Cochrane
Gynaecology and Fertility

Cochrane Gynaecology and Fertility Group
Guidance for Authors
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.
- Please also refer to Cochrane Author Guidelines, the detailed guidance for authors preparing Cochrane reviews.
- You can create a practice review in Revman Web
- You can now use Milestones and Task assignment in Revman Web to help you follow your review development process.
- **IMPORTANT – make sure all authors have read** Cochrane's Conflict of Interest **policy and are compliant!**

Need help?

If at any point you need help or advice with your protocol, review or update 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

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 most common problems with reviews involve the issues listed below:

Focused review format – all Cochrane reviews are now using focused review format. Find out more about the format. A summary of some important changes is provided below:

You need to create an 'Overview of included studies and syntheses' table – it is now mandatory for all Cochrane reviews.

Primary outcomes are now called Critical outcomes. Secondary outcomes are called Important outcomes.

Prior to submission:

Complete "Pre-submission checklist" -it is mandatory and must be submitted with your review.

Fill in the "Authorship change form" if there are any changes in authorship since protocol or review stage. The forms mandatory and must be submitted with your review.

Complete and gather signed Conflict of interest form from authors. The form is available to download as a Word document.

Before checking in your review for editorial approval check the RevMan Web Dashboard for Validation Errors and Warnings.

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 pre-specified based on clinical importance. We encourage using a core outcome set – refer to page 27.
- 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 you are updating a 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 participants; low-certainty evidence).
- Do not report I^2 in the abstract.
- 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 34 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".
- 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 the Synthesis methods section in this guidance and section 7.6.2 of the *Cochrane Handbook*.
- Ensure sensitivity analyses are conducted for the type of data synthesis you use as stated in your methods (see page 18).
- 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 same outcomes for these sections: this means including all **main outcomes** (critical outcomes, adverse events and pre-specified important 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 (SoF) table in the Methods section (see page 23).
- 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 (critical outcomes, adverse events and prespecified important 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 contact us. We would like to check the review one more time before you submit.

- 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 strict COI policy which applies to all review authors. **It is critical that ALL authors read and comply with the 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 make sure they are compliant with the COI policy (financial and non-financial) prior to and during the review process.** Reviews that do not comply with the COI Policy will not be assessed by Cochrane. Important additional restrictions to authorship include:

- The first and last authors **must not** have:
  - any relevant financial interests
  - 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**:
  - Determine the overall study inclusion and exclusion criteria
  - Make study eligibility decisions about, extract data from, carry out the risk of bias assessment for, or perform GRADE assessments of that study.
- Please note, review authors who have direct involvement in the conduct, analysis, and publication of a study that could be included in the review, cannot make study eligibility decisions about, extract data from, carry out the risk of bias assessment for, or perform GRADE assessments of that study.
- Please note, authors joining the author team of a Cochrane Review after the publication of the protocol or after publication of the full review (for an update) must be free of relevant financial conflicts of interest for 36 months before joining the team.

Part B – Abstract

KEY POINTS

- Report all main outcomes as explained in Part A - Important information to read before starting a review.
- do not include the I^2 when reporting results in the abstract
- 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.
- **Abstract word limit: 700 to 1000 words**
- **Recommended**: Insert dynamic analysis results when reporting results in the text

Example text:

Rationale

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.

Eligibility criteria

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

Outcomes

The critical review outcomes were cumulative live birth and multiple pregnancy. Other adverse effects were an important outcome.

Risk of bias

We used Risk of Bias 1.0 tool.

Synthesis methods

We conducted meta-analyses using fixed-effect models to calculate risk ratio (RR) and 95% confidence intervals (CI) for all but one outcome, which used random-effects models due to heterogeneity. Primary analyses of the critical outcomes were restricted to studies at low risk of selection bias. Sensitivity analysis including all studies was then performed.

Included studies

We included four RCTs (811 women analysed). All studies were conducted in Europe.

Synthesis of results

Treatment A vs Treatment B

The primary analysis was restricted to studies at low risk of selection bias, which left only one study. There may be little or no difference in cumulative live birth rate when comparing Treatment A to Treatment B (RR 1.11, 95% CI 0.78 to 1.59, 1 study, 210 participants; low-certainty evidence). 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 the primary analysis was restricted to studies at low risk of selection 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 participants; low-certainty evidence). 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%.

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. Treatment may have little or no effect on the cumulative live birth. 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.

Funding

This Cochrane review had no dedicated funding.

Registration

Protocol [and previous versions] available via DOIXXX, [DOIXXX and DOIXXX].

Applicable to Protocols only: Not registered

Part C - Plain language summary

KEY POINTS

- Format the plain language summary (PLS) using the format below
- Report all main outcomes in the abstract and PLS and summarise any evidence about adverse effects (including lack of data).
- **Word limit:** 400 to 850 words, including the title.

- **Under Key messages -**Add at least 2 and no more than 3 bullet points that summarize the main findings and implications of the review.
- Explain any technical terms that appear in the key messages. The key messages will likely be read first, and they might be the only part of the summary that some people read. Do not use any terms that your readers might not understand. Even if you explain those technical terms later in the summary, you should also explain them in the key messages.
- Do not make any recommendations about whether or not a treatment should be used.
- Do not use technical phrases like 'risk of bias' or 'low-certainty evidence'.

- 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 31.
- 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 III.S2 Supplementary material: Guidance for writing a Cochrane Plain language summary.
-
- If you have any questions 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)

Key messages

- *Based on data from one study, we concluded that Treatment A may not increase live birth compared to Treatment B*
- *Treatment A may decrease multiple pregnancy compared to treatment B, in women undergoing ART*

Tailored heading: for example, *What is the problem? What is infertility? What are the available treatments?*

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.

What did we want to find out?

We wanted to find out if Treatment A was better than Treatment B to improve live birth. We also wanted to find out if Treatment A was associated with increased risk of multiple pregnancy.

What did we do?

We reviewed the evidence about the effect of Treatment A versus Treatment B in women undergoing ART. We compared and summarized the results of the studies and rated our confidence in the evidence, based on factors such as study methods and sizes.

What did we find?

We found four randomised controlled trials comparing Treatment A with Treatment B in a total of 811 women undergoing ART.

Main 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 live birth.

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.

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 between 27% and 55%. It also suggests that if the chance of multiple pregnancy following Treatment B is 13%, then the chance following Treatment A would be between 0% and 3%.

Data were lacking on other adverse effects.

What are the limitations of the evidence?

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.

How up to date is this evidence?

The evidence is current to June 20XX.

Part D – Methods

The content of the previous section "Differences between protocol and review" is now located here. Any differences between protocol and review should be reported here.

- **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.
- If more than a few sentences are needed to detail the deviations, use an additional supplementary material. Alternatively, if there were no deviations to information provided at registration, in the protocol or the last update, please state this.

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 critical 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 Synthesis methods sections.

Example text:

Types of studies

We included randomised controlled trials (RCTs). We excluded quasi-RCTs.. Crossover trials were included but only data from the first phase was included in meta-analyses, as the crossover is not a valid design in this context.

Types of participants

Women/couples with unexplained infertility undergoing ART were 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 Treatment A vs Treatment B were included.

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

Outcome measures

KEY POINTS

- The critical 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 critical 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, endometriosis, and menopause reviews, we encourage you to use the core outcome set for infertility/endometriosis/menopause – refer to page 24 for the full list with definitions.** Not all trials use the core outcome definitions. For example, live birth may be reported as > 24 weeks gestation rather than > 20 weeks. For such trials, authors should still extract the data and in their methods section add "as reported by study authors" to the definition. In case different definitions were used in each trial, authors could add a footnote in the forest to clarify what definition was used.
- 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 "Certainty of the evidence assessment" '.
- You can number the outcomes as in the example below and follow the same order in Synthesis of results.
- Focus on clinical outcomes
- 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:

Critical outcomes

1. *Live birth or ongoing pregnancy*
   - *Live birth: if live birth is reported using the core outcome set definition (delivery of a live foetus after 20 completed weeks of gestation), we will use this definition. Otherwise, we will use the definition used by the study authors, and will report the definition used in each case.*

- *Ongoing pregnancy, defined as evidence of a gestational sac with foetal heart motion at 12 weeks, confirmed with ultrasound.*

2. *Multiple pregnancy*

3. ***Important outcomes*** *Clinical pregnancy defined as evidence of a gestational sac, confirmed by ultrasound.*
4. *Miscarriage*
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) below. 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 https://trialsearch.who.int/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 searched for all published and unpublished RCTs of XX, without language restriction, in consultation with the Gynaecology and Fertility Group (CGF) Information Specialist. All search strategies are presented in Supplementary material 1.

Electronic searches

(1) We searched 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 was 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 included:

1. *LILACS and other Spanish and Portuguese language 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)*
2. *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]*
3. *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 handsearched 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 handsearched relevant journals and conference abstracts that were 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

Using Covidence, at least two of the review authors (from XX, YY and ZZ) independently screened titles and abstracts of trials for eligibility and obtained the full texts of all potentially eligible studies. Two review authors (XX and YY) independently examined these full text articles for compliance with the inclusion criteria and selected eligible studies. We corresponded with study investigators as required, to clarify study eligibility. Disagreements were resolved by discussion in the first instance, followed by consultation with a third review author (ZZ) if required. We documented the selection process with a PRISMA flow chart.

Research integrity

Cochrane has published a policy for managing potentially problematic studies with an accompanying implementation guidance. Importantly, studies that have been retracted, withdrawn, or have an expression of concern should be excluded from our reviews**. We encourage our authors to make use of a screening checklist to assess the trustworthiness of all RCTs to be included in the review**. Examples of screening checklists include:

Mol BW, Lai S, Rahim A, Bordewijk EM, Wang R, van Eekelen R, et al

Checklist to assess Trustworthiness in RAndomised Controlled Trials (TRACT checklist): concept proposal and pilot. Research integrity and peer review. 2023 Jun 20;8(1):6. doi: 10.1186/s41073-023-00130-8.

Alfirevic Z, Kellie FJ, Stewart F, Jones L, Hampson L, on behalf of Pregnancy and Childbirth Editorial Board. Identifying and handling potentially untrustworthy trials in Pregnancy and Childbirth Cochrane Reviews. Download from Policy for managing potentially problematic studies: implementation guidance.

Grey A, Bolland MJ, Avenell A, Klein AA, Gunsalus CK. Check for publication integrity before misconduct. Nature2020;577:167–9. https://doi.org/10.1038/d41586-019-03959-6

Studies that are assessed as potentially problematic should be put in Awaiting assessment.

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, data from arms that do not meet eligibility criteria should be excluded.

Example text:

Two of the review authors (XX with YY or ZZ) independently extracted data from the included studies using a data extraction form. Any disagreements were resolved by discussion or by consultation with a third review author who was not involved in data extraction for that particular study. Data extracted includes study characteristics and outcome data (see data extraction table for details, Supplementary material X). Where studies have multiple publications, we collated multiple reports of the same under a single study ID with multiple references. We corresponded with study investigators for further data on methods and/or results, as required.

Assessment of risk of bias in included studies

KEY POINTS

- For review updates authors should continue to use RoB 1, read section 8.5 of the 2011 *Cochrane Handbook* for detailed guidance on assessing each type of bias If you wish to use

RoB 2 for your update, contact Helen Nagels or Elena Kostova (h.nagels@auckland.ac.nz; e.b.kostova@amsterdamumc.nl).

- If likely sources of "other bias" can be identified in advance, these should be specified in this section and the number of domains increased accordingly. You can add the trustworthiness assessment to Other bias.
- Assessment of risk of bias involves considering the potential impact of each domain in the context of individual studies (or individual outcomes). For example, lack of blinding in RoB 1 may not increase the risk of bias if follow-up is complete and outcomes are unequivocal (e.g. live birth).
- Selective reporting in RoB 1 is a type of reporting bias that affects the internal validity of an individual study. It refers to the selective reporting of some outcomes (e.g. positive outcomes) and the failure to report others (e.g. adverse events). Study authors should report all pre-stated outcomes, which should include all outcomes that you would expect, such as adverse events. Reporting in a paper what they say they will report in the methods section of the same paper is not sufficient to get a low-risk rating. If a study has no separate published protocol, you can use the information from clinical trial registration (e.g. clinicaltrials.gov or the WHO portal). Otherwise, unless trial authors confirm relevant detail, this will generally be unclear risk.
- Reference when using RoB 1:
    - Higgins JPT, Altman DG, Sterne JAC (editors). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.2.0 (updated June 2017), Cochrane, 2017. Available from www.training.cochrane.org/handbook.

Example text:

Two of the review authors (XX and YY) independently assessed the included studies for risk of bias using the Cochrane risk of bias assessment tool (Higgins 2017) to assess: selection (random sequence generation and allocation concealment); performance (blinding of participants and personnel); detection (blinding of outcome assessors); attrition (incomplete outcome data); reporting (selective reporting); and other bias. Judgements were assigned as recommended in the Cochrane Handbook Section 8.5 (Higgins 2017). Disagreements were resolved by discussion or consultation with a third review author as required. All judgements and conclusions are presented in the risk of bias table and incorporated into the interpretation of the review findings by means of sensitivity analyses (see below). With respect to within-trial selective reporting, where identified studies fail to report the critical outcome of live birth, but do report interim outcomes such as pregnancy, we planned to assess whether the interim values are similar to those reported in studies that also report live birth.

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 Part B, Part C and Part 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.)

Example text:

For dichotomous data (e.g. live birth rates), we took the numbers of events in the control and intervention groups of each study to calculate Mantel-Haenszel odds ratios (ORs). Mean difference (MDs) between treatment groups were calculated for continuous data (e.g. quality of life) where all studies reported the outcome in the same way. If similar outcomes had been reported on different scales (e.g. quality of life) we planned to calculate the standardised mean difference (SMD). We planned to reverse the direction of effect of individual studies, if required, to ensure consistency across trials. We planned to treat ordinal data (e.g. quality of life scores) as continuous data. We presented 95% confidence intervals (CIs) for all outcomes. Where data to calculate ORs or MDs were not available, we utilised the most detailed numerical data available. For example, if dichotomous data supplies percentages with sample numbers, we used this to calculate OR's; for continuous data, if alternate measurement of error (e.g. test statistics, p values) were supplied we used 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) but they should also be presented per woman randomised.
- 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 was per woman randomised; we also included per pregnancy data for some outcomes (e.g. miscarriage). Data that did not allow valid analysis (e.g. "per cycle" data) were briefly summarised in an additional table and were not meta-analysed. We counted multiple births as one live birth event. Only first-phase data from crossover trials was included. [*OR: *Statistical advice was sought regarding the analysis of crossover trials, to facilitate the appropriate inclusion of crossover data in meta-analysis* and say what you done*].*

Dealing with missing data

KEY POINTS

- For fertility reviews it is reasonable to assume missing participants did not have the outcome of interest (e.g. when data is missing on participants for the outcome of live birth, we assume they did not have a live birth).
- 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 see text below.

Example text:

We analysed the data using 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). We attempted to obtain missing data from the original study authors. Where these were unobtainable, we undertook imputation of individual values for the critical outcome of live birth only. Live birth was assumed not to have occurred in participants without a reported outcome. For other outcomes, we analysed only the available data. We planned that any imputation undertaken would be subjected to sensitivity analysis (see Sensitivity analysis).

When studies reported sufficient detail to calculate MDs but no information on associated standard deviation (SD), we assumed the outcome to have a standard deviation equal to the highest SD from other studies within the same analysis.

Reporting bias assessment

KEY POINTS

- This section on reporting biases refers to review-wide reporting bias (e.g. publication bias, multiple publication bias, language bias etc), whereby 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.

Example text:

In view of the difficulty of detecting and correcting for publication bias and other reporting biases, we aimed to minimise their potential impact by ensuring a comprehensive search for eligible studies and by being alert for duplication of data. If there are ten or more studies in an analysis, we planned to 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).

Synthesis methods

KEY POINTS

- Statistical analysis will be performed using Review Manager Web (RevMan Web).
- 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 Investigation of heterogeneity and subgroup analysis

Example text for strategy 1:

We presented the primary analysis including trials to be judged at [low risk of selection bias/ low risk of bias on all domains (choose one)*].*

If the studies were sufficiently similar, we combined the data using a [fixed effect/random effects (choose one)*] model in the following comparisons:*

4. *Factor X versus placebo, (subgrouped by dose or mode of administration if appropriate [see **Error! Reference source not found.** section]). We plan to pool the data for the included studies.*
5. *Factor X versus Factor G*
6. *Factor X versus Factor H*

Example text for strategy 2:

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

If the studies were sufficiently similar, we combined the data using a [fixed effect/random effects (choose one)*] model in the following comparisons:*

7. *Factor X versus placebo, (subgrouped by dose or mode of administration if appropriate [see* ***Error! Reference source not found.*** *section). We plan to pool the data for the included studies.*
8. *Factor X versus Factor G*
9. *Factor X versus Factor H*

Example text for strategy 3:

If the studies were sufficiently similar, we combined the data using a [fixed effect/random effects (choose one)*] model in the following comparisons:*

10. *Factor X versus placebo, (subgrouped by dose or mode of administration if appropriate [see* ***Error! Reference source not found.****section). We plan to pool the data for the included studies.*
11. *Factor X versus Factor G*
12. *Factor X versus Factor H*

- 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 low risk of bias. Authors can choose to use either low risk of selection bias or low risk of bias on any domain.**
- 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 (or stated in the *Differences between protocol and review* section) and have the agreement of the editorial base.

The main strategy options are:

1. Primary analyses restricted to studies at low risk bias **(preferred option)**
    - We recommend this strategy for all CGF reviews.
- Restrict the primary analyses to studies judged to be at low risk of bias. **Authors will need to define whether a specific domain (usually selection bias) or all domains must be at low risk.** 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 high and unclear risk of bias were included.

- This stratification should apply to at least to critical outcomes within the review.

2. Present multiple (stratified) analyses
    - Stratifying according to the risk of bias in predetermined domains will produce multiple estimates of the intervention effect: one based on all studies, one based on studies at low risk of bias, and one based on studies at unclear/high risk of bias.
    - 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. Present all studies in primary analyses

- 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 critical outcomes removing studies judged to be at high and unclear risk of selection bias.
    - 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 presented the primary analysis including trials to be judged at [low risk of selection bias/ low risk of bias on all domains (choose one)*].*

If the studies were sufficiently similar, we combined the data using a [fixed effect/random effects (choose one)*] model in the following comparisons:*

13. *Factor X versus placebo, (subgrouped by dose or mode of administration if appropriate [see* ***Error! Reference source not found.*** *section]). We plan to pool the data for the included studies.*
14. *Factor X versus Factor G*
15. *Factor X versus Factor H*

Example text for strategy 2:

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

If the studies were sufficiently similar, we combined the data using a [fixed effect/random effects (choose one)*] model in the following comparisons:*

16. *Factor X versus placebo, (subgrouped by dose or mode of administration if appropriate [see* ***Error! Reference source not found.*** *section). We plan to pool the data for the included studies.*
17. *Factor X versus Factor G*
18. *Factor X versus Factor H*

Example text for strategy 3:

If the studies were sufficiently similar, we combined the data using a [fixed effect/random effects (choose one)*] model in the following comparisons:*

19. *Factor X versus placebo, (subgrouped by dose or mode of administration if appropriate [see **Error! Reference source not found.**section). We plan to pool the data for the included studies.*
20. *Factor X versus Factor G*
21. *Factor X versus Factor H*

Investigation of heterogeneity and subgroup analysis

KEY POINTS

- Keep subgroups to an absolute minimum.
- Subgroups can be defined either by characteristics of the study or by those of the participants. 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. Interpretation of the statistical analysis for subgroups is problematic.
- 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 low risk of bias, then subgroup analysis should only be performed among these studies.
- If you are unable to implement subgroup analysis because there are insufficient studies (the general rule is you need at least 10 studies for a meaningful analysis) or insufficient information available about them, report this as a difference between the protocol and the review (at the beginning of the Methods section).

Example text:

To determine whether findings differed in studies, we planned to perform the following subgroup analyses for the primary effectiveness outcome if substantial heterogeneity existed (I^2 statistic value > 50%) and if enough data were available.

- *Studies with low dose versus studies of high dose: benefit from treatment may vary depending on dosage received.*
- *Including only women with a high BMI (> 32kg/m2): benefit from treatment may vary in this group compared to those with a lower BMI.*

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

From July 2025 the default in RevMan will be the restricted maximum likelihood (REML) model. There are 2 options for calculating the CI: the Wald method has been the default in RevMan and should still be used where there are ≤ 3 studies in an analysis or heterogeneity is 0%. Otherwise, authors should use the Hartung Knapp Sidik and Jonkman (HKSJ) method (See handbook Chapter 10).

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 low risk of selection 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 at low risk of bias, then their sensitivity analysis should include all studies regardless of risk of bias.

Example text:

Sensitivity analyses was conducted on the critical outcomes to determine whether the conclusions were robust to arbitrary decisions we made regarding eligibility and analysis. These analyses included consideration of whether the review conclusions would have differed if:

- *We included all studies in the analysis (i.e. no restriction to studies considered to be at low risk of selection bias).*
- *A [fixed effect/random effects* (delete one)*] model had been adopted.*
- *The summary effect measure had been odds ratio rather than relative risk.*

Certainty of the evidence assessment

KEY POINTS

- Summary of Findings (SoF) tables will appear at the front of the published review.
- You can now **create** and **edit** SoF tables directly in GRADEpro as all SoF tables are linked to GRADEpro.
- 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 critical outcomes and other pre-specified important outcomes
- 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 (consistency of reporting).
- The outcomes in SoF tables will be those in primary analyses pre-specified using one of the three options presented in the Synthesis methods 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* Chapter 14.2 and How to grade.

- 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 www.training.cochrane.org/handbook.
  - 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 prepared summary of findings tables using GRADEpro and Cochrane methods (Schünemann 2021; GRADEpro GDT 2015). These tables evaluate the overall certainty of the body of evidence for the critical review outcomes (live birth and multiple pregnancy) for the main review comparison (Treatment A versus Treatment B). We also assessed the important outcomes ongoing pregnancy and OHSS. Additional summary of findings tables were also prepared for other important comparisons (Treatment A versus placebo, and Treatment B versus Treatment C). We assessed 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) were made by at least two review authors (XX with YY and ZZ) working independently, and disagreements were resolved by discussion. All judgements are justified, documented, and incorporated into the reporting of results for each outcome.

Consumer involvement

Example text: *We did not involve consumers in this review due to limited resources, but we used the core outcome set for infertility which has been developed with consumer involvement.*

Core outcome sets and definitions

CGF strongly encourages the use of the core outcome sets below 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 table on the next page: '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 foetus 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 foetus 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 foetus 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 foetal 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 foetus 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).

COMMA (Core Outcomes in Menopause)

A core outcome set for vasomotor symptoms associated with menopause

Frequency of vasomotor symptoms
Severity of vasomotor symptoms
Impact on sleep
Distress, bother or interference caused by vasomotor symptoms
Satisfaction with treatment
Side-effects of treatment

References for COMMIT, endo:outcomes, and COMMA

- 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, 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.
- Lensen S, Archer D, Bell RJ, Carpenter JS, Christmas M, Davis SR, Giblin K, Goldstein SR, Hillard T, Hunter MS, et al. A core outcome set for vasomotor symptoms associated with menopause: the COMMA (Core Outcomes in Menopause) global initiative. Menopause. 2021 Apr 26;28(8):852-858. doi: 10.1097/GME.0000000000001787.

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.

Included studies

- Summarise the characteristics of the included studies.
- !! Go to Tables > Add table to create your overview of synthesis and included studies table !!
- All included studies must be referenced in the overview of synthesis and included studies table or in this section.
- Insert four subheadings (as below) and briefly summarise important points. Include full details of individual studies in Supplementary material 2 (Characteristics of included studies) (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 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:

- *Three were not RCTs*
- *Two did not compare Treatment A versus Treatment B*
    - *No studies were classified as awaiting classification or ongoing.*
    - Link to the Supplementary material 3 (Characteristics of excluded studies), which gives reasons for exclusion.
    - Only studies that initially appeared potentially 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

- If funding source is of concern, it could be mentioned under other bias
- Consider the degree to which blinding is likely to influence specific outcomes.
- For examples on criteria for judging risk of bias in all domains, see Criteria for judging risk of bias in the 'Risk of bias' assessment tool.
- For RoB 1 (the original risk of bias tool for RCTs) link to the Characteristics of included studies supplementary material (Supplementary material 2).
- Note that using subheadings for each domain is not mandatory

Example text:

Allocation (selection bias)

Random sequence generation

Two studies were rated as at low risk of selection bias related to sequence generation, as they used computer randomisation or a random numbers table. The other two studies did not describe the method used and were rated as at unclear risk of this bias.

Allocation concealment

Two studies were rated as at low risk of selection bias related to allocation concealment as they used sequentially labelled, sealed, opaque envelopes. The other two studies failed to describe methods of allocation concealment and we rated these as at unclear risk of bias for this domain.

Blinding of participants and personnel (performance bias)

We did not consider that blinding of participants and personnel was likely to influence findings for the primary review outcomes (live birth and multiple pregnancy). Blinding might influence outcomes for other adverse events, but no studies reported relevant data for this outcome.

Blinding of outcome assessors (detection bias)

We did not consider blinding of outcome assessors likely to influence the primary review outcomes (live birth and multiple pregnancy). Blinding might influence outcomes for other adverse events as these could be observer-reported outcome measures, but no studies reported relevant data for this outcome.

Incomplete outcome data (attrition bias)

All four studies analysed all or most (>95%) of the women randomised and we judged them to be at low risk of attrition bias.

Selective reporting (reporting bias)

We rated all four studies as at low risk of selective reporting bias. All outcomes planned in the protocols were reported and these included live birth and multiple pregnancy (i.e. the most clinically relevant outcomes).

We rated xx studies as at unclear risk of bias although they reported our review's critical outcomes; we could not obtain a study protocol and the study was not prospectively registered so there was no information we could use to verify study details.

Other potential sources of bias (other bias)

In one study there was a statistically significant baseline difference in age between the two groups and the risk of bias was deemed unclear. We found no potential sources of within-study bias in the other three studies. All studies were assessed as having no concerns using the TRACT checklist.

Characteristics of included studies tables –now supplementary material

- Participants: include here the numbers 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.

Synthesis of results

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 critical and important outcome 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 studies, 811 participants; low-certainty evidence; Figure 5).
- 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 comment's section of the SoF table.
- 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".
- 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 critical outcomes 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).
- Please note that numbering of Comparisons and Outcomes is not mandatory.
- 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 the percentages

Outcome	Assumed risk (control)	Corresponding risk (intervention)	Relative effect	No of participants (studies)	Quality 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 if the clinical pregnancy rate using monitoring with TVUS plus serum oestradiol is 35%, then the clinical pregnancy rate using TVUS alone will be between 29% and 44%."

Example text using strategy 1:

1. Comparison of Treatment A versus Treatment B

Critical outcome

1.1 Live birth

Three studies reported live birth and for one study, we obtained this information after we contacted study authors.

1.1.1 Primary analysis (low risk of bias only)

Due to the high risk of bias associated with some studies, we conducted a primary analysis excluding studies at high or unclear risk of bias. This analysis yielded one study (Blogs 2020). There may be little or no difference in live birth rate when comparing Treatment A to Treatment B (RR 1.11, 95% CI 0.78 to 1.59; I^2 = 21%, 1 study, 210 participants; ; low-certainty

evidence; Figure 4.) 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; I^2 = 42%, 3 studies, 621 participants; very low-certainty evidence).

1.2 Multiple pregnancy
Four studies reported multiple pregnancy rate (Blogs 2020; XX 2019; YY 2015; ZZ 2019).

1.2.1 Primary analysis (low risk of bias only)
When the primary analysis was restricted to studies at low risk of selection 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; I^2 = 19%; 2 studies, 379 participants, low-certainty evidence). This suggests that if the risk of multiple pregnancy following Treatment B is 13%, then the risk 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, I^2 = 31%;4 studies, 811 participants; ; moderate-certainty evidence;.

Important 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 levels of risk of bias e.g. 1.1.1 studies with low risk of bias, 1.1.2 studies with unclear risk of bias, 1.1.3 studies at high risk of bias, 1.1.4 all studies regardless of 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.

Don't forget to report all pre-specified sensitivity and subgroup analysis.

Equity assessment

If the review does not consider health inequity, leave this section blank and it will not publish.

Reporting biases

Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed. If not applicable for the review, leave this section empty, and it will not be published.

Summary of findings tables – suggested statements

The following wording is recommended for interpretation of findings:

Table 1. Final list of informative statements to communicate results of systematic reviews

Size of the effect estimate	Suggested statements (replace X with intervention, replace 'reduce/increase' with direction of effect, replace 'outcome' with name of outcome, include 'when compared with Y' when needed)
HIGH Certainty of the evidence	
Large effect	X results in a large reduction/increase in outcome
Moderate effect	X reduces/increases outcome X results in a reduction/increase in outcome
Small important effect	X reduces/increases outcome slightly X results in a slight reduction/increase in outcome
Trivial, small unimportant effect or no effect	X results in little to no difference in outcome X does not reduce/increase outcome
MODERATE Certainty of the evidence	
Large effect	X likely results in a large reduction/increase in outcome X probably results in a large reduction/increase in outcome
Moderate effect	X likely reduces/increases outcome X probably reduces/increases outcome X likely results in a reduction/increase in outcome X probably results in a reduction/increase in outcome
Small important effect	X probably reduces/increases outcome slightly X likely reduces/increases outcome slightly X probably results in a slight reduction/increase in outcome X likely results in a slight reduction/increase in outcome
Trivial, small unimportant effect or no effect	X likely results in little to no difference in outcome X probably results in little to no difference in outcome X likely does not reduce/increase outcome X probably does not reduce/increase outcome
LOW Certainty of the evidence	
Large effect	X may result in a large reduction/increase in outcome The evidence suggests X results in a large reduction/increase in outcome
Moderate effect	X may reduce/increase outcome The evidence suggests X reduces/increases outcome X may result in a reduction/increase in outcome The evidence suggests X results in a reduction/increase in outcome
Small important effect	X may reduce/increase outcome slightly The evidence suggests X reduces/increases outcome slightly X may result in a slight reduction/increase in outcome The evidence suggests X results in a slight reduction/increase in outcome
Trivial, small unimportant effect or no effect	X may result in little to no difference in outcome The evidence suggests that X results in little to no difference in outcome X may not reduce/increase outcome The evidence suggests that X does not reduce/increase outcome
VERY LOW Certainty of the evidence	
Any effect	The evidence is very uncertain about the effect of X on outcome X may reduce/increase/have little to no effect on outcome but the evidence is very uncertain

The range of effects sizes compatible with the confidence interval is important. Consider explicitly describing the range of effects compatible with the CI to convey a clear message.

Santesso N, Glenton C, Dahm P, Garner P, Akl EA, Alper B et al. GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. Journal of Clinical Epidemiology. 2020 Mar;119:126-135. doi: 10.1016/j.jclinepi.2019.10.014. See table on page 131.

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 quality/certainty of the evidence from the summary of findings (SoF) table.

Limitations of the evidence included in the review

- Present an assessment of how well the evidence identified in the review addressed the review question.
- It should indicate whether the studies identified were sufficient to address all the objectives of the review, and whether all relevant types of participants, interventions and outcomes have been investigated
- Information presented under Description of studies will be useful to draw on in writing this part of the discussion.
- This section should summarize the considerations that led to downgrading or upgrading the certainty of the evidence in the implementation of GRADE. This information can be based on explanations for downgrading decisions alongside the summary of findings tables in the review.
-

Limitations of the review process

- Discuss any limitations of the review processes used and comment on the potential impact of each limitation, such as incomplete identification of studies, completeness of data collection processes, any completed studies that have been identified as potentially eligible but have not been incorporated into the review (see item above), assumptions made regarding classification of interventions, outcomes or subgroups, and methods used to account for missing results in specific syntheses. In particular, if the review methods do not allow for detection of serious or rare adverse events, or either, the review authors must explicitly state this as a limitation.

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

- You can make 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.
- Avoid making recommendations for practice.

Equity-related implications for practice

- If the review considered equity, discuss the equity-related implications for practice and policy. Otherwise, leave this section empty and it will not be published.

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.

Equity-related implications for research

- If the review considered equity, discuss the equity-related implications for research. Otherwise, leave this section empty and it will not be published.

Acknowledgements

Cochrane [GROUP/UNIT NAME] supported the authors in the development of this review. We are grateful to all authors who developed the protocol for this review, and contributed to previous review versions. The following people conducted the editorial process for this review: [NAME, AFFILIATION] (Sign-off Editor); [NAME, AFFILIATION] (Managing Editor); [NAME, AFFILIATION] (Editorial Assistant); [NAME, AFFILIATION] (clinical/content peer review), [NAME, AFFILIATION] (consumer peer review), [NAME, AFFILIATION] (methods peer review), [NAME, AFFILIATION] (search peer review). [NUMBER] of additional peer reviewers provided [CLINICAL/CONTENT/CONSUMER/METHODS/SEARCH] peer review but chose not to be publicly acknowledged. [NAME, AFFILIATION] copyedited the [ARTICLE TYPE] during the production process.*

Contributions of authors

AB, CD, and EF conducted the literature searches for the review, selected relevant trials, procured data and information about studies, assessed the validity and checked the data

extraction for each trial, entered all study information, data, and text into Revman Web, performed the analyses, wrote the abstract, background, methods, results, and conclusion sections of the review, and gave approval to the final version.

Declarations of interest

The statements entered here must accurately reflect all interests declared by the review authors in their individual Declaration of Interest forms.

Example format:

AB: Employee of the X and no commercial or non-commercial conflicts of interest relevant to this review; CD and EF: No commercial or non-commercial conflicts of interest relevant to this review.

Registration and protocol

Recommended format in protocols: *Cochrane approved the proposal for this review in MONTH YEAR.*

Recommended format in reviews and updates:

Protocol (YEAR) DOI

Original Review (YEAR) DOI

Review Update (YEAR) DOI

Data, code and other materials

Recommended format in protocols: *Data sharing is not applicable to this article as it is a protocol, so no datasets were generated or analysed.*

Recommended format in reviews and updates:

As part of the published Cochrane review, the following is made available for download for users of the Cochrane Library (see [INCLUDE LINK TO DATA PACKAGE, e.g. Supplementary material 2*]: full search strategies for each database; [DELETE IF NOT RELEVANT: full citations of each unique report for all studies (included, ongoing or awaiting classification, or excluded at the full text screen) in the final review; study data, including study information, study arms, and study results or test data; consensus risk of bias assessments; and analysis data, including overall estimates and settings, subgroup estimates, and individual data rows.] Appropriate permissions have been obtained for such use. Analyses and data management were conducted within Cochrane's authoring tool, RevMan, using the inbuilt computation methods. The following scripts and artefacts were used to generate analyses outside of RevMan: [list each including the public archive and citation]. Template data extraction forms from [Covidence, Excel, etc.] are available [from the authors on reasonable request/publicly available XXX].*

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. Go to Cochrane Style Essentials for a summary of most essential points to keep in mind in terms of 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 and risk of bias tables

- These sections are now supplementary data and they will not be copy-edited.
- Authors must make sure that all their supplementary materials comply with Cochrane editorial policies and follow Cochrane style.
- Don't add full stops at ends of incomplete sentences, e.g. 'Unclear' not 'Unclear.'
- Follow colons with lower case letters (except when followed by a name).
- RoB tables – make sure that you have used quotation marks around any text copied directly from study reports Once allocated, the treatment was revealed to both the investigator and the patient."
- 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.

References

- You can automatically update all your references to Cochrane style in RevMan.
- If needed, 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'.

See how to format the references RevMan Web, Covidence, and GRADEpro.

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).