National Institute for Health and Care Excellence

Polycystic ovary syndrome

Adaptation report 5 – Assessment and treatment of infertility

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 All NICE guidance is subject to regular review and may be updated or withdrawn.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

5 Assessment and treatment of infertility

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.

5.1 Preconception risk factors

Review question 5.1: In women with PCOS with infertility, what are the preconception risk factors associated with poor or negative fertility outcomes?

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

Evidence-based recommendation:

5.1.1 Women with PCOS should be counselled on the adverse impact of excess weight on clinical pregnancy, miscarriage and live birth rates, following infertility treatment.

Consensus recommendation:

5.1.2 Consistent with routine preconception care, in women with PCOS planning pregnancy, weight, blood pressure, smoking, alcohol, diet and nutritional status, folate supplementation (higher dose in those with BMI > 30), exercise, sleep and mental, emotional and sexual health should be considered and optimised to improve reproductive and pregnancy outcomes and overall health.

Practice points:

5.1.3 A reproductive life plan and age appropriate education on optimising reproductive health, is recommended in adolescents and women with PCOS, including healthy lifestyle, prevention of excess weight gain, and optimising preconception risk factors.

5.1.4 Healthcare professionals are encouraged to seek permission and if given, to assess weight and body mass index and initiate a dialogue on the importance of weight and lifestyle on women's health before pregnancy. This requires caution to avoid weight stigma and needs to consider the cultural, social and environmental determinants of health (see 3.6).

5.1.5 Chronic conditions such as diabetes, high blood pressure, anxiety, depression and other mental health conditions, should be optimally managed and women should be counselled regarding the risk of adverse pregnancy outcomes.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Criteria well described, studies appear to meet this)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Detailed PICO available which is suitable for the review)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(Restrictions on date are present but not well described)*
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 *(Language and study types selected are appropriate for review question)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All areas appear to be well covered with relevant information and appropriate limits)*

Section	Question	Answer
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(Reporting is very unclear, in the search details they only mention Medline and Embase, but the search strategy report suggests that in addition CINAHL and EBM (including the Cochrane Database of SRs) was included in the search. No unpublished sources mentioned)*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	No information *(No information is given about additional methods such as manual searches. PRISMA diagram shows 124 references from other sources, but no explanation is given as to what these are)*
Identification and selection of sources	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(There are no terms in the strategy for preconception. The subject headings have not been translated for different databases.* *The exclusion of certain study types is unclear and some reported CINAHL terms cannot be found(possibly not translated properly from Medline)* *An example of the lack of confidence in the search robustness is the larger number of hits for Mesh (Medline thesaurus) terms in CINAHL which has different thesaurus than in all the other databases searched put together)*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably no *(Not explicitly stated but appears to have been appropriate. This is a new SR, and the protocol does not mention date limits, but they have limited to after 2000)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(Study selection was completed by five reviewers)*

Section	Question	Answer
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(The search strategy is flawed, lack of preconception terms, no translation of subject headings between different databases, confusing use of the BOOLEAN operator "NOT" when filtering by study type)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(Three reviewers at title and abstract should be sufficient)*
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 table available for 56/65 of the studies for interpretation of results)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Unclear *(65 studies included in review, 26 included in descriptive analysis, 23 in meta-analysis, 26 in GRADE evidence tables, no discussion as to why some studies were not included in GRADE or meta-analysis)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(Risk of bias assessments completed for each outcome detailed table with risk of bias assessments for each individual study)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Yes *(3 reviewers completed the assessment)*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review, more details required for why some studies not included in meta-analysis)*
Synthesis and findings	Did the synthesis include all studies that it should?	Unclear

Section	Question	Answer
		(Some of the studies do not appear in the study characteristics table, and there is no written explanation for this)
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(No information was given regarding the planned analysis of the results)*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(Meta-analyses and descriptive analyses were used appropriately for this review question, study designs were appropriate for analyses)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Yes *(For comparison 1 weight comparisons serious inconsistency was found for pregnancy rate per patient, clinical pregnancy rate per patient, live birth rate per patient, OHSS and clinical pregnancy rate per patient (younger vs older PCOS age groups).* *Subgroup analyses were conducted for the outcomes that had showed high heterogeneity, which did not resolve the heterogeneity but was downgraded once for inconsistency in GRADE)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Yes *(Funnel plots present but lack of descriptive summary)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Probably yes *(Risk of bias was downgraded in GRADE)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(There were no pre-defined subgroups in the protocol, however inconsistency was addressed through subgroup analyses and downgrading in GRADE)*

Section	Question	Answer
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(All areas appear to be well covered with relevant information and appropriate limits)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(The search strategy is flawed, lack of preconception terms, no translation of subject headings between different databases, confusing use of the BOOLEAN operator "NOT" when filtering by study type)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review, more details required for why some studies not included in meta-analysis)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(There were no pre-defined subgroups in the protocol, however inconsistency was addressed through subgroup analyses and downgrading in GRADE)*
Overall review ratings	Overall risk of bias	Low *(Low grading for each section except for identification and searching of studies, no major issues noted)*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: twenty studies were included in this new systematic review. These studies reported on preconception BMI and/or obesity as potential risk factors for fertility outcomes for people with PCOS. Studies generally had large sample sizes with a low or moderate risk of bias, although the BMI cut offs used varied between studies. Nine different outcomes were assessed for this question, for which the IG determined a BMI range of 18.5 – 24.9 kg/m^2 were assigned to the lean/normal weight group, with a BMI of 25 kg/m^2 or higher assigned

to the overweight/obese group. The IG describes that people in the lean/normal weight group had higher pregnancy rates, higher live birth rates, higher ovulation rates and lower miscarriage rates than the overweight/obese group. The outcomes of clinical pregnancy rate per patient and ovarian hyperstimulation syndrome (OHSS) in lean/normal versus overweight/obese patients showed no difference between groups. A comparison of four outcomes in younger age versus older age PCOS patients found no difference between groups for 5 pregnancy-related outcomes. However, the authors note that certainty was very low for most outcomes due to the observational nature of the included studies. Most outcomes assessed also had a serious risk of bias. One EBR was made from this evidence which was a strong recommendation for the option. The recommendation states that women with PCOS should be counselled on the adverse impact of excess weight on clinical pregnancy, miscarriage and live birth rates following infertility treatment, this is consistent with the evidence found, which showed lean women with PCOS had higher pregnancy rates, live birth rates and lower miscarriage rates. The clinical pregnancy rate showed no difference between the overweight and lean groups for women with PCOS. However, the IG panel discussed that the evidence for the impact of excess weight on clinical pregnancy in the general population was strong, and the benefits of counselling to women with PCOS would be similar. Concerns around weight stigma were raised however it appears that the potential benefits of counselling women about the impact of excess weight outweighed this concern. The recommendations in this section took into account the available evidence from the new systematic review, alongside published guidelines from the WHO and FIGO, it is also highlighted that preconception risk factors is an area that would benefit from further research.

IG economic evidence

No health economic evidence was identified in the IG for review question 5.1 on preconception risk factors.

5.1.2 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 for this review question.

Economic model

Review question 5.1 was not prioritised for original health economic modelling as this review question was not concerned with comparing alternative courses of action, but instead determining preconception risk factors. Original health economic modelling was therefore not deemed appropriate for this review question.

5.1.3 NICE recommendations

The relevant recommendations for this section are Rec 1.20.1, 1.20.2 and 1.21.1 to 1.21.4.

5.1.4 The committee's discussion and interpretation of the evidence

Clinical

The committee discussed the importance of using sensitive language in discussions about overweight/obesity, avoiding weight stigma. The committee noted that the NICE guideline on overweight and obesity management has guidance on using appropriate language and how to have sensitive conversations. This area overlaps with the NICE guideline on overweight and obesity management, and as such the committee agreed that a cross referral would be appropriate. Committee members

revised the wording to better support women attempting weight loss and to acknowledge the challenges involved. The committee advised that excess fat accumulation can lead to insulin resistance and higher androgen levels, which can disrupt ovulation and reduce fertility. As such they felt the evidence for the recommendation was acceptable and should be included but with more appropriate wording. The committee decided to contextualise IG recommendation 5.1.1 (EBR) and CR 5.1.2 into two recommendations. They also decided to add a recommendation from the NICE fertility guideline which advises people that weight loss alone may restore ovulation and have a positive impact on pregnancy outcomes.

Health economic

No health economic evidence was identified in either the IG or in NICE's health economic literature search for review question 5.1 on preconception risk factors.

As the recommendations associated with this review question cross-refer to existing NICE guidance and are concerned with identifying risk factors – which does not impact on resource use – no significant resource impact is associated with this review question

5.2 Tubal patency testing

Review question 5.2: Should women with PCOS and infertility due to anovulation alone with normal semen analysis undergo tubal patency testing prior to starting ovulation induction with timed intercourse or intrauterine treatment?

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

Consensus recommendation:

5.2.1 In women with PCOS and infertility due to anovulation alone with normal semen analysis, the risks, benefits, costs and timing and techniques of tubal patency testing in relation to the cost and complexity of the treatment, should be considered on an individual basis, depending on personal history and population prevalence, prior to starting ovulation induction with timed intercourse or intrauterine insemination.

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

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Clear and detailed PICO)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Detailed PICO available which is suitable for the review)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(Restrictions on date are present but not well described)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(Study types selected and language are appropriate for review question)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All parameters are appropriate to review)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(5 named databases used, all appear appropriate to review question. Reasonable range of databases but PsycINFO seems superfluous. No other methods mentioned)*

Section	Question	Answer
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably no *(No mention of manual searches or other methods being used, no additional sources mentioned in PRISMA chart or narrative of methods)*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(Translations of subject headings not done between databases. The search result numbers seem very low and the structure very convoluted. It is PCOS terms and fallopian tube diseases and terms for patency tube tests or PCOS terms and mesh headings for fallopian tube patency tests and fallopian tube diseases with diagnosis subheading n.b. the Embase equivalents have not been used (tubal patency test and uterine tube disease and the correct Embase subheading for diagnosis). A simpler more comprehensive structure would have been PCOS terms and tubal patency tests terms. Whilst the last line of the CINAHL strategy is duplicated it appears simpler with concepts of PCOS and fallopian tube diseases and the tubal patency tests.* *The application of human limits by autoindexing is not a usual method)*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably no *(Date limits have been applied. This is not stated in the protocol)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(One reviewer in addition to evidence team- however no studies met inclusion criteria so progressed as a narrative review of evidence)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Narrow strategy and lack of translation of subject headings in the OVID Medline, All EBM, PsycINFO, EMBASE set)*

Section	Question	Answer
Data collection and study appraisal	Were efforts made to minimise error in data collection?	No information *(N/A – progressed as narrative review)*
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 *(N/A – progressed as narrative review)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	No information *(N/A – progressed as narrative review)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	No information *(N/A – progressed as narrative review)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	No information *(N/A – progressed as narrative review)*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(N/A – progressed as narrative review)*
Synthesis and findings	Did the synthesis include all studies that it should?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(N/A – progressed as narrative review)*
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 *(N/A – progressed as narrative review)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	No information *(N/A – progressed as narrative review)*

Section	Question	Answer
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(N/A – progressed as narrative review)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(All parameters are appropriate to review)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Narrow strategy and lack of translation of subject headings in the OVID Medline, All EBM, PsycINFO, EMBASE set)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(N/A – progressed as narrative review)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(N/A – progressed as narrative review)*
Overall review ratings	Overall risk of bias	Low *(Note issues with narrow search strategy and lack of translation of subject headings in the databases used)*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: no evidence was found for this updated systematic review so there were no EBRs, only one CR. The CR made suggestions regarding when to offer tubal patency testing for women with PCOS who have anovulation alone. The CR states that these decisions should be made on an individual patient basis, taking into account risks, benefits, available techniques and personal history.

IG economic evidence

No health economic evidence was identified in the IG for review question 5.2 on tubal patency testing.

5.2.2 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

One health economic study was identified for this review question but was excluded due to a combination of applicability and methodological concerns.

Economic model

Review question 5.2 from the IG on tubal patency testing was not prioritised for original health economic modelling as this review question was identified as an area which is covered by existing NICE guidance. Therefore, no original health economic work was conducted for this review question.

5.2.3 NICE recommendations

No recommendations have been contextualised from this section of the IG.

5.2.4 The committee's discussion and interpretation of the evidence

Clinical

No EBRs were made for review question 5.2 on tubal patency testing and one CR was made in the IG. The committee decided not to contextualise this CR on tubal patency testing (recommendation 5.2.1 in the IG) as the committee concluded that there was insufficient clinical evidence to justify the recommendation. The committee

did, however, cross-refer to NICE's existing guideline on fertility problems, specifically referring to the section on investigation of fertility problems and management strategies. The committee acknowledged that NICE's fertility guideline provides information on when to conduct tubal patency testing as part of investigations for fertility problems.

Health economic

No health economic evidence was identified in the IG, or in NICE's health economic literature search.

As no new recommendations were made on tubal patency testing, there were not any health economic considerations for the committee to assess.

5.3 Letrozole

Review question 5.3: In women with PCOS, are aromatase inhibitors effective for improving fertility outcomes?

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

Evidence-based recommendation:

5.3.1 Letrozole should be the first-line pharmacological treatment for ovulation induction in infertile anovulatory women with PCOS, with no other infertility factors.

Practice point(s):

5.3.2 The use of letrozole is still off label in many countries. Where it is not allowed, clinicians should use other ovulation induction agents.

5.3.3 Letrozole should not be given where there is any possibility of a pre-existing pregnancy, though there is no evidence for increased teratogenicity compared to other ovulation induction agents.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Clear and detailed PICO)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Detailed PICO available which is suitable for the review)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*

Section	Question	Answer
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(Restrictions on date are present but not well described)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(Study types selected and language used are appropriate for review question)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All areas appear to be well covered with relevant information)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(Appropriate number (5) and range of databases used. Reasonable range but PsycINFO seems superfluous.*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably no *(PRISMA flowchart mentions number of new systematic reviews, RCTs identified from those reviews, new RCTs and studies from the previous version of the guideline)*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(Full search strategy is detailed in the GDG 5 methodology appendix. Subject terms not translated between databases. This strategy covers questions addressing a range of fertility technologies. Terms for a few are included but not all. It may have been considered that using subject headings for reproductive techniques/fertility may suffice but free text terms for specific technologies would be expected to be there e.g. aromatase inhibitors. No terms at all for ovarian surgery etc.* *Use of autoindexing to identify human studies is irregular. no systematic review terms used to specify required publication types.)*

Section	Question	Answer
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(Date limits not mentioned in the protocol but it is clear that this updates a 2017 review)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(Studies screened by one reviewer with evidence team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(Study appraisal completed by 2 reviewers plus evidence team)*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(Study table with study characteristics present allowing interpretation of results)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	*Yes* *(14 studies met inclusion criteria for review, all appear to be included in analysis.13 of the studies were used in various meta-analyses for the various comparisons, 1 study was not included in the meta-analysis but was used as an individual study with a descriptive analysis given for comparison 4. One study is described in the study characteristics table as partially randomised, on further inspection, this study had 3 groups, 2 of which were randomised and 1 which was not. The international guideline has only used data from the 2 randomised groups in the meta-analysis for comparison 1, as such there is unlikely to be any impact of the study being described as only partially randomised)*
Data collection and	Was risk of bias (or methodological quality)	Yes

Section	Question	Answer
study appraisal	formally assessed using appropriate criteria?	*(Detailed integrity assessment in GDG 5 methodology appendix for all studies, individual quality appraisals also present)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Yes *(2 reviewers plus evidence team)*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(All 14 studies appear present in characteristics table and results)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(No information was given regarding the planned analysis of the results)*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(Meta-analysis was used appropriately for this review question, study design is appropriate for analysis, however odds ratios were used where typically risk ratios would be used)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Probably yes *(Serious inconsistency noted for the outcomes of multiple pregnancy rate per patient, live birth rate per patient, multiple pregnancy rate per pregnancy, miscarriage rate per patient. Miscarriage rate per patient was heterogeneous I^2=52% but no subgroup analysis was conducted. GRADE was downgraded for this outcome. Outcomes were downgraded in GRADE either once or twice depending on the severity of the inconsistency. This is clearly detailed by footnotes in each comparison where needed.* *No serious inconsistency seen for outcomes of live birth rate per patient, clinical pregnancy rate per patient, pregnancy rate*

Section	Question	Answer
		per patient, ovulation rate per patient, clinical pregnancy rate per patient, ovulation rate per cycle and miscarriage rate per patient, all other outcomes were not applicable for inconsistency due to only one study being included for that outcome. *Subgroup comparisons included BMI, naïve to medication or failed previous use of medication, ovulation rates per cycle)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Yes *(Funnel plots present for described outcomes and appear suitable)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Yes *(Risk of bias assessed per outcome in section 5.3 and per study in GDG 5 methodology appendix, appears to have taken bias into account but further write up about this could have been useful RoB included in study characteristics table and GRADE assessment also completed)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(Limited details on how conclusions reached but synthesis seems appropriate)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(All areas appear to be well covered with relevant information)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(Limited details on how conclusions reached but synthesis seems appropriate)*
Overall review ratings	Overall risk of bias	Low

Section	Question	Answer
		(Note lack of terms used and issues with translating subject terms. More written discussion of conclusions would have been helpful)
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: eleven RCTs were included in this updated systematic review, resulting in one EBR (strong recommendation for the option) and two practice points. The EBR was made following review of the evidence in the IG, a Cochrane review on aromatase inhibitors (2022) and a systematic review and NMA (2022). The evidence found letrozole to be the most suitable first line treatment, however noted that although letrozole is widely available it may be an off label use for this indication in many countries including the UK.

The IG compared letrozole to 6 other combinations of medications including clomiphene citrate, metformin and recombinant follicle stimulating hormone (rFSH) in various combinations. Six of the studies had a high risk of bias, one had a moderate risk of bias and four had a low risk of bias. Some studies included UK and USA populations.

Letrozole was superior to clomiphene citrate for 2 live birth outcomes, clinical pregnancy rate per patient, pregnancy rate per patient and two ovulation outcomes. There was no difference between letrozole and clomiphene citrate for outcomes related to multiple pregnancy rates and miscarriage rates. Participant numbers in the studies ranged from n=280 to n=1870.

For the comparison of letrozole and metformin compared to clomiphene citrate and metformin, 2 studies were included comparing ovulation rate per cycle, pregnancy rate and miscarriage rate. Full term pregnancy, ovulation rate and clinical pregnancy rate were higher with in the letrozole and metformin group. The studies had moderate and high risk of bias due to lack of participant blinding and allocation concealment issues.

For letrozole alone versus letrozole and metformin, 2 studies were included comparing live birth rate and pregnancy rate. Meta-analysis was only possible for clinical pregnancy rate, finding no significant difference between treatment options.

For letrozole versus clomiphene citrate and rFSH one study was used, finding that the letrozole group was more effective for the outcomes of clinical pregnancy and ovulation rate with no difference between treatment groups for the outcome of miscarriage rate.

For letrozole and metformin compared to clomiphene citrate, one study found clinical pregnancy rate and ovulation rate per cycle were higher in the letrozole group with no difference seen between groups for the outcomes of live birth rate per patient and miscarriage rate.

For letrozole versus clomiphene citrate, one study found no difference between groups for the outcomes of live birth rate, clinical pregnancy rate, ovulation rate or miscarriage rate between groups, however the authors state that certainty was very low for this analysis due to the study having low participant numbers and a high risk of bias.

The guideline authors stated that letrozole should be considered for first line treatment due to the studies used having a cumulative large sample size, favourable outcomes in the network meta-analysis and consistent benefits across all outcomes including live births, improved ovulation and clinical pregnancy rates. The panel stated that this showed clear definitive evidence for letrozole to be used as a first line treatment despite some noted side effects. The panel noted that a combination of side effects occurred with letrozole and clomiphene citrate. It was difficult to determine which of these side effects would be more acceptable, and as such was not a suitable reason for not recommending letrozole. However, it is noted that patient preference should be taken into account when discussing the potential adverse events. The panel also discussed that although the cost of letrozole might be higher, it was likely to result in savings in other areas due to its higher efficiency compared to other medication options. As such the IG appears to have good

evidence for recommending letrozole, however it is important to note that letrozole is off label in the UK for infertility.

The IG highlights several research priorities for this section, including how long to use letrozole to ovulation induction before moving on to another medication, validation of prediction models, combination therapies and the best treatments to offer to drug naïve and drug-resistant populations.

IG economic evidence

No health economic evidence was identified in the IG for review question 5.3 on letrozole.

On costs, the IG noted that the cost of drugs and monitoring are comparable – noting, that the greater effectiveness of letrozole would result in cost savings. However, no additional information was provided in the IG to support this statement.

The IG identified clinical evidence for the comparators listed below. The IG searched for evidence on the effectiveness of aromatase inhibitors compared to, placebo, no intervention or other infertility treatment interventions. It therefore can be inferred that the IG's statement on comparable costs and greater efficacy, relates to the evidence for the included interventions and comparators listed below.

- Letrozole vs Clomiphene Citrate
- Letrozole + Metformin vs Clomiphene Citrate + Metformin
- Letrozole vs Letrozole + Metformin
- Letrozole vs Clomiphene Citrate + rFSH vs continuous rFSH
- Letrozole + Metformin vs Clomiphene Citrate
- Letrozole vs Clomiphene Citrate + Metformin

5.3.2 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 therefore excluded for this review question.

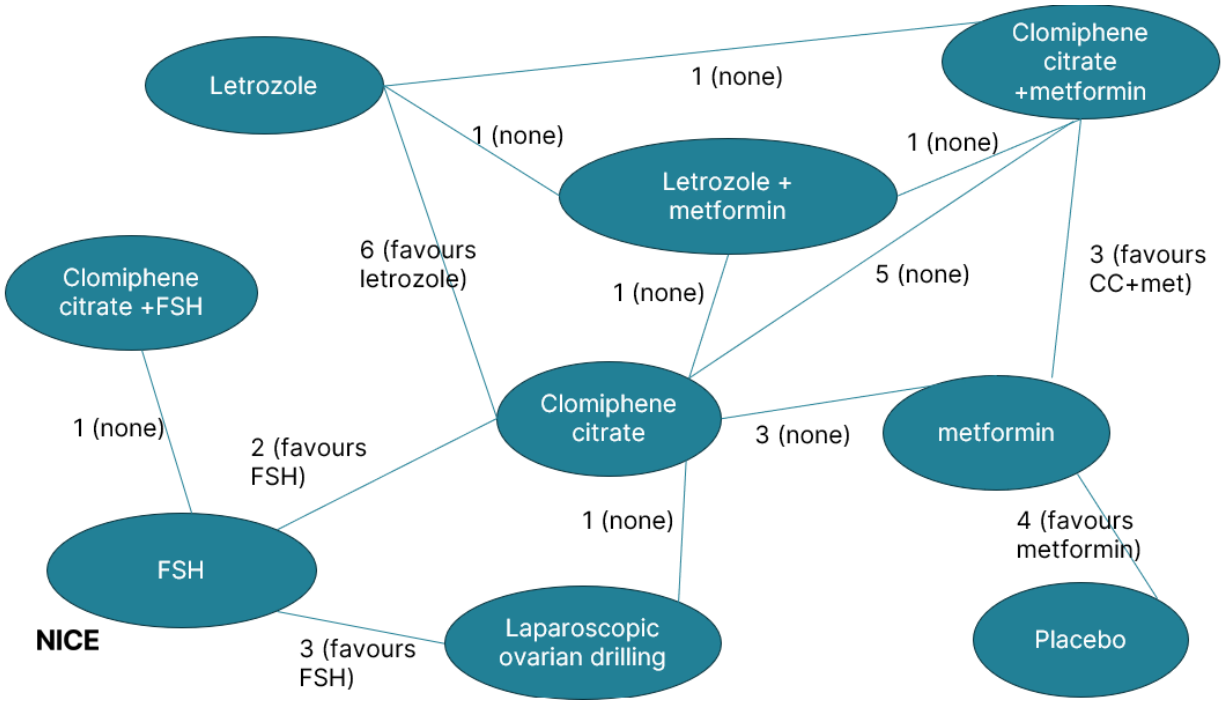
Economic model

No original health economic modelling was conducted for this review question.

Given the availability of multiple randomised controlled trials and comparators an NMA was initially considered as potentially providing the best source of evidence for this review question. In addition, substantial cost differences between ovulation induction strategies meant this topic was prioritised for original economic modelling.

The feasibility of an NMA for the topic of ovulation induction strategies for PCOS was explored by drawing the network based on the clinical studies included in the IG for the outcome of live birth per patient. A network was established for the IG's review questions 5.3 – 5.6 and therefore included the clinical evidence for ovulation induction strategies and laparoscopic ovarian drilling. The network schematic is presented in Figure 1 below.

Figure 1: Network for ovulation induction strategies and laparoscopic ovarian drilling based on the included evidence in the IG for the outcome of live birth per patient



Abbreviations: CC = clomiphene citrate; FSH = Follicle-stimulating hormone, me = metformin

In the schematic above, the numbers presented alongside the lines connecting two interventions denote the number of studies available. The description in the brackets that sits alongside this number represents the overall conclusion of the evidence for these interventions. For example, (none) represents that there was no clinically important difference between the two comparators being compared (as depicted by the ovals where the intervention is listed).

The committee noted that letrozole is cheaper than clomiphene citrate and clinically superior meaning letrozole is the dominant strategy (cheaper and more effective). The committee also noted the lack of direct clinical evidence comparing letrozole to follicle-stimulating hormone (FSH) – gonadotrophins. It was discussed that an estimate of effectiveness for FSH and letrozole could be estimated by conducting an NMA, but the committee considered that the results of this indirect estimate would likely be subject to a high a degree of uncertainty.

More generally when considering the cost-effectiveness of ovulation induction strategies this schematic (**Figure 1**) and the recommendations from the IG were presented to the guideline committee. In addition, in the context of the pre-existing assessment of the clinical evidence from the IG, the committee concluded that conducting an NMA was not required. The committee noted that an NMA, and any subsequent cost-effectiveness analysis, would unlikely alter the conclusions of the recommendations from the IG. Therefore, conducting an NMA would be a high resource intensive way to reach the same conclusions as the IG.

Unit costs

Unit costs for letrozole, clomiphene citrate and gonadotrophins were presented to the guideline committee to aid their consideration of cost-effectiveness and are presented in **Table 1** below.

Table 1: Unit costs

Resource	Unit costs per tablet / vial [(a)]	Unit costs per month based on average monthly dose
Letrozole		
Letrozole 2.5mg tablet	£0.05	£0.26[(b)]
Clomiphene citrate		
Clomiphene citrate 50mg tablet	£0.34	£1.70[(c)]
Metformin		
Metformin 500mg tablet	£0.02	£0.79[(d)]
Metformin 500mg modified-release tablet	£0.03	£2.05[(d)]
Metformin 1g tablet	£0.22	£6.56[(d)]
Metformin 1g modified-release tablet	£0.03	£0.99[(d)]
Gonadotrophins		
Choriogonadotropin alfa: Ovitrelle 250micrograms/0,5ml solution for injection pre-filled syringes	£38	£527[(e)]
Follitropin delta: Rekovelle 12micorgrams/0.36 solution	£118	£1,065.[(f)]
Follitropin delta: Rekovelle 36micorgrams/1.08 solution	£355	£3,194[(f)]
Follitropin delta: Rekovelle 72micorgrams/2.16 solution	£710	£6,389[(f)]
Follitropin alfa: Bemfola 150units/0.25ml solution for injection pre-filled pens	£52	£724[(e)]
Follitropin alfa: Bemfola 225units/0.375ml solution for injection pre-filled pens	£78	£1,086[(e)]
Follitropin alfa: Bemfola 300units/0.5ml solution for injection pre-filled pens	£103	£1,448[(e)]

Resource	Unit costs per tablet / vial [a]	Unit costs per month based on average monthly dose
Follitropin alfa: Bemfola 450units/0.75ml solution for injection pre-filled pens	£155	£2,171[e]
Follitropin alfa: Bemfola 75units/0.125ml solution for injection pre-filled pens	£26	£362[e]
Follitropin alfa: Ovaleap 300units/0.5ml solution for injection cartridges	£94	£1,316[e]
Follitropin alfa: Ovaleap 450units/0.75ml solution for injection cartridges	£141	£1,974[e]
Follitropin alfa: Ovaleap 900units/1.5ml solution for injection cartridges	£282	£3,984[e]
Follitropin alfa: Gonal-f 300units/0.48ml solution for injection pre-filled pens	£113	£1,579[e]
Follitropin alfa: Gonal-f 300units/0.72ml solution for injection pre-filled pens	£169	£2,369[e]
Follitropin alfa: Gonal-f 300units/1.44ml solution for injection pre-filled pens	£338	£4,738[e]
Menotrophin: Menopur / Meriofert 75unit powder and solven for solution injection vials	£18	£252[e]
Menotrophin: Meriofert 150unit powder and solven for solution injection vials	£56	£781[e]
Menotrophin: Menopur 600unit powder and solven for solution injection vials	£144	£2,018[e]
Menotrophin: Meriofert PFS 900unit powder and solven for solution injection vials	£335	£4,687[e]
Menotrophin: Menopur 1200unit powder and solven for solution injection vials	£288	£4,036[e]
Lutropin alfa: Luveris 75unit powder and solvent for solution for injection vials	£31	£439[e]
Urofollitropin: Fostimon 75unit powder and solvet for solution for injection vials	£28	£391[e]
Urofollitropin: Fostimon 150unit powder and solvet for solution for injection vials	£56	£781[e]
Follitropin alfa with lutropin alfa: Pregoveris 300units/150units/0.48ml solution for injection pre-filled pens	£145	£2,026[e]
Follitropin alfa with lutropin alfa: Pregoveris 450units/225units/0.72ml solution for injection pre-filled pens	£217	£3,039[e]
Follitropin alfa with lutropin alfa: Pregoveris 900units/450units/1.44ml solution for injection pre-filled pens	£434	£6,077[e]
Follitropin alfa with lutropin alfa: Pregoveris 150unit/75unit powder and solvent for solution for injection vials	£72	£1,013[e]
Cetroelix: Cetrotide 250microgram powder and solvent for solution for injection vials (adjunct to gonadotrophins)	£27	£136[g]
Ganirelix: Fyremadel / Ovamex 250micrograms/0.5ml solution for injection pre-filled syringes (adjunct to gonadotrophins)	£19	£97[g]

Resource	Unit costs per tablet / vial [a]	Unit costs per month based on average monthly dose
Laparoscopic ovarian surgery		
Laparoscopic ovarian surgery	£2,587[h]	

(a) Source of costs; British National Formulary () date accessed 29/04/26
(b) Assuming a dose of 2.5mg per day, for five days. Dose can be increased by 2.5mg increments if conception is not successful.
(c) Assuming a dose of 50mg per day, for five days. Dose can be increased by 50mg increments if conception is not successful (up to maximum dose of 150mg).
(d) Assuming a dose of one tablet per day (assuming 30.44 days per month)
(e) Assuming a dose of one injection a day for fourteen days
(f) Assuming a dose of on injection a day for nine days
(g) Assuming a dose of one injection a day for five days
(h) Source of cost; NHS Cost Collection 2024/25 (date accessed 29/04/26), currency code MA10Z, day case

5.3.3 NICE recommendations

The relevant recommendations for this section are Rec 1.21.9 and 1.21.10.

5.3.4 The committee's discussion and interpretation of the evidence

Clinical

The committee decided to contextualise EBR 5.3.1 from the IG, following discussions that this is the current first line treatment being used in practice in women with PCOS, and the evidence base supported this decision. The committee agreed to provide further dosing information for letrozole as it is currently off label for this indication. They agreed that 2.5mg starting dose for the first cycle was common practice for ovulation induction.

Health economic

No health economic evidence was identified in either the IG, or in NICE's health economic literature search, for review question 5.3 on letrozole.

The guideline committee contextualised the IG's EBR (5.3.1), to offer letrozole as a first-line ovulation induction therapy to women with PCOS who have anovulatory infertility and no other infertility factors. Unit costs for letrozole and other potential first line-treatment options (for example, clomiphene citrate and gonadotrophins) were

presented to the committee. These costs were discussed alongside the clinical evidence, of which the highest quantity and quality of clinical evidence was for letrozole versus clomiphene citrate. Overall, the evidence showed clinical superiority for letrozole compared to clomiphene citrate. Therefore, the committee concluded that letrozole was dominant (less costly and more effective) compared to clomiphene citrate. The committee also acknowledged that offering letrozole as a first-line treatment option for people with PCOS is reflective of current UK practice, and therefore no significant resource impact is anticipated for this recommendation.

As part of their recommendations, the committee also added an additional recommendation, reflective of current practice, which provided information on the dosing for letrozole.

In instances where letrozole is contraindicated or not tolerated, the committee made a recommendation to offer either clomiphene citrate and metformin or clomiphene citrate alone. This recommendation was made based on the costs and clinical evidence presented to the committee and further discussion on clomiphene citrate and metformin (in combination and alone) can be found in the committee's discussion of the evidence in section 5.4.4.As this recommendation is also reflective of current practice, there are no anticipated significant resource impact implications associated with this recommendation.

5.4 Clomiphene citrate and metformin

Review question 5.4: In women with PCOS, is clomiphene citrate effective for improving fertility outcomes? In women with PCOS, is metformin effective for improving fertility outcomes?

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

Metformin versus placebo

Evidence-based recommendation:

5.4.1.1 Metformin could be used alone, in women with PCOS with anovulatory infertility and no other infertility factors, to improve clinical pregnancy and live birth rates, whilst informing women that there are more effective ovulation agents.

Practice point(s):

5.4.1.2 Women should be counselled as to potential mild gastrointestinal side-effects with metformin.

5.4.1.3 Healthcare and resource burden including monitoring, travel and costs are lower with metformin.

5.4.1.4 Consideration of age and screening for other fertility factors needs to be discussed before prescribing metformin.

Clomiphene citrate versus metformin

Evidence-based recommendation:

5.4.2.1 Clomiphene citrate could be used in preference to metformin in women with PCOS with anovulatory infertility and no other infertility factors, to improve ovulation, clinical pregnancy and live birth rates.

Practice point(s):

5.4.2.2 The risk of multiple pregnancy is increased with clomiphene citrate use (alone or in combination with metformin) and therefore clomiphene cycles may require ultrasound monitoring.

Clomiphene citrate and metformin versus clomiphene citrate alone

Evidence-based recommendation:

5.4.3.1 Clomiphene citrate combined with metformin could be used rather than clomiphene citrate alone in women with PCOS with anovulatory infertility and no other infertility factors to improve ovulation and clinical pregnancy rates.

Clomiphene citrate and metformin versus metformin alone

Evidence-based recommendation:

5.4.4.1 Clomiphene citrate combined with metformin could be used rather than metformin alone in women with PCOS with anovulatory infertility and no other infertility factors to improve live birth rates.

Practice point(s):

5.4.4.2 Monitoring of combined cycles will need to be equivalent to clomiphene citrate alone.

Clomiphene citrate versus letrozole

Evidence-based recommendation:

5.4.5.1 Letrozole should be used rather than clomiphene citrate in women with PCOS with anovulatory infertility and no other infertility factors to improve ovulation, clinical pregnancy, and live birth rates.

Practice point(s):

5.4.5.2 Current evidence demonstrates no difference in fetal abnormality rates between letrozole or clomiphene citrate ovulation induction or natural conception.

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

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Clear and detailed PICO)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Detailed PICO available which is suitable for the review)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(Restrictions on date are present but not well described)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(Study types selected and language restrictions are appropriate for review question)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All areas appear to be well covered with relevant information)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(Appropriate number (5) and range of databases used. Reasonable range but PsycINFO seems superfluous.*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably no *(PRISMA flowchart mentions number of new systematic reviews, RCTs identified from those reviews, new RCTs and studies from the previous version of the guideline)*

Section	Question	Answer
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(Full search strategy is detailed in the GDG 5 methodology appendix. Subject terms not translated between databases. This strategy covers questions addressing a range of fertility technologies. Terms for a few are included but not all. It may have been considered that using subject headings for reproductive techniques/fertility may suffice but free text terms for specific technologies would be expected to be there e.g. aromatase inhibitors. No terms at all for ovarian surgery etc)*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(Date limits not mentioned in the protocol but it is clear that this updates a 2017 review)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(Studies screened by one reviewer with evidence team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(Study appraisal completed by 2 reviewers independently with discussion to resolve any discrepancies)*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(Study table with study characteristics present allowing interpretation of results)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(34 studies of the 57 includes for section 5 were relevant to section 5.4)*
Data collection and	Was risk of bias (or methodological quality)	Yes

Section	Question	Answer
study appraisal	formally assessed using appropriate criteria?	*(Detailed integrity assessment in GDG 5 methodology appendix for all studies, individual quality appraisals also present)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Yes *(2 reviewers for analysis should have minimised error)*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(All 34 included studies for section 5.4 have been included in the study characteristics table and appear to have been included across the 16 different comparisons)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(No information was given regarding the planned analysis of the results)*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(Meta-analysis was used appropriately for this review question, study design is appropriate for analysis, however odds ratios were used where typically risk ratios would be used)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Yes *(Heterogeneity appears to have been accounted for, there are detailed tables for each of the 16 outcomes which detail inconsistency for individual outcomes highlighted where each outcome has been downgraded once or twice for inconsistency. Random effects were used. several subgroup analyses were also undertaken, mostly for BMI, however they do not state whether this resolved the inconsistency seen in the GRADE tables. Forest plots are also present for each of the 16 outcomes and all of the subgroup analyses within each outcome)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through	Yes

Section	Question	Answer
	funnel plot or sensitivity analyses?	*(Funnel plots available for each subgroup analyses, appears appropriate)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Yes *(Risk of bias assessed per outcome in section 5.4 and per study in GDG 5 methodology appendix, appears to have taken bias into account but further write up about this could have been useful. RoB included in study characteristics table and GRADE assessment also completed)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(Synthesis seems appropriate, all areas had no or low concern with heterogeneity being addressed within the outcomes and subgroups used)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(All areas appear to be well covered with relevant information)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(Synthesis seems appropriate, all areas had no or low concern with heterogeneity being addressed within the outcomes and subgroups used)*
Overall review ratings	Overall risk of bias	Low *(Methods used are largely appropriate to review question, however issues described with search terms and subject terms)*
Overall review ratings	Applicability as a source of data	Fully applicable

1

IG evidence to recommendations justification: This section was part of one literature search for section 5. For this review question, 34 studies provided evidence for section 5.4, covering the following comparisons:

- Comparison 1. Metformin vs Placebo
    - 7 studies, 6 had moderate risk of bias, 1 had low risk of bias
    - Metformin was better than placebo for live birth rate, clinical pregnancy rate, and pregnancy rate, odds ratios all >1.8, moderate certainty
    - no difference was seen for other outcomes
    - Comparison 2. Metformin vs Clomiphene Citrate
- 3 studies, 1 moderate risk of bias, 2 low risk of bias
    - Relatively large sample sizes (>400)
    - One study found ovulation rate per cycle was lower with metformin (OR=0.43)
    - 3 studies found miscarriage rate higher with metformin (OR=2.44)
    - No difference was seen for other outcomes
- Comparison 3. Metformin + Clomiphene Citrate vs Metformin
    - Three studies, one low, 2 medium risk of bias
    - Metformin + clomiphene citrate had higher live birth and ovulation rates compared to metformin alone (OR 2.44 and 3.72 respectively)
    - No difference seen for other outcomes
- Comparison 4. Metformin + Letrozole vs Letrozole
    - 2 studies, both high risk of bias due to lack of blinding/allocation concealment issues
    - No difference seen between groups for any outcome
- Comparison 5. Metformin + Letrozole vs Metformin + Clomiphene Citrate
    - 2 studies, moderate risk of bias due to lack of blinding/allocation concealment issues
    - Metformin + letrozole group showed significantly better outcomes for clinical pregnancy, ovulation rate and full- term pregnancy rate per patient
    - No difference seen for other outcomes
- Comparison 6. Metformin vs gonadotropins (hMG)
    - One study, moderate risk of bias due to lack of blinding

    - No difference seen between groups for any outcome
- Comparison 7. Metformin + gonadotropins vs gonadotropins
    - One study, moderate risk of bias due to lack of blinding
    - No difference seen between groups for any outcome
- Comparison 8. Clomiphene Citrate vs Metformin + Clomiphene Citrate
    - Eight studies, 1 low risk, 3 moderate and 4 high risk of bias
    - 2 studies specifically reported on women with clomiphene citrate resistance
    - Metformin + CC was significantly better for clinical pregnancy rate per patient, ovulation rate per patient and per cycle
    - No difference seen for other outcomes
- Comparison 9. Clomiphene Citrate vs Letrozole
    - 11 RCTs, 1 moderate, 4 low, 6 high risk of bias
    - Letrozole was significantly better for live birth rate per patient, clinical pregnancy rate per patient, pregnancy rate per patient, ovulation rate both per patient and per cycle
    - No difference seen for other outcomes
- Comparison 10. Clomiphene Citrate vs Metformin + Letrozole
    - One study high risk of bias due to lack of blinding/allocation concealment issues
    - Clinical pregnancy rate and ovulation rate per cycle were significantly better in the metformin + letrozole group
    - No difference seen for other outcomes
- Comparison 11. Clomiphene Citrate + Metformin vs Letrozole
    - One study, high risk of bias due to lack of blinding/allocation concealment
    - No difference seen between groups for any outcome
- Comparison 12. Clomiphene Citrate vs gonadotropins
    - 2 studies, 1 moderate, 1 high risk of bias
    - FSH group had better outcomes for live birth rate, clinical pregnancy rate and ovulation rate per cycle
    - No difference seen for other outcomes
- Comparison 13. Clomiphene Citrate + gonadotropins vs Clomiphene Citrate

    - 2 studies, moderate risk of bias due to single blind design
    - No difference seen between groups for any outcome
- Comparison 14. Clomiphene Citrate + gonadotropins (FSH) vs Letrozole
    - 1 study, moderate risk of bias due to single blind design
    - Clinical pregnancy rate and ovulation rate per patient was higher with the FSH + letrozole group
    - No difference seen for miscarriage rate
- Comparison 15. Clomiphene Citrate + gonadotropins vs Metformin
    - 1 study, high risk of bias due to assumption that participants were not blind to their medication
    - Clinical pregnancy rate higher with FSH group
    - No difference for miscarriage rate
- Comparison 16. Clomiphene Citrate vs Laparoscopic ovarian drilling (LOD)
    - 1 UK study, moderate risk of bias due to allocations not being concealed
    - No difference seen between groups for any outcomes

Five EBRs were made describing the preferred use of medication for the above outcomes in women with PCOS. One or more practice point(s) accompanies each EBR giving additional information on risks and monitoring requirements. All comparisons had low or very low GRADE outcomes. For recommendation 5.4.1.1, a CR was made for the use of metformin alone in women with PCOS and anovulatory infertility, this was justified as the evidence for 2 critical outcomes (live birth, clinical pregnancy rate) and one important outcome (pregnancy rate per patient) all had moderate certainty evidence. The panel discussion details that metformin is much cheaper than other medications and appears effective. However, side effects were noted in around 20% of people, though it was discussed that these could be lowered by advice on medicine timing. As the evidence for the three outcomes highlighted was of moderate certainty, the EBR for 5.4.1.1 seems appropriate.

Recommendation 5.4.2.1 highlights that clomiphene citrate could be used in preference to metformin, this was based on good evidence from high quality studies for the outcome of live birth rate. Clinical pregnancy rate was twice that of metformin,

however the odds ratios' confidence intervals were wide. There was no difference in multiple pregnancy rate between the two, with ovulation rate per cycle strongly associated with clomiphene citrate. The IG mentions that the high confidence in the evidence for this recommendation relies on 3 studies, with one study in particular dominating the results. They also discussed that adding clomiphene citrate to metformin may increase side effects but there was no evidence found for this. It is also highlighted that combination treatment may incur extra costs as additional monitoring may be required. The evidence discussed supports recommendation 5.4.2 and allows personal preference of treatment based on the side effects and information available to patients, however a suggested order of preference of medications may be helpful for patients and non-clinical service users.

Recommendation 5.4.3.1 suggests that a combination of clomiphene citrate and metformin could be used rather than clomiphene citrate alone, this is supported by evidence from 5 studies which found clinical pregnancy rates were higher in the combination group with a narrow odds ratio range, with minimal heterogeneity. Issues were raised with the potential for more side effects as both metformin and clomiphene citrate have known side effects, however no specific evidence was found for this. The panel commented that women may value the extra benefit of the combined treatment program, but did not provide any further details on this. The panel also noted there was little to no increased cost for this. As such the recommendation is permissive of the use of the combined treatment, and highlights to service users that this is an evidence-based option for treatment of anovulatory fertility.

For recommendation 5.4.4.1, this is the first of the medication recommendations in this section with a strong recommendation for the option, stating that letrozole should be used in preference to clomiphene citrate. This is due to high quality evidence for live birth rate, moderate evidence for pregnancy and low evidence for multiple pregnancies. The panel discussed that the evidence for letrozole compared to clomiphene citrate was superior for live birth and pregnancy rate, supported by a good number of high-quality studies with narrow confidence intervals, increasing

certainty in this outcome. It is noted that no data on side effects was included in this review. As such, recommendation 5.4.4.1 is suitable given the strength of evidence for the outcomes of live birth rate and pregnancy, however this would be better placed at the start of the EBRs for section 5.4 as it is the only strong recommendation that gives a preference on which medication to use first.

The IG highlights potential implementation issues regarding drug side effects and highlights that research priorities should focus on when the drugs should be stopped if pregnant.

IG economic evidence

No health economic evidence was identified in the IG for review question 5.4 on clomiphene citrate and metformin.

5.4.2 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

One health economic study was identified for this review question but was excluded due to a combination of applicability and methodological concerns.

See Appendix D of this document for a list of excluded economic studies, with reason for exclusion.

Economic model

No original health economic modelling was conducted for this review question.

Further information on the health economic prioritisation for this review question, and other ovulation induction strategy review questions, can be found in the Letrozole sub-section of this report (section 5.3).

Unit costs

For the unit costs applicable to this review question, please see the unit costs presented in **Table 1**, which can be found in the Letrozole sub-section of this report (section 5.3).

5.4.3 NICE recommendations

The relevant recommendations for this section are Rec 1.21.11 to 1.21.13.

5.4.4 The committee's discussion and interpretation of the evidence

Clinical

The committee discussed the use of medication for fertility outcomes and agreed that letrozole should be the first line treatment. However, the committee felt it important to offer the use of metformin and clomiphene citrate for those patients who could not tolerate letrozole, or prefer to use these medications. As such the committee agreed to contextualise two EBRs from the IG into two new NICE recommendations. The committee discussed the dosing options for clomiphene citrate plus metformin. They agreed that the starting dose for each drug individually was suitable when the two were combined, and as such the dosing recommendations in the NICE recommendation were the most appropriate for the PCOS population. The committee also added that the doses stated were the most commonly reported in RCTs for PCOS management. The committee added a recommendation to highlight that continued treatment with either clomiphene citrate alone, or in combination should not be continued for longer than 6 cycles to be consistent with the NICE guideline on fertility. The 6 cycle limit was applied in the fertility guideline as it was discussed that 6 cycles of treatment should be sufficient to determine whether treatment would be effective, and also to be in line with manufacturer advice that treatment should not normally be continued for longer than 6 cycles due to a possible increase in the risk

of developing ovarian cancer. The committee did not think it was necessary to contextualise metformin alone as they agreed with the IG that there were more effective and evidence-based alternatives where letrozole cannot be used.

Health economic

No health economic evidence was identified in either the IG, or in NICE's health economic literature search, for review question 5.4.

As discussed in section 5.3.4, the committee concluded that letrozole dominated clomiphene citrate (letrozole was less costly and more effective). Letrozole was therefore recommended as a first-line treatment option for ovulation induction. Clomiphene citrate and metformin or clomiphene citrate alone was also recommended for those when letrozole is contraindicated or not tolerated. The committee considered the clinical evidence and the costs of clomiphene citrate and metformin and clomiphene citrate alone. There was limited clinical evidence which showed limited differences in effectiveness for clomiphene citrate and metformin and clomiphene citrate alone. The committee discussed that different treatment options performed better for different outcomes. Therefore, due to the small increase in costs associated with prescribing metformin in addition to clomiphene citrate, the committee concluded that it was important to provide the patient and healthcare professional a choice of treatment which could be decided upon depending on patient preferences, individual risk factors and the problem which the person with PCOS was presenting with.

Clomiphene citrate and metformin or clomiphene citrate alone was not recommended in any other instance apart as a first-line therapy option for those people where letrozole is contraindicated or not tolerated.

The committee acknowledged that these recommendations are reflective of current practice and therefore no significant resource impact is anticipated.

5.5 Gonadotrophins

Review question 5.5: In women with PCOS, are gonadotrophins effective for improving fertility outcomes?

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

Evidence-based recommendations:

5.5.1 Gonadotrophins alone could be considered rather than clomiphene citrate in therapy naïve women with PCOS with anovulatory infertility and no other infertility factors to improve ovulation, clinical pregnancy and live birth rates (refer to PP 5.5.6).

5.5.2 Gonadotrophins alone could be used over gonadotrophins combined with clomiphene citrate in women with PCOS who are anovulatory and infertile with clomiphene citrate resistance or failure, and no other infertility factors.

5.5.3 Gonadotrophins could be considered rather than the combination of clomiphene citrate and metformin in women with PCOS who are anovulatory and infertile, with clomiphene citrate-resistance and no other infertility factors.

5.5.4 Either gonadotrophins or laparoscopic ovarian surgery could be used in women with PCOS who are anovulatory and infertile, with clomiphene citrate-resistance and no other infertility factors, following counselling on higher live birth rate and higher multiple pregnancy rates with gonadotrophins.

5.5.5 Gonadotrophins could be second-line pharmacological therapy for women with PCOS who are anovulatory and infertile, with no other infertility factors and who have failed first-line oral ovulation induction.

Practice points:

5.5.6 Where gonadotrophins are to be prescribed, the following should be considered:

- Cost of the intervention for ovulation induction.

- Expertise required for the use of the intervention for ovulation induction.
- The degree of intensive ultrasound monitoring that is required.
- A low dose step-up gonadotrophin protocol should be used to optimise the chance of monofollicular development.
- Implications of potential multiple pregnancy.

5.5.7 There appears to be no difference in the clinical efficacy of the available gonadotrophin preparations.

5.5.8 When using gonadotrophins, best clinical practice is to avoid multiple pregnancy. Considerations here include cancelling cycles when there is more than a total of two follicles greater than 14 mm in diameter and advising avoidance of unprotected intercourse.

5.5.9 Live birth rate, clinical pregnancy rate per patient and ovulation rate per cycle are higher with gonadotrophins than with clomiphene citrate.

5.5.10 A low dose gonadotrophin protocol should be used to optimise the chance of monofollicular growth and minimise multiple pregnancy.

5.5.11 Cycle monitoring and drug costs coupled with multiple injection will influence choice in gonadotrophin use.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Clear and detailed PICO)*

Section	Question	Answer
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Detailed PICO available which is suitable for the review)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(Restrictions on date are present but not well described)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(Study types selected and language restrictions used are appropriate for review question)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All parameters are appropriate to review)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(Appropriate number (5) and range of databases used. Reasonable range but PsycINFO seems superfluous.*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably no *(PRISMA flowchart mentions number of new systematic reviews, RCTs identified from those reviews, new RCTs and studies from the previous version of the guideline)*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(Full search strategy is detailed in the GDG 5 methodology appendix. Subject terms not translated between databases. This strategy covers questions addressing a range of fertility technologies. Terms for a few are included but not all. It may have*

Section	Question	Answer
		been considered that using subject headings for reproductive techniques/fertility may suffice but free text terms for specific technologies would be expected to be there e.g. aromatase inhibitors. No terms at all for ovarian surgery etc)
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(Date limits not mentioned in the protocol but it is clear that this updates a 2017 review)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(Studies screened by one reviewer with evidence team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(Study appraisal completed by 2 reviewers plus evidence team)*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(Study table with study characteristics present allowing interpretation of results)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(12 studies met inclusion criteria for review, all appear to be included in analysis)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(Detailed integrity assessment in GDG 5 methodology appendix for all studies, individual quality appraisals also present)*

Section	Question	Answer
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Yes *(2 reviewers plus evidence team)*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(All 12 included studies for section 5.5 have been included in the study characteristics table and appear to have been included across the 7 different comparisons)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(No information was given regarding the planned analysis of the results)*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(MA was used appropriately for this review question, study design is appropriate for analysis, however odds ratios were used where typically risk ratios would be used.)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Yes *(Heterogeneity appears to have been accounted and appropriately downgraded for, there are detailed tables for each of the 7 outcomes which detail inconsistency for individual outcomes. Forest plots are also present for each of the 7 outcomes and all the subgroup analyses within each outcome)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Yes *(Funnel plots available for each subgroup analyses, appears appropriate)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Yes

Section	Question	Answer
		(Risk of bias assessed per outcome in section 5.5 and per study in GDG 5 methodology appendix, appears to have taken bias into account but further write up about this could have been useful. RoB included in study characteristics table and GRADE assessment also completed)
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(More detailed write up would be helpful to confirm conclusions, but otherwise no notable issues)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(All parameters are appropriate to review)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(More detailed write up would be helpful to confirm conclusions, but otherwise no notable issues)*
Overall review ratings	Overall risk of bias	Low *(Low overall, though note issues with search terms and subject terms translation between databases)*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: in this updated systematic review, 12 studies were included covering 7 comparisons.

- Comparison 1. FSH vs CC

- 2 studies, moderate and high risk of bias
- FSH was better than CC for the outcomes of live birth rate per patient, clinical pregnancy rate per patient and ovulation rate per cycle
- No difference was seen between groups for all other outcomes
- Comparison 2. FSH + CC vs FSH
- 2 studies, moderate risk of bias
- No difference seen for any outcome

- Comparison 3. FSH + CC vs LET
  - One study, moderate risk of bias due to single blind design
  - Clinical pregnancy rate and ovulation rate per patient were significantly higher in the LET group
  - No difference seen for miscarriage rate
- Comparison 4. FSH vs LOD
  - 3 studies, all high risk of bias due to small sample size, lack of blinding and high dropout rates
  - FSH was better than LOD for live birth rate
  - Multiple pregnancy was higher in the FSH group, (favours LOD)
  - No difference seen for other outcomes
- Comparison 5. FSH + Metformin vs FSH +/- Placebo
  - 3 studies, 2 high and 1 moderate risk of bias due to lack of reporting on key elements and blinding issues
  - No differences seen between groups for any outcomes
- Comparison 6. FSH vs Clomiphene Citrate + Metformin
  - 1 study, high risk of bias, due to lack of information about blinding/unblinding
  - FSH was significantly better for clinical pregnancy rate
  - No difference seen for miscarriage rate
- Comparison 7. hMG vs Metformin
  - 1 study, moderate risk of bias due to lack of blinding
  - No difference seen for any outcome

The evidence for gonadotrophin therapy for IG recommendation 5.5.1 shows better ovulation rate per cycle, pregnancy and live birth rates compared to oral anti-oestrogens in anovulatory women with no other fertility issues, however this evidence comes from 1 (low certainty) or 2 (moderate certainty) studies. There is an increased risk of multiple pregnancy, and an impact on women's daily lives as they have daily injections to take, increased monitoring, more frequent travel to clinics and there is an increased cost of the medication compared to oral medications, however the panel discussed that the likely benefits outweigh any undesirable effects. As this is a weak recommendation stating that gonadotrophins could be used over clomiphene citrate, it is reasonable to make this recommendation as it highlights additional treatment options, despite the evidence being low to moderate in this area. The evidence base for IG recommendation 5.5.2 is similar to that for 5.5.1, with a difference in population from medication naïve to those who have experienced medication failure on clomiphene citrate. As such the evidence generally supports the recommendation to use gonadotrophins alone rather than gonadotrophin plus clomiphene citrate, again highlighting the potential added costs of medication and clinic visits, and the additional impact of the injections on women. For IG recommendation 5.5.3, gonadotrophins are suggested over metformin plus clomiphene citrate due to the primary outcomes of live birth being higher with gonadotrophins. This is well supported by the evidence which found gonadotropins to have much higher rates of live births, despite the need for increased monitoring and increased medication costs. For IG recommendation 5.5.4, it gives the option of either gonadotrophins or laparoscopic ovarian surgery (LOS) for women with PCOS who are anovulatory and infertile, and who have experienced clomiphene citrate resistance. The evidence highlights that gonadotrophins have better live birth rates compared to LOS, however the guideline highlights the potential for regional variation for the availability of these two options, and highlights that the increased risk of multiple pregnancy with gonadotrophins must be explained. The evidence here supports the recommendation for gonadotrophins, however there is little discussion as to why the option for laparoscopic surgery was included, other than to briefly mention availability of medication/procedures. As such the evidence stated does not

fully support IG recommendation 5.5.4. The Evidence for IG recommendation 5.5.5 was noted to be very low to moderate certainty with very serious risk of bias and imprecision, and serious risk of inconsistency and indirectness. The recommendation is a weak recommendation which makes this suitable given the evidence for the use of gonadotrophins in IG recommendations 5.5.1-5.5.4, however there was no obvious discussion about the use of gonadotrophins as a second line option when another treatment (other than clomiphene citrate) had failed. As such this recommendation appears superfluous, considering there are already 3 other EBRs that suggest gonadotrophins are superior. Further discussion in this area would have been useful.

The IG describes gonadotrophin as an effective treatment for anovulatory women with PCOS.. Research priorities include more data comparing letrozole and gonadotrophins in women with anovulatory PCOS.

IG economic evidence

No health economic evidence was identified in the IG for review question 5.5 on gonadotrophins.

In terms of costs and cost-effectiveness the IG discussed the costs of gonadotrophins compared to laparoscopic ovarian surgery – noting that surgery requires access to a day surgery centre with endoscopic equipment, in addition to access to skilled surgeons.

The IG also noted that laparoscopic ovarian surgery is a single costly procedure but without need for further medications and monitoring. It was also noted that gonadotropins have a higher probability of multiple pregnancies and that the cost of the medication is an important consideration for implementation.

It was also acknowledged that in certain healthcare systems access to surgery may be restricted or unaffordable.

5.5.2 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

Two health economic studies were identified for this review question but excluded due to a combination of applicability and methodological concerns.

See Appendix D of this document for a list of excluded economic studies, with reason for exclusion.

Economic model

No original health economic modelling was conducted for this review question.

Further information on the health economic prioritisation for this review question, and other ovulation induction strategy review questions, can be found in the Letrozole sub-section of this report (section 5.3).

Unit costs

For the unit costs applicable to this review question, please see the unit costs presented in **Table 1**, which can be found in the Letrozole sub-section of this report (section 5.3).

5.5.3 NICE recommendations

The relevant recommendation for this section is rec 1.21.14.

5.5.4 The committee's discussion and interpretation of the evidence

Clinical

The committee discussed the use of gonadotrophins for ovulation induction. It was queried whether increased monitoring would be required when using gonadotrophins, however the committee concluded that all aspects of ovulation induction were likely to require a degree of monitoring, and this was not specific to gonadotrophins. As such the committee decided to adapt EBR 5.5.5 from the IG, to allow the option of gonadotrophins where ovulation induction with first line medications has been unsuccessful.

Health economic

No health economic evidence was identified in either the IG, or in NICE's health economic literature search, for review question 5.5 on gonadotrophins.

The committee considered the costs presented to them and the available clinical evidence in the IG and concluded that gonadotrophins should be prescribed as a second-line treatment option for ovulation induction. The committee agreed that letrozole was dominant (less costly and more effective) compared to clomiphene citrate but acknowledged that the clinical evidence for gonadotrophins was sparse (see section 5.3.4 for further discussion on the cost-effectiveness of letrozole).

Cost-effectiveness was therefore considered qualitatively, with the committee concluding that there was only a low probability of conceiving with clomiphene citrate after treatment with letrozole (for those where letrozole is tolerated and not contraindicated). The committee acknowledged that although gonadotrophins are significantly more expensive than clomiphene citrate, up to 6-months of treatment with clomiphene citrate and subsequent treatment with gonadotrophins, is more expensive and resource intensive than offering gonadotrophins straight away. Therefore, the committee agreed that gonadotrophins were the most cost-effective second-line therapy.

The committee also made a recommendation to consider laparoscopic ovarian surgery if an alternative to gonadotrophins is needed. Further discussion on the cost-effectiveness of laparoscopic ovarian surgery can be found in the committee's discussion of the evidence related to this review question (section 5.6.4).

The committee acknowledged that the recommendations made for this review question are reflective of UK current practice and therefore no significant resource impact is anticipated.

5.6 Laparoscopic ovarian surgery

Review question 5.6: In women with PCOS, is ovarian surgery effective for improving fertility outcomes?

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

Evidence-based recommendation:

5.6.1 Laparoscopic ovarian surgery could be second-line therapy for women with PCOS who are anovulatory and infertile, with clomiphene citrate resistance and no other infertility factors.

Practice point:

5.6.2 When using laparoscopic ovarian surgery, the following should be considered:

- Comparative cost of the intervention for ovulation induction.
- Expertise required for the safe use of the intervention for ovulation induction.
- Both intraoperative and postoperative risks, which are higher in women who are above healthy weight.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Clear and detailed PICO)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Detailed PICO available which is suitable for the review)*

Section	Question	Answer
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(Restrictions on date are present but not well described)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(Study types selected and language used are appropriate for review question)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All parameters are appropriate to review)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(Appropriate number (5) and range of databases used. Reasonable range but PsycINFO seems superfluous)*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably no *(PRISMA flowchart mentions number of new systematic reviews, RCTs identified from those reviews, new RCTs and studies from the previous version of the guideline)*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(Full search strategy is detailed in the GDG 5 methodology appendix. Subject terms not translated between databases. This strategy covers questions addressing a range of fertility technologies. Terms for a few are included but not all. It may have been considered that using subject headings for reproductive techniques/fertility may suffice but free text terms for specific technologies would be expected to be there e.g. aromatase inhibitors. No terms at all for ovarian surgery etc)*

Section	Question	Answer
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(Date limits not mentioned in the protocol but it is clear that this updates a 2017 review)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(Studies screened by one reviewer with evidence team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(Study appraisal completed by 2 reviewers plus evidence team)*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(Study table with study characteristics present allowing interpretation of results)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(6 studies met inclusion criteria for review, all appear to be included in across the 3 comparisons for this section.)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(Detailed integrity assessment in GDG 5 methodology appendix for all studies, individual quality appraisals also present)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Yes *(2 reviewers plus evidence team)*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(All 6 included studies for section 5.6 have been included in the study characteristics*

Section	Question	Answer
		table and appear to have been included across the 3 different comparisons)
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(No information was given regarding the planned analysis of the results)*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(MA was used appropriately for this review question, study design is appropriate for analysis, however odds ratios were used where typically risk ratios would be used)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Yes *(Heterogeneity appears to have been accounted for and downgraded appropriately, there are detailed tables for each of the 3 outcomes which detail inconsistency for individual outcomes. Forest plots are also present for each of the 3 outcomes and all the subgroup analyses within each outcome)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Yes *(Funnel plots available for each subgroup analyses, appears appropriate)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Yes *(Risk of bias assessed per outcome in section 5.5 and per study in GDG 5 methodology appendix, appears to have taken bias into account but further write up about this could have been useful. RoB included in study characteristics table and GRADE assessment also completed)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(More detailed write up would be helpful to confirm conclusions, but otherwise well conducted)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(All parameters are appropriate to review)*

Section	Question	Answer
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(More detailed write up would be helpful to confirm conclusions, but otherwise well conducted)*
Overall review ratings	Overall risk of bias	Low *(Search terms and subject terms had issues and a further written discussion on results would have been helpful)*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: in this updated systematic review, 6 studies were included to cover 3 comparisons.

- Comparison 1. Laporoscopic ovarian drilling (OD) vs gonadotrophins (FSH)
    - 3 studies, all in clomiphene citrate resistant PCOS, all high risk of bias due to lack of blinding, small sample sizes and high dropout rates
    - FSH was better than LOD for the outcome of live birth rate
    - Multiple pregnancy rate was higher in FSH than LOD
    - No difference was seen for clinical pregnancy rate, ovulation rate per patient or miscarriage rate per patient
- Comparison 2. LOD vs CC
    - 1 study (UK based), women with clomiphene citrate resistant PCOS, moderate risk of bias due to lack of blinding following randomisation
    - No difference was seen for any outcome
- Comparison 3. Unilateral LOD vs Bilateral LOD

- 2 studies, high risk of bias due to lack of information provided on blinding and allocation concealment
- No difference was seen for any outcome

The IG discussed how FSH was superior to LOD for live birth rates 0.45 (95% CI 0.27 to 0.76) however it also had a higher risk of multiple pregnancies which raised concerns. Evidence for outcomes other than live birth rates were low or very low due to the high risk of bias, varied effect estimates and wide confidence intervals (clinical pregnancy rate 0.44 (95% CI 0.16 to 1.23), ovulation rate 0.66 (95% CI 0.21 to 2.07), miscarriage 0.73 (95% CI 0.32 to 1.66). As such the authors note that evidence in this area should be interpreted with caution. Three studies support recommendation 5.6.1, one was UK based and two were India based, however these had moderate to high risk of bias and had small sample sizes. LOS was highlighted as being an important single intervention that leads to a preferred outcome in a significant proportion of women with PCOS. Compared to treatment with gonadotrophins, there was no need for ongoing monitoring or daily injections which was felt might be preferred by patients. It was highlighted that access to skilled surgeons and appropriate clinics might be an issue for some areas, however despite the low-quality evidence, the authors felt this was an important treatment option to be offered. As such the recommendation is reasonable given the improvement in live birth outcomes, the wording is appropriately cautious and limits the use of LOS to those with clomiphene citrate resistance. As the evidence in this area is limited and based on either 1 or 2 studies, with little evidence of effect, the wording of the recommendation could include information on the limited evidence available to help people make a more informed decision. An implementation issue regarding ensuring surgeons are adequately trained to carry out treatments was raised, with research priorities regarding how LOS restores ovulatory function suggested

IG economic evidence

No health economic evidence was identified in the IG for review question 5.6 on ovarian surgery.

The IG noted the large costs associated with ovarian surgery, stating that ovarian surgery is a single costly procedure, but omits the need for further medications and monitoring, also reducing the probability multiple pregnancies when compared to gonadotrophins. The IG also noted that ovarian surgery requires access to a day surgery centres with endoscopic equipment, in addition to access to skilled surgeons who can conduct the procedure.

5.6.2 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

Two health economic studies were identified for this review question but excluded due to a combination of applicability and methodological concerns.

See Appendix D of this document for a list of excluded economic studies, with reason for exclusion.

Economic model

No original health economic modelling was conducted for this review question.

Further information on the health economic prioritisation for this review question, and other ovulation induction strategy review questions, can be found in the Letrozole sub-section of this report (section 5.3).

Unit costs

For the unit costs applicable to this review question, please see the unit costs presented in **Table 1**, which can be found in the Letrozole sub-section of this report (section 5.3).

5.6.3 NICE recommendations

The relevant recommendation for this section is Rec 1.21.15.

5.6.4 The committee's discussion and interpretation of the evidence

Clinical

The committee discussed the use of laparoscopic ovarian surgery and felt that this area would benefit from further research. However, as it was an option for those patients who could not use gonadotrophins and/or had experienced failure of induction of ovulation with first line medications, the committee agreed to contextualise the IG recommendation as a consider recommendation. The committee discussed that alternative options to medications are sometimes needed and this should remain an option open to patients who fit these criteria.

Health Economic

No health economic evidence was identified in either the IG, or in NICE's health economic literature search, for review question 5.6 on laparoscopic ovarian surgery.

The committee discussed the clinical evidence noting that in terms of live birth rates gonadotrophins are more effective than laparoscopic ovarian surgery. The committee also noted that specialist health care professionals are required to conduct laparoscopic ovarian surgery. The costs of laparoscopic ovarian surgery and various gonadotrophins were presented to the committee, and overall, the committee concluded that gonadotrophins are likely to be the dominant (less costly and more effective) treatment option. The committee did, however, make a recommendation to consider laparoscopic ovarian surgery when an alternative to gonadotrophins is needed, noting that in certain situations, where the direct cause of infertility is targeted, laparoscopic ovarian surgery could be more cost-effective than

gonadotrophins. For example, for those with significant pelvic pain, suspected endometriosis, fallopian tube pathology or ovarian cysts. The committee noted that in these instances laparoscopic ovarian surgery would likely be indicated, as in line with current UK practice.

Overall, the committee concluded that as the recommendations made for this review question are reflective current practice, no significant resource impact is anticipated.

5.7 (0) In vitro fertilisation/intracytoplasmic sperm injection (IVF/ICSI)

Review question 5.7: In women with PCOS is stimulated In vitro fertilisation/intracytoplasmic sperm injection (IVF/ICSI) effective for improving fertility outcomes?

5.7.1 5.7.0 Recommendations from the International evidence-based guideline for PCOS*

Consensus recommendation:

5.7.0.1 In the absence of an absolute indication for in vitro fertilisation (IVF)/intracytoplasmic sperm injection (ICSI), IVF could be offered in women with PCOS and anovulatory infertility, if first- or second-line ovulation induction therapies have failed.

Practice point:

5.7.0.2 In women with anovulatory PCOS, the use of IVF is effective and when elective single embryo transfer is used, multiple pregnancies can be minimised.

5.7.0.3 Women with PCOS undergoing IVF/ICSI treatment should be counselled prior to starting treatment about the increased risk of ovarian hyperstimulation syndrome and options to reduce the risk should be offered.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

There is no ROBIS assessment for this section as there was no evidence review in the IG for recommendations 5.7.0.1-5.7.0.3

IG evidence to recommendations justification: The overview for section 5.7 contains one CR and 3 practice points. The Monash technical document does not give any search history, study selection, or a narrative review as to how these recommendations have been created. Sections 5.7.1-5.7.6 all have individual ROBIS assessments and subsequent evidence to recommendation discussions following this section. There is a GRADE certainty of moderate for the comparison of IVF versus in vitro maturation (IVM) which resulted in a conditional weak recommendation against the option, leading to the CR stating that IVF could be offered if first- or second-line ovulation induction therapies had failed. The GRADE considerations note that live birth rate per patient and clinical pregnancy rate per patient is significantly higher with stimulated IVF compared to IVM + IVF. Live birth rate per pregnancy and miscarriage per pregnancy were similar when the two groups were compared. The IG notes that although IVM is no longer considered experimental there may still be cost and expertise issues related to implementation of IVM. It is difficult to determine whether or not the strength of this recommendation is appropriate given the evidence is not available for review.

IG economic evidence

No health economic evidence was identified in the IG for review question 5.7 on in-vitro fertilisation.

5.7.2 NICE Economic evidence

Included studies

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

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

Excluded studies

One health economic study was identified for this review question but excluded due to a combination of applicability and methodological concerns.

See Appendix D of this document for a list of excluded economic studies, with reason for exclusion.

Economic model

No original health economic modelling was conducted for this review question.

Unit costs

Unit costs for IVF and ICSI are presented below in **Table 2**.

Unit costs for letrozole, clomiphene citrate, metformin, and laparoscopic drilling are presented in **Table 1**, which can be found in the Letrozole sub-section of this report.

Table 2: Unit costs for IVF and ICSI

Resource	Unit costs for women aged 37 and under	Unit costs for women aged 38 or older, or previous non-responder
IVF (price to include one fresh and one frozen cycle)	£3,649	£4,120
ICSI (price to include one fresh and one frozen cycle)	£4,120	£4,708
Subsequent frozen cycles	£1,777	£1,777

Source: 2024/25 NHS pay award prices, date accessed 21/05/2026

5.7.3 NICE recommendations

The relevant recommendations for this section are Rec 1.21.16 to 1.21.17.

5.7.4 The committee's discussion and interpretation of the evidence

Clinical

The committee discussed the use of IVF for fertility outcomes in women with PCOS, highlighting that this topic is also included in the NICE guideline on fertility. The NICE fertility guideline has a section on women at high risk of a high ovarian response but does not specify women with PCOS. The IG has only a CR and practice points that cover IVF, as such the committee felt it would be helpful to include two new recommendations for women with PCOS. It was agreed that both of the new NICE recommendations should refer to the NICE fertility guideline for more detailed information on IVF use.

Health economic

No health economic evidence was identified in either the IG, or in NICE's health economic literature search, for review question 5.7 on IVF.

As the guideline committee decided to cross-refer to NICE's existing guidance on fertility problems for their recommendations relating to IVF, both in terms of when IVF can be accessed and how to perform IVF, no significant resource impact is anticipated due to this cross-referral because it does not impact on the population eligible for IVF.

5.7 (1) Gonadotrophin releasing hormone protocol

Review question 5.7.1: In women with PCOS undergoing IVF/ICSI treatment, is the gonadotrophin releasing hormone (GnRH) antagonist protocol or GnRH agonist long protocol the most effective for improving fertility outcomes?

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

Practice point(s):

5.7.1.1 Gonadotrophin releasing hormone protocol (GnRH) antagonist protocol cannot be recommended over GnRH agonist long protocol for women with PCOS undergoing IVF/ICSI to improve clinical pregnancy or live birth rate.

5.7.1.2 The use of a GnRH antagonist protocol for women with PCOS undergoing IVF/ICSI is recommended as it enables the use of an agonist trigger, with the freezing of all embryos generated if required, without compromising the cumulative live birth rate, to reduce the risk of significant ovarian hyperstimulation syndrome.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Clear and detailed PICO)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes

Section	Question	Answer
		(Detailed PICO available which is suitable for the review)
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(Restrictions on date are present but not well described)*
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 *(Language and study types selected are appropriate for review question)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All areas appear to be well covered with relevant information)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(Appropriate number (5) and range of databases used. Reasonable range but PsycINFO seems superfluous)*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably no *(PRISMA flowchart mentions number of new systematic reviews, RCTs identified from those reviews, new RCTs and studies from the previous version of the guideline)*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(Full search strategy is detailed in the GDG 5 methodology appendix. Subject terms not translated between databases. This strategy covers questions addressing a range of fertility technologies. Terms for a few are included but not all. It may have been considered that using subject headings for reproductive*

Section	Question	Answer
		techniques/fertility may suffice but free text terms for specific technologies would be expected to be there e.g. aromatase inhibitors. No terms at all for ovarian surgery etc)
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(Date limits not mentioned in the protocol but it is clear that this updates a 2017 review)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(Studies screened by one reviewer with evidence team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(Study appraisal completed by 2 reviewers plus evidence team)*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(Study table with study characteristics present allowing interpretation of results)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(7 studies met inclusion criteria for review, all appear to be included in analysis with between 3 and 7 studies were used for each outcome or subgroup)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(detailed integrity assessment in GDG 5 methodology appendix for all studies, individual quality appraisals also present)*
Data collection and	Were efforts made to minimise error in risk of bias assessment?	Yes *(2 reviewers plus evidence team)*

Section	Question	Answer
study appraisal		
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(All 7 included studies for section 5.7.1 have been included in the study characteristics table and appear to have been included across the 7 different comparisons)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(No information was given regarding the planned analysis of the results)*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(MA was used appropriately for this review question however odds ratios were used where typically risk ratios would be used. Study design is appropriate for analysis)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Yes *(Heterogeneity appears to have been accounted for, there is a detailed table for the single outcome which details inconsistency for the individual outcomes. Inconsistency where present was considered, downgrading noted where required for GRADE assessments. Forest plots are also present for this outcome and all of the subgroup analyses within this outcome)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Yes *(Funnel plots available for each subgroup analyses, appears appropriate)*

Section	Question	Answer
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Yes *(Risk of bias assessed per outcome in section 5.5 and per study in GDG 5 methodology appendix, appears to have taken bias into account but further write up about this could have been useful RoB included in study characteristics table and GRADE assessment also completed)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(More detailed write up would be helpful to confirm conclusions, but otherwise well-conducted)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(All areas appear to be well covered with relevant information)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(More detailed write up would be helpful to confirm conclusions, but otherwise ok)*
Overall review ratings	Overall risk of bias	Low *(More detailed written discussion would have been helpful, some issues with search terms and translation of subject terms across databases)*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: 7 studies were included in

this updated systematic review. All 7 studies compared GnRH antagonist protocol with GnRH agonist long protocol. Duration of ovarian stimulation favoured the GnRH antagonist protocol; however, no difference was seen for any other outcome. Two practice points were made for this section explaining why one protocol should not be used preferentially over the other. Due to the lack of evidence, the practice points are suitable for this section. The IG highlights that more research is required in this area with studies that are adequately powered and conducted and added a research recommendation in order to encourage further work in this area.

IG economic evidence

No health economic evidence was identified in the IG for review question 5.7.1 on gonadotrophin releasing hormone protocol.

5.7.2 NICE economic evidence

Included studies

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

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

Excluded studies

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

Economic model

No original health economic model was developed for this review question.

5.7.3 NICE recommendations

No recommendations have been contextualised from this section of the IG.

5.7.4 The committee's discussion and interpretation of the evidence

Clinical

The committee discussed the use of IVF/ICSI with a GnRH antagonist and felt that this level of detail was more suited to the NICE guideline on fertility and did not need to be repeated in the PCOS guideline.

Health economic

No health economic evidence was identified in either the IG, or in NICE's health economic literature search, for review question 5.7.1 on gonadotrophin releasing hormone protocol.

The recommendations made in the IG for this review question were practice point recommendations. The committee discussed these practice point recommendations but also acknowledged that information on the type of gonadotrophin releasing hormone protocol can be found in NICE's existing guideline on fertility problems. The committee therefore decided to cross-refer to NICE's existing guideline on fertility problems in lieu of contextualising the IG's practice point recommendations. As no recommendations from the IG were contextualised, no resource implications associated with this review question.

5.7 (2) Trigger type

Review question 5.7.2: In women with PCOS undergoing GnRH antagonist IVF/ICSI treatment, is the use of hCG trigger or GnRH agonist trigger the most effective for improving fertility outcomes?

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

Consensus recommendation:

5.7.2.1 Triggering final oocyte maturation with a GnRH agonist and freezing all suitable embryos is recommended, in an IVF/ICSI cycle with a GnRH antagonist protocol, where a fresh embryo transfer is not intended or where there is an increased risk of ovarian hyperstimulation syndrome.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Clear and detailed PICO)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Detailed PICO available which is suitable for the review)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*

Section	Question	Answer
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(Restrictions on date are present but not well described)*
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 *(Language and study types selected are appropriate for review question)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All parameters are appropriate to review)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(Appropriate number (5) and range of databases used. Reasonable range but PsycINFO seems superfluous)*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably no *(PRISMA flowchart mentions number of new systematic reviews, RCTs identified from those reviews, new RCTs and studies from the previous version of the guideline)*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(Full search strategy is detailed in the GDG 5 methodology appendix. Subject terms not translated between databases. This strategy covers questions addressing a range of fertility technologies. Terms for a few are included but not all. It may have been considered that using subject headings for reproductive techniques/fertility may suffice but free text terms for specific technologies would be expected to be there e.g.*

Section	Question	Answer
		aromatase inhibitors. No terms at all for ovarian surgery etc)
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(Date limits not mentioned in the protocol but it is clear that this updates a 2017 review)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(Studies screened by one reviewer with evidence team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	No information *(N/A – progressed as narrative review)*
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 *(N/A – progressed as narrative review)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	No information *(N/A – progressed as narrative review)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	No information *(N/A – progressed as narrative review)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	No information *(N/A – progressed as narrative review)*

Section	Question	Answer
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Progressed as narrative review as no studies included)*
Synthesis and findings	Did the synthesis include all studies that it should?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(N/A – progressed as narrative review)*
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 *(N/A – progressed as narrative review)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(N/A – progressed as narrative review)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(all parameters are appropriate to review)*

Section	Question	Answer
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(N/A – progressed as narrative review)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(N/A – progressed as narrative review)*
Overall review ratings	Overall risk of bias	Low *(Issues noted with search terms and translation of subject terms but otherwise well-conducted)*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: no studies were found that were suitable for inclusion in this updated systematic review. As such it progressed as a narrative review, forming one CR. The IG noted that additional research is required in this area, specifically regarding the use of a gonadotrophin-releasing hormone (GnRH) agonist in Ovarian Hyperstimulation Syndrome (OHSS) high risk patients. Due to the lack of evidence found, the CR is appropriate. There is little discussion as to how the CR was formed, however the guideline does highlight that GnRH is preferred over human chorionic gonadotropin (hCG) because hCG increases the risk of ovarian hyperstimulation syndrome.

IG economic evidence

No health economic evidence was identified for review question 1.7.2 on trigger type for improving fertility outcomes.

The IG noted that the choice to trigger final oocyte maturation with GnRH-agonist instead of hCG is important for the prevention of OHSS as hCG alone induces oocyte maturation but is associated with OHSS. The IG also noted that GnRH- agonist triggers are associated with lower pregnancy rates, primarily in fresh embryo transfers, which can be overcome in frozen cycles.

5.7.2 NICE economic evidence

Included studies

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

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

Excluded studies

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

Economic model

No original health economic model was developed for this review question.

5.7.3 NICE recommendations

No recommendations have been contextualised from this section of the IG.

5.7.4 The committee's discussion and interpretation of the evidence

Clinical

The committee discussed the option of triggering final oocyte maturation with a GnRH agonist in an IVF/IVSI cycle. The committee felt that this level of detail was more suited to the NICE guideline on fertility and did not need to be repeated in the PCOS guideline. As such no recommendations have been contextualised from this section of the IG, nor any new recommendations made.

Health economic

No health economic evidence was identified in either the IG, or in NICE's health economic literature search, for review question 5.7.2 on trigger type for oocyte maturation.

As no specific recommendation on trigger type for IVF/ICSI was made there are no resource implications associated with this review question.

5.7 (3) Choice of follicle stimulating hormone

Review question 5.7.3: In women with PCOS undergoing (controlled) ovarian (hyper) stimulation for IVF/ICSI, does the choice of FSH effect fertility outcomes?

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

Consensus recommendation:

5.7.3.1 Either urinary or recombinant follicle stimulating hormone (FSH) could be used in women with PCOS undergoing (controlled) ovarian (hyper) stimulation for IVF/ICSI, with insufficient evidence to recommend a particular type of FSH preparation.

IG clinical evidence

Critical Appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Clear and detailed PICO)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Detailed PICO available which is suitable for the review)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics	Yes

Section	Question	Answer
	appropriate (e.g. date, sample size, study quality, outcomes measured)?	*(Restrictions on date are present but not well described)*
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 *(Language and study types selected are appropriate for review question)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All areas appear to be well covered with relevant information)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(Appropriate number (5) and range of databases used. Reasonable range but PsycINFO seems superfluous)*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably no *(PRISMA flowchart mentions number of new systematic reviews, RCTs identified from those reviews, new RCTs and studies from the previous version of the guideline)*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(Full search strategy is detailed in the GDG 5 methodology appendix. Subject terms not translated between databases. This strategy covers questions addressing a range of fertility technologies. Terms for a few are included but not all. It may have been considered that using subject headings for reproductive techniques/fertility may suffice but free text terms for specific technologies would be expected to be there e.g. aromatase inhibitors. No terms at all for ovarian surgery etc)*

Section	Question	Answer
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(Date limits not mentioned in the protocol but it is clear that this updates a 2017 review)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(Studies screened by one reviewer with evidence team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	No information *(N/A – progressed as narrative review)*
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 *(N/A – progressed as narrative review)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	No information *(N/A – progressed as narrative review)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	No information *(N/A – progressed as narrative review)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	No information *(N/A – progressed as narrative review)*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Progressed as narrative review as no studies included)*

Section	Question	Answer
Synthesis and findings	Did the synthesis include all studies that it should?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(N/A – progressed as narrative review)*
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 *(N/A – progressed as narrative review)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(Consensus rec made that was rated as a conditional weak recommendation for either the option or the comparison)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(All areas appear to be well covered with relevant information)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases*

Section	Question	Answer
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Progressed as narrative review as no studies included)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(Consensus rec made that was rated as a conditional weak recommendation for either the option or the comparison)*
Overall review ratings	Overall risk of bias	Low *(some issues noted with search terms and lack of translation of subject terms between databases)*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: no studies were found that were suitable for inclusion in this updated systematic review. As such it progressed as a narrative review, forming one CR. The CR also highlights the lack of evidence in this area which is appropriate. The IG highlights availability, convenience and cost as implementation considerations, and adds research should prioritise comparing different types of FSH preparation for women with PCOS.

IG economic evidence

No health economic evidence was identified in the IG for review question 5.7.3 for the choice of follicle stimulating hormone.

The IG noted that availability, convenience and cost considerations may be an important factor in the decision for which follicle stimulating hormone, however no additional information was provided.

5.7.2 NICE economic evidence

Included studies

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

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

Excluded studies

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

Economic model

No original health economic model was developed for this review question.

5.7.3 NICE recommendations

No recommendations have been contextualised from this section of the IG.

5.7.4 The committee's discussion and interpretation of the evidence

Clinical

The guideline committee concluded that this review question was more specific to fertility in general, as opposed to being PCOS specific, and therefore cross-referred to NICE's existing guidance on fertility problems that provides information on the procedures and treatments that should be used during in vitro fertilisation.

The committee did not feel that there was enough evidence to contextualise the CR from the IG regarding the use of either urinary or recombinant follicle stimulation hormone for women with PCOS who were undergoing ovarian stimulation.

Health economic

No health economic evidence was identified in either the IG, or in NICE's health economic literature search, for review question 5.7.3 on the choice of follicle stimulating hormone to supplement IVF.

5.7 (4) Exogenous luteinising hormone

Review question 5.7.4: In women with PCOS undergoing (controlled) ovarian (hyper) stimulation for IVF/ICSI, is exogenous luteinising hormone (LH) treatment during IVF ± ICSI effective for improving fertility outcome?

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

Consensus recommendation:

5.7.4.1 Exogenous recombinant luteinising hormone treatment should not be routinely used in combination with follicle stimulating hormone therapy in women with PCOS undergoing controlled ovarian hyperstimulation for IVF/ICSI.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Clear and detailed PICO)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Detailed PICO available which is suitable for the review)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date,	Yes *(Restrictions on date are present but not well described)*

Section	Question	Answer
	sample size, study quality, outcomes measured)?	
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 *(Language and study types selected are appropriate for review question)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All areas appear to be well covered with relevant information)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(Appropriate number (5) and range of databases used. Reasonable range but PsycINFO seems superfluous.)*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably no *(PRISMA flowchart mentions number of new systematic reviews, RCTs identified from those reviews, new RCTs and studies from the previous version of the guideline* *Note – although the review found no studies to include for this section, the narrative review references several systematic reviews but there is no description as to where these studies have come from. They could be from other sections of the guideline, or previous versions of the guideline, but this is not stated)*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(Full search strategy is detailed in the GDG 5 methodology appendix. Subject terms not translated between databases. This strategy covers questions addressing a range of fertility technologies. Terms for a few are included but not all. It may have been considered that using subject headings for reproductive*

Section	Question	Answer
		techniques/fertility may suffice but free text terms for specific technologies would be expected to be there e.g. aromatase inhibitors. No terms at all for ovarian surgery etc)
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(Date limits not mentioned in the protocol but it is clear that this updates a 2017 review)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(Studies screened by one reviewer with evidence team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases. Studies are referenced in the narrative review with no description as to where these studies are from)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	No information *(N/A – progressed as narrative review)*
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 *(N/A – progressed as narrative review)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	No information *(N/A – progressed as narrative review)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	No information *(N/A – progressed as narrative review)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	No information *(N/A – progressed as narrative review)*

Section	Question	Answer
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Progressed as narrative review as 0 studies included)*
Synthesis and findings	Did the synthesis include all studies that it should?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(N/A – progressed as narrative review)*
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 *(N/A – progressed as narrative review)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	No information *(N/A – progressed as narrative review)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(Consensus recommendation made that was rated as a conditional weak recommendation against the option)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(All areas appear to be well covered with relevant information)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not*

Section	Question	Answer
		translated between databases. Studies are referenced in the narrative review with no description as to where these studies are from)
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Progressed as narrative review as 0 studies included)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(Consensus recommendation made that was rated as a conditional weak recommendation against the option)*
Overall review ratings	Overall risk of bias	Low *(Some issues with search methods, terms and unknown studies being referenced in the narrative review, other areas ok)*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: no studies were found that were suitable for inclusion in this updated systematic review. As such it progressed as a narrative review, forming one CR. The narrative review determined there would be no anticipated benefit to adding exogenous LH supplements as such the CR concludes that LH should not be routinely used with FSH, however as there was no evidence found for this outcome, it might be more appropriate to word the recommendation as such. The IG suggest that additional research should focus on the dose of exogenous LH used in addition to FSH, and whether or not this would improve live birth rates.

IG economic evidence

No health economic evidence was identified in the IG for review question 5.7.4 on exogenous luteinising hormone treatment.

5.7.2 NICE economic evidence

Included studies

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

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

Excluded studies

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

Economic model

No original health economic model was developed for this review question as the committee were aware of existing NICE guidance on fertility problems and concluded that this guideline would likely cover this topic.

5.7.3 NICE recommendations

No recommendations have been contextualised from this section of the IG.

5.7.4 The committee's discussion and interpretation of the evidence

Clinical

The committee agreed with IG's recommendation that exogenous recombinant luteinising hormone treatment should not be routinely used in combination with follicle stimulating hormone therapy in women with PCOS undergoing controlled ovarian hyperstimulation for IVF/ICSI. This recommendation was not, however, contextualised directly as the guideline committee concluded that this information was sufficiently covered in the

fertility problems guideline. As such no recommendations were contextualised from this section of the IG.

Health economic

No health economic evidence was identified in either the IG, or in NICE's health economic literature search, for review question 5.7.4 on exogenous luteinising hormone.

As NICE's guideline on fertility problems was cross-referred to, and no recommendations were contextualised, no resource implications associated with this review question.

5.7 (5) Adjunct metformin

Review question 5.7.5: In women with PCOS undergoing (controlled) ovarian (hyper) stimulation for IVF ± ICSI, is adjunct metformin effective for improving fertility outcomes?

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

Evidence-based recommendation:

5.7.5.1 Adjunct metformin therapy could be used before and/or during FSH ovarian stimulation in women with PCOS undergoing IVF/ICSI treatment with GnRH agonist long protocol, to reduce the risk of developing ovarian hyperstimulation syndrome and miscarriage.

Practice point(s):

5.7.5.2 Good practice in PCOS and IVF is the use of a GnRH antagonist protocol as it gives the flexibility of using a GnRH agonist trigger, freeze all strategy to reduce the risk of ovarian hyperstimulation syndrome. However, if using a GnRH agonist long protocol then metformin could be considered. If using metformin, the following could be considered:

- Commence metformin at the start of GnRH agonist treatment.
- Gradually titrate metformin up to a dose of between 1000 mg to 2500 mg daily in order to minimise side-effects.
- Stopping metformin therapy at the time of the pregnancy test or period, unless the metformin therapy is otherwise indicated.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Clear and detailed PICO)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Detailed PICO available which is suitable for the review)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(Restrictions on date are present but not well described)*
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 *(Language and study types selected are appropriate for review question)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All areas appear to be well covered with relevant information)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(Appropriate number (5) and range of databases used. Reasonable range but PsycINFO seems superfluous)*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably no *(PRISMA flowchart mentions number of new systematic reviews, RCTs identified from those reviews, new RCTs and studies from the previous version of the guideline)*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(Full search strategy is detailed in the GDG 5 methodology appendix. Subject terms not translated between*

Section	Question	Answer
		databases. This strategy covers questions addressing a range of fertility technologies. Terms for a few are included but not all. It may have been considered that using subject headings for reproductive techniques/fertility may suffice but free text terms for specific technologies would be expected to be there e.g. aromatase inhibitors. No terms at all for ovarian surgery etc)
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(Date limits not mentioned in the protocol but it is clear that this updates a 2017 review)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(Studies screened by one reviewer with evidence team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(Study appraisal completed by 2 reviewers plus evidence team)*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(Study table with study characteristics present allowing interpretation of results)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(12 studies met inclusion criteria for review, all appear to be included in analysis)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(Detailed integrity assessment in GDG 5 methodology appendix for all*

Section	Question	Answer
		studies, individual quality appraisals also present)
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Yes *(2 reviewers plus evidence team)*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(All 12 included studies for section 5.5 have been included in the study characteristics table and appear to have been included across the 7 different comparisons. Between 2 and 7 studies were used for each subgroup/outcome analysis.)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(No information was given regarding the planned analysis of the results)*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(Meta-analysis was used appropriately for this review question however odds ratios were used where typically risk ratios would be used. Study design is appropriate for analysis)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Yes *(Heterogeneity appears to have been accounted for, there are detailed tables for each of the 7 outcomes which detail inconsistency for individual outcomes. Forest plots are also present for each of the 7 outcomes and all of the subgroup analyses within each outcome)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Yes *(Funnel plots available for each subgroup analyses, appears appropriate)*

Section	Question	Answer
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Yes *(Risk of bias assessed per outcome in section 5.5 and per study in GDG 5 methodology appendix, appears to have taken bias into account but further write up about this could have been useful RoB included in study characteristics table and GRADE assessment also completed)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(More detailed write up would be helpful to confirm conclusions, but otherwise ok)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(All areas appear to be well covered with relevant information)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(More detailed write up would be helpful to confirm conclusions, but otherwise ok)*
Overall review ratings	Overall risk of bias	Low *(Some issues with search terms and subject terms but otherwise ok)*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: 8 studies were included in this updated systematic review. Two were low, 2 moderate and 4 high risk of bias. No difference was seen for any of the 11 outcomes explored, except for

side effects which were higher with metformin than with the placebo. This was given a low GRADE rating, leading to one EBR which was a conditional weak recommendation for the option. Recommendation 5.7.5.1 states that adjunct metformin can be used during FSH ovarian stimulation as it may reduce the risk of OHSS. This is supported by the evidence which found there was a reduction in OHSS and miscarriage rate when adjunct metformin was used, however there was no difference in outcomes for clinical pregnancy rate per patient. It is appropriate that the recommendation specifies the use of adjunct metformin is to reduce the risk of OHSS and miscarriage so that it may be targeted to women who have higher risks in these areas. One practice point was also added to cover good practice. Implementation barriers were also raised due to the availability of metformin and the costs which may vary due to the healthcare setting. More adequately powered studies were recommended as a research priority to cover the areas of adjunct metformin treatment and when to cease metformin use.

IG economic evidence

No health economic evidence was identified in the IG for review question 5.7.5 on adjunct metformin

The IG noted that the cost of oral metformin therapy for 4 months is small, but the cost of managing OHSS can be moderate to large depending on severity OHSS and whether or not outpatient management or hospitalisation required. No further information was provided.

5.7.2 NICE economic evidence

Included studies

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

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

Excluded studies

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

Economic model

No original health economic model was developed for review question as other areas of the guideline where deemed a higher priority area for original health economic modelling, especially due to the fact there was limited clinical evidence in this area. The committee therefore concluded that cost-effectiveness could be considered through qualitative discussion of the evidence alongside the presentation of unit costs.

Unit costs

Table 3: Unit costs of metformin

Resource	Unit costs
Metformin 500mg	£0.03
Metformin 500mg modified release	£0.07
Metformin 1g	£0.22
Metformin 1g modified release	£0.03

Source of costs; British National Formulary
date accessed 29/04/2026. Dosing is typically up to a maximum dose of 2g a day.

5.7.3 NICE recommendations

No recommendations have been contextualised from this section of the IG.

5.7.4 The committee's discussion and interpretation of the evidence

Clinical

The committee discussed the use of adjunct metformin for women with PCOS undergoing IVF/ICSI. The committee felt there was insufficient evidence at this time to contextualise a recommendation from this section of the IG.

Health economic

No health economic evidence was identified in either the IG, or in NICE's health economic literature search, for review question 5.7.5 on adjunct metformin.

Although the IG made one evidenced-based recommendation in this area, noting adjunct metformin could be considered, the committee decided not to contextualise this recommendation. The committee acknowledged that the cost of metformin is relatively cheap and that OHSS can be expensive to treat. However, the committee concluded that there was insufficient clinical evidence to demonstrate clinical effectiveness and therefore infer cost-effectiveness based on the unit costs presented.

As no recommendation was made for this review, there are no associated resource implications for the NHS.

5.7 (6) In vitro maturation

Review question 5.7.6: In women with PCOS, is In Vitro Maturation (IVM) effective for improving fertility outcomes?

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

Evidence-based recommendation:

5.7.6.1 The use of in vitro maturation (IVM) and ICSI could be considered in women with PCOS, as an alternative to a stimulated IVF/ICSI cycle, where an embryo is frozen and replaced in a subsequent embryo transfer cycle, acknowledging there is no risk of ovarian hyperstimulation syndrome, but a lower cumulative live birth rate.

Consensus recommendation:

5.7.6.2 The use of IVM and ICSI could be considered prior to stimulated IVF/ICSI cycles acknowledging both benefits and limitations.

Practice point(s):

5.7.6.3 IVM should only be considered in services with sufficient expertise, and advocacy is needed for regional or national centres of expertise.

5.7.6.4 IVM could be offered as an option in women with prior severe ovarian hyperstimulation syndrome and where the risk of severe ovarian hyperstimulation syndrome is deemed unacceptably high, provided that expertise in IVM techniques exists.

5.7.6.5 Evidence suggests that IVM/ICSI is less effective than standard IVF/ICSI in terms of clinical pregnancy per patient and live birth rate per patient.

IG clinical evidence

Critical appraisal - ROBIS systematic review checklist

Section	Question	Answer
Study eligibility criteria	Did the review adhere to pre-defined objectives and eligibility criteria?	Yes *(Clear and detailed PICO)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Detailed PICO available which is suitable for the review)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(Restrictions on date are present but not well described)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(Study types selected are appropriate for review question, language restrictions are appropriate)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(No issues noted, methods appear appropriate for review question)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(Appropriate number (5) and range of databases used. Reasonable range but PsycINFO seems superfluous)*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably no

Section	Question	Answer
		(PRISMA flowchart mentions number of new systematic reviews, RCTs identified from those reviews, new RCTs and studies from the previous version of the guideline)
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	No *(Full search strategy is detailed in the GDG 5 methodology appendix. Subject terms not translated between databases. This strategy covers questions addressing a range of fertility technologies. Terms for a few are included but not all. It may have been considered that using subject headings for reproductive techniques/fertility may suffice but free text terms for specific technologies would be expected to be there e.g. aromatase inhibitors. No terms at all for ovarian surgery etc)*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(Date limits not mentioned in the protocol but it is clear that this updates a 2017 review)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(Studies screened by one reviewer with evidence team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(Study appraisal completed by 2 reviewers plus evidence team)*
Data collection and study appraisal	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Yes *(Study table with study characteristics present allowing interpretation of results)*

Section	Question	Answer
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(One study met the inclusion criteria for the review, and appears to be included in the analysis)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(Detailed integrity assessment in GDG 5 methodology appendix for all studies, individual quality appraisals also present)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Yes *(2 reviewers plus evidence team)*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(Methods used are appropriate to review question, no issues noted)*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(The one included study for section 5.7.6 has been included in the study characteristics table and has been included in the comparison)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	No information *(No information was given regarding the planned analysis of the results)*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(Meta-analysis was not possible, descriptive analysis was used for this review question, study design is appropriate for analysis however odds ratio was used instead of risk ratio.)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Yes *(Inconsistency not possible to determine as only one study included imprecision downgrade once due to this)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated	Yes

Section	Question	Answer
	through funnel plot or sensitivity analyses?	*(Funnel plots available for each subgroup analyses, appears appropriate)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Yes *(Risk of bias assessed per outcome in section 5.5 and per study in GDG 5 methodology appendix, appears to have taken bias into account but further write up about this could have been useful RoB included in study characteristics table and GRADE assessment also completed)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(Further written discussion of results would have been helpful, no other issues noted)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(No issues noted, methods appear appropriate for review question)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Lack of terms for specific technologies. Subject terms not translated between databases)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Methods used are appropriate to review question, no issues noted)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(Further written discussion of results would have been helpful, no other issues noted)*
Overall review ratings	Overall risk of bias	Low *(Some issues with subject and search terms, otherwise ok)*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: one study met the inclusion criteria for this new systematic review. The study had a low risk of bias with high quality methodology. Live birth rate per patient and clinical pregnancy rate per patient both improved with IVF compared to IVM, no difference was seen for the outcomes of live birth rate per pregnancy or miscarriage rate per pregnancy, however IVM avoids the risk of OHSS. The GRADE outcome for this comparison was moderate, downgraded once for imprecision due to evidence being from a single study. The EBR was a conditional weak recommendation for the option, as the panel felt that IVM was a valid treatment option for women with PCOS prior to IVF/ICSI, stating that IVM was safe, simple, less costly and still had reasonable fertility options. The four outcomes for fertility options all had only one study included with the recommendation being a conditional weak recommendation for either option. The recommendation appears to be appropriate based on the level of evidence found, as it states IVM and ICSI “could be used”, highlighting that it does have a lower cumulative live birth rate than IVF/ICSI. The previously experimental use of IVM was noted as a potential implementation consideration. The IG highlighted that future research should focus on the optimal IVM protocol for women with PCOS and look at any long-term health effects for their children.

IG economic evidence

No health economic was evidence was identified in the IG for review question 5.7.6 in the IG on In Vitro Maturation (IVM).

The IG noted that while IVM is no longer considered experimental, potential limitations in offering IVM are the costs of implementation and the expertise related to both the surgical procedure and the laboratory consumables and processes. The IG also noted that IVM might have lower cost compared to

stimulated IVF, however no additional information was provided to demonstrate this lower cost.

5.7.2 NICE economic evidence

Included studies

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

No economic studies were included 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 this review question.

5.7.3 NICE recommendations

No recommendations have been made or contextualised from this section of the IG.

5.7.4 The committee's discussion and interpretation of the evidence

Clinical

The committee discussed the use of IVM, however felt that this was sufficiently covered by the NICE fertility guideline, and as such did not require repeating here. As such no recommendations were made or adapted from this section of the IG.

Health economic

No health economic evidence was identified in either the IG, or in NICE's health economic literature search, for review question 5.9 on IVM.

The committee concluded that the topic IVM was sufficiently covered by NICE's fertility problems guideline – that being that IVF is the assisted reproductive technology of choice when compared to IVM.

No specific recommendations on IVM were made as part of this guideline. Therefore there are no significant resource implications associated with this review question.

5.8 Inositol

Review question 5.8: In adolescents and adults with PCOS, is inositol alone or in combination, effective for management of reproductive outcomes?

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

Evidence-based recommendation:

5.8.1 Inositol in any form alone, or in combination with other therapies, should be considered experimental therapy in women with PCOS with infertility, with benefits and risks currently too uncertain to recommend the use of these agents as fertility therapies.

Practice point:

5.8.2 There is limited evidence with uncertain results, on the effect of inositol on ovulation, clinical pregnancy and live birth rates.

5.8.3 Side-effects and safety are not known for inositol.

5.8.4 Women need to be aware that these agents can have limited regulation with variable dose, quality, consistency and combination with other agents.

5.8.5 Women's personal goals and preferences should be considered when discussing complimentary therapies.

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 *(No predefined criteria identified. PICO is well defined)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Clearly defined PICO. Appropriate to the review question)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Probably yes *(Restrictions on date are present but not well described)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Yes *(Study inclusion limited to English language. No publication date or format limits applied)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(Note that this research question was completed as part of the search for section 4.7 and as such methods and total included study numbers reflect both sections 4.7 and 5.8.* *All areas appear to be well covered with relevant information)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably yes *(Searches were carried out for 5 databases - Medline, PsycINFO, EMBASE, All EMB, CINAHL. Reasonable range of databases, PsycINFO seems superfluous)*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably no

Section	Question	Answer
		(Nothing in narrative and no other sources quoted in PRISMA diagram for 4.7)
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 strategy for 4.7 has been used to identify evidence relating to fertility outcomes. The structure of the search (inositol plus PCOS terms) would do this (no outcome terms used).* *However, the subject headings used in Embase have not been translated correctly e.g. the MeSH polycystic ovary syndrome corresponds to the Emtree heading of ovary polycystic disease used in Embase. The Mesh has been used in the Embase strategy. The CINAHL strategy has not been reported)*
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(No date limits applied as per the protocol)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Probably yes *(Studies screened by one reviewer with evidence team)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(CINAHL strategy not reported. Subject headings not translated correctly between databases)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Probably yes *(Does not clearly state how many reviewers conducted data extraction)*
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 sufficiently detailed)*

Section	Question	Answer
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Yes *(Appears suitable)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Probably yes *(ROB conducted using older version of Cochrane's ROB tool)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Probably yes *(Does not clearly states how many reviewers conducted quality appraisal)*
Data collection and study appraisal	Concerns regarding methods used to collect data and appraise studies	Low *(methods seem appropriate for review)*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(29 studies were included in the review and 19 in meta-analysis. Note that the evidence processing statement mentioned 26 studies were included, however the included studies table lists 29 separate studies, PRISMA diagram is not visible as half is missing. This covers the total number of studies included for both section 5.8 and 4.7 as the reviews were conducted together)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	Probably yes *(All analyses mentioned and addressed in the results section. All necessary analysis (narrative analyses and meta-analyses) was carried out to include all of the studies)*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(Study design and analysis are appropriate for review, however several outcomes have only single studies, and it is not clear from the*

Section	Question	Answer
		discussion whether these have been used to aid decision making.)
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Probably yes *(Appears to have been taken into account, outcomes with serious inconsistency were downgraded appropriately but this is not well described)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Probably yes *(No funnel plots or sensitivity analysis)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Probably yes *(GRADE assessment is carried out and biases are addressed, however, they are not addressed in the discussions)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(Appears to have taken bias into account but further write up about this could have been useful)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low *(All areas appear to be well covered with relevant information)*
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(CINAHL strategy not reported. Subject headings not translated correctly between databases)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(Appears to have taken bias into account but further write up about this could have been useful)*
Overall review ratings	Overall risk of bias	Low

Section	Question	Answer
		(Some issues with search terms and subject headings, this question was part of the search for section 4.7)
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: four comparisons were made for this question, Comparison 1: DCI vs placebo, Comparison 2: MI+FA vs FA Comparison 3: MI vs. Metformin Comparison 4: MI + DCI (higher dose) vs. MI + DCI (lower dose). All were either low or very low GRADE certainty. One EBR was made which was a conditional weak recommendation for the option. The only comparison to include fertility outcomes was for myoinositol plus folic acid compared to folic acid alone, which had high heterogeneity due to the different types of infertility interventions used. There was no difference in pregnancy rate between the two groups. Four practice points were also included to provide supporting information. Most studies were noted as having a high to moderate risk of bias due to their small study populations and heterogeneous nature. As the evidence for this question was mixed and lacked any clear conclusion, it is appropriate for the EBR to highlight the experimental nature of inositol for fertility issues in women with PCOS. No implementation considerations were noted. Future research priorities included determination of optimal formulations and side effects.

IG economic evidence

No health economic evidence was identified in the IG for review question 5.8 on inositol. The IG also noted that inositol is an over-the-counter product with low to moderate cost.

5.8.2 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 included which were applicable to this review question (see economic study selection flow chart in Appendix B).

Excluded studies

One health economic study was initially identified for this review question but was excluded due to a combination of applicability and methodological concerns.

See Appendix D of this document for a list of excluded economic studies, with reason for exclusion.

Economic model

No original health economic modelling was conducted for this review question.

Unit costs

Unit costs were not sourced for inositol as inositol is an over-the-counter supplement.

5.8.3 NICE recommendations

The relevant recommendation for this section is Rec 1.21.18.

5.8.4 The committee's discussion and interpretation of the evidence

Clinical

The committee discussed the use of inositol for fertility indications in women with PCOS. The committee felt that the wording in the IG around inositol was not directive and felt that the recommendation should be contextualised with clear language to ensure users are aware that inositol is not effective for fertility indications in women with PCOS.

Health economic

No health economic evidence was included in either the IG, or in NICE's health economic literature search, for review question 5.8 on inositol.

One health economic study was initially identified that was applicable to this review question but subsequently excluded due to a combination of applicability and methodological limitations. Further details on this study, including the reason for exclusion, can be found in **appendix C.**

The committee noted that inositol is an over-the-counter supplement and therefore no costs would be incurred to the NHS if a positive recommendation for inositol was made. The committee agreed with the IG's recommendation on inositol; agreeing that inositol is not an effective treatment for targeting fertility related problems for people with PCOS.

5.9 Anti-obesity pharmacological agents

Review question 5.9: Are anti-obesity pharmacological agents alone or in combination, effective for management of reproductive outcomes in adolescents and adults with PCOS?

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

Consensus based recommendation:

5.9.1 We recommend using anti-obesity agents in PCOS for reproductive outcomes only in research settings to establish the efficacy and safety.

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 *(PICO is well defined)*
Study eligibility criteria	Were the eligibility criteria appropriate for the review question?	Yes *(Clearly defined PICOS)*
Study eligibility criteria	Were eligibility criteria unambiguous?	Yes *(Eligibility criteria clearly described in PICO)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Yes *(Appropriate inclusion and exclusion criteria according to the research question)*
Study eligibility criteria	Were any restrictions in eligibility criteria based on sources of information	Yes

Section	Question	Answer
	appropriate (e.g. publication status or format, language, availability of data)?	*(This is a new SR. Limits are applied based on the language in the last 10 years)*
Study eligibility criteria	Concerns regarding specification of study eligibility criteria	Low *(All areas appear to be well covered with relevant information)*
Identification and selection of studies	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Probably no *(Searches were carried out for 4 databases - Medline, PsycINFO, EMBASE and CINAHL. No search of Cochrane databases)*
Identification and selection of studies	Were methods additional to database searching used to identify relevant reports?	Probably yes *(PRISMA shows 1 record from other sources. not clear what this is or whether it relates to the fertility outcomes)*
Identification and selection of studies	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Yes *(The search for 4.5 was used to identify evidence relating to fertility outcomes. The structure of this does not include outcome terms so this is sound. The CINAHL strategy has not been reported.* *Subject headings have not been translated from Medline to Embase. There is a comprehensive list of drugs included but this is restricted to generic names. However, terms for the general drug class/approach seem thin and underdeveloped.* *The floating subheading for drug therapy and the terms for publication types (and with PCOS and the drugs) would usually be used in an OR combination. Terms used in this section do not include ones for systematic reviews or crossover trials as mentioned in the protocol).*

Section	Question	Answer
Identification and selection of studies	Were restrictions based on date, publication format, or language appropriate?	Probably yes *(Whilst not following ROBIS guidance they reflect the protocol i.e. restricted to 10 years)*
Identification and selection of studies	Were efforts made to minimise error in selection of studies?	Yes *(Study selection was carried out by two reviewers)*
Identification and selection of studies	Concerns regarding methods used to identify and/or select studies	High *(Subject headings not translated between databases. Underdeveloped intervention section. Use of the drug therapy floating subheading)*
Data collection and study appraisal	Were efforts made to minimise error in data collection?	Yes *(Report suggests that 2 independent reviewers conducted data extraction)*
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 is sufficiently detailed)*
Data collection and study appraisal	Were all relevant study results collected for use in the synthesis?	Probably yes *(11 studies were included in the review question for section 4.5 which also includes review question 5.9. The methods used appear appropriate for review)*
Data collection and study appraisal	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Yes *(Suitable analysis for this question)*
Data collection and study appraisal	Were efforts made to minimise error in risk of bias assessment?	Yes *(2 independent reviewers appraised the studies alongside the reviewing team)*
Data collection and	Concerns regarding methods used to collect data and appraise studies	Low

Section	Question	Answer
study appraisal		*(Methods seem appropriate for review)*
Synthesis and findings	Did the synthesis include all studies that it should?	Yes *(All studies appear to be present)*
Synthesis and findings	Were all pre-defined analyses reported or departures explained?	Yes *(All analyses mentioned and addressed in the results section. All necessary analysis (narrative analyses and meta-analyses) was carried out to include all of the studies. Meta analysis was done for 4 out of 11 studies, reasons for not including rest of the studies is explained in the report)*
Synthesis and findings	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Yes *(Combined with section 4.5, appears appropriate)*
Synthesis and findings	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Probably yes *(Narrative synthesis was carried out where possible which addressed heterogeneity as statistical combination was not possible.)*
Synthesis and findings	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Probably yes *(Results combined with section 4.5, appears appropriate)*
Synthesis and findings	Were biases in primary studies minimal or addressed in the synthesis?	Yes *(GRADE assessment is carried out and biases are addressed in the evidence summary section)*
Synthesis and findings	Concerns regarding the synthesis and findings	Low *(Limited details on how conclusions reached but synthesis seems appropriate)*
Judging risk of bias	Concerns regarding specification of study eligibility	Low

Section	Question	Answer
		(All areas appear to be well covered with relevant information)
Judging risk of bias	Concerns regarding methods used to identify and/or select studies	High *(Subject headings not translated between databases. Underdeveloped intervention section. Use of the drug therapy floating subheading)*
Judging risk of bias	Concerns regarding methods used to collect data and appraise studies	Low *(Methods seem appropriate for review)*
Judging risk of bias	Concerns regarding the synthesis and findings	Low *(Limited details on how conclusions reached but synthesis seems appropriate)*
Overall review ratings	Overall risk of bias	Low *(Some issues with search and subject terms, further discussion on results would have been helpful. The searches for this section were part of section 4.5, with outcomes for section 5.9 extracted and reported here)*
Overall review ratings	Applicability as a source of data	Fully applicable

IG evidence to recommendations justification: one CR was made for this review question; however, it is listed as a conditional weak recommendation for either option. None of the study outcomes for section 4.5 listed any fertility outcomes, as such it is suitable that no EBR was made for this review question, however the CR is worded quite strongly when there is no available evidence and no details of the consensus discussion to explain why the recommendation was made. No implementation or research considerations were noted.

IG economic evidence

No health economic evidence was identified in the IG for review question 5.9 on anti-obesity agents.

5.9.2 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 therefore excluded from this review.

Economic model

No original health economic modelling was conducted for this review question as for this review question the committee cross-referred to NICE's guideline on overweight and obesity management (NG246).

5.9.3 NICE recommendations

The relevant recommendation for this section is Rec 1.21.19.

5.9.4 The committee's discussion and interpretation of the evidence

Clinical

The committee discussed the impact of weight loss medications on fertility outcomes and agreed that this was currently uncertain and not well supported by the evidence. The committee decided to contextualise the IG CR which

states that for fertility outcomes alone, that anti-obesity mediations should not be used outside of a research setting. The committee did however highlight that anti-obesity medication is a highly evolving area of research, noting that more research might now be available given the increased use of GLP-1 medications. The committee discussed the cost of GLP-1 medication compared to the cost of medications that induce ovulation, however it was questioned whether the GLP-1 medications improved ovulation, or whether it was a consequence of the weight loss achieved. The committee opted to include a cross referral to the NICE guideline on overweight and obesity management (NG246), which highlights (section 1.17) that tirzepatide, semaglutide and liraglutide should not be used in pregnancy or in women of childbearing potential without the use of effective contraception. NICE guideline NG246 was unable to comment on the use of the GLP-1 medications for the purpose of fertility as the evidence was insufficient for specific subgroups, including for fertility outcomes. The committee agreed that until further research was made available, it was sensible to contextualise the recommendation from the IG.

Health economic

No health economic evidence was identified in either the IG, or in NICE's health economic literature search, for review question 5.11 on anti-obesity pharmacological agents for fertility outcomes.

As no positive recommendation was made for the use of anti-obesity pharmacological agents for fertility outcomes, no resource implications for the NHS are associated with the recommendations made as part of this guideline.

1

2

3

4

5

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.

1 10 or/1-9 55812

2 11 Economics/

3 12 Value of life/

4 13 exp "Costs and Cost Analysis"/

5 14 exp Economics, Hospital/

6 15 exp Economics, Medical/

7 16 Economics, Nursing/

8 17 Economics, Pharmaceutical/

9 18 exp "Fees and Charges"/

10 19 exp Budgets/

11 20 budget*.ti,ab.

12 21 cost*.ti.

13 22 (economic* or pharmaco?economic*).ti.

14 23 (price* or pricing*).ti,ab.

15 24 (cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or
16 variable*)).ab.

17 25 (financ* or fee or fees).ti,ab.

18 26 (value adj2 (money or monetary)).ti,ab.

19 27 or/11-26

20 28 10 and 27

1 29 letter.pt. or letter/

2 30 note.pt.

3 31 editorial.pt.

4 32 case report/ or case study/

5 33 (letter or comment*).ti.

6 34 or/29-33

7 35 randomized controlled trial/ or random*.ti,ab.

8 36 34 not 35

9 37 animals/

10 38 exp Animals, Laboratory/

11 39 exp Animal Experimentation/

12 40 exp Models, Animal/

13 41 exp Rodentia/

14 42 (rat or rats or mouse or mice or rodent*).ti.

15 43 or/37-42

16 44 43 not humans/

17 45 36 or 44

18 46 28 not 45

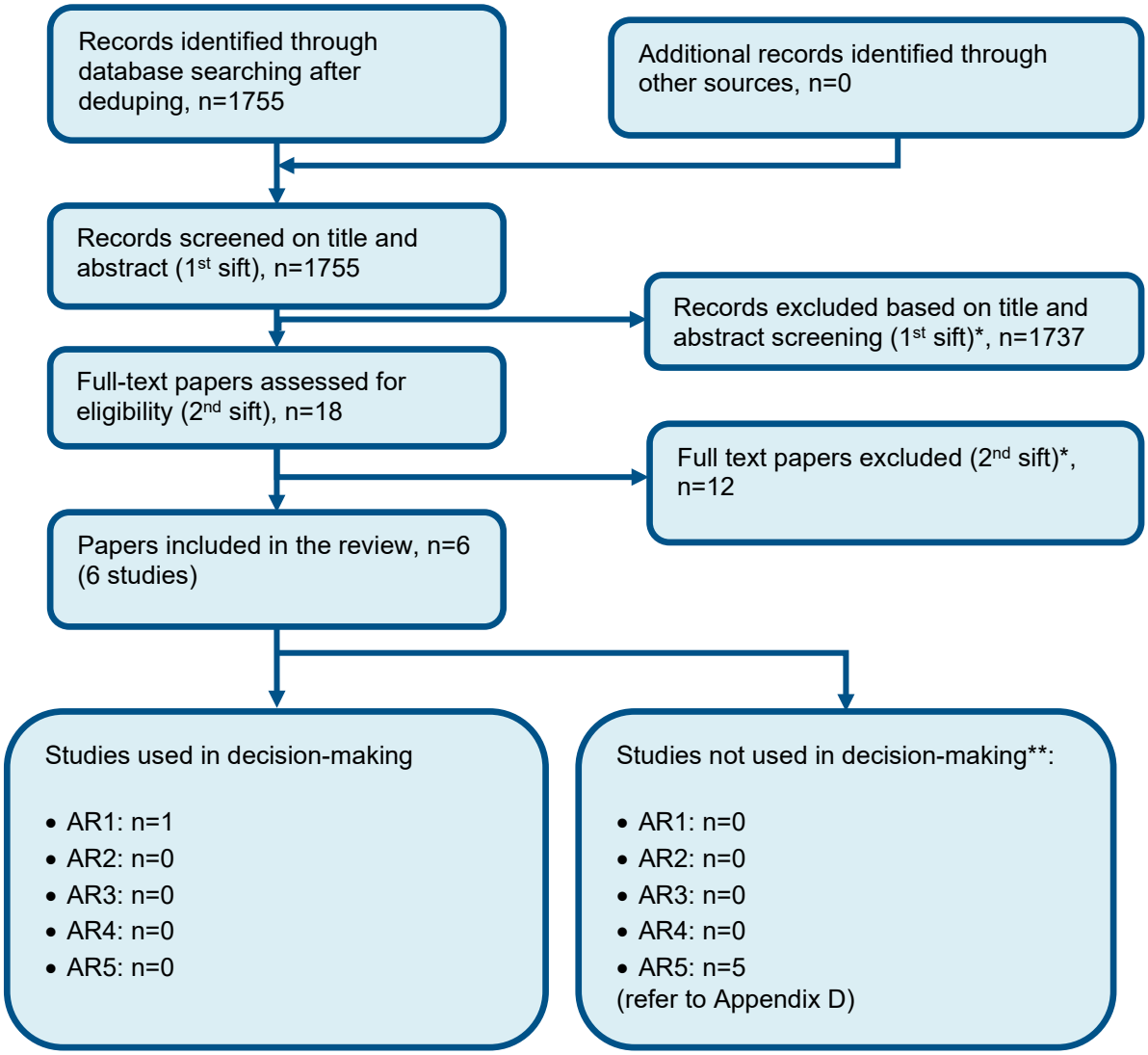
19 47 limit 46 to english language/

20 48 limit 47 to ed=20090101-20241

1 49 limit 47 to dt=20090101-20241205

2 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

Table 4: Studies excluded from the economic review for Section 5

Study	Reason for exclusion
Beresniak 2023	CEA comparing Myo-inositol and gonadotrophin to gonadotrophin alone. Excluded due to a combination of applicability and methodological limitations. This study was assessed as partially applicable (no QALYs reported; Italian setting does not reflect current NHS context) and judged to have very serious limitations (effectiveness evidence either from indirect population or for incorrect comparator; effectiveness studies not included in International Guideline; unclear how effectiveness probabilities used in the model were calculated from clinical evidence; some authors are employees of myo-inositol manufacturer).
De Frene 2015	CEA comparing laparoscopic ovarian surgery to gonadotrophin. Excluded due to a combination of applicability and methodological limitations. This study was assessed as partially applicable (no QALYs reported; Belgian setting does not reflect current NHS context) and judged to have very serious limitations (Resource use from 2000 to 2009 and unit costs from 2007 does not reflect current NHS context; effectiveness based on retrospective study not included in International Guideline).
Moolenaar 2014	CEA comparing multiple ovulation induction strategies including in vitro fertilisation, clomiphene citrate and gonadotrophin, alone or sequentially. Excluded due to a combination of applicability and methodological limitations. This study was assessed as partially applicable (no QALYs reported; Dutch setting does not reflect current NHS context; no discounting despite 2-year time horizon) and judged to have very serious limitations (Unit costs and resource use pre 2010 and does not reflect current NHS context; effectiveness based on different studies to those included in International

Study	Reason for exclusion
	Guideline as well as assumption and a prospective cohort study).
Nahuis 2012	CEA comparing laparoscopic ovarian surgery to gonadotrophin. Excluded due to a combination of applicability and methodological limitations. This study was assessed as partially applicable (no QALYs reported; Dutch setting does not reflect current NHS context; 5% discounting applied) and judged to have very serious limitations (Unit costs and resource use from 2009 and does not reflect current NHS context; effectiveness based on a single study Bayram 2004 which was included in International Guideline but economic analysis was based on a post randomisation follow up that is not reported in the original RCT and therefore may not reflect full body of evidence from the International Guideline).
Piróg 2026	CEA comparing primary 3 cycles of ovulation induction with hysterosalpingo-foam sonography followed by 3 cycles of ovulation induction if at least one tube was patent. Excluded due to a combination of applicability and methodological limitations. This study was assessed as partially applicable (no QALYs reported; Polish setting and costs presented in US dollars which does not reflect current NHS context; no discounting reported) and judged to have very serious limitations (Unit costs and resource use poorly reported and only reference from 2005 therefore does not reflect current NHS context; effectiveness based on a single study, included in this paper, which compared tubal patency testing in women with and without PCOS, which is outside of the protocol for this question, finally this study was not included in International Guideline and therefore may not reflect full body of evidence from the International Guideline).

Abbreviations: CEA: cost-effectiveness analysis, NHS: National Health Service, QALY: quality adjusted life year, RCT: randomised control trial.

Appendix E Health economic model

No original health economic modelling was undertaken as part of the contextualisation process for section 5 of the IG.

Appendix F References

Beresniak, Ariel, Russo, Michele, Forte, Gianpiero et al. (2023) A Markov-model simulation of IVF programs for PCOS patients indicates that coupling myo-Inositol with rFSH is cost-effective for the Italian Health System. Scientific reports 13(1): 17789

De Frene, Veerle, Gerris, Jan, Weyers, Steven et al. (2015) Gonadotropin Therapy versus Laparoscopic Ovarian Drilling in Clomiphene Citrate-Resistant Polycystic Ovary Syndrome Patients: A Retrospective Cost-Effectiveness Analysis. Gynecologic and obstetric investigation 80(3): 164-9

Moolenaar, Lobke M, Nahuis, Marleen J, Hompes, Peter G et al. (2014) Cost-effectiveness of treatment strategies in women with PCOS who do not conceive after six cycles of clomiphene citrate. Reproductive biomedicine online 28(5): 606-13

Nahuis, M J, Oude Lohuis, E, Kose, N et al. (2012) Long-term follow-up of laparoscopic electrocautery of the ovaries versus ovulation induction with recombinant FSH in clomiphene citrate-resistant women with polycystic ovary syndrome: an economic evaluation. Human reproduction (Oxford, England) 27(12): 3577-82

Pirog, Magdalena; Chrostowski, Bartosz; Jach, Robert (2026) Tubal patency testing in women with polycystic ovary syndrome - is it worth before ovarian induction: Cost-effectiveness analysis. European Journal of Obstetrics, Gynecology, & Reproductive Biology 318: 114888

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