UNIVERSITYOF BIRMINGHAM

Tocolytics for delaying preterm birth

Wilson, Amie; Hodgetts-Morton, Victoria A.; Marson, Ella J.; Markland, Alexandra D.; Larkai, Eva; Papadopoulou, Argyro; Coomarasamy, Arri; Tobias, Aurelio; Chou, Doris; Oladapo, Olufemi T.; Price, Malcolm J.; Morris, Katie; Gallos, Ioannis D.

[Link to publication on Research at Birmingham portal](#)

Cochrane Library

Cochrane Database of Systematic Reviews

Tocolytics for delaying preterm birth: a network meta-analysis (Protocol)

Wilson A, Hodgetts-Morton VA, Marson EJ, Markland AD, Larkai E, Papadopoulou A, Coomarasamy A, Tobias A, Chou D, Oladapo OT, Price MJ, Morris K, Gallos ID

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Protocol]

Tocolytics for delaying preterm birth: a network meta-analysis

Amie Wilson[1], Victoria A Hodgetts-Morton[1], Ella J Marson[1], Alexandra D Markland[1], Eva Larkai[2], Argyro Papadopoulou[2], Arri Coomarasamy[2], Aurelio Tobias[2], Doris Chou[3], Olufemi T Oladapo[3], Malcolm J Price[4], Katie Morris[1], Ioannis D Gallos[2]

[1]Institute of Applied Health Research, University of Birmingham, Birmingham, UK. [2]Tommy's National Centre for Miscarriage Research, Institute of Metabolism and Systems Research (IMSR), WHO Collaborating Centre for Global Women's Health Research, University of Birmingham, Birmingham, UK. [3]UNDP/UNFPA/UNICEF/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP), Department of Sexual and Reproductive Health and Research, World Health Organization, Geneva, Switzerland. [4]Test Evaluation Research Group, Institute of Applied Health Research, University of Birmingham, Birmingham, UK

Contact address: Ioannis D Gallos, i.d.gallos@bham.ac.uk, gallosioannis@gmail.com.

A B S T R A C T

Objectives

This is a protocol for a Cochrane Review (intervention). The objectives are as follows:

To conduct a meta-analysis to compare the effectiveness of tocolytic drugs for delaying preterm birth and to generate a ranking among all available tocolytic drugs according to their relative effectiveness and side effects.

BACKGROUND

Description of the condition

In 2019, five million children under five years of age died. Almost half of these deaths occurred in the first month of life (UNIGME 2020). Preterm birth is the most important contributing factor for high newborn death rates, and is the leading cause of death in under-fives (Liu 2016). Preterm birth is defined as birth before 37 completed weeks of pregnancy. In addition to altering the survival chances of newborns, preterm birth also causes significant morbidity (WHO 1977). Preterm infants are at increased risk of short-term complications such as breathing complications and difficulties with feeding and body temperature regulation, and long-term complications including neurodevelopmental, respiratory, and gastrointestinal complications (Escobar 2006; Kinney 2006; Wang 2004). Despite advances in medicine, the number of preterm births appears to be rising in most countries (WHO 2018).

The multifactorial aetiology of preterm birth means that it is difficult to predict and prevent. Several risk factors have been identified, including multiple pregnancy, infection, cervical insufficiency, maternal medical conditions, and previous history of miscarriage and preterm birth (Blondel 2006; Lee 2008). Preterm birth can either be spontaneous (occurring without medical intervention) or iatrogenic (when the pregnancy is interrupted before term due to maternal or fetal compromise). The cause of spontaneous preterm labour often remains uncertain (Menon 2008). Iatrogenic preterm birth occurs only in cases where the continuation of the pregnancy poses great risks to the mother or the fetus (or both), and its prevention should focus on preventing contributing conditions such as pre-eclampsia (Kalra 2008; Mukhopadhaya 2007).

Description of the intervention

Tocolytic drugs have been used for delaying preterm birth since the 1950s. Tocolytic drugs aim to delay preterm birth by suppressing uterine contractions. Specifically, they induce smooth muscle relaxation by engaging slightly different mechanisms of action, and as a result each has different side-effects and different administration challenges. Even within individual drug classes there is significant variation in administration regimens. There are many different types of tocolytic drugs, however most fall within the following tocolytic drug classes.

1. Betamimetics (e.g. ritodrine)
2. Calcium channel blockers (e.g. nifedipine)
3. Magnesium sulphate
4. Oxytocin receptor antagonists (e.g. atosiban)
5. Nitric oxide donors (e.g. glyceryl trinitrate)
6. Cyclo-oxygenase (COX) inhibitors (e.g. indomethacin)
7. Combinations of tocolytics (e.g. betamimetics plus magnesium sulphate)

Betamimetics (e.g. ritodrine, terbutaline, and salbutamol) have been widely used, especially in resource-poor countries. Their use has declined over time due to their side effects (NICE 2015). Betamimetics are beta receptor agonists and as a result can cause heart palpitations, tremor, nausea, vomiting, headaches, nervousness, anxiety, chest pain, shortness of breath, and biochemical disturbances such as hyperglycaemia. Rarely, they can cause heart failure and pulmonary oedema (Medicines.org.uk 2020). Betamimetics cross the placenta and cause fetal tachycardia and neonatal hypoglycaemia (Medicines.org.uk 2020). They can be administered orally, subcutaneously, intramuscularly, and intravenously.

Calcium channel blockers (e.g. nifedipine) are used for the treatment of hypertension in pregnancy, and are increasingly also used as tocolytic drugs. Calcium channel blockers are administered orally. They are generally better tolerated than betamimetics, but are associated with cardiovascular side effects, such as headache, hypotension, dyspnoea, pulmonary oedema, and even myocardial infarction (Medicines.org.uk 2020).

Magnesium sulphate is used widely in obstetrics for the prevention and treatment of eclampsia. It is also an established fetal neuroprotective drug, and is given to women with imminent preterm birth for the prevention of cerebral palsy in infants and children (WHO 2015). It can also be used as a tocolytic drug as it decreases the frequency of depolarisation of smooth muscle, which in turn inhibits uterine contractions. Magnesium sulphate can be administered intravenously or intramuscularly. In current clinical practice, intramuscular administration regimens are recommended only if intravenous access is not possible. Side effects are dose-dependent and include nausea, vomiting, headache, heart palpitations, and, rarely, pulmonary oedema (Medicines.org.uk 2020). Concentrations above the recommended therapeutic range can lead to respiratory depression, respiratory arrest, and cardiac arrest (Crowther 2014).

Oxytocin receptor antagonists (e.g. atosiban) are the only drugs that have been purposefully developed to delay preterm birth. They block oxytocin receptors, and by blocking the action of oxytocin they are able to prevent uterine contractions and relax the uterus. They can only be administered intravenously, and are associated with side effects such as nausea, vomiting, headache, chest pain, and hypotension (Medicines.org.uk 2020). However, their side effect profile is considered more favourable compared to other tocolytic drugs. Important disadvantages of the oxytocin receptor antagonists are their cost and availability.

Nitric oxide donors (e.g. glyceryl trinitrate) have also been used as tocolytic drugs. Nitric oxide is a free radical that induces smooth muscle relaxation, cervical ripening, and vasodilation. The effect of nitric oxide donors on the uterus is fast, which can be of great value in obstetric emergencies. They can be administered intravenously, transdermally or sublingually, and are typically associated with maternal adverse effects related to vasodilation, such as headache, flushing, hypotension and tachycardia (Duckitt 2014). Nitric oxide donors could adversely affect the developing fetus because they induce changes to the uterine blood flow (Duckitt 2014).

Cyclo-oxygenase (COX) inhibitors (e.g. indomethacin) are frequently used and can easily be administered orally or rectally. They have a more favourable side-effect profile compared with betamimetics (Babay 1998). However, COX inhibitors easily cross the placenta and can interfere with the fetal prostaglandin homeostasis. A meta-analysis published in 2006 found that even short-term use of COX inhibitors in late gestations is associated with a 15-fold increase of premature ductal closure (Koren 2006). Because of these concerns, COX inhibitors are currently contraindicated in the third trimester. In view of this

contraindication, and because COX inhibitors are rarely used as tocolytic drugs, they are not of direct interest to this review and are only included as a historical intervention to improve the inferences among the other interventions.

Combinations of tocolytic drugs from different classes (e.g. betamimetics plus magnesium sulphate) have been used together to delay preterm birth. Using tocolytic drugs from different classes suppress uterine contractions by targeting different pathways in the myometrium. Using a combination of tocolytic drugs could have the benefit of improving the desirable effects while using lower doses of the drugs resulting in fewer side effects.

How the intervention might work

Tocolytics can potentially delay preterm birth by suppressing uterine contractions (Haas 2009). The rationale for tocolysis is that the delay in preterm birth can allow time for antenatal optimisation; this includes the administration of corticosteroids for fetal lung maturation, magnesium sulphate for neuroprotection, antibiotics for Group B Streptococcus prophylaxis, and time for the pregnant person to be transported to a facility with appropriate neonatal care facilities.

Why it is important to do this review

With the increasing contribution of neonatal deaths to overall child mortality, it is critical to address the determinants of poor outcomes related to preterm birth to achieve further reductions in child mortality. Infant mortality and morbidity can be reduced through interventions delivered to the mother before or during pregnancy, and to the infant after birth. The most beneficial set of maternal interventions are those that are aimed at improving outcomes for preterm infants when preterm birth is inevitable (e.g. antenatal corticosteroids, magnesium sulphate, and antibiotic prophylaxis) (WHO 2015). The success of the antenatal optimisation is dependent on appropriate timing of these interventions. For example, corticosteroids are more beneficial when administered more than 24 hours before birth, but no more than seven days before birth; magnesium sulphate needs to be administered no more than 24 hours prior to birth; antibiotics are administered in labour; and transfer takes time to arrange. Therefore, once a diagnosis of preterm labour is made, prompt action is vital for maximising survival and reducing complications for the infant.

Tocolytics potentially delay preterm birth, which in turn could enhance the beneficial effects of the interventions mentioned above. However, there is still uncertainty about whether they are effective in improving health outcomes. Current guidelines state that tocolytic drugs are not recommended for women at risk of imminent preterm birth for the purpose of improving newborn outcomes (WHO 2015). The evidence informing these guidelines was based on low-certainty evidence from several individual Cochrane Reviews containing small- to medium-sized trials (Bain 2013; Crowther 2014; Duckitt 2014; Flenady 2014a; Flenady 2014b; Neilson 2014; Reinebrant 2015; Su 2014).

The comparisons of interest for this review are those of tocolytic drugs versus placebo or no treatment, to determine if tocolytics are effective in delaying preterm birth and improving neonatal outcomes. The comparison of tocolytic drugs with each other

is also of interest, the aim being to determine which tocolytic drug is the most effective. Where several competing drug options exist, not all of which have been directly compared, a network meta-analysis may allow for more comparisons to be made and a more comprehensive synthesis of relative effects for all available tocolytic drugs (Caldwell 2005; Caldwell 2010). A network meta-analysis, unlike conventional Cochrane Reviews, simultaneously pools all direct and indirect evidence into one single coherent analysis. Indirect evidence is obtained by inferring the relative effectiveness of two competing drugs through a common comparator, even when these two drugs have not been compared directly. A network meta-analysis also calculates the probability for each competing drug to constitute the most effective drug with the least side effects, thereby allowing ranking of the available tocolytic drugs (Caldwell 2005).

O B J E C T I V E S

To conduct a meta-analysis to compare the effectiveness of tocolytic drugs for delaying preterm birth and to generate a ranking among all available tocolytic drugs according to their relative effectiveness and side effects.

M E T H O D S

Criteria for considering studies for this review

Types of studies

All randomised controlled trials or cluster-randomised trials comparing tocolytic drugs with other tocolytic drugs, placebo or no treatment will be eligible for inclusion. Cross-over trials and quasi-randomised trials will be excluded. The cross-over study design is inappropriate to investigate the effectiveness of tocolytic drugs, and quasi-randomisation rather than true randomisation introduces an elevated risk of bias that we wish to eliminate for the purpose of this review. Randomised trials published only as abstracts will be eligible only if sufficient information can be retrieved.

Types of participants

This review will include trials involving women with live fetus(es), with signs and symptoms of preterm labour defined as uterine activity with or without ruptured membranes; or ruptured membranes, with or without cervical dilatation or shortening or biomarkers consistent with a high risk of preterm birth. We will consider studies conducted in all settings.

Types of interventions

Trials will be eligible if they administered tocolytic drugs of any dosage, route, or regimen for delaying preterm birth, and compared them with other tocolytic drugs, placebo, or no treatment. We will exclude trials which exclusively compared different dosages, routes or regimens of the same tocolytic drug. Eligible interventions include the tocolytic classes listed below and shown in Figure 1; if we identify in the included studies interventions that we are not aware of, we will consider them as eligible and include them in the network after assessing their comparability with those tocolytic classes named below.

Figure 1. Network meta-analysis diagram.
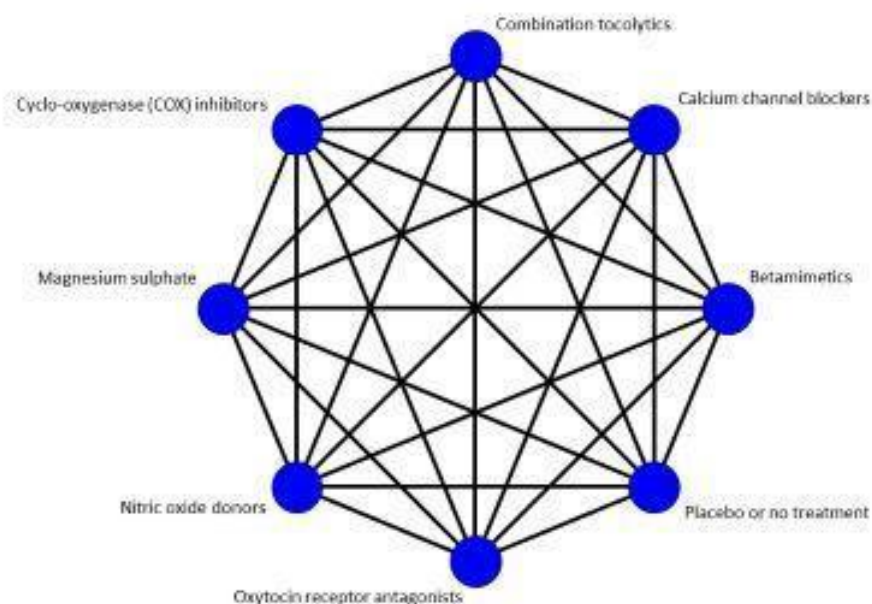
1. Betamimetics (e.g. ritodrine, terbutaline, and salbutamol)
2. COX inhibitors (e.g. indomethacin)
3. Calcium channel blockers (e.g. nifedipine)
4. Magnesium sulphate
5. Oxytocin receptor antagonists (e.g. atosiban)
6. Nitric oxide donors (e.g. glyceryl trinitrate)
7. Combinations of tocolytics (e.g. betamimetics plus magnesium sulphate)

Participants in the network could in principle be randomised to any of the tocolytic drugs being compared. All tocolytic drugs are of direct interest to this review except COX inhibitors. COX inhibitors are only included as a historical intervention to improve the inferences among the other interventions. We will include trials in which adjuvant co-interventions such as progesterone or cervical cerclarge (inserting a stitch around the cervix) were administered in combination with tocolytic drugs; the effects of such co-interventions will be tested through sensitivity analyses. We will include information about co-interventions of interest in the table of study characteristics.

Types of outcome measures

Outcomes will be based on World Health Organization critical outcomes for preterm birth and will include both neonatal and maternal outcomes (WHO 2015).

Primary outcomes

The main (primary) outcomes are as follows. These outcomes will feature in the 'Summary of findings' tables.

1. Delay in birth by 48 hours
2. Delay in birth by seven days
3. Neonatal death before 28 days
4. Pregnancy prolongation (time from trial entry to birth)
5. Serious adverse effects of drugs
6. Maternal infection
7. Cessation of treatment due to side effects

Secondary outcomes

1. Birth prior to 28/40 weeks of gestation
2. Birth prior to 32/40 weeks of gestation
3. Birth prior to 34/40 weeks of gestation
4. Birth prior to 37/40 weeks of gestation
5. Maternal death
6. Pulmonary oedema
7. Dyspnoea
8. Palpitation
9. Headaches
10. Nausea or vomiting
11. Tachycardia
12. Maternal cardiac arrhythmias
13. Maternal hypotension
14. Perinatal mortality
15. Stillbirth
16. Neonatal death before seven days
17. Neurodevelopmental morbidity
18. Gastrointestinal morbidity
19. Respiratory morbidity
20. Mean birthweight
21. Birthweight less than 2000 g
22. Birthweight less than 2500 g

23. Gestational age at birth
24. Neonatal infection

Search methods for identification of studies

Electronic searches

We will search Cochrane Pregnancy and Childbirth's Trials Register by contacting their Information Specialist. The Register is a database containing over 25,000 reports of controlled trials in the field of pregnancy and childbirth. It represents over 30 years of searching. For full current search methods used to populate Pregnancy and Childbirth's Trials Register, including the detailed search strategies for CENTRAL, MEDLINE, Embase and CINAHL; the list of handsearched journals and conference proceedings; and the list of journals reviewed via the current awareness service, please follow this link.

Briefly, Cochrane Pregnancy and Childbirth's Trials Register is maintained by their Information Specialist and contains trials identified from:

1. monthly searches of the Cochrane Central Register of Controlled Trials (CENTRAL);
2. weekly searches of MEDLINE (Ovid);
3. weekly searches of Embase (Ovid);
4. monthly searches of CINAHL (EBSCO);
5. handsearches of 30 journals and the proceedings of major conferences;
6. weekly current awareness alerts for a further 44 journals plus monthly BioMed Central email alerts.

Search results are screened by two people and the full text of all relevant trial reports identified through the searching activities described above is reviewed. Based on the intervention described, each trial report is assigned a number that corresponds to a specific Pregnancy and Childbirth review topic (or topics), and is then added to the Register. The Information Specialist searches the Register for each review using this topic number rather than keywords. This results in a more specific search set that will be fully accounted for in the relevant review sections ('Included studies', 'Excluded studies', 'Studies awaiting classification' or 'Ongoing studies').

In addition, we will search ClinicalTrials.gov and the WHO International Clinical Trials Registry Platform (ICTRP) for unpublished, planned, and ongoing trial reports. The search terms we use will be detailed in an appendix in the full review.

Searching other resources

We will search the reference lists of retrieved studies. We will not apply any language or date restrictions.

Data collection and analysis

Selection of studies

Two review authors will independently assess for inclusion all the potential studies we identify as a result of the search strategy. We will resolve any disagreement through discussion or, if required, we will consult a third person. We will create a study flow diagram to present the number of records identified, included and excluded.

Screening eligible studies for scientific integrity/trustworthiness

All studies meeting our inclusion criteria will also be evaluated by two review authors against predefined criteria to select studies that, based on available information, were deemed to be sufficiently trustworthy to be included in the analysis. These criteria have developed by Cochrane Pregnancy and Childbirth (see Appendix 1). The criteria are as follows.

Research governance

1. No prospective trial registration for studies published after 2010 without plausible explanation
2. When requested, trial authors refuse to provide/share the protocol or ethics approval letter (or both)
3. Trial authors refuse to engage in communication with the Cochrane Review authors
4. Trial authors refuse to provide trial data upon request with no justifiable reason

Baseline characteristics

1. Characteristics of the study participants being too similar (distribution of mean (standard deviation (SD)) excessively narrow or excessively wide, as noted by Carlisle 2017)

Feasibility

1. Implausible numbers (e.g. 500 women with severe cholestasis of pregnancy recruited in 12 months)
2. (Close to) zero losses to follow-up without plausible explanation

Results

1. Implausible results (e.g. massive risk reduction for main outcomes with small sample size)
2. Unexpectedly even numbers of women 'randomised', including a mismatch between the numbers and the methods, e.g. if it is stated that no blocking was used but there are still equal numbers, or it is stated that blocks of four were used but the final numbers differ by six

Where a study is classified as being at 'high risk' for one or more of the above criteria, we will attempt to contact the study authors to address any possible lack of information and concerns. If adequate information remains unavailable, the study will be categorised as 'awaiting classification', and the concerns and communications with the author (or lack thereof) will be described in detail. The process is described fully in Figure 2.

Figure 2. Process for using the Cochrane Pregnancy and Childbirth criteria for assessing the trustworthiness of a study
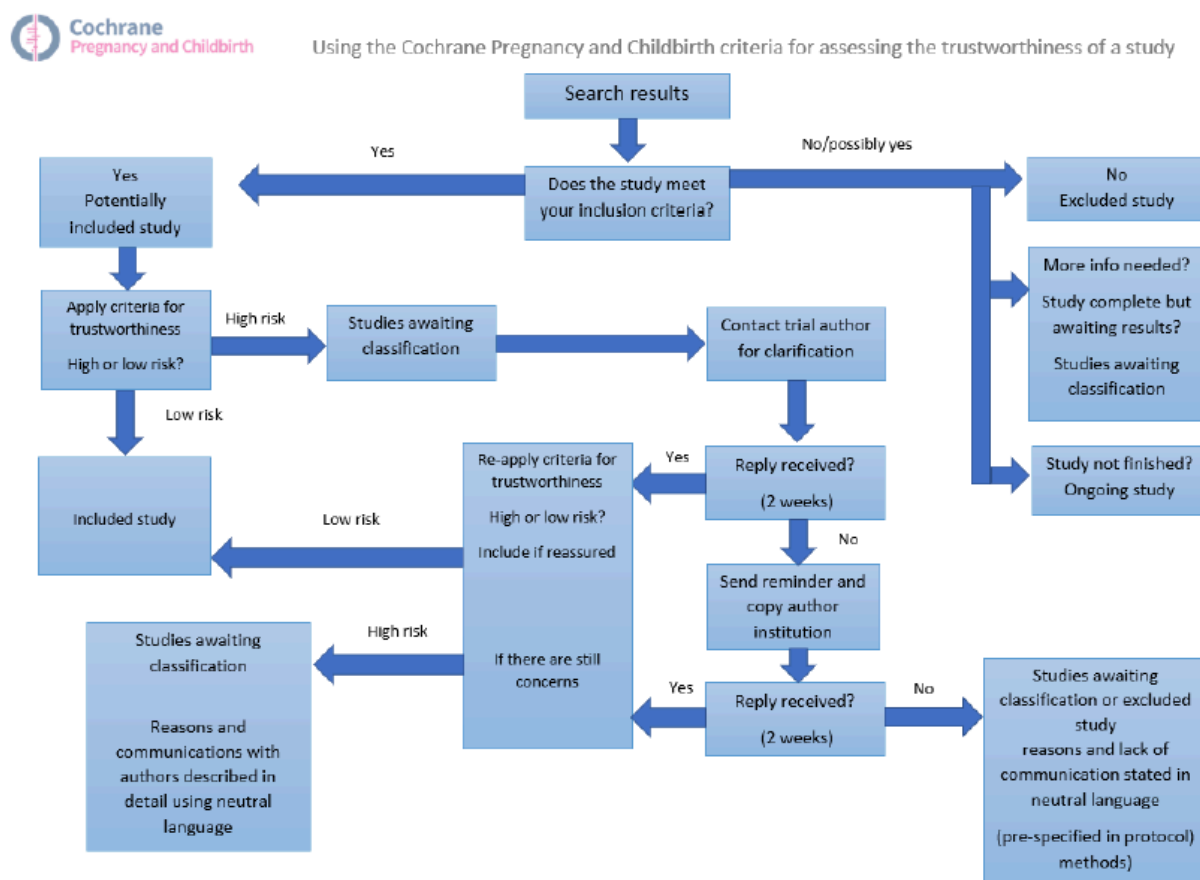
Data extraction and management

We will design a form to extract data. For eligible studies, at least two review authors will independently extract the data using the agreed form. We will resolve discrepancies through discussion, or, if required, through consultation with a third person. We will enter data into Review Manager 5 (Review Manager 2020) and check them for accuracy. When information regarding any of the above is unclear, we will attempt to contact authors of the original reports to provide further details.

Assessment of risk of bias in included studies

Two review authors will independently assess risk of bias for each study using the criteria outlined in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). We will resolve any disagreement by discussion or by involving a third assessor.

(1) Random sequence generation (checking for possible selection bias)

We will describe for each included study the method used to generate the allocation sequence, in sufficient detail to allow an assessment of whether it should produce comparable groups. We will assess the method as being at:

1. low risk of bias (any truly random process, e.g. random number table; computer random number generator);

2. high risk of bias (any non-random process, e.g. odd or even date of birth; hospital or clinic record number); or

3. unclear risk of bias.

(2) Allocation concealment (checking for possible selection bias)

We will describe for each included study the method used to conceal allocation to interventions prior to assignment and will assess whether intervention allocation could have been foreseen in advance of or during recruitment, or changed after assignment. We will assess the methods as being at:

1. low risk of bias (e.g. telephone or central randomisation; consecutively numbered, sealed, opaque envelopes);

2. high risk of bias (open random allocation; unsealed or non-opaque envelopes; alternation; date of birth); or

3. unclear risk of bias.

(3.1) Blinding of participants and personnel (checking for possible performance bias)

We will describe for each included study the methods used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. We will consider that studies are at low risk of bias if they were blinded, or if we judge that the lack of blinding would be unlikely to affect results. We will assess blinding separately for different outcomes or classes of outcomes. We will assess the methods as being at:

1. low, high, or unclear risk of bias for participants; and
2. low, high, or unclear risk of bias for personnel.

(3.2) Blinding of outcome assessment (checking for possible detection bias)

We will describe for each included study the methods used, if any, to blind outcome assessors from knowledge of which intervention a participant received. We will assess blinding separately for different outcomes or classes of outcomes. We will assess the methods used to blind outcome assessment as being at:

1. low, high, or unclear risk of bias.

(4) Incomplete outcome data (checking for possible attrition bias due to the amount, nature, and handling of incomplete outcome data)

We will describe for each included study, and for each primary outcome, the completeness of data, including attrition and exclusions from the analysis. We will state whether attrition and exclusions were reported and the numbers included in the analysis at each stage (compared with the total number of randomised participants), reasons for attrition or exclusion where reported, and whether missing data were balanced across groups or were related to outcomes. Where sufficient information is reported, or can be supplied by the trial authors, we will re-include missing data in the analyses which we undertake. We will assess methods as being at:

1. low risk of bias (e.g. no missing outcome data; missing outcome data are balanced across groups);
2. high risk of bias (e.g. numbers or reasons for missing data are imbalanced across groups; 'as treated' analysis done with substantial departure of intervention received from that assigned at randomisation; number of drop-outs exceeding 10%); or
3. unclear risk of bias.

(5) Selective reporting (checking for reporting bias)

We will describe for each included study how we investigated the possibility of selective outcome reporting bias and what we found. We will assess the methods as being at:

1. low risk of bias (where it is clear that all of the study's pre-specified outcomes and all expected outcomes of interest to the review have been reported);
2. high risk of bias (where not all the study's pre-specified outcomes have been reported; one or more reported primary outcomes were not pre-specified; outcomes of interest are reported incompletely and so cannot be used; or the study fails to include results of a key outcome that would have been expected to have been reported); or
3. unclear risk of bias.

(6) Other bias

We will describe for each included study any important concerns we have about other possible sources of bias not covered by (1) to (5) above. We will assess whether each study is at:

1. low risk of other bias;
2. high risk of other bias; or
3. unclear risk of other bias.

(7) Overall risk of bias

We will make explicit judgements about whether studies are at high risk of bias, according to the criteria given in the *Cochrane Handbook* (Higgins 2011). With reference to (1) to (6) above, we will assess the likely magnitude and direction of the bias and whether we consider it is likely to impact on the findings. We will explore the impact of the level of bias through undertaking sensitivity analyses (see: Sensitivity analysis).

Measures of treatment effect

Dichotomous data

For dichotomous data, we will present results as summary risk ratio (RR) with 95% confidence intervals.

Continuous data

For continuous data, we will use the mean difference (MD) if outcomes are measured in the same way between trials.

Unit of analysis issues

Cluster-randomised trials

We will include cluster-randomised trials in the analyses along with individually randomised trials. We will adjust their sample sizes using the methods described in the *Cochrane Handbook* (Higgins 2011), using an estimate of the intracluster correlation coefficient (ICC) derived from the trial (if possible), from a similar trial, or from a study of a similar population. If we use ICCs from other sources, we will report this and conduct sensitivity analyses to investigate the effect of variation in the ICC. If we identify both cluster-randomised trials and individually randomised trials, we plan to synthesise the relevant information. In cluster-randomised trials, particular biases to consider include:

1. recruitment bias;
2. baseline imbalance;
3. loss of clusters;
4. incorrect analysis; and
5. comparability with individually randomised trials.

We will consider it reasonable to combine the results from both cluster-randomised trials and individually randomised trials if there is little heterogeneity between the study designs and the interaction between the effect of intervention and the choice of randomisation unit is considered to be unlikely. We will also acknowledge heterogeneity in the randomisation unit and perform a sensitivity analysis to investigate the effects of the randomisation unit.

Dealing with missing data

For included studies, we will note levels of attrition. We will explore the impact of including studies with high levels of missing data in the overall assessment of treatment effect by using sensitivity analysis. For all outcomes, we will carry out analyses on an intention-to-treat basis as far as possible, i.e. we will attempt to include all participants randomised to each group in the analyses, and all participants will be analysed in the group to which they were allocated, regardless of whether or not they received the allocated intervention. The denominator for each outcome in each trial will be the number randomised minus any participants whose outcomes are known to be missing.

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Assessment of heterogeneity

Assumptions when estimating heterogeneity

In standard pairwise meta-analyses we assume there to be different heterogeneity for each pairwise comparison. In network meta-analyses we assume a common estimate for heterogeneity across the different comparisons.

Measures and tests for heterogeneity

To evaluate the presence of clinical heterogeneity, we will describe the study population characteristics across all included trials. We will assess the presence of clinical heterogeneity by comparing these characteristics. In pairwise meta-analyses, we will estimate the heterogeneity for each comparison. In network meta-analysis we will assume a common estimate for the heterogeneity variance across all of the different comparisons. We will assess statistically the presence of heterogeneity within each pairwise comparison using the I^2 statistic and its 95% confidence interval, which measures the percentage of variability that cannot be attributed to random error. We will base the assessment of statistical heterogeneity in the entire network on the magnitude of the heterogeneity variance parameter estimated from the network meta-analysis models. For dichotomous outcomes we will compare the magnitude of the heterogeneity variance with the empirical distribution, as derived by Turner 2012.

Assessment of transitivity across treatment comparisons

We will assess the assumption of transitivity by comparing the distribution of potential effect modifiers across the different pairwise comparisons. We consider that the assumption of transitivity will be likely to hold given that: the common drug used to compare different tocolytic drugs indirectly is likely to be similar in different trials (e.g. calcium channel blockers will be administered similarly in studies of calcium channel blockers versus oxytocin receptor antagonists, and studies of calcium channel blockers versus betamimetics); and pairwise comparisons are unlikely to differ in respect of the distribution of effect modifiers (e.g. all trial designs and characteristics are similar). The assumption of transitivity will be evaluated by comparing the clinical and methodological characteristics of sets of studies grouped by drug comparisons.

Assessment of statistical inconsistency

The statistical agreement between the various sources of evidence in a network of interventions (consistency) will be evaluated by global and local approaches to complement the evaluation of transitivity.

Local approaches for evaluating inconsistency

To evaluate the presence of inconsistency locally we will use the loop-specific approach. This method evaluates the consistency assumption in each closed loop of the network separately as the difference between direct and indirect estimates for a specific comparison in the loop (inconsistency factor). Then, the magnitude of the inconsistency factors and their 95% confidence intervals can be used to infer about the presence of inconsistency in each loop. We will assume a common heterogeneity estimate within each loop.

Global approaches for evaluating inconsistency

To check the assumption of consistency in the entire network we will use the "design-by-treatment" model, as described by Higgins and colleagues (Higgins 2012). This method accounts for different sources of inconsistency that can occur when studies with different designs (two-arm trials versus three-arm trials) give different results as well as disagreement between direct and indirect evidence. Using this approach, we will infer about the presence of inconsistency from any source in the entire network based on a Chi² test. We will perform the design-by-treatment model in Stata using the "mvmeta" command (StataCorp 2019). Inconsistency and heterogeneity are interweaved; to distinguish between these two sources of variability we will employ the I^2 statistic for inconsistency, which measures the percentage of variability that cannot be attributed to random error or heterogeneity (within-comparison variability).

Assessment of reporting biases

In view of the difficulty of detecting and correcting for publication bias and other reporting biases, we aim to minimise the potential impact of these biases by ensuring a comprehensive search for eligible studies and by being alert to duplication of data. If there are 10 or more studies in the pair-wise meta-analysis, we will investigate reporting biases (such as publication bias) using funnel plots. We will visually assess the funnel plots for asymmetry. If asymmetry is suggested, we will perform exploratory analyses to investigate it.

Data synthesis

The comparisons of direct interest are shown in Figure 1. We plan to compare all tocolytic drugs with placebo or no treatment to determine their effectiveness, and all tocolytic drugs against each other to determine which tocolytic drug is the most effective.

Methods for direct treatment comparisons

We will perform standard pairwise meta-analyses using a random-effects model for every drug comparison with at least two trials, using Review Manager 5 software (Review Manager 2020). The random-effects method (DerSimonian 1986) is preferred as it incorporates an assumption that the different studies are estimating different, yet related, intervention effects. The standard errors of the study-specific estimates are adjusted to incorporate a measure of the extent of heterogeneity. This results in wider confidence intervals in the presence of heterogeneity, and corresponding claims of statistical significance are more conservative.

Methods for indirect and network comparisons

We will extract the sample size and number of outcome events per trial arm, to be used in the Stata network suite of commands (White 2015). Once extracted, we will set up the data using the augmented format, where all drugs are compared with a reference treatment (placebo or no treatment), and studies without the reference treatment have a reference treatment arm created with a small amount of data. The augmentation process using arm-based values will calculate the risk estimates of the comparisons with reference treatment and their variances and covariances (White 2015). We plan to generate and assess the network diagrams to determine if a network meta-analysis is feasible. Then we will perform the network meta-analysis within a frequentist framework using

multivariate random-effects meta-analysis estimated by restricted maximum likelihood. We will use the "mvmeta" command within the network suite of commands for network meta-analysis (White 2015), and other Stata commands for visualising and reporting results in network meta-analysis (Chaimani 2015).

Relative treatment ranking

For each intervention we will estimate cumulative probabilities for each tocolytic drug being at each possible rank and obtain a hierarchy using the surface under the cumulative ranking curve (SUCRA); the larger the SUCRA, the higher its rank among all available (Salanti 2011). Uncertainty intervals (95% confidence intervals) around the ranking of each drug will be reported and considered when interpreting the results. Each outcome will be evaluated to determine confidence in the output of the network meta-analysis, as described by Salanti and colleagues (Salanti 2011). The probabilities to rank the drugs are estimated under a Bayesian model with flat priors, assuming that the posterior distribution of the parameter estimates is approximated by a normal distribution with mean and variance equal to the frequentist estimates and variance–covariance matrix. Rankings are constructed drawing 1000 samples from their approximate posterior density. For each draw, the linear predictor is evaluated for each study, and the largest linear predictor is noted. All analyses will be done using Stata statistical software, release 16 (StataCorp 2019). We will use the network suite of Stata commands designed for this purpose.

Multi-arm trials

Multi-arm trials will be included and we will account for the correlation between the effect sizes in the network meta-analysis. We will treat multi-arm studies as multiple independent comparisons in pairwise meta-analyses.

Subgroup analysis and investigation of heterogeneity

If we identify substantial heterogeneity, we will investigate it using subgroup analyses. We will consider whether an overall summary is meaningful, and if it is, use random-effects analysis to produce it. Regardless of heterogeneity or inconsistency, in respect of the primary outcomes we will perform subgroup analyses by evaluating the relative effects and assessment of model fit for the following.

Population

1. Gestational age at study entry (less than 32/40 completed weeks versus 32/40 completed weeks or more)
2. Amniotic membranes (women with ruptured membranes versus women with intact membranes)
3. Multiple pregnancy (singleton versus multiple pregnancy)

Intervention

1. Duration of tocolysis (suppression alone versus suppression plus long-term maintenance)

We will assess subgroup differences by firstly comparing the network diagram for each subgroup. Next, we will perform a pairwise and network meta-analysis for each subgroup and compare their relative treatment effects and their relative treatment ranking. We will examine the subgroups for qualitative interactions where the direction of effect could be reversed, which

is if an intervention was beneficial in one subgroup but harmful in another.

Sensitivity analysis

For the primary outcomes, we will perform sensitivity analyses for the following factors. Differences will be assessed by evaluating the relative effects and assessment of model fit.

1. Risk of bias (restricted to studies with low risk of bias only): studies will be ranked as 'low risk of bias' if they are double-blinded and have allocation concealment with little loss to follow-up (less than 10%). We will consider protocol publication in advance of the results to be an unsuitable criterion for sensitivity analyses, because protocol publication only became widespread in recent years.
2. Co-intervention (we will remove trials where participants received co-interventions such as progesterone)
3. Choice of relative effect measure (risk ratio versus odds ratio)
4. Use of fixed-effect versus random-effects model
5. Randomisation unit (cluster versus individual)

Summary of findings and assessment of the certainty of the evidence

We will produce 'Summary of findings' tables to present the evidence comparing all tocolytic drugs with the reference comparator, placebo, or no tocolysis. Each table will describe key features of the evidence relating to a single outcome, with one table for each of our seven main outcomes, in accordance with the GRADE approach. These outcomes include delay in birth by 48 hours, delay in birth by seven days, neonatal death before 28 days, pregnancy prolongation, serious adverse effects of the drugs, maternal infection, and cessation of treatment due to side effects.

We will use the GRADE working group's approach (Brignardello-Petersen 2018; Puhan 2014) for rating the certainty of the network meta-analysis effect estimates for all the comparisons and all outcomes. We will appraise the certainty of the direct, indirect, and network evidence sequentially (in the following order). First, we will assess the certainty of the direct evidence (where available) for a given outcome, and rate the evidence using the standard GRADE approach based on consideration of: study design limitations (risk of bias); inconsistency; imprecision; indirectness; and publication bias (Higgins 2011). On the network diagram for all the comparisons and all outcomes we will display the certainty of the direct evidence. We will then rate the certainty of the indirect evidence for the same outcome; this will be determined based on the lower of the certainty ratings of the two arms forming the dominant 'first-order' loop in the network diagram for this outcome. Our final step will be to determine the quality of network evidence based on: 1) the higher certainty rating of the direct and indirect evidence; 2) whether the relevant network diagram exhibits 'transitivity' (i.e. whether all the comparisons contributing data to the estimate are directly consistent with the PICO question); 3) consideration of coherence between direct and indirect effect estimates; and 4) precision of the network effect estimate. At each of these stages, two review authors will independently appraise the certainty ratings for the direct, indirect, and network evidence. Disagreements between authors will be resolved through discussion and consultation with a third review author where necessary. The quality of network evidence for each

outcome will be rated as 'high', 'moderate', 'low' or 'very low', in accordance with the GRADE approach and explained below.

1. High certainty: we are very confident that the true effect lies close to that of the effect.
2. Moderate certainty: we are moderately confident in the effect estimate. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
3. Low certainty: our confidence in the effect estimate is limited. The true effect may be substantially different from the estimate of the effect.
4. Very low certainty: we have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect.

For ease of comparison when interpreting the relative findings of all tocolytic drugs, the 'Summary of findings' tables will include all effect estimates and certainty judgements for the direct evidence, indirect evidence, and the network meta-analysis; all the findings for a single outcome will be described in each table. The anticipated absolute effects will also be included, based on the network effect estimate for each tocolytic drug in comparison with placebo or no treatment.

ACKNOWLEDGEMENTS

REFERENCES

Additional references

Babay 1998

Babay Z, Lange IR, Amin H. Neonatal and maternal complications after administration of indomethacin for preterm labour between 24 and 30 weeks gestation. *Journal of the Society of Obstetricians and Gynaecologists of Canada* 1998;**20**:505-11.

Bain 2013

Bain E, Heatley E, Hsu K, Crowther CA. Relaxin for preventing preterm birth. *Cochrane Database of Systematic Reviews* 2013, Issue 8. Art. No: CD010073. [DOI: 10.1002/14651858.CD010073.pub2]

Blondel 2006

Blondel B, Macfarlane A, Gissler M, Breart G, Zeitlin J, PERISTAT Study Group. Preterm birth and multiple pregnancy in European countries participating in the PERISTAT project. *BJOG: an International Journal of Obstetrics & Gynaecology* 2006;**113**:528-35.

Brignardello-Petersen 2018

Brignardello-Petersen R, Bonner A, Alexander PE, Siemieniuk RA, Furukawa TA, Rochwerg B, et al. Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. *Journal of Clinical Epidemiology* 2018;**93**:36-44.

Caldwell 2005

Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;**331**(7521):897-900.

Caldwell 2010

Caldwell DM, Welton NJ, Ades AE. Mixed treatment comparison analysis provides internally coherent treatment effect estimates based on overviews of reviews and can reveal inconsistency. *Journal of Clinical Epidemiology* 2010;**63**(8):875-82.

Carlisle 2017

Carlisle JB. Data fabrication and other reasons for non-random sampling in 5087 randomised, controlled trials in anaesthetic and general medical journals. *Anaesthesia* 2017;**72**:944–52.

Chaimani 2015

Chaimani A, Salanti G. Visualizing assumptions and results in network meta-analysis: the network graphs package. *Stata Journal* 2015;**15**:4.

Crowther 2014

Crowther CA, Brown J, McKinlay CJD, Middleton P. Magnesium sulphate for preventing preterm birth in threatened preterm labour. *Cochrane Database of Systematic Reviews* 2014, Issue 8. Art. No: CD001060. [DOI: 10.1002/14651858.CD001060.pub2]

DerSimonian 1986

DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986;**7**:77–88.

Duckitt 2014

Duckitt K, Thornton S, O'Donovan OP, Dowswell T. Nitric oxide donors for treating preterm labour. *Cochrane Database of Systematic Reviews* 2014, Issue 5. Art. No: CD002860. [DOI: 10.1002/14651858.CD002860.pub2]

Escobar 2006

Escobar GJ, McCormick MC, Zupancic JA, Coleman-Phox K, Armstrong MA, Greene JD, et al. Unstudied infants: outcomes of moderately premature infants in the neonatal intensive care unit. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2006;**91**:F238–44.

Flenady 2014a

Flenady V, Wojcieszek AM, Papatsonis DNM, Stock OM, Murray L, Jardine LA, et al. Calcium channel blockers for inhibiting preterm labour and birth. *Cochrane Database of Systematic Reviews* 2014, Issue 6. Art. No: CD002255. [DOI: 10.1002/14651858.CD002255.pub2]

Flenady 2014b

Flenady V, Reinebrant HE, Liley HG, Tambimuttu EG, Papatsonis DNM. Oxytocin receptor antagonists for inhibiting preterm labour. *Cochrane Database of Systematic Reviews* 2014, Issue 6. Art. No: CD004452. [DOI: 10.1002/14651858.CD004452.pub3]

Haas 2009

Haas DM, Imperiale TF, Kirkpatrick PR, Klein RW, Zollinger TW, Golichowski AM. Tocolytic therapy: a meta-analysis and decision analysis. *Obstet Gynecol* 2009;**113**:585-594.

Higgins 2011

Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from training.cochrane.org/handbook.

Higgins 2012

Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network metaanalysis: concepts and models for multi-arm studies. *Research Synthesis Methods* 2012;**3**:98-110.

Kalra 2008

Kalra SK, Molinaro TA. The association of in vitro fertilization and perinatal morbidity. *Seminars in Reproductive Medicine* 2008;**26**:423-35.

Kinney 2006

Kinney HC. The near-term (late preterm) human brain and risk for periventricular leukomalacia: a review. *Seminars in Perinatology* 2006;**30**:81–8.

Koren 2006

Koren G, Florescu A, Costei AM, Boskovic R, Moretti ME. Nonsteroidal anti-inflammatory drugs during third trimester and the risk of premature closure of the ductus arteriosus: a meta-analysis. *Annals of Pharmacotherapy* 2006;**40**:824-9.

Lee 2008

Lee SE, Romero R, Park CW, Jun JK, Yoon BH. The frequency and significance of intraamniotic inflammation in patients with cervical insufficiency. *American Journal of Obstetrics and Gynecology* 2008;**198**:e1-8.

Liu 2016

Liu L, Oza S, Hogan D, Chu Y, Perin J, Zhu J, et al. Global, regional, and national causes of under-5 mortality in 2000-15: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet* 2016;**388**:3027-35.

Medicines.org.uk 2020

Medicinesorguk. Summary of Product Characteristics (SPC). www.medicines.org.uk/emc/ (accessed 02/12/2020).

Menon 2008

Menon R. Spontaneous preterm birth, a clinical dilemma: etiologic, pathophysiologic and genetic heterogeneities and racial disparity. *Acta Obstetricia et Gynecologica Scandinavica* 2008;**87**:590-600.

Mukhopadhaya 2007

Mukhopadhaya N, Arulkumaran S. Reproductive outcomes after in-vitro fertilization. *Current Opinion in Obstetrics and Gynecology* 2007;**19**:113-9.

Neilson 2014

Neilson JP, West HM, Dowswell T. Betamimetics for inhibiting preterm labour. *Cochrane Database of Systematic Reviews* 2014, Issue 2. Art. No: CD004352. [DOI: 10.1002/14651858.CD004352.pub3]

NICE 2015

National Institute for Health and Clinical Excellence. Preterm labour and birth. NICE guideline 2015.

Puhan 2014

Puhan MA, Schünemann HJ, Murad MH, Li T, Brignardello-Petersen R, Singh JA, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ* 2014;**349**:5630.

Reinebrant 2015

Reinebrant HE, Pileggi-Castro C, Romero CLT, dos Santos RAN, Kumar S, Souza JP, et al. Cyclo-oxygenase (COX) inhibitors for treating preterm labour. *Cochrane Database of Systematic Reviews* 2015, Issue 6. Art. No: CD001992. [DOI: 10.1002/14651858.CD001992.pub3]

Review Manager 2020 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration Review Manager (RevMan). Version 5.4. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2020.

Salanti 2011

Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *Journal of Clinical Epidemiology* 2011;**64**(2):163-71.

StataCorp 2019 [Computer program]

CollegeStation, TX Stata Statistical Software: Release 16. StataCorp. CollegeStation, TX, 2019.

Su 2014

Su LL, Samuel M, Chong YS. Progestational agents for treating threatened or established preterm labour. *Cochrane Database of Systematic Reviews* 2014, Issue 1. Art. No: CD006770. [DOI: 10.1002/14651858.CD006770.pub3]

Turner 2012

Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JPT. Predicting the extent of heterogeneity in meta-analysis,using empirical data from the Cochrane Database of Systematic Reviews. *International Journal of Epidemiology* 2012;**41**:818–27.

UNIGME 2020

United Nations Inter-agency Group for Child Mortality Estimation (UNIGME). Levels & Trends in Child Mortality: Report 2020, Estimates developed by the United Nations Inter-agency Group for Child Mortality Estimation. United Nations Children's Fund 2020.

Wang 2004

Wang ML, Dorer DJ, Fleming MP, Catlin EA. Clinical outcomes of near-term infants. *Pediatrics* 2004;**114**:372-6.

White 2015

White IR. Network meta-analysis. *Stata Journal* 2015;**15**:4.

WHO 1977

World Health Organization. Recommended definitions, terminology and format for statistical tables related to the perinatal period and use of a new certificate for cause of perinatal deaths. Modifications recommended by FIGO as amended October 14, 1976. *Acta Obstetricia et Gynecologica Scandinavica* 1977;**56**:247-53.

WHO 2015

World Health Organization. WHO recommendations on interventions to improve preterm birth outcomes. WHO recommendations 2015.

WHO 2018

World Health Organization. Preterm birth: key facts. www.who.int/news-room/fact-sheets/detail/preterm-birth (accessed 02/12/202).

APPENDICES

Appendix 1. Screening eligible studies for scientific integrity/trustworthiness

All studies meeting the inclusion criteria will undergo further independent evaluation by two review authors against the criteria below.

Criteria questions	Assessment		Comments and concerns
	High risk	**Low risk**	
Research governance			
Was the study prospectively registered (for those studies published after 2010)?			
When requested, did the trial authors refuse to provide/share the protocol and/or ethics approval letter?			
Did the trial authors refuse to engage in communication with the Cochrane Review authors within the agreed timelines?			
Did the trial authors refuse to provide individual participant data upon request, with no justifiable reason?			
Baseline characteristics			
Is there anything about the characteristics of the study participants that appear too similar? (E.g. distribution of the mean (standard deviation (SD)) excessively narrow or excessively wide, as noted by Carlisle 2017)			
Feasibility			
Is there anything about the study characteristics that, in your opinion, could be implausible? (E.g. large numbers of women with a rare condition (such as severe cholestasis in pregnancy) recruited within 12 months).			
Results			
Is there anything about the reported results of the study that could be implausible? (E.g. massive risk reduction for the main study outcomes with a small sample size?)			
Do you have any concerns about the methods of randomisation such as unexpectedly even numbers of women 'randomised' including a mismatch between the numbers and the methods? (E.g. if the authors say 'no blocking was used' but still end up with equal numbers, or if the authors say they used 'blocks of 4' but the final numbers differ by 6.)			
Are there (close to) zero losses to follow up without plausible explanation?			
For abstracts only:			
Have the study authors confirmed in writing that the data to be included in the review have come from the final analysis and will not change?			

Assessment after applying trustworthiness criteria high risk (awaiting classification) OR low risk (include)

Decision after attempting to contact authors high risk (awaiting classification) OR low risk (include)

H I S T O R Y

Protocol first published: Issue 4, 2021

C O N T R I B U T I O N S   O F   A U T H O R S

Ioannis D Gallos (IDG) and Olufemi T Oladapo (OTO) conceived the idea for this protocol.
IDG, Amie Wilson (AW), Victoria A Hodgetts-Morton (VAH), Ella Marson (EM), Alexandra Markland (AM), Eva Larkai (EL), and Rachel K Morris (RKM) designed the review.
Malcolm J Price (MJP), Aurelio Tobias (AT), and Argyro Papadopoulou (AP) provided statistical advice and input.
Doris Chou (DC), Arri Coomarasamy (AC), and OTO reviewed the protocol and provided critical feedback.
IDG is the guarantor for this review.

D E C L A R A T I O N S   O F   I N T E R E S T

Ioannis D Gallos: The World Health Organization provided payment to Ioannis Gallos for working on this review.

Amie Wilson has no declarations of interest.

Victoria A Hodgetts-Morton works as a health professional, Clinical O&G ST5 in training, Birmingham Women's Hospital.

Ella Marson has no declarations of interest.

Alexandra Markland has no declarations of interest.

Eva Larkai has no declarations of interest.

Argyro Papadopoulou: I am currently a PhD student at the University of Birmingham, UK. My tuition fees are paid by Tommy's charity, Tommy's National Centre for Miscarriage Research. Tuition fees are directly paid to the University of Birmingham.

Arri Coomarasamy has no declarations of interest.

Aurelio Tobias has no declarations of interest.

Doris Chou: in terms of guideline and recommendation synthesis, I manage the maternal/perinatal living guideline process within the World Health Organization. The technical group may consider this review in deliberations related to the use of tocolytics. During these meetings, I do not carry any voting capacity.

Olufemi T Oladapo has no declarations of interest.

Malcolm J Price has no declarations of interest.

Katie Morris has received expenses for attendance at meetings as a Board/Committee member on the Royal College of Obstetricians and Gynaecologists (RCOG) Scientific Advisory Committee, Tommy's Scientific Advisory Board, Wellbeing of Women Scientific Advisory Committee, RCOG Academic Committee, RCOG Research Committee, and British Maternal & Fetal Medicine Society; published review articles related to preterm birth in PubMed journals as a member of their academic institution; and is a Consultant in Maternal Fetal Medicine at Birmingham Women's and Children's Hospital NHS Foundation Trust.

S O U R C E S   O F   S U P P O R T

Internal sources

- World Health Organization, Switzerland
- Birmingham Women's Hospital, UK
- Tommy's National Centre for Miscarriage Research, Institute of Metabolism and Systems Research (IMSR), WHO Collaborating Centre for Global Women's Health Research, University of Birmingham, Birmingham, UK

External sources

- No sources of support supplied