluded that the mechanisms for increased weight gain are unclear. Subsequently, given the assumptions made, and the lack of a clear explanation of the specific mechanisms underpinning their assumptions, the NICE committee decided not to contextualise this recommendation. The NICE committee concluded that the most important message to convey in the guideline is that obesity and being overweight are common in people with PCOS. This was therefore incorporated into the contextualised practice point recommendation to form a single awareness-raising recommendation.

Health economic

No health economic evidence was included in the IG for evidence review 3.5 on weight gain. In addition, no health economic evidence was identified in the health economic literature search conducted by NICE.

The guideline committee contextualised the practice point recommendation associated with this review question to explain that a healthy diet and physical activity can help with preventing and managing excess weight but can also have other health benefits regardless of BMI. The committee acknowledged that this recommendation is concerned with providing information and noted that this will likely be provided at existing healthcare contacts, for example upon a diagnosis of PCOS. No significant resource implications are therefore associated with review question.

3.6 Weight stigma

Review question 3.6: What is the burden of weight stigma in PCOS?

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

Evidence-based recommendation:

3.6.1 Many women with PCOS experience weight stigma in healthcare and other settings and the negative biopsychosocial impacts of this should be recognised.

Consensus recommendation:

3.6.2 Health professionals should be aware of their weight biases and the impact this has on their professional practice and on women with PCOS.

3.6.3 Health policy makers, managers and educators should promote awareness of weight stigma and invest in weight stigma education and minimisation strategies.

Practice points(s):

3.6.4 Health professionals should be aware of weight-inclusive practices which promote acceptance of and respect for body size diversity and focus on improvement of health behaviours and health outcomes for people of all sizes. In PCOS this includes:

- acknowledging that whilst higher weight is a risk factor for PCOS and its complications, it is only one indicator of health and broader factors should be assessed
- asking permission to discuss and measure weight and using strategies to minimise discomfort (e.g. blind weighing)
- recognising that the terms “overweight” and “obese/obesity” can be stigmatising with suggested alternatives including “higher weight”

- if weighing, explaining how weight information will be used to inform risks, prevention and treatment and how not knowing may impact on recommendations
- ensuring appropriate equipment are available for women of all sizes
- offering options of weight-centric care (promoting intentional weight loss) or weight-inclusive care (promoting healthy lifestyle change without intentional weight loss) tailored to women's goals and preferences
- offering all women best-practice assessment, treatment and support regardless of weight, acknowledging that weight may be a non-modifiable risk factor when using lifestyle modification alone.

3.6.5 Increasing awareness of weight stigma among family members of women and adolescents with PCOS should be considered.

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?	Probably yes *(This is an updated systematic review, however there are no date limits. The protocol has not a lot of detail)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Probably yes *(This was a qualitative review, but it leaves the study type open as any qualitative or quantitative, however it does not state that it would be a cross-sectional study for the quantitative study type)*.
Study eligibility criteria	Were eligibility criteria unambiguous?	Probably no *(The question 'What is the burden of weight stigma?' is very open, would be better to be specific what type of burden. There is also a clinical practice point: how do we alleviate weight stigma in PCOS in and outside*

Section	Question	Answer
		healthcare settings? It is not clear how this will be researched in the review).
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 limits on the date).*
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 *(Inclusion is appropriately limited to English language).*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(The eligibility criteria was rated 'yes' or 'probably yes' for most of the questions. There are not a lot of details, but few exclusions or limits made. They include qualitative or quantitative study types but does not explain what type of quantitative studies were included)*.
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Yes *(Six databases were searched. Adequate list of sources. Trial registries less of an issue for this type of question)*.
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Yes *(8 references identified through "expert data source". The methods section of the technical report says that GDG members were consulted, bibliographies of relevant reviews/meta-analysis studies identified by the search strategy were also searched for identification of additional studies, but this is not specified in the review).*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(Minimal use of subject headings including leaving out MesH term weight prejudice and Emtree term weight bias, both of which retrieve additional, potentially relevant material. Focus very much on terminology about "weight stigma" but the same concept is also referred to as obesity stigma, fat*

Section	Question	Answer
		phobia/shame etc – these terms are not included in the searches).
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably *no* *(Exclusion of dissertations carried out as part of sifting. This is per the exclusion criteria but there is no rationale for why this type of document is excluded).*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(Studies were selected by one reviewer alongside the reviewing team. Although it is one reviewer it says that it was done 'in consultation with the evidence team').*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(6 databases were searched but studies were selected by one reviewer alongside the reviewing team. The search strategy is missing key terminology).*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Probably no *(Data is collected by one reviewer alongside the 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 *(There were study characteristics available).*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Probably yes *(The results were collected which were relevant to the review however there was no details provided about the themes within the studies).*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Probably yes *(GRADE CerQual was used to assess the quality of the outcomes).*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Probably no *(Data is appraised by one reviewer alongside the evidence team)*.
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Unclear *(Most of the questions were answered with probably yes or probably no. Study characteristics, risk of bias and CerQual-*

Section	Question	Answer
		GRADE are provided, however only one reviewer collected data and appraised data. Also, there could be more details provided regarding the themes within the original studies and the study extractions were more relating to other type of studies than qualitative studies).
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(It included all 7 studies that were found).*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(Details on how themes were generated is not given, for example, how many reviewers helped in generating themes).*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(Qualitative studies were included and themes were identified).*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Probably yes *(Coherence is provided in the GRADE-CerQual grading as no/minor concerns but no footnotes).*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Not applicable *(This is not relevant to a qualitative review).*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Yes *(GRADE CerQual was conducted but there were no footnotes for GRADE-CerQual table so can't tell what they were downgraded for. Studies in various countries Iran, Canada. The risk of bias for each study was not qualitative based).*
Synthesis and findings	Concerns regarding the synthesis and findings	Unclear *(The questions were varied from yes, probably yes, to No information to not applicable. GRADE CerQual was conducted but there were no footnotes for GRADE-CerQual table so can't tell what they were downgraded for. Studies in various countries Iran, Canada so relevance to UK may need to be considered. The risk of bias for each study was not qualitative based. Details on*

Section	Question	Answer
		how themes were generated is not provided).
Judging risk of bias	Concerns regarding specification of study eligibility criteria	Low *(The eligibility criteria was rated 'yes' or ''probably yes for most of the questions. There are not a lot of details, but few exclusions or limits made. They include qualitative or quantitative study types but should explain what type of quantitative studies as it is in a PICO format for intervention reviews rather than qualitative).*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	Low *(6 databases were search but there was no information about additional methods used for searching and studies were selected by one reviewer alongside the reviewing team).*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Unclear *(Most of the questions were answered with probably yes or probably no. Study characteristics, risk of bias and GRADE-CerQual are provided, however only one reviewer collected data and appraised data. Also, there could be more details on the themes within the original studies, and the study extractions related more to other study types than qualitative studies).*
Judging risk of bias	Concerns regarding the synthesis and findings	Unclear *(The questions were varied from 'yes', 'probably yes', to 'No information' to 'not applicable'. GRADE-CerQual was conducted but there were no footnotes for GRADE-CerQual table so cannot tell what they were downgraded for. Studies in various countries Iran, Canada so relevance to UK may need to be considered. The risk of bias for each study was not qualitative based. Details on how themes were generated is not provided).*
Overall review ratings	Overall risk of bias	No, most of the domains were rated low or unclear. The reasons for unclear for the 2 domains were not addressed in the interpretation of findings.
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: the IG conducted a qualitative study which included 7 studies. The IG noted that there was a lack of studies that report instances of weight stigma in women with PCOS. They therefore felt that the findings were only preliminary for women with PCOS, but were consistent in other populations (community, clinical, student) as pervasive, detrimental to mental health, affected by societal attitudes and perpetuated every day. The evidence does not justify the EBR that many women with PCOS experience weight stigma in healthcare in other settings, as there was not enough evidence. They state that the findings were consistent with other populations, but they did not analyse the evidence from other populations within this review, however they provide a reference for one study. The second recommendation is consistent with the limited evidence that found some healthcare professionals were sources of weight stigma. The IG stated that weight stigma is pervasive and entrenched and therefore, many healthcare professionals are unaware of their own weight-stigmatising beliefs and behaviours. The IG also noted that even healthcare professionals who are aware of weight stigma and its impacts may find it difficult to provide weight-neutral care. As a result of the lack of evidence a CR was made that health professionals should be aware of their weight biases and the impact of this.

IG economic evidence

No health economic evidence was included in the IG relating to review question 3.6 on the burden of weight stigma.

The IG concluded that practicing weight-inclusive care will likely require increasing consultation time, training, education and may also increase allied healthcare referrals to assist with weight-inclusive interventions.

3.6.2 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 model was developed for review question 3.6 on weight stigma as the committee concluded that there would be limited health economic implications associated with the recommendations made for this review question.

3.6.3 NICE recommendations

The relevant recommendations for this section are Rec 1.1.2 and 1.1.3.

3.6.4 The committee's discussion and interpretation of the evidence

Clinical evidence

The committee's contextualisation of the IG recommendations on weight stigma was informed primarily by a comparison of the IG's recommendations with those in NICE's guideline on the prevention and management of obesity and overweight (NG246), which, with the exception of pregnant women, applies to all women with PCOS. The committee noted that there was overlap between the two, however, concluded that it was important the NICE guideline should include recommendations on weight stigma from a PCOS perspective. The committee therefore contextualised the one EBR associated with this review question (3.6.1) to raise awareness of the

negative experience many women with PCOS experience in terms of weight stigma and the negative consequences this can have.

The committee noted that some women with PCOS may be reluctant to be weighed in a healthcare setting and that this reluctance can have consequences for their further care, including decisions about management and access to certain services, e.g. fertility services. For this reason, the committee made a CR detailing the benefits of weight measurement for those with PCOS who are reluctant to be weighed.

The committee noted that the NICE guideline on overweight and obesity contained further useful guidance connected with weight stigma and chose to cross-refer to that guidance. In particular, cross-referring to the recommendations on non-stigmatising messaging, awareness, training, planning and commissioning services and interventions for all ages.

The committee decided not to contextualise the other recommendations in the IG – those being the two CRs (3.6.2 & 3.6.3) and one practice point recommendations (3.6.5). Part of practice point 3.6.4 was adapted including asking permission to discuss and measure weight and if weighing and explaining how weight information will be used to inform risks prevention and treatment. Overall, concluding that the cross-referral to the existing NICE guideline (NG246) and adoption of the EBR would provide healthcare practitioners sufficient information on weight stigma in PCOS.

Health economic

No health economic evidence was included in the IG for review question 3.6 on the burden of weight stigma. In addition, no health economic evidence was identified in the NICE health economic literature search.

The committee contextualised the one EBR from the IG (3.6.1) and one of the practice point recommendations (3.6.4). The committee noted that these recommendations are concerned with providing information on how weight and weight stigma should be managed for those with PCOS and therefore no resource

implications are associated with these recommendations as they do not involve a choice between competing alternatives.

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

28 10 and 27

29 letter.pt. or letter/

30 note.pt.

1 31 editorial.pt.

2 32 case report/ or case study/

3 33 (letter or comment*).ti.

4 34 or/29-33

5 35 randomized controlled trial/ or random*.ti,ab.

6 36 34 not 35

7 37 animals/

8 38 exp Animals, Laboratory/

9 39 exp Animal Experimentation/

10 40 exp Models, Animal/

11 41 exp Rodentia/

12 42 (rat or rats or mouse or mice or rodent*).ti.

13 43 or/37-42

14 44 43 not humans/

15 45 36 or 44

16 46 28 not 45

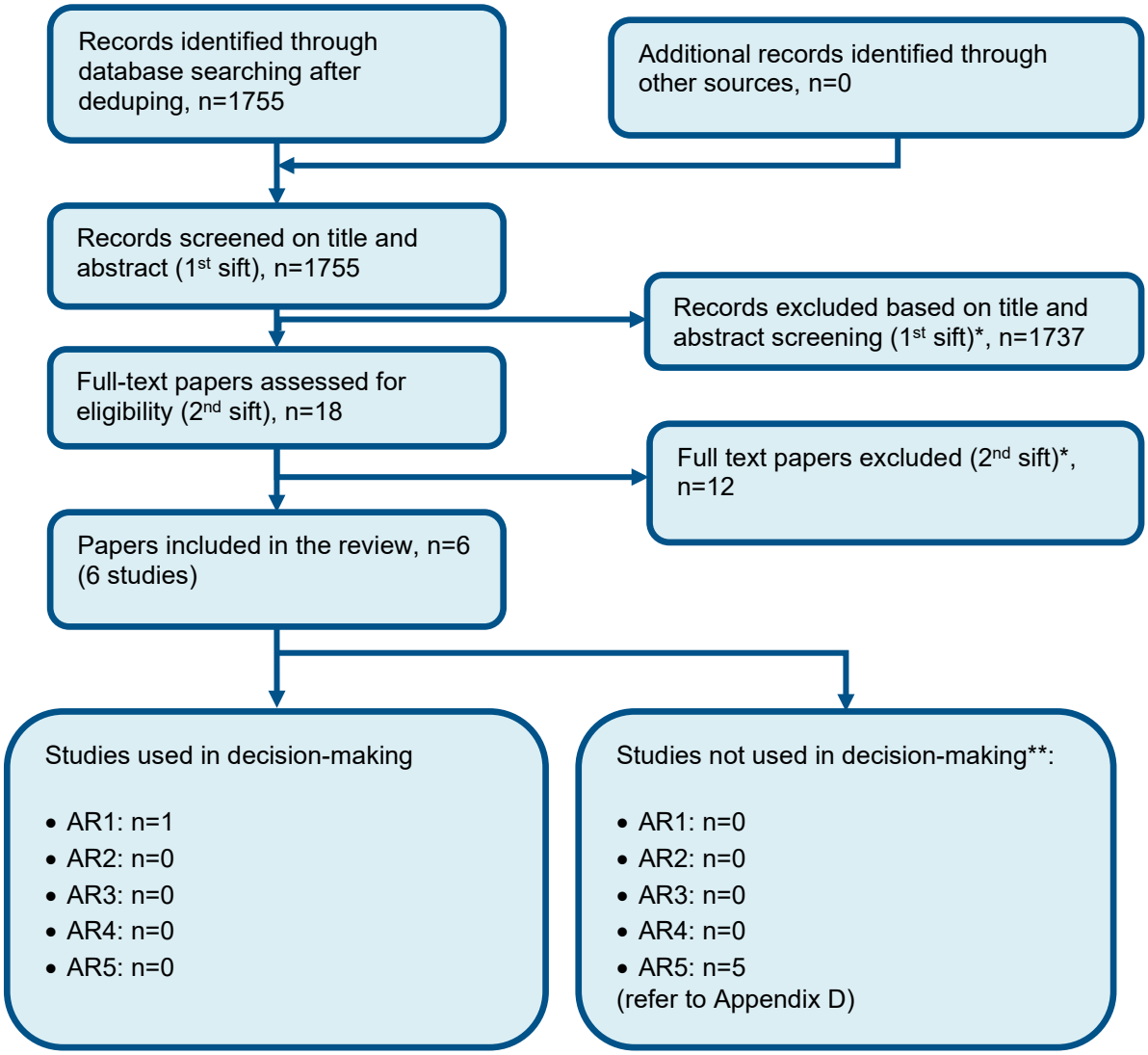
17 47 limit 46 to english language/

18 48 limit 47 to ed=20090101-20241

19 49 limit 47 to dt=20090101-20241205

20 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

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