Cochrane Gynaecology and Fertility Group
Guidance for Authors
New Protocols and Reviews

- This guidance is in eight sections, in the order that they appear in a published review.
- For a new protocol, read Parts A, D and H.
- For a full review, read all eight sections.
- This guidance only covers RoB 2. For help with RoB 1 see the guidance for *Review Updates and New Reviews using RoB 1*.
- Please read the KEY POINTS sections where available.
- Example text are largely based on analysis strategy 1 and are in greyed out italics. Notes and comments from us are in plain text.

Need help?

If at any point you need help or advice with your protocol or new review please contact any of our friendly and knowledgeable Managing Editors (h.nagels@auckland.ac.nz; e.b.kostova@amsterdamumc.nl; cochrane.MDSG@auckland.ac.nz). We are here to help you successfully complete your review.

Common abbreviations:

- SoFs – Summary of findings
- PLS – Plain language summary
- RoB – Risk of bias
- CGF – Cochrane Gynaecology and Fertility

CONTENTS

Cochrane
Gynaecology and Fertility

Part A - Important information to read before starting a review

All authors please read this section before starting your review - it will save you time! Refer back to this list of helpful tips and common errors as you go through each section of your review. By far the <u>most common problems</u> with reviews involve the issues listed below:

Methods

- In general, do not exclude studies on the basis of their reported outcome measures; take care to determine if relevant outcomes are not available because they have not been measured rather than simply not reported.
- Define in advance details of what are acceptable outcome measures (e.g. differing scales, time-points), and state a preference order when there are several possible measures.
- Outcomes must be prespecified based on clinical importance. We encourage using a core outcome set – refer to page 22.
- Keep subgroup analyses to an absolute minimum and explain the rationale for these.
- If planning any departures from the Methods specified in the protocol, you must check with editorial base first. Once changes are approved by editorial base, document and justify any changes in the review section 'Differences between the protocol and review'.
- It is now mandatory for authors with our group to contact study authors in order to obtain or confirm data. This includes details to inform risk of bias assessments, details of interventions and outcomes, and study results.
- If there is a substantial lapse in time since the protocol was published, or you have taken over a protocol or review, you must check if the methods are still clinically and methodologically sound before you start. Contact your ME to discuss this before starting.

Reporting results

- Include all pre-specified comparisons and outcomes: if there are no relevant data, say so.
- Use the same order of comparisons and outcomes throughout the review as listed in the methods. Do not reorder comparisons and outcomes on the basis of results.
- Use this format for presenting results: (RR 0.89, 95% CI 0.75 to 1.05, I^2 = 31%, 3 studies, 811 women; low-certainty evidence; Analysis 1.1).
- Interpret the main findings in absolute terms e.g. if 10% of women taking placebo experience pain, between 2% and 5% of those using XX will do so (see page 27 key points on how to format your findings).
- Do not confuse lack of evidence of an effect with evidence of a lack of effect: say something like "There was insufficient evidence to determine whether there was a difference" not "There was no significant difference".
- As "no evidence of a difference" may imply equivalence, we do not use this anymore (see page 27 for key points on how to format your findings).
- Consider clinical rather than statistical significance.
- Interpret subgroups very cautiously. In general, do not report them in the abstract.
- If investigating subgroup differences, present an overall plot or figure containing all subgroups, rather than multiple forest plots.
- Include the total sample size and number of participants by arm for each included study in the Characteristics of Included Studies table.
- There are options for presenting the primary analyses, including restricting analyses to low risk of bias studies: see Data synthesis below and <u>section 7.6.2</u> of the *Cochrane Handbook*.

- Ensure mandatory sensitivity analyses are conducted for the type of data synthesis you use as stated in your methods (see page 17).
- If the odds of an outcome (beneficial e.g. live birth or detrimental e.g. adverse effects) increase with the intervention it will be displayed graphically in the meta-analyses to the right of the centreline. If the odds of an outcome decrease with the intervention, it will be displayed on the left of the centreline.

Conclusions

In summary parts of the review abstract, plain language summary and summary of findings:
- Include the <u>same</u> outcomes for these sections: this means including all **main outcomes** (primary outcomes, adverse events and prespecified secondary outcomes).
- Include the same comparisons: those that are clinically most important specified at the protocol stage, not on the basis of the results.
- Incorporate the findings of the GRADE assessment.
- Be 100% consistent with wording across all these sections.

Summary of findings table

- Specify detailed plans for the summary of findings (SoFs) table in the Methods section (see page 21).
- The SoF table should be prepared once the study data have been entered, and before the results section is written.
- Include max 7 main outcomes (primary outcomes, adverse events and prespecified secondary outcomes).
- Clearly explain SoF evidence downgrades in footnotes: e.g. We downgraded the evidence by two levels, due to very serious imprecision: only 29 events.
- When you reach this stage, before proceeding any further, please submit your review to our Managing Editor for an editorial check.
- For help with preparing a summary of findings table, see Chapter 14 of the *Handbook*, otherwise contact us.

Conflict of Interest Policy

Cochrane has a new COI policy which applies to all review authors. **It is critical that ALL authors read and comply with the new policy**. The Quick Guide for Authors provides an overview of the policy and should be read in conjunction with the full COI policy. **It is the responsibility of all authors to advise CGF of any potential conflicts (financial and non-financial) prior to and during the review process.** Reviews that do not comply with the COI Policy may not be published in the Cochrane Library. Important additional restrictions to authorship include:
- The first and last authors **must not** have:
  o any relevant financial interests
  o been involved in industry-controlled studies eligible for inclusion in the review.
- **Overall,** 67% (two-thirds) of the authors must be free of relevant conflicts.
- Anyone who has been involved in the conduct, analysis, and publication of a study that could be included in the review **cannot**:
  o Determine the overall study inclusion and exclusion criteria
  o Make study eligibility decisions about, extract data from, carry out the risk of bias assessment for, or perform GRADE assessments of that study.

Cochrane
Gynaecology and Fertility

Part B – Abstract

KEY POINTS

- Report all main outcomes as explained in Part A - Important information to read before starting a review.
- Note that the effect estimate in the abstract should not include the I^2.
- Do not report subgroup analyses in the abstract.
- Structure the abstract by comparison rather than by outcome.
- Always report outcomes in the same order as reported in the methods.
- Do not reorder comparisons and outcomes on the basis of results.
- Aim to limit the abstract to fewer than 700 words. Absolute maximum is 1000 words.

Example text:

Background

Treatment B is commonly used to optimise the chance of live birth in women undergoing assisted reproductive technology (ART). However, it is known to increase multiple pregnancy rates, potentially causing maternal and perinatal morbidity. Treatment A is an alternative intervention which may reduce the risk of multiple pregnancy. We compared the benefits and risks of the two treatments.

Objectives

To evaluate the effectiveness and safety of Treatment A in women undergoing ART.

Search Methods

We searched the Cochrane Gynaecology and Fertility (CGF) Group trials register, CENTRAL (now containing output from two trials registers and CINAHL), MEDLINE, Embase and PsycINFO on XXX 20YY together with reference checking and contact with study authors and experts in the field to identify additional studies.

Selection criteria

We included randomised controlled trials (RCTs) of the use of Treatment A compared with Treatment B for infertile women.

Data collection and analysis

We used standard methodological procedures recommended by Cochrane, including assessing risk of bias of the included studies using RoB 2. The primary review outcomes were cumulative live birth and multiple pregnancy. Other adverse effects were a secondary outcome. Due to high risk of bias associated with some of the studies, primary analyses of all review outcomes were restricted to studies at overall low risk of bias. Sensitivity analysis including all studies was then performed.

Main results

We included four RCTs with 811 women in the review. Using the GRADE method, we assessed the certainty of evidence as moderate to low across measured outcomes.

The primary analysis was restricted to studies at overall low risk of bias, which left only one study included. We are uncertain whether Treatment A compared to Treatment B has an effect on cumulative live birth rate, as only one study is included in the analysis and the confidence interval is wide (RR 1.11, 95% CI 0.78 to 1.59, 1 study, 210 women; low-certainty evidence; Analysis 1.1). Evidence suggests that if the chance of live birth following Treatment B is assumed to be 34%, then the chance with Treatment A would be 27% to 55%. When all studies were included in the sensitivity analysis, we are uncertain of the effect of Treatment A compared to Treatment B on cumulative live birth rate (RR 1.00, 95% CI 0.92 to 1.09, 3 studies, 621 women; low-certainty evidence; Analysis 1.1.2). The evidence suggests that if the chance of live birth following Treatment B is assumed to be 37%, the chance following Treatment A would be between 34% and 40%.

When the primary analysis was restricted to studies at overall low risk of bias, two studies were included. Treatment A may reduce multiple pregnancy rates compared with Treatment B (RR 0.10, 95% CI 0.02 to 0.28; 2 studies, 379 women; low-certainty evidence; Analysis 1.2). This suggests that if the chance of multiple pregnancy following Treatment B is 13%, then the chance following Treatment A would be 0% to 3%. When all studies were included in the sensitivity analysis, Treatment A probably reduces multiple pregnancy rates compared with Treatment B (RR 0.04, 95% CI 0.01 to 0.15, 4 studies, 811 women; moderate-certainty evidence; Analysis 1.2.2). This suggests that if the risk of multiple pregnancy following Treatment B is assumed to be 13%, the risk following Treatment A would be between 0% and 2%.

There was insufficient evidence to reach a conclusion regarding other adverse effects, as no studies reported data suitable for analysis.

Authors' conclusions

Treatment A may reduce the risk of multiple pregnancy in women undergoing ART but the effect on the cumulative live birth rate is uncertain. Data were lacking on other adverse effects. The pooled results should be interpreted with caution, as the evidence was of low-certainty due to high risk of bias present in most of the included studies and an overall low level of precision.

Cochrane
Gynaecology and Fertility

Part C - Plain language summary

KEY POINTS

- Format the plain language summary (PLS) under the five headings used in the example text below.
- Report all main outcomes in the abstract and PLS and summarise any evidence about adverse effects (including lack of data).
- Aim to limit the PLS to fewer than 400 words. Absolute maximum is 700.
- Translate the effect estimates for important clinical outcomes in a user-friendly way. Rates per 100 (as used in the example text below) are easily extracted from the review SoF table – see page 27.
- Report and interpret the units used for continuous outcomes (e.g. a VAS scale of 0-10 where 0 is pain-free and 10 is unbearable pain)
- For more detailed information on the standards for PLS go to: Standards for the reporting of Plain language summaries in new Cochrane Intervention Reviews 2013.
- New PLS guidance is under development through Cochrane's Plain Language Summaries pilot project – please see an example of the new structure here; contact Helen Nagels or Elena Kostova for further information (h.nagels@auckland.ac.nz; e.b.kostova@amsterdamumc.nl; cochrane.MDSG@auckland.ac.nz).

Example text:

Title
Treatment A versus treatment B for women undergoing assisted reproductive technology (ART)

Review question
Cochrane authors reviewed the evidence about the effect of Treatment A versus Treatment B in women undergoing ART.

Background
Treatment B is commonly used to increase the chance of live birth in women undergoing ART. However, it is known to increase multiple pregnancy rates, which can cause serious health risks for both mother and baby. Treatment A is an alternative approach which may reduce the risk of multiple pregnancy. We compared the benefits and risks of the two treatments.

Study characteristics
We found four randomised controlled trials comparing Treatment A with Treatment B in a total of 811 women undergoing ART. The evidence is current to June 20XX.

Key results
Only one trial comparing Treatment A with Treatment B was well designed and was included in the analysis for live birth. This study did not provide enough evidence to show whether there is a difference in the chance of live birth; the quality of the evidence was low.

Two well designed studies were included in the analysis for multiple pregnancy. This showed that Treatment A compared to Treatment B may reduce the risk of multiple pregnancy in women undergoing ART. Data were lacking on other adverse effects.

The evidence suggests that if the chance of live birth following Treatment B is assumed to be 34%, then the chance following Treatment A would be 27% to 55%. It also suggests that if the chance of multiple pregnancy following Treatment B is 13%, then the chance following Treatment A would be 0% to 3%.

Evidence on other adverse events was poorly reported and inconclusive.

Certainty of the evidence

There remains uncertainty about whether Treatment A compared to Treatment B increases the chance of having a baby, but it may reduce the risk of multiple pregnancy in women undergoing ART. The certainty of the evidence was assessed as low. The reason for this is that the studies included in this review were not very well designed and did not recruit a large enough number of women to provide meaningful results. This means that results must be treated cautiously, and further studies are needed to confirm findings.

Part D – Methods

Criteria for considering studies for this review

KEY POINTS

- Usually, all randomised trial designs should be included though quasi-RCTs should not be included. You may also wish to include:
    - Crossover trials - This type of RCT may or may not be appropriate for inclusion and these can be dealt with in 3 ways:
        - Crossover trials will be excluded, as the design is not valid in this context. For example, crossovers are not valid for long term outcomes in women undergoing HRT.
        - Crossover trials will be included but only data from the first phase will be included in meta-analyses, as the crossover is not a valid design in this context. For example, crossovers cannot be used in full in fertility research where the primary outcome is live birth.
        - Crossover trials will be included, as the crossover is a valid design in this context. For example, cross overs are a valid design for chronic stable conditions where the research is assessing short term outcomes such as dysmenorrhoea.
    - Cluster-randomised trials - There is no reason to exclude cluster-randomised trials. However, these type of RCT's are unlikely to have been performed in this research area. If the review includes interventions randomised by clinic rather than individuals, then these are likely to be included. Please address this in the Unit of analysis issues and Data synthesis sections.

Example text:

Types of studies

Randomised controlled trials (RCTs) will be included. Quasi-RCTs will not be included. Crossover trials will be included but only data from the first phase will be included in meta-analyses, as the crossover is not a valid design in this context.

Types of participants

Women/couples with unexplained infertility will be eligible for inclusion.

- Definitions of the condition (e.g. unexplained infertility) belong in the Background section but diagnostic criteria maybe included if appropriate

Types of interventions

Trials comparing Factor X via any route versus any other active intervention or placebo will be included.

- Consider if the intervention is used alone or in combination with other intervention(s).

Types of outcome measures

KEY POINTS

- The primary outcomes should normally include one measure of effectiveness (e.g. live birth) and one of potential harm (e.g. multiple pregnancy).
- You can combine ongoing pregnancy data with live birth data in one primary outcome live birth/ongoing pregnancy only if ongoing pregnancy is directly reported in the trial (not by calculating it from clinical pregnancy). In the absence of live birth data, you should report ongoing pregnancy separately.
- For infertility and endometriosis reviews, we encourage you to use the core outcome set for infertility/endometriosis – refer to page 22 for the full list with definitions.
- Try to limit to the most relevant outcomes as only seven outcomes (max) will be included in the summary of findings table which should be listed in the section 'Summary of findings and assessment of the certainty of the evidence'.
- Number outcomes as in the example text below and use the same numbers in the Effects of the interventions.
- Keep secondary outcomes to a minimum. Focus on clinical outcomes and try to avoid lab outcomes (e.g. implantation rate).
- Define in advance details of what are acceptable outcome measures (e.g. differing definitions, assessors, scales, time-points) and state a preference order when there are several possible measures.

Example text:

Primary outcomes

1. *Live birth or ongoing pregnancy*
   - *Live birth is defined as delivery of a live fetus after 20 completed weeks of gestation*
   - *Ongoing pregnancy is defined as evidence of a gestational sac with fetal heart motion at 12 weeks, confirmed with ultrasound.*
2. *Multiple pregnancy*

Secondary outcomes

3. *Clinical pregnancy, defined as evidence of a gestational sac, confirmed by ultrasound.*
4. *Any adverse event (including miscarriage, bleeding, drug reactions), reported either as a composite measure or separately.*
5. *Quality of life. If studies report more than one scale, preference will be given to the SF-36, then other validated generic scales, and finally condition-specific scales.*

Search methods for identification of studies

KEY POINTS

- The Gynaecology and Fertility Information Specialist (IS) (Marian Showell) will help design your search and will run a search in the electronic databases listed under (1) above. It is the responsibility of the review authors to run, document and date (with day, month and

year) the searches of other sources (i.e. those listed under (2) and (3) above, as appropriate).

- Please note that http://www.clinicaltrials.gov and http://ictrptest.azurewebsites.net/Default.aspx (the World Health Organisation International Trials Registry Platform search portal) are now indexed in CENTRAL with 1-month delay. Authors can search these databases if they wish to cover the 1-month lag, but this is not mandatory.
- As from July 2020 CINAHL records are automatically downloaded to CENTRAL so the IS no longer searches CINAHL as an individual database.
- Review authors MUST search reference lists of articles retrieved and contact experts in the field in order to obtain any additional studies.
- Reviews of Traditional Chinese Medicine or Chinese complementary therapies must search at least one Chinese database.
- Epistemonikos database is a good source of systematic reviews, for reference checking.
- Liaise with the CGF Information Specialist to avoid duplication of handsearching and for other advice on searching.
- List all sources searched in the Methods section of the review (as in the example above).
- The search process should be summarised in a PRISMA flow diagram in the full review.
- Full search strategies for all sources searched (including those for Other electronic sources of trials) must be copied and pasted into the appendices (not in the body of the text) of the review along with dates and the platforms used for each database.
- The MEDLINE, Embase, PsycINFO, and AMED searches are on the OVID platform, CENTRAL is now searched via CRS ONLINE via the Web.
- In the protocol the numbers of hits per search line (i.e. the numbers in brackets after the keywords) are removed from the strategies, however at the review stage the numbers of hits per keyword remain.
- It is mandatory to run/update searches for all relevant databases no more than six months (maximum 12 months) before publication of the full review.
- Ideally any new studies should be fully incorporated. As a minimum, potentially eligible studies should be referenced under "awaiting classification" or "ongoing", but authors have to show that they are waiting for information from trial authors.
- References for this section:
  - Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, Noel-Storr A, Rader T, Shokraneh F, Thomas J, Wieland LS. Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.2 (updated February 2021). Cochrane, 2021. Available from *www.training.cochrane.org/handbook*.
  - Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, Noel-Storr A, Rader T, Shokraneh F, Thomas J, Wieland LS. Technical Supplement to Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston MS, Li T, Page MJ, Welch VA (eds). *Cochrane Handbook for Systematic Reviews of Interventions* Version 6.2 (updated February 2021). Cochrane, 2021. Available from: *www.training.cochrane.org/handbook*.

Example text:

We will search for all published and unpublished RCTs of XX, without language restriction, in consultation with the Gynaecology and Fertility Group (CGF) Information Specialist.

Electronic searches

(1) We will search the following electronic databases for relevant trials:

1. *The Cochrane Gynaecology and Fertility Group (CGF) Specialised Register of Controlled Trials, PROCITE platform (from inception onwards)*
    2. *CENTRAL (now containing output from two trials registers and CINAHL), via the Cochrane Register of Studies Online (CRSO), Web platform searched from inception onwards;*
    3. *MEDLINE (Epub Ahead of Print, In-Process & Other Non-Indexed Citations), Ovid platform, searched from 1946 to present;*
    4. *Embase, Ovid platform, searched from 1980 to present;*
    5. *PsycINFO, Ovid platform searched from 1806 to present;*
    6. *AMED, Ovid platform, searched from 1985 to present (for any complementary therapy review topics).*

The MEDLINE search will be combined with the Cochrane highly sensitive search strategy for identifying randomised trials which appears in the Cochrane Handbook of Systematic Reviews of Interventions (Version 6.2 chapter 4, 4.4.7; 4.S1). The Embase search is combined with trial filters developed by the Scottish Intercollegiate Guidelines Network (SIGN) (https://www.sign.ac.uk/what-we-do/methodology/search-filters/).

(2) Other electronic sources of trials may include:

1. *International trial registers: the ClinicalTrials database, a service of the US National Institutes of Health (clinicaltrials.gov/ct2/home) and the World Health Organization International Trials Registry Platform search portal (http://ictrptest.azurewebsites.net/Default.aspx);*
2. *LILACS and other Spanish/Portuguese regional databases (Latin American and Caribbean Health Science Information database, Web platform, searched from 1982 to present; found in the Virtual Health Library Regional Portal (VHL) http://pesquisa.bvsalud.org/portal/ (the right-hand drop-down box allows you to filter out MEDLINE records);*
3. *Google Scholar, Web platform (for recent trials not yet indexed in the major databases) [The searches above will be simple short keyword searches and only checking the top few hits]*
4. *Epistemonikos database https://www.epistemonikos.org/, a multilingual database of health evidence, Web platform (the largest source of systematic reviews and also other scientific evidence).*

Searching other resources

(3) We will handsearch reference lists of relevant trials and systematic reviews retrieved by the search and contact experts in the field to obtain any additional trials. We will also handsearch relevant journals and conference abstracts that are not covered in the CGF register, in liaison with the Information Specialist.

Data collection and analysis

There should not be any text under this heading.

Selection of studies

After an initial screen of titles and abstracts retrieved by the search, conducted by two of the review authors (XX and YY) we will retrieve the full texts of all potentially eligible studies. Two of the review authors (XX and YY) will independently examine these full text articles for compliance with the inclusion criteria and select eligible studies. We will correspond with study investigators as required, to clarify study eligibility. Disagreements will be resolved by discussion. If any reports require translation, we will describe the process used for data collection. We will document the selection process with a PRISMA flow chart.

Data extraction and management

KEY POINTS

- Data are often presented in a non-standardised format. Studies should be included irrespective of whether outcomes are reported in a "usable" way. In multi-arm studies, only data from arms that do meet eligibility criteria will be included.

Example text:

Two of the review authors (XX with YY or ZZ) will independently extract data from eligible studies using a data extraction form designed and pilot-tested by the authors. Any disagreements will be resolved by discussion. Data extracted will include study characteristics and outcome data (see data extraction table for details, Appendix XX). Where studies have multiple publications, the authors will collate multiple reports of the same under a single study ID with multiple references. We will correspond with study investigators for further data on methods and/or results, as required.

Assessment of risk of bias in included studies

KEY POINTS

- All new protocols and reviews should use RoB 2, contact Helen Nagels or Elena Kostova at editorial base (h.nagels@auckland.ac.nz; e.b.kostova@amsterdamumc.nl) for support and relevant resources. If you are using RoB 1 for a new review, and this has been approved by the editorial team, read section 8.5 of the 2011 *Cochrane Handbook* for detailed guidance on assessing each type of bias, and see our separate guidance on *Review Updates and New Reviews using RoB 1*.
- Authors are required to use the supplied excel RoB 2 tool to conduct RoB 2 judgements which will then be transferred into RevMan Web. The excel spreadsheet will need to be submitted for editorial.
- RoB 2 assesses bias for specific outcomes within an RCT, rather than for the RCT as a whole. Each outcome will have an overall risk of bias across all domains.

- References when using RoB 2:
  - Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng H-Y, Corbett MS, Eldridge SM, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019; **366**: l4898.
  - Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.

Example text:

Two of the review authors (XX with YY or ZZ) will independently assess the included studies for risk of bias using the Cochrane RoB 2 assessment tool (Sterne 2019). We will carry out the risk of bias assessments for the included outcomes: live birth rate/ongoing pregnancy rate, clinical pregnancy rate, cycle cancellation rate, number of days of stimulation and other adverse effects. The effect of interest will be the effect of assignment to the intervention at baseline, regardless of whether the interventions are received as intended (the 'intention-to-treat' effect).

RoB 2 assesses five domains:
- *bias arising from the randomisation process;*
- *bias due to deviations from intended interventions;*
- *bias due to missing outcome data;*
- *bias in measurement of the outcome; and*
- *bias in selection of the reported result.*

These domain-level judgements provide the basis for an overall risk of bias judgement for the specific trial result being assessed. We will classify judgements as 'low risk of bias', 'some concerns' and 'high risk of bias', as suggested by the tool in response to answers to signalling questions for each domain and an algorithm in the tool. We will then reach our overall judgement following the algorithm in the tool, as recommended in Chapter 8 of The Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2021). We will resolve any disagreements by discussion. We will describe all judgements fully and present the consensus judgements in the main review document (as a table, a figure, and within a forest plot of the results).

We will use the RoB 2 Excel tool to implement RoB 2 (available on the riskofbias.info website) and will store data for access in an online repository [please state where this will be stored and give link].

Measures of treatment effect

KEY POINTS

- There are three available statistics to analyse binary (dichotomous) outcomes - the odds ratio, risk ratio and risk difference. The odds ratio further divides into the Mantel-Haenszel and Peto estimates. Any analysis compatible with the *Cochrane Handbook* (see Chapter 6) is acceptable although it is rarely appropriate to use a risk difference. We recommend use of the odds ratio (Mantel Haenszel by default, Peto if events are very rare) because of its superior mathematical properties. Whichever statistic you use, we encourage 'translation' of the result to actual percentages for a typical population to maximise understanding. You will find examples of this in Parts B, C and E.
- Only include information relevant to the review (e.g. many fertility reviews contain only binary outcomes, so you do not need to provide for continuous outcomes.)
- If you have used either crossover or cluster randomised trials these will need to be included as appropriate. If crossover is used the overall result from paired analysis would be included using generic inverse variance (GIV). If in the rare circumstance cluster randomised trials are included these will also need to be analysed using GIV taking into account, the correlation in the cluster. If using these types of trials, please make sure you have statistical support or contact the editorial base.

Example text:

For dichotomous data (e.g. live birth rates), we will use the number of events in the control and intervention groups of each study to calculate Mantel-Haenszel odds ratios (ORs). For continuous data (e.g. quality of life), if all studies report the outcome in exactly the same way, we will calculate the mean difference (MDs) between treatment groups. If similar outcomes are reported on different scales (e.g. quality of life) we will calculate the standardised mean difference (SMD). We will reverse the direction of effect of individual studies, if required, to ensure consistency across trials if scales have different directions of effect. We will treat ordinal data (e.g. quality of life scores) as continuous data. We will present 95% confidence intervals (CIs) for all outcomes. Where data to calculate ORs or MDs are not available, we will utilise the numerical data available. For example, if dichotomous data supplies percentages with sample numbers, we can use this to calculate OR's. For continuous data, if alternate measurement of error (e.g. test statistics, p values) are supplied we will use these to calculate CIs.

Unit of analysis issues

KEY POINTS

- Report all outcomes per randomised woman, as this is the unit of randomisation. Rates per clinical pregnancy may be used as the denominator for a sensitivity analysis, as this will help give the full picture.

- If studies report only "per cycle" data, contact authors and request "per woman" data.
- Some outcomes can only occur in women who reach clinical pregnancy (e.g. multiple pregnancy, miscarriage, etc).
- If including crossover or cluster randomised trials. Both of these need to have been analysed correctly to take into account the patient correlations.

Example text:

The primary analysis will be per woman randomised; per pregnancy data may also be included for some outcomes (e.g. miscarriage). Data that do not allow valid analysis (e.g. "per cycle" data) will be briefly summarised in an additional table and will not be meta-analysed. Multiple births will be counted as one live birth event.

Dealing with missing data

KEY POINTS

- For fertility reviews it is reasonable to assume missing participants did not have the outcome of interest.
- For other topics where the main outcomes are not pregnancy related, we suggest you only report data for participants on which measurements have been taken. It is not advised that you impute data for these.
- If measurement of error data is unavailable it is acceptable to impute SD's from other included studies, please text below.

Example text:

We will analyse the data on an intention-to-treat basis as far as possible (i.e. including all randomised participants in analysis, in the groups to which they were randomised). Attempts will be made to obtain missing data from the original trialists. Where these are unobtainable, we will undertake imputation of individual values for live birth only: live birth will be assumed not to have occurred in participants without a reported outcome. For other outcomes, we will analyse only the available data. Any imputation undertaken will be subjected to sensitivity analysis (see Sensitivity analysis section below).

If studies report sufficient detail to calculate mean differences but no information on associated standard deviation (SD), we will assume the outcome to have a standard deviation equal to the highest SD from other studies within the same analysis.

Assessment of heterogeneity

We will consider whether the clinical and methodological characteristics of the included studies are sufficiently similar for meta-analysis to provide a clinically meaningful summary. We will assess statistical heterogeneity by the measure of the I^2. An I^2 measurement greater than 50% will be taken to indicate substantial heterogeneity (Deeks 2021).

Assessment of reporting biases

KEY POINTS

- This section on reporting biases refers to data missing from analyses either as a result of selective reporting of trial outcomes or as a result of studies that have not been able to be included (e.g. outcome reporting bias, publication bias, multiple publication bias, language bias etc), where the dissemination of research findings is influenced by the nature and direction of results, reducing the likelihood that all studies eligible for a review will be retrieved.
- For each outcome consider which study results would be eligible for inclusion in each synthesis but are not present either as the results have not been presented in such a way as to allow inclusion or are not reported (this can be determined by looking at trial registrations).
- Record whether any of the studies identified are missing from each synthesis because results known (or presumed) to have been generated are unavailable: the 'known unknowns'.
- Consider whether each synthesis is likely to be biased because of the missing results in the studies identified.
- Consider whether results from additional studies are likely to be missing from each synthesis: the 'unknown unknowns'.
- RoB-ME, a tool for assessing risk of bias due to Missing Evidence in a synthesis, is under development and follows the above process. To discuss the option of using this tool please contact Helen Nagels or Elena Kostova at editorial base (h.nagels@auckland.ac.nz; e.b.kostova@amsterdamumc.nl).

Example text:

In view of the difficulty of detecting and correcting for reporting and publication bias, we will aim to minimise their potential impact by ensuring a comprehensive search for eligible studies and by being alert for duplication of data. For missing outcomes, we will report known missing outcome results in the narrative and consider whether the outcome is likely to be biased as a result of the missing data. For missing studies, if there are ten or more studies in an analysis, we will use a funnel plot to explore the possibility of small study effects (a tendency for estimates of the intervention effect to be more beneficial in smaller studies).

Data synthesis

KEY POINTS

- All data needs to be narratively discussed.
- Statistical analysis will be performed using Review Manager Web (RevMan Web). RoB 2 is only available using this platform.

- Analyses need to be defined by comparison and outcome. Studies should be grouped where the participants and interventions/comparisons are similar enough to be combined in order to answer the review objectives.
- You need to prespecify which model (i.e. random or fixed effect) will be used to underpin the meta-analysis. Fixed effect is the most common model if studies are expected to be similar in the participants and interventions. A random effects model should be chosen if participants are likely to be very different or interventions very heterogeneous.
- If analyses are subgrouped (as in the example text below), state whether it is planned to pool the subgroups (see Subgroup analysis and investigation of heterogeneity section).
- If the odds of an outcome (beneficial e.g. live birth or detrimental e.g. adverse effects) increase with the intervention it will be displayed graphically in the meta-analyses to the right of the centreline. If the odds of an outcome decrease with the intervention, it will be displayed on the left of the centreline.
- We now encourage authors to restrict the primary analyses to studies judged to be at overall low risk of bias, or studies judged to be at overall low risk and with some concerns with regard to bias.
- Three different strategies for data synthesis are presented below.

Strategies for Data Synthesis

When risk of bias varies across studies in a meta-analysis, the *Cochrane Handbook* sets out broad strategies to incorporate these assessments into the analysis (see section 7.6.2). The strategy you choose will influence how you present your main findings for a particular outcome. The choice between strategies (1, 2 or 3) should be based to a large extent on the balance between the potential for bias and the loss of precision when studies at higher risk of bias are excluded. This choice must be prespecified in the protocol and have the agreement of the editorial base.

The main strategy options are:

1. <u>Primary analyses restricted to studies at low risk or with some concerns with regard to bias</u> **(preferred option)**
   - We recommend this strategy for all CGF reviews.
   - Restrict the primary analyses to studies judged to be at overall low risk of bias or with some concerns i.e. exclude studies with an overall high risk of bias. Review authors who restrict their primary analyses in this way are encouraged to perform **sensitivity analyses** to show how conclusions might be affected if studies with an overall high risk of bias were included.
   - It may be reasonable to present the sensitivity analyses in the abstract, PLS and SoFs in this instance.
   - This stratification should apply to all outcomes within the review.
2. <u>Present multiple (stratified) analyses</u>
   - Stratifying according to the overall risk of bias will produce multiple estimates of the intervention effect: for example, one on low risk studies, one on studies with some concerns and one with studies at high risk of bias and an overall total for all studies.

- o   This stratification should apply to all outcomes within the review. All stratified groups should be presented in the Effects of the interventions, abstract, SoFs and PLS.

3. <u>Present all studies in primary analyses</u>
   - o   Historically this strategy has been the most commonly used in past CGF reviews. If using this option, it is mandatory to conduct sensitivity analyses for the primary outcomes removing studies judged to be at an overall high risk of bias.
   - o   Sensitivity analyses should only be presented in the Effects of the interventions, not in the abstract, PLS or SoFs.

Example text for strategy 1:

We will present the primary analysis including trials to be judged at [overall low risk of bias/overall low risk bias and with some concerns with regard to bias (<u>choose one</u>)].

If the studies are sufficiently similar, we will combine the data using a [fixed effect/random effects (<u>choose one</u>)] model in the following comparisons:

1. *Factor X versus placebo, (subgrouped by dose or mode of administration if appropriate [see Subgroup analysis below]). We plan to pool the data for the included studies.*
2. *Factor X versus Factor G*
3. *Factor X versus Factor H*

Example text for strategy 2:

We will present the primary analysis stratified by the overall risk of bias with trials judged to be at low risk, some concerns or high risk of bias presented separately but also combined to give an overall total including all studies.

If the studies are sufficiently similar, we will combine the data using a [fixed effect/random effects (<u>choose one</u>)] model in the following comparisons:

1. *Factor X versus placebo, (subgrouped by dose or mode of administration if appropriate [see Subgroup analysis below]). We plan to pool the data for the included studies.*
2. *Factor X versus Factor G*
3. *Factor X versus Factor H*

Example text for strategy 3:

If the studies are sufficiently similar, we will combine the data using a [fixed effect/random effects (<u>choose one</u>)] model in the following comparisons:

1. *Factor X versus placebo, (subgrouped by dose or mode of administration if appropriate [see Subgroup analysis below]). We plan to pool the data for the included studies.*
2. *Factor X versus Factor G*
3. *Factor X versus Factor H*

Subgroup analysis and investigation of heterogeneity

KEY POINTS

- Keep subgroups to an absolute minimum.
- Subgroups can be defined either by characteristics of the study or by those of the participants if studies have reported results stratified by patient characteristics. In practice, the latter are unlikely to be available in reported data. Subgroups should be explicit and few. Preferably the rationale for each will be clear from the Background section. If not, it must be explained here.
- If subgroups are to be compared, this should be done with a formal statistical test which is available on RevMan plots. Interpretation of the statistical analysis for subgroups must be conservative as subgroups are based on observation evidence.
- Subgroup analysis should be performed on the primary analysis for each outcome. For instance, if the review team has restricted the primary analysis to studies at overall low risk of bias, then subgroup analysis should only be performed among these studies.

Example text:

Where data are available, we will conduct subgroup analyses to determine the separate evidence within the following subgroups:

1. *Studies of low dose versus studies of high dose*
2. *Studies including only women with a high BMI (> 32kg/m2)*

If we detect substantial heterogeneity, we will explore possible explanations in subgroup analyses (e.g. differing populations) and/or sensitivity analyses (see below). We will take any statistical heterogeneity into account when interpreting the results, especially if there is any variation in the direction of effect.

Sensitivity analysis

KEY POINTS

- Investigating whether the conclusions are robust is of utmost importance.
- If authors chose to present all studies in the primary analyses, then they should perform sensitivity analyses restricting inclusion to studies at overall low risk or with some concern with regard to bias for outcomes as prespecified and report this in the Effects of the interventions.
- If authors chose to restrict the studies presented in primary analyses to those with overall low risk or with some concern with regard to bias, then their sensitivity analysis should include all studies regardless of risk of bias.

Example text:

We will conduct sensitivity analyses on all outcomes to determine whether the conclusions are robust to arbitrary decisions made regarding the eligibility and analysis. These analyses will include consideration of whether the review conclusions would have differed if:

1. *Inclusion of studies [was not restricted to those with overall low risk or with some concern with regard to bias/was restricted to those with overall low risk or with some concern with regard to bias (*<u>delete one</u>*)].*
2. *A [fixed effect/random effects (*<u>delete one</u>*)] model had been adopted.*
3. *Alternative imputation strategies had been implemented.*
4. *The summary effect measure had been relative risk rather than odds ratio.*

Summary of findings and assessment of the certainty of the evidence

KEY POINTS

- SoF tables will appear at the front of the published review.
- Additional SoFs can be prepared for other important comparisons (those reported in full in the abstract), but it is not essential to have SoFs for every comparison.
- Include a maximum of seven outcomes in each SoF table i.e. the main outcomes which include your primary outcomes and all other outcomes unless there are more than seven in which case please prespecify which secondary outcomes will be included.
- You must include the same outcomes for each comparison.
- The same comparisons and outcomes should be reported in the abstract and PLS as in the SoF tables.
- The outcomes in SoF tables will be those in primary analyses prespecified using one of the three options presented in the Data synthesis section.
    - If using strategy 1 authors can consider including the sensitivity analyses (including all studies) in the SoFs.
    - If using strategy 2 or 3 authors should only include the primary analyses in the SoFs. Sensitivity analyses can be presented in narrative in the Effects of the interventions.
- For examples of how to grade evidence see the *Cochrane Handbook* <u>Chapter 14.2</u> and <u>How to grade</u>.
- References for this section:
    - Schünemann HJ, Higgins JPT, Vist GE, Glasziou P, Akl EA, Skoetz N, Guyatt GH. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.2 (updated February 2021). Cochrane, 2021. Available from <u>www.training.cochrane.org/handbook</u>.
    - GRADEpro GDT 2015 - GRADEpro GDT [Computer program]. Version accessed XXX (for full review or update insert date that applies). Hamilton (ON): GRADE Working Group, McMaster University, 2015. Available at gradepro.org.

Example text:

We will prepare summary of findings tables using GRADEpro and Cochrane methods (Schünemann 2021; GRADEpro GDT 2015). These tables will evaluate the overall certainty of the body of evidence for the main review outcomes (live birth, ongoing pregnancy, multiple pregnancy, OHSS) for the main review comparison (Treatment A versus Treatment B). Additional summary of findings tables will be also prepared for the main review outcomes for other important comparisons (Treatment A versus placebo, and Treatment B versus Treatment C). We will assess the certainty of the evidence using GRADE criteria: risk of bias, consistency of effect, imprecision, indirectness, and publication bias). Judgements about evidence certainty (high, moderate, low or very low) will be made by at least two review authors working independently, with disagreements resolved by discussion. Judgements will be justified, documented, and incorporated into reporting of results for each outcome.

We plan to extract study data, format our comparisons in data tables and prepare a summary of findings table before writing the results and conclusions of our review.

Core outcome sets and definitions

CGF strongly encourages the use of the following core outcome sets for infertility and endometriosis reviews.

Embedding the core outcome sets within RCTs and systematic reviews should ensure the comprehensive selection, collection and reporting of core outcomes. The generic reporting tables below should provide clear guidance to researchers and improve the reporting of their results.

Core outcomes for endometriosis (endo:outcomes)

A core outcome set for future endometriosis research

Trials evaluating treatments for pain and other symptoms associated with endometriosis
Overall pain
Improvement in most troublesome symptom
Quality of life

Trials evaluating treatments for infertility associated with endometriosis*
Viable intrauterine pregnancy confirmed by ultrasound
Pregnancy loss
Live birth
Time to pregnancy leading to live birth
Gestational age at delivery
Birthweight
Neonatal mortality
Major congenital abnormalities

All trials
Adverse events
Patient satisfaction with treatment

---

*see definitions for infertility outcomes in the previous table 'A core outcome set and standardised definitions for future infertility research'

Core outcomes for infertility trials (COMMIT)

A core outcome set and standardised definitions for future infertility research

Viable intrauterine pregnancy confirmed by ultrasound	A pregnancy diagnosed by ultrasonographic examination of at least one fetus with a discernible heartbeat.	- Researchers should report at which gestation the ultrasound examination was performed. - Pregnancies are counted as pregnancy events, for example, a twin pregnancy is counted as one pregnancy event. - Effect size estimates and 95% confidence interval should be reported for pregnancy events. The denominator should be per participant randomized. - Singleton, twin and higher multiple pregnancy should be reported separately.
Pregnancy loss		When considering twin and higher multiple pregnancies, pregnancy loss should be explicitly accounted for.
Ectopic pregnancy	A pregnancy outside the uterine cavity, diagnosed by ultrasound, surgical visualization or histopathology.	
Miscarriage	The spontaneous loss of an intrauterine pregnancy prior to 20 completed weeks of gestational age.	Miscarriage should be reported after a viable pregnancy has been confirmed by ultrasound.
Stillbirth	The death of a fetus prior to the complete expulsion or extraction from its mother after 20 completed weeks of gestational age. The death is determined by the fact that, after such separation, the fetus does not breathe or show any other evidence of life, such as heartbeat, umbilical cord pulsation or definite movement of voluntary muscles.	When considering stillbirth involving twins and higher multiple births they should be reported as a single event.
Termination of pregnancy	Intentional loss of an intrauterine pregnancy, through intervention by medical, surgical or unspecified means.	Selective embryo or fetal reduction should be reported.
Live birth	The complete expulsion or extraction from a woman of a product of fertilization, after 20 completed weeks of gestational age; which, after such separation, breathes or shows any other evidence of life, such as heart beat, umbilical cord pulsation or definite movement of	- Live births are counted as birth events, for example, twin live birth is counted as one live birth event. - Effect size estimates and 95% confidence interval should be reported for live birth events. The denominator should be per participant randomized.

	voluntary muscles, irrespective of whether the umbilical cord has been cut or the placenta is attached. A birth weight of 350 g or more can be used if gestational age is unknown.	• Singletons, twin and higher multiple births should be reported separately.
Gestational age at birth	The age of a fetus is calculated by the best obstetric estimate determined by assessments which may include early ultrasound, and the date of the last menstrual period, and/or perinatal details. In the case of assisted reproductive techniques, it is calculated by adding 14 days to the number of completed weeks since fertilization.	• The gestational age of both live births and stillbirths should be reported. • Gestational age at birth should be reported as a median and interquartile range. Reporting the mean and standard deviation in addition would support future meta-analysis.
Birthweight	Birth weight should be collected within 24 h of birth and assessed using a calibrated electronic scale with 10-g resolution.	• The birthweight of singletons, twins and higher multiples should be reported separately. • Birthweight for each newborn infant of the multiple birth set should be reported. • Birthweight should not be adjusted for gestational age. • The birthweight of stillbirths should be reported.
Neonatal mortality	Death of a live born baby within 28 days of birth. This can be sub-divided into early neonatal mortality, if death occurs in the first 7 days after birth and late neonatal, if death occurs between 8 and 28 days after birth.	• Mortality related to preterm infants should be collected up to 28 days beyond their estimated due date. • If a member of a multiple birth set dies in the neonatal period this should be explicitly reported.
Major congenital anomaly	Structural, functional and genetic anomalies, that occur during pregnancy, and identified antenatally, at birth, or later in life, and require surgical repair of a defect, or are visually evident, or are life-threatening, or cause death.	• Major congenital anomalies should be classified using a standardized taxonomy. • Major congenital anomaly should be reported as an infant with at least one major congenital anomaly detected. • If a major congenital anomaly is identified in a member of a multiple set this should be explicitly reported.
*When applicable – **Time to pregnancy leading to live birth**		• Detailed guidance regarding the collection, analysis and reporting of time to pregnancy leading to live birth was approved by the meeting participants (see supplementary data file).

References for COMMIT and endo:outcomes:

- Duffy JM, AlAhwany H, Bhattacharya S, Collura B, Curtis C, Evers JL, Farquharson RG, Franik S, Giudice LC, Khalaf Y, Knijnenburg JM. Developing a core outcome set for future infertility research: an international consensus development study. Human Reproduction. 2020 Dec;35(12):2725-34. https://doi.org/10.1093/humrep/deaa241.
- Duffy JM, Bhattacharya S, Bhattacharya S, Bofill M, Collura B, Curtis C, Evers JL, Giudice LC, Farquharson RG, Franik S, Hickey M. Standardizing definitions and reporting guidelines for the infertility core outcome set: an international consensus development study. Human Reproduction. 2020 Dec;35(12):2735-45. https://doi.org/10.1093/humrep/deaa243.
- Duffy JM, Hirsch M, Vercoe M, Abbott J, Barker C, Collura B, Drake R, Evers JL, Hickey M, Horne AW, Hull ML. A core outcome set for future endometriosis research: an international consensus development study. BJOG: An International Journal of Obstetrics & Gynaecology. 2020 Jul;127(8):967-74. https://doi.org/10.1111/1471-0528.16157.

Related social media platforms:

- https://twitter.com/EndoOutcomes
- https://twitter.com/FertilityTop10
- https://twitter.com/CoreOutcomes

Cochrane
Gynaecology and Fertility

Part E – Results

Example text:

Description of studies

No text should appear directly under this heading

Results of the search

The search retrieved 97 articles. Eleven studies (13 articles) were potentially eligible and were retrieved in full text. Four studies (five articles) met our inclusion criteria. We excluded five studies and two are ongoing. See the PRISMA flow chart and study tables: Characteristics of included studies, Characteristics of excluded studies, Characteristics of studies awaiting classification.

Included studies

- Insert four subheadings (as below) and briefly summarise important points. Include full details of individual studies in Characteristics of included studies table (not this section).

Study design and setting

Four parallel design randomised controlled trials (RCTs) were included. All were multicentre studies conducted in ART clinics in the Netherlands or the UK.

Participants

The four studies included 811 infertile women undergoing ART. Their mean age ranged across studies from 32 to 37 years.

Interventions

All four studies compared Treatment A with Treatment B

Outcomes

All four studies reported live birth and multiple pregnancy

All four studies also reported other adverse effects, but none included data suitable for analysis

Excluded studies

Five studies were excluded from the review, for the following reasons:
- *3/5 were not RCTs*
- *2/5 did not compare Treatment A versus Treatment B*

    - Only studies that initially appeared eligible should be listed as excluded studies (i.e. if you had to read the full-text publication in order to determine that it was ineligible).

Risk of bias in included studies

KEY POINTS

- RoB 2 assesses risk of bias at an outcome level so results for the risk of bias for the appropriate outcome should be listed in Effects of the interventions.

- This section should include information on how to find these judgements and some overarching comments with regard to the included studies' overall risk of bias.
- Funding source, power calculations or ethics approval are not part of the RoB 2 tool. These issues should be reported in the Characteristics of Included Studies table. If issues such as funding are of concern, they can be reported in the Results section (Description of studies), and the Discussion section (Quality of the evidence).
- For examples on criteria for judging risk of bias in all domains, see Domains of bias and how they are addressed.

Example text:

Risk of bias assessments for each outcome, including all domain judgements and support for judgement, are located in the risk of bias section (located after the Characteristics of included studies) and at the side of all forest plots.

Risk of bias of outcomes across all studies was similar and predominantly of 'some concerns'. Study authors poorly reported the details of allocation concealment and pre-agreed statistical analysis plans in sufficient detail.

Across most outcomes risk of bias was similar: we judged it as 'some concerns'. The only exception was live birth which we judged to be at high risk of bias due to deviations from the intervention that were unbalanced between groups.

Characteristics of included studies tables
- Participants: include here the number randomised to each intervention and control group.
- Use the Notes section to record the following: clinical trial registration number – check this was prospective registration; study dates; funding; conflicts of interest; whether trial authors were contacted for missing information relating to data, RoB, etc. as necessary.

Effects of the interventions

KEY POINTS: HOW TO FORMAT YOUR RESULTS

- Include all pre-specified comparisons and outcomes: if there are no relevant data, say so. A finding of no evidence is in itself an important finding and can form the basis of the Implications for research section.
- Use the same order of comparisons and outcomes (separating primary and secondary outcomes) and numbering system as in the Methods section and data tables.
- Present results using the prespecified analysis strategy chosen from the three options in the methods section.
- Use this format for presenting results: (RR 0.89, 95% CI 0.75 to 1.05, I^2 = 31%, 3 RCTs, 811 women; low-certainty evidence; Analysis 1.1).
- Report and interpret the units used for continuous outcomes (e.g. a VAS scale of 0-10 where 0 is pain-free and 10 is unbearable pain): report this in the abstract, main text and the comments section of SoF table/s.

- Do not confuse lack of evidence of an effect with evidence of a lack of effect: say something like "There was insufficient evidence to determine whether there was a difference" not "There was no significant difference".
- As "no evidence of a difference" may imply equivalence, we do not use this unless the evidence does indeed suggest equivalence.
- Consider clinical rather than statistical significance.
- If there are multi-arm studies, avoid double-counting of controls.
- Do not describe the results of individual studies unless there is only one study in the comparison.
- If presenting multiple sensitivity analyses or different ways of subgrouping the same studies, present these in summary form (e.g. a single Table or Figure) and not in multiple forest plots.
- Report all pre-specified sensitivity and subgroup analyses at the end of each comparison. If there were too few studies to conduct the analyses, state this.
- Interpret subgroups very cautiously. In general, do not report them in the abstract.
- Report any post-hoc analyses at the end of each comparison, noting that they were not pre-specified and that they require extra caution in interpretation.
- Report the results of funnel plots e.g. "Funnel plots for the primary outcomes (live birth and ongoing pregnancy) did not suggest reporting bias".
- If there were too few studies to construct a funnel plot, state this in the results section.
- Acknowledge any substantial statistical heterogeneity detected and explore it (e.g. by means of subgroup analyses).
- Translate the effect estimates for important clinical outcomes in a user-friendly way. We suggest using percentages (as in the example below), derived from the SoF table.

Example: summary of findings table highlighting how to derive percentages

Outcome	Assumed risk (control)	Corresponding risk (intervention)	Relative effect	No of participants (studies)	Certainty of evidence	Comment
	TVUS plus serum oestradiol	TVUS alone				
Clinical pregnancy	349 per 1000	360 per 1000 (287 to 442)	OR 1.05 (0.75 to 1.48)	602 (4 studies)	⊕⊕⊕⊖ moderate	-

- Using the numbers in the 'assumed' and 'corresponding' risks columns of the SoFs, convert into a percentage and round up e.g. 349 becomes 35%, and 287 to 442 becomes 29% to 44% .
- The translation for this would be: "This suggests that in women with a 35% chance of clinical pregnancy using monitoring with TVUS plus serum oestradiol, the clinical pregnancy rate in women using TVUS alone will be between 29% and 44%."

![Cochrane Gynaecology and Fertility logo]

Example text using strategy 1:

1. Comparison of Treatment A versus Treatment B

Primary outcomes

1.1 Live birth

The studies included in this outcome were mostly judged to be of low risk for bias arising from the randomisation process and had some concerns with regard to all other bias domains leading to an overall judgement of some concerns with regard to these studies.

1.1.1 Primary analysis (overall low risk of bias only)

Due to the high risk of bias associated with some studies, we conducted a primary analysis excluding studies with a high risk or some concerns with regard to bias overall. This analysis yielded one study (Blogs 2020). We are uncertain whether Treatment A compared to Treatment B has an effect on cumulative live birth rate (RR 1.11, 95% CI 0.78 to 1.59, 1 study, 210 women; low-certainty evidence; Analysis 1.1). Evidence suggests that if the chance of live birth following Treatment B is assumed to be 34%, then the chance with Treatment A would be 27% to 55%.

1.1.2 Sensitivity analysis

When all studies reporting cumulative live birth were included in the sensitivity analysis, we are uncertain of the effect of Treatment A compared to Treatment B on cumulative live birth rate (RR 1.00, 95% CI 0.92 to 1.09, 3 studies, 621 women; low-certainty evidence; Analysis 1.1.2). The evidence suggests that if the chance of live birth following Treatment B is assumed to be 37%, the chance following Treatment A would be between 34% and 40%.

1.2 Multiple pregnancy

The studies included in this outcome were mostly judged to be of low risk for bias arising from the randomisation process and had some concerns with regard to bias in selection of the reported result leading to an overall judgement of some concerns with regard to these studies.

1.2.1 Primary analysis (overall low risk of bias only)

When the primary analysis was restricted to studies at overall low risk of bias, two studies were included. Treatment A may reduce multiple pregnancy rates compared with Treatment B (RR 0.10, 95% CI 0.02 to 0.28; 2 studies, 379 women; low-certainty evidence; Analysis 1.2). This suggests that if the chance of multiple pregnancy following Treatment B is 13%, then the chance following Treatment A would be 0% to 3%.

1.1.2 Sensitivity analysis

When all studies reporting multiple pregnancy were included in the sensitivity analysis, Treatment A probably reduces multiple pregnancy rates compared with Treatment B (RR 0.04, 95% CI 0.01 to 0.15, 4 studies, 811 women; moderate-certainty evidence; Analysis 1.2.2). This suggests that if the risk of multiple pregnancy following Treatment B is assumed to be 13%, the risk following Treatment A would be between 0% and 2%.

Secondary outcomes

1.3 Other adverse events

No studies reported on other adverse events.

- When using strategy 1: if there are no studies at low risk of bias to include in the primary analysis, ensure you report this and report the sensitivity analysis including all studies reporting the outcome of interest.
  For example:

  1.4.1 Primary analysis (overall low risk of bias only)
  This analysis was not performed, as no studies were at low risk of bias.

- If you are using strategy 2, you will need to list results for the different overall levels of risk of bias e.g. 1.1.1 studies with overall low risk of bias, 1.1.2 studies with some concerns with regard to bias, 1.1.3 studies at high risk of bias, 1.1.4 all studies regardless of overall bias.

- If using strategy 3: present results under each outcome and add details on sensitivity analysis under the relevant outcome as defined in the methods.

GRADE evidence table – suggested statements

The following wording should be used for interpretation of findings:

Level (quality/certainty) of evidence	Important benefit or harm# (e.g. absolute increase or decrease of on average at least 5% in live birth; 1 not in CI)	Less important benefit or harm# (e.g. absolute increase or decrease between 1 and 5% in live birth; 1 not in CI)	No important benefit/harm or null effect
High	improves*	improves slightly	little or no difference in [outcome]
Moderate	probably improves	probably improves slightly	probably little or no difference in [outcome]
Low	may improve	may improve slightly	In case of a wide confidence interval with an absolute difference to the left or right of at least 5% for pregnancy outcomes: **We are uncertain of the effect of [intervention]**** When the 95% CI of the absolute difference is 5% or less a better statement would be: "The analysis ruled out a clinically relevant difference but the quality of the evidence was low" (rare for low quality evidence).
Very low	We are uncertain whether [intervention] improves [outcome]***		
No events or rare events	Use comments in SoF table in a plainer language or summarise the results		
No studies	No studies were found that looked at [outcome]		

* Substitute the appropriate verb for 'improves' throughout the table, depending on the results: for example, 'increases', 'reduces', 'leads to', 'prevents''.

** You can also say "due to wide confidence interval (CI), we are uncertain / it is unclear". Use this this wording for moderate quality evidence as well in case of wide CI.

*** There is a debate about whether results which are rated as 'very low quality' should present numbers or not. Both approaches are currently used.

What is considered a clinically important benefit or harm? Predefine this in the protocol or update. For instance, for live birth we may consider an increase or decrease of at least 5% to be of important benefit or harm.

Part F- Discussion

No text should appear directly under this heading

Summary of main results

- Briefly summarise the main review findings, directly addressing the objectives. Highlight any outstanding uncertainties, balancing important benefits against important harms. Express results in the most consumer-friendly way possible. Refer to certainty of the evidence from the summary of findings table.

Overall completeness and applicability of evidence

- This section addresses the external validity of the review.
  - Did the included studies answer the review question?
  - Were relevant participants, interventions and outcomes investigated?
  - Do the review findings support current practice?
  - Comment on studies that measured outcomes but had no 'usable' data.

Quality of the evidence

- This section addresses the internal validity of the review.
  - How robust are the conclusions?
  - Use the GRADE ratings from the SoF table to describe the certainty of the evidence for each comparison and use the footnotes from the SoF table to summarise the limitations of the evidence.
  - Discuss limitations of the review at study and outcome level (e.g. regarding risk of bias), and at review-level (e.g. incomplete identification of studies, reporting bias).

Potential biases in the review process

- Comment on the strengths and limitations of the review process
  - Were all relevant studies identified?
  - Could review authors' methods have introduced bias?

Agreements and disagreements with other studies or reviews

- How do the review findings fit into the wider research context?

Part G - Authors' Conclusions

Implications for practice

- We suggest making this identical to the conclusions in the abstract (copy and paste)
- Do not go beyond the evidence reviewed, mention GRADE ratings
- If relevant, summarise the likely benefits and risks of the intervention and for whom it should be considered.

Implications for research

- Which questions have been well answered (no further trials needed?)
- Which questions remain unanswered (further trials needed?)
- Whether further trials in selected populations are warranted
- Identify any new research areas (dose modification, combined therapies etc)
- If recommending further research, structure the implications for research to address the nature of evidence required, including population, intervention comparison, outcome, and type of study.

Differences between protocol and review

- If planning any departures from the Methods specified in the protocol, check with editorial base first.
- If changes are approved, you must document and justify the changes in the 'Differences between protocol and review' section.
- Before checking in your review for editorial approval check the RevMan Web Dashboard for Validation Errors and Warnings.

Part H - Copy-Editing

All Cochrane Reviews (protocols, full reviews, and updates) are copy-edited before publication. Prior to copy-editing, all submissions are checked for readiness for copy-editing. This assessment focuses on areas such as structure and content of key sections, in-text citations, consistency and formatting of outcomes, table formatting, standard of English, and quality of references. Here we provide some tips to help you prepare your review according to the guidelines. To find out more about Cochrane's copy-editing policy and guideline go to Editorial and Publishing Policy Resource Copy-Editing.

Plain language summary
Provide explanations in plain English for all medical and scientific terms.

PRISMA diagram
Make sure that the numbers of trials included in the diagram add up and match the numbers in the text.

In-Text citations
- Ensure that citations in brackets are located just ahead of punctuation (see Citing References in the Cochrane Style Manual).

Characteristics of included studies
- Don't add full stops at ends of incomplete sentences.
- Follow colons with lower case letters (except when followed by a name).
- Use abbreviations used in text throughout the table.
- Insert a list of abbreviations used in the tables at the end of the tables in the Footnotes section.

RoB 2
- We suggest that RoB 2 assessments are presented as follows.
    o Present the domain-level judgements in the main review document (e.g. as a table, a figure, and within a forest plot of the results). Only consensus judgements across multiple assessors should be presented. If space permits, abridged free text justifications for each judgement would be an attractive supplement to this within the main review document.
    o For full transparency of the process, review authors may wish to present the answers to the signalling questions, free text supports and judgements for each assessor separately. Since these may be confusing to the reader, we recommend that they are not presented prominently, so should be included as a supplementary document uploaded to an online repository such as Figshare.
- Refer to the full guidance document for help with the RoB 2 tool.

References
- Check the formatting of all references (see Reference types) including full stops, capitalisation in titles, italics, page numbers.
- For references with more than six authors, list the first six authors followed by 'et al'.

General

- Ensure that all abbreviations are defined in full on first use separately in abstract, PLS, Main text, and Authors' conclusions.

Formatting of symbols

- Use 'to' instead of '-' to denote ranges in text (OK to use '-' in tables).
- Use 'mL' and not 'ml' or 'mLs' throughout.
- Use the correct spacing around symbols such as = < > (see Symbols and special characters).
- Use round brackets (find out more about Punctuation).