

National Institute for Health and Care Excellence

1

2

Polycystic ovary syndrome (PCOS)

3

4

NICE guideline NGXXX

Methods

5

July 2026

6

7

Draft

8

9

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 All NICE guidance is subject to regular review and may be updated or withdrawn.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

© NICE 2026 All rights reserved. Subject to [Notice of rights](#).

ISBN: [XXX]

Contents

1		
2	Development of the guideline	4
3	Remit	4
4	Methods and process	4
5	Contextualisation	4
6	Reviewing economic evidence.....	9
7	Updating the guideline	13
8	Glossary of terms	14
9	References.....	21
10	Appendix A Appraising the quality of external guideline	22
11	First stage assessment using the AGREE II instrument:	22
12	AGREE II results.....	23
13	Second stage assessment – risk of bias using ROBIS assessment tool	25
14	Appendix B Economic review protocol	28

1 **Development of the guideline**

2 **Remit**

3 To see “What this guideline covers” and “What this guideline does not cover” please
4 see the [NICE guideline scope for Polycystic ovary syndrome \(PCOS\)](#).

5 **Methods and process**

6 The guideline was developed using the methods and process described in
7 [Developing NICE guidelines: the manual](#).

8 Declarations of interest were recorded according to [NICE's conflicts of interest policy](#).

9 **Contextualisation**

10 The NICE recommendations were developed through a new NICE process called
11 contextualisation, to improve timeliness of guidelines and avoid duplication of work.
12 Contextualisation makes use of externally produced guidance or recommendations,
13 either by adapting relevant evidence reviews to formulate NICE recommendations or
14 contextualise external recommendations to fit NICE’s requirements. The external
15 guideline and its recommendations must be publicly available and original (not
16 derived from another guideline). In this instance the International evidence-based
17 Guideline for the assessment and management of polycystic ovary syndrome
18 (PCOS) (IG) conducted by the Monash Centre for Health Research and
19 Implementation was identified at the pre-scoping phase of the NICE guideline and
20 their recommendations contextualised for the NICE Guideline on PCOS. In
21 discussion with the clinical Topic Advisor a scope was produced following the
22 standard NICE approach described in [Developing NICE guidelines: the manual](#), to
23 include all topic areas from the International Guideline. All systematic reviews
24 included in the IG were assessed for their risk of bias. The recommendations were
25 assessed for their applicability to a UK context and contextualised, with expert
26 Guideline Committee consensus, into NICE recommendations. The Guideline
27 Committee (and co-opted members) included patient lay members, gynaecologist,
28 obstetrician, pharmacist, dermatologist, dietitian, general practitioners, fertility
29 specialists, and consultant endocrinologists, Reviewing clinical evidence

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

Overlap with existing NICE guidance

During development of the guideline, any overlap with existing NICE guidance was identified. Where NICE guidance covered the same areas and was relevant to people with PCOS, it took precedence over the International Guideline. In such cases, a cross-reference to the existing NICE guidance was included.

Review protocols

Review protocols outline the inclusion and exclusion criteria used to select studies for each evidence review. As part of the contextualisation process the review protocols were included from the International Guideline. No further protocols were developed by NICE.

Searching for evidence

The evidence came from searches conducted by the International Guideline. No further searches for the clinical reviews were completed. NICE Information Scientists assessed the search criterion from the International Guideline and completed the ROBIS assessment for Identifying evidence.

Selecting studies for inclusion

NICE accepted the studies included by the International Guideline.

Appraisal of the quality of the external guideline

The first step in appraising the quality of the external guideline involved using the [Appraisal of Guidelines for Research and Evaluation II \(AGREE II\) instrument](#) to evaluate the International Guideline. The AGREE II instrument provides an assessment of the robustness of the process used to develop the guideline, see appendix A for the completed Agree II score sheet for the IG. The International Guideline was then contextualised in line with the [NICE guideline manual](#), section 8.4.

Assessment of the risk of bias in evidence reviews

1 When an external guideline has been assessed to be of suitable quality, the evidence
2 reviews informing the recommendations should be assessed for risk of bias.
3 Evidence reviews are assessed using the [ROBIS critical appraisal tool for systematic](#)
4 [reviews](#) following the standard NICE approach, including quality assurance by a
5 Senior Technical Analyst or Technical Adviser. The Information Specialist team were
6 also involved with the ROBIS assessment for the section relevant to the literature
7 searching undertaken (identification and selection of studies). The International
8 Guideline evidence reviews were appraised using the following four domains in
9 ROBIS: eligibility criteria; identification and selection of studies; data collection and
10 study appraisal; and synthesis and findings. An overall assessment of risk of bias is
11 then conducted, and all appraisals are included in the relevant adaptation reports.
12 For more details on the domains and questions within ROBIS see [Appendix A](#).

13 Where evidence reviews used different methods to NICE's standard approach (e.g.
14 non-English papers, search strategies, data analysis), This was assessed by the
15 technical team and discussed with the committee to determine whether this would
16 have significant impact on the results and whether this reduced the confidence of the
17 committee in the evidence being presented. For example, the risk of bias
18 assessment often indicated concerns related to study identification and selection
19 methods due to limitations in the search strategy. However, the committee reviewed
20 the IG technical reports and confirmed that no important studies appeared to be
21 missing. The GC concluded that it is unlikely any additional relevant evidence from
22 primary studies would alter the reviews' findings, that no important studies were
23 overlooked and this would not be a reason to exclude that part of the guideline.

24 **Justification for the recommendations**

25 Following ROBIS assessment the evidence to recommendations from the
26 International Guideline were appraised. This involved assessing whether the
27 description of the process for moving from evidence to any recommendations (or for
28 not making recommendations) was provided; whether a clear rationale has been
29 provided to justify the recommendation and strength of recommendation, based on
30 the evidence that underpins it; a clear description of the consensus process used,
31 and a rationale have been provided where no evidence has been identified, and a

1 consensus recommendation has been made appropriately. This is detailed in the
2 individual Adaptation reports.

3 **Developing NICE recommendations from external guidelines**

4 A committee is usually recruited to develop the NICE guideline following the standard
5 approach described in [Developing NICE guidelines: the manual](#). The International
6 guideline recommendations, the ROBIS assessment, and for some areas required for
7 health economics a summary of the evidence, and the evidence to recommendations
8 decision was presented to the committee.

9 **Writing up and validation**

10 The resultant discussion and a rationale for how and why the recommendations were
11 made in relation to the International Guideline was provided under ‘committee
12 discussion’ for each question. This included whether the committee agreed with the
13 recommendations from the IG and the justification they gave based on the evidence;
14 whether there was overlap with other NICE guidelines which took precedence and
15 were cross-referred to or whether the recommendations were made through
16 committee consensus. Any deviations from the International Guideline were
17 explained in the committee discussion.

18 **The International Guideline recommendation format**

19 The recommendations in the International Guideline consist of evidence-based and
20 consensus recommendations and practice points. Evidence-based recommendations
21 (EBR) are those that had evidence to support them, consensus recommendations
22 (CR) are based on the International Guidelines’ committee expertise and practice
23 points are those which arose from the International Guideline committee discussion
24 of the evidence and EBR and CR recommendations. The NICE recommendations
25 followed their normal format. Thus evidence-based and consensus recommendations
26 from the International Guideline were considered for inclusion in the NICE guideline.
27 Practice points were not included unless the guideline committee thought they were
28 particularly relevant and in line with current UK practice.

29 **Contextualisation of the International guideline’s recommendations**

- 1 Where possible recommendations were used in their entirety and only amended for
2 NICE style, unless there are concerns the recommendations:
- 3 • are likely to have a high resource impact (see below for health economic analysis)
 - 4 • not applicable to the settings within the guideline scope
 - 5 • is for something not currently available on the NHS, or implementable in the NHS
 - 6 • may not be applicable to the guideline population (e.g. specific population,
7 policies)
 - 8 • health inequality issues
 - 9 • contradicts another NICE recommendation (including Technical Appraisal (TA)
10 recommendations or guidance in development)
 - 11 • and there are significant concerns about the interpretation of the evidence
12 informing the recommendation.

1 **Reviewing economic evidence**

2 **Identifying and analysing evidence of cost effectiveness**

3 The committee is required to make decisions based on the best available evidence of
4 effectiveness and cost effectiveness. Guideline recommendations should be based
5 on the expected costs of the different options in relation to their expected benefits
6 (that is, their 'cost effectiveness') rather than the total implementation cost. However,
7 as the cost of implementation increases, the committee needs to be increasingly
8 confident in the cost effectiveness of a recommendation. Recommendations that are
9 expected to have a significant impact on resources (as defined in the [NICE](#)
10 [Assessing resource impact process manual](#)) need to be supported by robust
11 evidence on effectiveness and cost effectiveness; any uncertainties must be offset by
12 a compelling argument in favour of the recommendation. However, the cost impact or
13 savings potential of a recommendation should not be the sole reason for the
14 committee's decision ([Developing NICE Guidelines: the manual](#))

15 Health economic evidence was gathered by:

- 16 • Undertaking systematic reviews of published economic literature.
- 17 • Conducting new cost-effectiveness analysis in priority areas.

18 **Identifying economic evidence**

19 The International Guideline conducted a systematic review which included a search
20 for health economic evidence. No studies were identified. NICE conducted its own
21 systematic review of economic literature for the whole guideline using the PCOS
22 population and a health economic filter. Titles and abstracts of articles identified
23 through the systematic economic literature searches were assessed for inclusion
24 using predefined eligibility criteria reported in the economic review protocol (provided
25 in Appendix B of this document).

26 Once the screening of titles and abstracts was completed, full-text copies of
27 potentially relevant articles were acquired for detailed assessment, applying the
28 economic review protocol inclusion and exclusion criteria.

1 Details of economic evidence study selection and lists of excluded studies after full-
 2 text assessment (together with reasons for exclusion) are presented in respective
 3 appendices of the adaptation reports.

4 **Appraising the quality of economic evidence**

5 The applicability and methodological quality of economic evidence derived either
 6 from published studies meeting the inclusion criteria or from new economic analysis
 7 conducted for the guideline was assessed using the economic evaluations checklist
 8 specified in [Developing NICE guidelines: the manual, Appendix H](#). This process led
 9 to applicability and quality statements for each included study, made by the health
 10 economist, following the criteria shown in Table 1.

11 **Table 1: Criteria for developing applicability and quality statements of**
 12 **economic evidence**

Appraised element	Statement and criteria
Applicability	<ul style="list-style-type: none"> • Directly applicable – the study meets all applicability criteria or fails to meet 1 or more applicability criteria but this is unlikely to change the conclusions about cost effectiveness. • Partially applicable – the study fails to meet 1 or more applicability criteria, and this could change the conclusions about cost effectiveness. • Not applicable – the study fails to meet 1 or more of the applicability criteria, and this is likely to change the conclusions about cost effectiveness. Such studies would usually be excluded from the review.
Quality	<ul style="list-style-type: none"> • Minor limitations – the study meets all quality criteria, or fails to meet 1 or more quality criteria, but this is unlikely to change the conclusions about cost effectiveness. • Potentially serious limitations – the study fails to meet 1 or more quality criteria, and this could change the conclusions about cost effectiveness. • Very serious limitations – the study fails to meet 1 or more quality criteria, and this is highly likely to change the conclusions about cost effectiveness. Such studies would usually be excluded from the review.

13

14 All studies that fully or partially met the applicability and quality criteria described in
 15 the methodology checklist were considered by the committee during the guideline
 16 development process. However, high-quality studies in line with the NICE reference

1 case (such as recent UK NHS/PSS cost-utility analyses using the QALY as the
2 measure of outcome) were prioritised for use in decision making as these are the
3 most relevant to the NICE guideline development context. Therefore, not all
4 economic studies meeting the inclusion criteria were necessarily used in decision-
5 making. Further criteria for prioritising economic studies for use in decision making
6 are provided in the economic review protocols (under 'Review strategy'), within the
7 respective appendices of the adaptation reports.

8 Details on methods and results of economic studies that met inclusion criteria and
9 were subsequently used in decision making are shown in economic evidence study
10 extraction tables, provided in respective appendices of the adaptation reports.

11 Economic evidence study extraction tables are followed by lists of economic studies
12 that met inclusion criteria but were not used in decision making (with reasons for non-
13 use).

14 Characteristics and results (cost-effectiveness estimates) of economic studies used
15 in decision making, including applicability and quality statements, have been
16 summarised in economic evidence characteristics and summary tables, respectively,
17 provided in the economic sections of adaptation reports, as relevant.

18 **New economic analysis**

19 New cost-effectiveness analysis, based on decision-analytic modelling, was
20 undertaken by the guideline health economist in topic areas prioritised by the
21 committee. The rationale for prioritising topic areas and/or specific review questions
22 for economic modelling was set out in an economic plan agreed between members
23 of the NICE technical team developing the guideline, the committee, and members of
24 the NICE team quality assuring the guideline. Economic modelling was undertaken in
25 areas with likely major resource implications, where the current extent of uncertainty
26 over cost effectiveness was significant and economic analysis was expected to
27 reduce this uncertainty. The following economic questions were selected as key
28 issues to be addressed by economic modelling in the guideline:

- 29 • What is the cost effectiveness of anti-mullerian hormone compared to transvaginal
30 ultrasound for the identification of PCOM when required for a diagnosis of PCOS

- 1 • What is the cost effectiveness of alexandrite high fluence laser hair reduction
2 compared to no treatment for adults with PCOS and hirsutism

3 The following general principles were adhered to in developing the guideline cost-
4 effectiveness analyses:

- 5 • Methods were consistent with the NICE reference case for interventions with
6 health outcomes in NHS settings.
- 7 • The committee was involved in the design of the model and related assumptions,
8 selection of inputs and interpretation of the results.
- 9 • Model inputs were based on the systematic review of the effectiveness literature
10 supplemented with other published data sources where needed.
- 11 • When published data were not available for a model input, the committee's expert
12 opinion was used to inform it.
- 13 • Model inputs and assumptions were reported fully and transparently.
- 14 • The results were subject to sensitivity analysis and limitations were discussed.
- 15 • The model was peer-reviewed by another health economist who was independent
16 of the guideline development process.

17 Full methods and results of the cost-effectiveness analyses are described in
18 economic supplements A and B for anti-mullerian hormone and laser therapy
19 respectively.

20 **Cost effectiveness criteria**

21 NICE's [principles](#) set out criteria that committees should consider when judging
22 whether an intervention offers good value for money. In general, an intervention was
23 considered to be cost effective if any of the following criteria applied (provided that
24 the estimate was considered plausible):

- 25 • the intervention dominated other relevant strategies (that is, it was both less costly
26 in terms of overall resource use and more effective compared with all other
27 relevant alternative strategies)
- 28 • the intervention cost less than £20,000 per QALY gained compared with the next
29 best strategy.

1 If the committee recommended an intervention that was estimated to cost more than
2 £20,000 per QALY gained, or did not recommend one that was estimated to cost less
3 than £20,000 per QALY gained, then the reasons for this decision were provided and
4 recorded, with reference to issues around the plausibility of the estimate or to other
5 factors, for example the degree of uncertainty around the ICER, aspects that relate to
6 uncaptured benefits and non-health factors, or aspects that relate to health
7 inequalities, as set out in the [NICE health technology evaluations manual](#).

8 When an economic analysis used an outcome measure other than the QALY or the
9 number of life years gained, results were difficult to interpret unless one strategy
10 dominated the others, and a judgment needed to be made by the committee
11 regarding the monetary value placed on one unit of the outcome measure of the
12 analysis.

13 When new economic evidence was not available and new economic analysis was not
14 prioritised, the committee made a qualitative judgement about cost effectiveness by
15 considering expected differences in resource use and/or related UK NHS unit costs
16 between options, alongside respective effectiveness evidence. Where possible,
17 relevant UK NHS unit costs related to the compared interventions were presented to
18 the committee (and listed under a 'Unit costs' section in the respective evidence
19 review) to inform the possible economic implications of the recommendations.

20 The committee's considerations of cost effectiveness are discussed explicitly in the
21 section 'Committee discussion and interpretation of the evidence' under the
22 subheading 'Resources and cost-effectiveness', in each evidence review.

23 **Updating the guideline**

24 Following publication, and in accordance with the NICE guidelines' manual, NICE will
25 undertake a review of whether the IG has changed significantly to alter the guideline
26 recommendations and warrant an update.

1 **Glossary of terms**

2 Abstract: Summary of a study, which may be published alone or as an introduction to
3 a full scientific paper.

4 Applicability: How well the results of a study or NICE evidence review can answer a
5 clinical question or be applied to the population being considered.

6 Arm: Subsection of individuals within a study who receive one particular intervention,
7 for example placebo arm.

8 Association: Statistical relationship between 2 or more events, characteristics or
9 other variables. The relationship may or may not be causal.

10 Baseline: The initial set of measurements at the beginning of a study (after run-in
11 period where applicable), with which subsequent results are compared.

12 Bayesian analysis: A method of statistics, where a statistic is estimated by combining
13 established information or belief (the 'prior') with new evidence (the 'likelihood') to
14 give a revised estimate (the 'posterior').

15 Bias: Influences on a study that can make the results look better or worse than they
16 really are. (Bias can even make it look as if a treatment works when it does not.) Bias
17 can occur by chance, deliberately or as a result of systematic errors in the design
18 and execution of a study. It can also occur at different stages in the research
19 process, for example, during the collection, analysis, interpretation, publication or
20 review of research data. For examples see selection bias, performance bias,
21 information bias, confounding factor, and publication bias.

22 Carer (caregiver): Someone who looks after family, partners or friends in need of
23 help because they are ill, frail or have a disability.

24 Cochrane: The Cochrane Library consists of a regularly updated collection of
25 evidence-based medicine databases including the Cochrane Database of Systematic
26 Reviews (reviews of randomised controlled trials prepared by the Cochrane
27 Collaboration).

- 1 Confidence interval (CI): A range of values for an unknown population parameter with
2 a stated 'confidence' (conventionally 95%) that it contains the true value. The interval
3 is calculated from sample data and generally straddles the sample estimate. The
4 'confidence' value means that if the method used to calculate the interval is repeated
5 many times, then that proportion of intervals will actually contain the true value.
- 6 Confounding factor: Something that influences a study and can result in misleading
7 findings if it is not understood or appropriately dealt with. For example, a study of
8 heart disease may look at a group of people that exercises regularly and a group that
9 does not exercise. If the ages of the people in the 2 groups are different, then any
10 difference in heart disease rates between the 2 groups could be because of age
11 rather than exercise. Therefore, age is a confounding factor.
- 12 Control group: A group of people in a study who do not receive the treatment or test
13 being studied. Instead, they may receive the standard treatment (sometimes called
14 'usual care') or a dummy treatment (placebo). The results for the control group are
15 compared with those for a group receiving the treatment being tested. The aim is to
16 check for any differences. Ideally, the people in the control group should be as similar
17 as possible to those in the treatment group, to make it as easy as possible to detect
18 any effects due to the treatment.
- 19 Cost-effectiveness analysis (CEA): Cost-effectiveness analysis is one of the tools
20 used to carry out an economic evaluation. The benefits are expressed in non-
21 monetary terms related to health, such as symptom-free days, heart attacks avoided,
22 deaths avoided or life years gained (that is, the number of years by which life is
23 extended as a result of the intervention).
- 24 Cost-effectiveness model: An explicit mathematical framework, which is used to
25 represent clinical decision problems and incorporate evidence from a variety of
26 sources in order to estimate the costs and health outcomes.
- 27 Credible interval (CrI): The Bayesian equivalent of a confidence interval.
- 28 Economic evaluation: An economic evaluation is used to assess the cost
29 effectiveness of healthcare interventions (that is, to compare the costs and benefits

1 of a healthcare intervention to assess whether it is worth doing). The aim of an
2 economic evaluation is to maximise the level of benefits – health effects – relative to
3 the resources available. It should be used to inform and support the decision-making
4 process; it is not supposed to replace the judgement of healthcare professionals.
5 There are several types of economic evaluation: cost–benefit analysis, cost–
6 consequences analysis, cost-effectiveness analysis, cost minimisation analysis and
7 cost–utility analysis. They use similar methods to define and evaluate costs, but differ
8 in the way they estimate the benefits of a particular drug, programme or intervention.

9 **Effect:** A measure that shows the magnitude of the outcome in one group compared
10 with that in a control group. For example, if the absolute risk reduction is shown to be
11 5% and it is the outcome of interest, the effect size is 5%. The effect size is usually
12 tested, using statistics, to find out how likely it is that the effect is a result of the
13 treatment and has not just happened by chance (that is, to see if it is statistically
14 significant).

15 **Effectiveness:** How beneficial a test or treatment is under usual or everyday
16 conditions, compared with doing nothing or opting for another type of care.

17 **Evidence:** Information on which a decision or guidance is based. Evidence is
18 obtained from a range of sources including randomised controlled trials,
19 observational studies, expert opinion (of clinical professionals or patients).

20 **Exclusion criteria:** Explicit standards used to decide which studies should be
21 excluded from consideration as potential sources of evidence.

22 **GRADE:** A system developed by the GRADE Working Group to address the
23 shortcomings of present grading systems in healthcare. The GRADE system uses a
24 common, sensible and transparent approach to grading the quality of evidence. The
25 results of applying the GRADE system to clinical trial data are displayed in a table
26 known as a GRADE evidence profile.

27 **Hazard Ratio:** The hazard or chance of an event occurring in the treatment arm of a
28 study as a ratio of the chance of an event occurring in the control arm over time.

- 1 Health economics: Study or analysis of the cost of using and distributing healthcare
2 resources.
- 3 Health-related quality of life (HRQoL): A measure of the effects of an illness to see
4 how it affects someone's day-to-day life.
- 5 Heterogeneity: The term is used in meta-analyses and systematic reviews to
6 describe when the results of a test or treatment (or estimates of its effect) differ
7 significantly in different studies. Such differences may occur as a result of differences
8 in the populations studied, the outcome measures used or because of different
9 definitions of the variables involved. It is the opposite of homogeneity.
- 10 Imprecision: Results are imprecise when studies include relatively few patients and
11 few events and thus have wide confidence intervals around the estimate of effect.
- 12 Inclusion criteria: Explicit criteria used to decide which studies should be considered
13 as potential sources of evidence.
- 14 Incremental cost effectiveness ratio (ICER): The difference in the mean costs in the
15 population of interest divided by the differences in the mean outcomes in the
16 population of interest for one treatment compared with another.
- 17 Indirectness: The available evidence is different to the review question being
18 addressed, in terms of PICO (population, intervention, comparison and outcome).
- 19 Intervention: In medical terms this could be a drug treatment, surgical procedure,
20 diagnostic or psychological therapy. Examples of public health interventions could
21 include action to help someone to be physically active or to eat a healthier diet.
- 22 Life years gained: Mean average years of life gained per person as a result of the
23 intervention compared with an alternative intervention.
- 24 Meta-analysis: A method often used in systematic reviews. Results from several
25 studies of the same test or treatment are combined to estimate the overall effect of
26 the treatment.

- 1 Multivariate model: A statistical model for analysis of the relationship between 2 or
2 more predictor (independent) variables and the outcome (dependent) variable.
- 3 Odds ratio: A measure of treatment effectiveness. The odds of an event happening in
4 the treatment group, expressed as a proportion of the odds of it happening in the
5 control group. The 'odds' is the ratio of events to non-events
- 6 Outcome: The impact that a test, treatment, policy, programme or other intervention
7 has on a person, group or population. Outcomes from interventions to improve the
8 public's health could include changes in knowledge and behaviour related to health,
9 societal changes (for example, a reduction in crime rates) and a change in people's
10 health and wellbeing or health status. In clinical terms, outcomes could include the
11 number of patients who fully recover from an illness or the number of hospital
12 admissions, and an improvement or deterioration in someone's health, functional
13 ability, symptoms or situation. Researchers should decide what outcomes to
14 measure before a study begins.
- 15 Posterior distribution: In Bayesian statistics this is the probability distribution for a
16 statistic based after combining established information or belief (the prior) with new
17 evidence (the likelihood).
- 18 Prior distribution: In Bayesian statistics this is the probability distribution for a statistic
19 based on previous evidence or belief.
- 20 Primary outcome: The outcome of greatest importance, usually the one in a study
21 that the power calculation is based on.
- 22 Publication bias: Publication bias occurs when researchers publish the results of
23 studies showing that a treatment works well and don't publish those showing it did
24 not have any effect. If this happens, analysis of the published results will not give an
25 accurate idea of how well the treatment works. This type of bias can be assessed by
26 a funnel plot.
- 27 Quality-adjusted life year (QALY): A measure of the state of health of a person or
28 group in which the benefits, in terms of length of life, are adjusted to reflect the
29 quality of life. One QALY is equal to 1 year of life in perfect health. QALYS are

1 calculated by estimating the years of life remaining for a patient following a particular
2 treatment or intervention and weighting each year with a quality of life score (on a
3 scale of 0 to 1). It is often measured in terms of the person's ability to perform the
4 activities of daily life, freedom from pain and mental disturbance.

5 Randomised controlled trial (RCT): A study in which a number of similar people are
6 randomly assigned to 2 (or more) groups to test a specific drug or treatment. One
7 group (the experimental group) receives the treatment being tested, the other (the
8 comparison or control group) receives an alternative treatment, a dummy treatment
9 (placebo) or no treatment at all. The groups are followed up to see how effective the
10 experimental treatment was. Outcomes are measured at specific times and any
11 difference in response between the groups is assessed statistically. This method is
12 also used to reduce bias.

13 Rate ratio: The ratio of the rate of an event occurring among those exposed to certain
14 conditions compared with the rate for those who are not exposed to the same
15 conditions

16 Resource implication: The likely impact in terms of finance, workforce or other NHS
17 resources.

18 Review question: In guideline development, this term refers to the questions about
19 treatment and care that are formulated to guide the development of evidence-based
20 recommendations.

21 Risk ratio (RR): The ratio of the risk of disease or death among those exposed to
22 certain conditions compared with the risk for those who are not exposed to the same
23 conditions (for example, the risk of people who smoke getting lung cancer compared
24 with the risk for people who do not smoke). If both groups face the same level of risk,
25 the risk ratio is 1. If the first group had a risk ratio of 2, subjects in that group would
26 be twice as likely to have the event happen. A risk ratio of less than 1 means the
27 outcome is less likely in the first group. The risk ratio is sometimes referred to as
28 relative risk.

1 Sensitivity analysis: A means of representing uncertainty in the results of economic
2 evaluations. Uncertainty may arise from missing data, imprecise estimates or
3 methodological controversy. Sensitivity analysis also allows for exploring the
4 generalisability of results to other settings. The analysis is repeated using different
5 assumptions to examine the effect on the results. One-way simple sensitivity
6 analysis (univariate analysis): each parameter is varied individually in order to isolate
7 the consequences of each parameter on the results of the study. Multi-way simple
8 sensitivity analysis (scenario analysis): 2 or more parameters are varied at the same
9 time and the overall effect on the results is evaluated. Threshold sensitivity analysis:
10 the critical value of parameters above or below which the conclusions of the study
11 will change are identified. Probabilistic sensitivity analysis: probability distributions
12 are assigned to the uncertain parameters and are incorporated into evaluation
13 models based on decision analytical techniques (for example, Monte Carlo
14 simulation).

15 Sub-groups: Planned statistical investigations if heterogeneity is found in the meta-
16 analysis. Specified a priori in the review protocol.

17 Systematic review: A review in which evidence from scientific studies has been
18 identified, appraised and synthesised in a methodical way according to
19 predetermined criteria. It may include a meta-analysis.

20 Utility: In health economics, a 'utility' is the measure of the preference or value that an
21 individual or society places upon a particular health state. It is generally a number
22 between 0 (representing death) and 1 (perfect health). The most widely used
23 measure of benefit in cost– utility analysis is the quality-adjusted life year, but other
24 measures include disability-adjusted life years (DALYs) and healthy year equivalents
25 (HYEs).

26

27

28

1 **References**

2 Barger S, Iannicelli V, Castellini G, Cinquini M, Gianola S. AGREE II appraisals of
3 clinical practice guidelines in rehabilitation showed poor reporting and moderate
4 variability in quality ratings when users apply different cuff-offs: a methodological
5 study. *J Clin Epidemiol.* 2021 Nov; 139:222-231. doi: 10.1016/j.jclinepi.2021.08.021.
6 Epub 2021 Aug 23. PMID: 34437947.

7 Deeks JJ, Bossuyt PM, Leeflang MM, Takwoingi Y (editors). *Cochrane Handbook for*
8 *Systematic Reviews of Diagnostic Test Accuracy.* Version 2.0 (updated July 2023).
9 Cochrane, 2023. Available from [https://training.cochrane.org/handbook-diagnostic-](https://training.cochrane.org/handbook-diagnostic-test-accuracy/current)
10 [test-accuracy/current.](https://training.cochrane.org/handbook-diagnostic-test-accuracy/current)

11 AGREE Next Steps Consortium (2017). *The AGREE II Instrument* [Electronic
12 version]. Retrieved (26 June 2025), from <http://www.agreetrust.org>.

13

14

15

16

17

18

19

20

21

22

23

24

1

2 **Appendix A Appraising the quality of external** 3 **guideline**

4 The external guideline was appraised using a two-stage process. The first stage
5 assessed whether the external guideline development process was robust and high
6 quality. This was conducted using the Appraisal of Guidelines for Research and
7 Evaluation (AGREE) II instrument.

8 The second stage of the appraisal process assessed the risk of bias of the evidence
9 reviews within the external guideline using the ROBIS assessment tool.

10 **First stage assessment using the AGREE II instrument:**

11 The Appraisal of Guidelines for Research & Evaluation (AGREE) Instrument was
12 developed to address the issue of variability in guideline quality. To that end, the
13 AGREE instrument is a tool that assesses the methodological rigour and
14 transparency in which a guideline is developed. The original AGREE instrument has
15 been refined, which has resulted in the new AGREE II. AGREE II has six domains
16 and an overall assessment. The domains are listed below:

- 17 • Domain 1. Scope and Purpose is concerned with the overall aim of the guideline,
18 the specific health questions, and the target population.
- 19 • Domain 2. Stakeholder Involvement focuses on the extent to which the guideline
20 was developed by the appropriate stakeholders and represents the views of its
21 intended users.
- 22 • Domain 3. Rigour of Development relates to the process used to gather and
23 synthesise the evidence, the methods to formulate the recommendations, and to
24 update them.
- 25 • Domain 4. Clarity of Presentation deals with the language, structure, and format of
26 the guideline.
- 27 • Domain 5. Applicability pertains to the likely barriers and facilitators to
28 implementation, strategies to improve uptake, and resource implications of
29 applying the guideline.

- 1 • Domain 6. Editorial Independence is concerned with the formulation of
2 recommendations not being unduly biased with competing interests.
- 3 • Overall assessment includes the rating of the overall quality of the guideline and
4 whether the guideline would be recommended for use in practice.

5 For further details on each domain see the agree reporting checklist at agreetrust.org

6 Each of the 23 AGREE II items were rated on a 7-point scale (1 indicating strong
7 disagreement and 7 indicating strong agreement). An overall rating for each of the 6
8 AGREE II domains was then calculated by summing all the scores of the individual
9 items in a domain and then calculating the total as a percentage of the maximum
10 possible score for that domain, as follows:

$$11 \quad \frac{\text{Obtained score} - \text{Minimum possible score}}{\text{Maximum possible score} - \text{Minimum possible score}} \times 100$$

12 An overall rating for all domains was then determined (score 1 to 7) and finally an
13 overall percentage rating was calculated for each guidance document based on the
14 following equation: (overall score – 1)/6.

15 The AGREE II also suggests an overall assessment which includes a rating of the
16 overall quality of the guideline and whether the guideline would be recommended for
17 use in practice. The overall assessment requires the user to make a judgment as to
18 the quality of the guideline, taking into account the criteria considered in the
19 assessment process.

20 **AGREE II results**

21 The AGREE II instrument was completed by two reviewers and a score given for
22 overall confidence in the guideline. The IG was scored ‘highly’ in most of the AGREE
23 II tool questions, with an overall grading of 7 in most sections (highest possible
24 quality). In accordance with the AGREE II literature, NICE considers an evidence-
25 based guideline to be of sufficient quality if domain 3 (rigour of development) and at
26 least two other domains each achieve a scaled score of 60% or higher (Barger et al.
27 2021). The IG guideline scored 97% in domain 3 and >60% in all other domains.

- 1 This verified the overall confidence in the guideline as a well-conducted, high-quality
- 2 guideline.

1 AGREE II reviewer scoring of Monash University PCOS guideline

Reviewer	1. Scope and purpose				2. Stakeholder involvement				3. Rigour of development								4. Clarity of presentation				
	Objectives	Question	Population	Total	Group membership	Target population	Target users	Total	Search methods	Evidence selection criteria	Evidence strengths and limitations	Formulation of recommendations	Consideration of benefits/harms	Link between evidence and recs	External review	Updating procedure	Total	Specific and unambiguous recs	Management of options	Identifiable key recs	Total
1	7	7	7	21	7	7	7	21	7	7	7	7	7	7	7	7	56	7	7	1	15
2	7	7	6	20	7	6	7	20	6	7	7	6	7	7	7	7	54	6	7	12	15
Score				97%				97%									97%				67%

2

3 Second stage assessment – risk of bias using ROBIS assessment 4 tool

5 The second stage of appraisal of the guideline involved using the ROBIS assessment
6 tool.

7 For the external guideline assessed to be high quality using the AGREE II
8 instrument, the risk of bias of individual evidence reviews was assessed using the
9 ROBIS checklist.

10 Phase 2: identifying concerns with the review process

11 This included answering each question with Y/PY/PN/N/NI = yes, probably yes,
12 probably no, no, or no information. These were summed up as concerns for the
13 domain rated as: low, high or unclear, with an overall rationale given for the concern.

14 Domain 1: study eligibility criteria included the following questions:

15 1.1 Did the review adhere to pre-defined objective and eligibility criteria?

- 1 1.2 Were the eligibility criteria appropriate for the review question?
- 2 1.3 Were eligibility criteria unambiguous?
- 3 1.4 Were any restrictions in eligibility criteria based on study characteristics
- 4 appropriate (e.g. date, sample size, study quality, outcomes measured)?
- 5 1.5 Were any restrictions in eligibility criteria based on sources of information
- 6 appropriate (e.g. publication status or format, language, availability of data)?
- 7 Domain 2: identification and selection of studies:
- 8 2.1 Did the search include an appropriate range of databases/electronic sources for
- 9 published and unpublished reports?
- 10 2.2 Were methods additional to database searching used to identify relevant reports?
- 11 2.3 Were the terms and structure of the search strategy likely to retrieve as many
- 12 eligible studies as possible?
- 13 2.4 Were restrictions based on date, publication format, or language appropriate?
- 14 2.5 Were efforts made to minimise error in selection of studies?
- 15 Domain 3: data collection and study appraisal:
- 16 3.1 Were efforts made to minimise error in data collection?
- 17 3.2 Were sufficient study characteristics available for both review authors and
- 18 readers to be able to interpret the results?
- 19 3.3 Were all relevant study results collected for use in the synthesis?
- 20 3.4 Was risk of bias (or methodological quality) formally assessed using appropriate
- 21 criteria?
- 22 3.5 Were efforts made to minimise error in risk of bias assessment?
- 23 Domain 4: synthesis and findings:

- 1 4.1 Did the synthesis include all studies that it should?
- 2 4.2 Were all pre-defined analyses reported or departures explained?
- 3 4.3 Was the synthesis appropriate given the nature and similarity in the research
4 questions, study designs and outcomes across included studies?
- 5 4.4 Was between-study variation (heterogeneity) minimal or addressed in the
6 synthesis?
- 7 4.5 Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity
8 analyses?
- 9 4.6 Were biases in primary studies minimal or addressed in the synthesis?
- 10 Phase 3 involved judging risk of bias. This summarised the concerns regarding each
11 of the four domains with the rationales for the concerns.
- 12 Risk of bias in the review was then rated as to whether the conclusions were
13 supported by the evidence for:
- 14 A. Did the interpretation of findings address all of the concerns identified in
15 Domains 1 to 4?
- 16 B. Was the relevance of identified studies to the review's research question
17 appropriately considered?
- 18 C. Did the reviewers avoid emphasising results on the basis of their statistical
19 significance.
- 20 Each was rated Y/PY/PN/N/NI. The risk of bias overall in the review was then rated
21 as low, high or unclear with a rationale for the risk.

1 Appendix B Economic review protocol

2 Table 2: Health economic review protocol

ID	Field	Content
1.	Review title	All review questions
2.	Objective	To identify economic studies for all relevant guideline review questions (as detailed in the International Guideline [IG])
3.	Inclusion criteria	<ul style="list-style-type: none"> • Populations, interventions and comparators must be as specified in the review protocol detailed in the IG • Relevant comparative economic study design: cost–utility analysis, cost–effectiveness analysis, cost–consequences analysis, comparative cost analysis • Decision analytic model-based or within-trial economic analyses • OECD countries (except USA) • Healthcare and personal social services cost perspective • Studies published from 2010 – this cut off has been applied to restrict the review to more recent studies which will have more applicable resource use and costs. <p>High-quality studies in line with the NICE reference case (recent UK NHS/PSS cost-utility analyses using the QALY as the measure of outcome) are the most applicable to NICE decision making. Not all studies meeting the inclusion criteria will therefore necessarily be used in decision-making - see Review strategy below for details.</p>
4.	Exclusion criteria	<ul style="list-style-type: none"> • Conference posters or abstract only studies – these do not provide sufficient information for quality assessment. • Studies published before 2010 – this cut off has been applied to restrict the review to more recent studies which will have more applicable resource use and costs • Studies from non-OECD countries or the USA – these are considered unlikely to be applicable to the UK NHS setting due to substantial differences in healthcare delivery and unit costs. • Economic analyses utilising effectiveness data from studies with a mirror-image design and studies that recruited participants retrospectively, due to their lower methodological quality. • Analyses based on clinical evidence that is not included in the clinical evidence review • Within-trial economic analyses based on studies that have been excluded from the clinical review • Economic model analyses based exclusively on studies that have been excluded from the clinical review • Non-comparative economic analyses including cost-of-illness studies. • Letters, editorials or commentaries, study protocols or reviews of economic evaluations (recent reviews will be ordered and the bibliographies will be checked for relevant individual economic studies, which will then be ordered and checked for eligibility). • Non-English language papers.

		<ul style="list-style-type: none"> • Studies considering exclusively intervention costs, e.g. medicine acquisition costs, without considering wider healthcare costs associated with the management of PCOS. • Studies comparing costs of branded vs generic forms of the same medicine. • Studies only focussing on productivity losses or gains.
5.	Search strategy	<p>An economic study search will be undertaken covering all review questions using guideline population-specific terms.</p> <p>For search details see Appendix A in the Adaptation Reports.</p>
6.	Review strategy	<ul style="list-style-type: none"> • Studies meeting the inclusion and exclusion criteria will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist in Appendix H of Developing NICE guidelines: the manual. • The NICE economic evaluation checklist assesses: <ul style="list-style-type: none"> ○ Applicability to the NICE guideline decision making context with consideration of the NICE reference case relevant to the guideline. Recent UK studies that use the NICE reference case methods are the most applicable when considering cost effectiveness. ○ Methodological limitations. • The aim is to present the best available economic evidence to inform committee decision-making in the context of the guideline, the current UK NHS setting and NICE methods. Therefore, the health economist may not present all studies that meet inclusion criteria. Studies that are deemed not applicable or have very serious methodological limitations should not inform committee decision-making. If recent high quality, UK cost-utility analyses are available for a question, it is often not deemed informative to present studies that are less applicable or lower quality such as older UK analyses or analyses from other countries. A similar principle is deemed to apply more generally when considering applicability and methodological limitations. Some specific examples are given below: <ul style="list-style-type: none"> ○ If multiple versions of a model are available for the UK and other countries it is usually reasonable to only present the UK version. ○ If multiple versions of the same UK model are available, it is usually reasonable to present only the most recent. ○ If there has been a NICE MTA or guideline model that informs current NHS practice it is usually reasonable not to present older studies, unless they address a different subpopulation or other specific issue. ○ If a UK model that includes all interventions in the decision space is available it may be reasonable not to present studies that only include individual or fewer interventions, if the analysis is sufficiently applicable and of good methodological quality. • Quality and relevance of effectiveness data used in the economic analysis: the more closely the clinical effectiveness data used in the economic analysis match with the outcomes of the studies included in the clinical review the more useful the analysis will be for decision-making in the guideline. • Hierarchy of economic evaluation evidence based on quality assessment <ul style="list-style-type: none"> ○ 'Directly applicable' and 'Minor limitations' (only recent UK CUAs can get this rating). Usually presented and used in decision-making. ○ Directly or partially applicable combined with minor or potentially serious limitations (other than 1). Discretion over whether these are presented and used in decision-making, depending on the availability of more relevant evidence.

		<ul style="list-style-type: none"> ○ 'Not applicable' or 'Very serious limitations'. Typically not presented and not used in decision-making. <p>The health economist will make a decision based on the relative applicability and quality of the available evidence for each question, in discussion with the guideline committee if required. All decisions will be transparently reported in the evidence report. Studies that are presented to the committee and used in decision-making when formulating recommendations will be included in the summary tables and will have an evidence extraction. Other studies may not be presented to the committee in detail but will be listed, with the reason for not being presented to the committee and thus not used in decision-making being provided. Committee members can review and query the decision not to present studies with the health economist and will be provided with full details of these studies where requested.</p>
--	--	---

1

2 ISBN: XX

3

4

5

6

7

8

9

10

11

12

13

14

15

