RESEARCH

Tocolytic therapy for preterm delivery: systematic review and network meta-analysis

OPEN ACCESS

David M Haas *associate professor of obstetrics and gynecology*[1], Deborah M Caldwell *MRC fellow population health science*[2], Page Kirkpatrick *research associate*[1], Jennifer J McIntosh *medical resident*[1], Nicky J Welton *MRC research fellow*[2]

[1]Department of Obstetrics and Gynecology, Indiana University School of Medicine, Indianapolis, IN, USA; [2]School of Social and Community Medicine, University of Bristol, Bristol, UK

Abstract

Objective To determine the most effective tocolytic agent at delaying delivery.

Design Systematic review and network meta-analysis.

Data sources Cochrane Central Register of Controlled Trials, Medline, Medline In-Process, Embase, and CINAHL up to 17 February 2012.

Study selection Randomised controlled trials of tocolytic therapy in women at risk of preterm delivery.

Data extraction At least two reviewers extracted data on study design, characteristics, number of participants, and outcomes reported (neonatal and maternal). A network meta-analysis was done using a random effects model with drug class effect. Two sensitivity analyses were carried out for the primary outcome; restricted to studies at low risk of bias and restricted to studies excluding women at high risk of preterm delivery (those with multiple gestation and ruptured membranes).

Results Of the 3263 titles initially identified, 95 randomized controlled trials of tocolytic therapy were reviewed. Compared with placebo, the probability of delivery being delayed by 48 hours was highest with prostaglandin inhibitors (odds ratio 5.39, 95% credible interval 2.14 to 12.34) followed by magnesium sulfate (2.76, 1.58 to 4.94), calcium channel blockers (2.71, 1.17 to 5.91), beta mimetics (2.41, 1.27 to 4.55), and the oxytocin receptor blocker atosiban (2.02, 1.10 to 3.80). No class of tocolytic was significantly superior to placebo in reducing neonatal respiratory distress syndrome. Compared with placebo, side effects requiring a change of medication were significantly higher for beta mimetics (22.68, 7.51 to 73.67), magnesium sulfate (8.15, 2.47 to 27.70), and calcium channel blockers (3.80, 1.02 to 16.92). Prostaglandin inhibitors and calcium channel blockers were the tocolytics with the best probability of being ranked in the top three medication classes for the outcomes of 48 hour delay in delivery, respiratory distress syndrome, neonatal mortality, and maternal side effects (all cause).

Conclusions Prostaglandin inhibitors and calcium channel blockers had the highest probability of delaying delivery and improving neonatal and maternal outcomes.

Introduction

Tocolytic therapy to delay preterm delivery is an important intervention in obstetrics. Although tocolytics have not been shown to improve neonatal outcomes, they can delay preterm delivery long enough for antenatal corticosteroids to be administered or for the mother to be transported to a tertiary care facility.[1] In premature neonates, antenatal corticosteroids reduce morbidity and mortality.[2] Tocolytic therapy may therefore have an important role in improving outcomes from preterm delivery. With over 500 000 preterm births in the United States alone (12.3% of all births in 2008)[3] and 29% of these being less than 34 weeks' gestation, preterm delivery is an important public health issue.

Many different classes of drugs have been used for tocolytic therapy.[4] These include beta mimetics such as ritodrine and terbutaline; magnesium sulfate; prostaglandin inhibitors (for example, indomethacin, ketorolac); calcium channel blockers such as nifedipine; nitrates (for example, nitroglycerine); oxytocin receptor blockers (for example, atosiban), and others. Each tocolytic has a unique mechanism of action, side effects, and degree of complexity to administer.[5] Several Cochrane reviews have compared individual tocolytic drugs with placebo

Correspondence to: D M Haas Wishard Memorial Hospital, 1001 W 10th Street, F5102, Indianapolis, IN 46202, USA dahaas@iupui.edu

or other tocolytics.[6-10] A recent pooled meta-analysis and decision analysis of trials on tocolytics showed that to delay delivery for 48 hours and seven days, prostaglandin inhibitors were the best first line tocolytic.[1] A standard pairwise meta-analysis, however, can only compare two treatments (or classes) that have been directly compared in head to head trials (direct evidence). Consequently, trials comparing two treatments from the same class are often excluded from class level meta-analyses. In the absence of a single high quality, randomized controlled trial comparing all tocolytic therapies, uncertainty remains about which is the most effective at delaying preterm delivery.[11]

For a complex condition such as preterm delivery with multiple competing treatment options, not all of which have been directly compared, a network meta-analysis may be better able to allow for comparisons and conclusions about which tocolytic is most effective. A network meta-analysis refers to networks of trial evidence in which all the available direct and indirect evidence on relative treatment effects are pooled simultaneously in a single coherent analysis.[12 13] Indirect evidence is obtained when the relative effectiveness of treatment B versus treatment C is inferred through a common comparator A (see supplementary file for equation). Thus a network meta-analysis produces estimates of the relative effects of each treatment compared with every other in a network, even though some pairs may not have been directly compared, and has the potential to reduce the uncertainty in treatment effect estimates.[12] It also allows for the calculation of the probability that each treatment, or class, is the best for any given outcome. Network meta-analysis can also be used to identify gaps in the evidence base.[14] In an active area such as tocolysis for preterm delivery, with six trials published since 2009, a network meta-analysis has the potential to inform future research agendas. We systematically reviewed and analysed trials on tocolytics and carried out a network meta-analysis to determine the most effective agent for delaying preterm delivery.

Methods

Using the search terms "preterm labor", "tocolytic", and "obstetric labor, premature" we systematically searched the Cochrane Central Register of Controlled Trials (February 2012), Medline (1950-present), Medline In-Process/Daily Update (17 February 2012), Embase (1988-2012), and CINAHL (1982-2012) for published randomized controlled trials of tocolytic therapy. We limited the search to articles reporting trials in humans, and excluded duplicate trial entries. To ensure completeness, we cross referenced our search results with the Cochrane reviews of tocolytic medications, hand searching for additional titles. We did not register a protocol for the review. Based on the titles we read the abstracts of potentially relevant papers and obtained the full text articles for those that seemed pertinent. Included trials were those that reported a comparison between different medications or between a medication and a placebo or usual care for delaying preterm delivery. Trials were excluded if they were not randomized controlled trials, did not study women at risk of preterm delivery (defined by trial), did not study at least one tocolytic drug, used combination drug therapies for tocolysis, or did not report maternal or neonatal outcomes in relation to preterm delivery. As published abstracts did not contain enough information for complete data to be extracted we did not include them. We also excluded personal communications cited in Cochrane reviews. At least two reviewers read the articles and extracted data from the trials. Discordance between the reviewers was resolved by consensus. Abstracts of articles in non-English languages were reviewed.

If the article was considered relevant, we obtained the full text and had it translated for possible data extraction. Published abstracts from conferences were not included.

Study quality was assigned utilizing the methodology and categories described in the *Cochrane Collaboration Handbook*.[15] The Cochrane collaboration's recommended tool for assessing risk of bias is neither a scale nor a checklist but a domain based evaluation, in which critical assessments are made separately for different domains. Briefly, the tool for assessing risk of bias addresses seven specific domains: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective reporting, and other sources of bias. Each domain is assigned a judgment relating to the risk of bias for that study: low risk, high risk, and unclear (or unknown). For the purpose of a planned sensitivity analysis, we counted the total number of low risk scores (out of 7) for each study. When at least four domains had a low risk score, with at least one of the domains needing to be sequence generation or allocation concealment, we considered the overall study quality to be high.

At least two reviewers extracted data on study design, characteristics, number of participants, and outcomes reported. Relevant studies were those that reported on pregnant women being treated for preterm delivery or at risk of preterm delivery. Information extracted on the mothers included age, estimated gestational age at entry into the study and at delivery, and number with previous preterm births. Primary data extracted on maternal outcomes included the numbers of participants with delivery delayed by 48 hours and the number of side effects (all causes) from the tocolytics. Secondary maternal outcomes extracted included the number of participants with delivery delayed by seven days and until 37 weeks' gestation, and the mean number of days by which delivery was delayed. These secondary outcomes are not considered in this paper owing to concerns about multiple statistical testing. We assigned a quality score to each study, the results of which are reported separately.[16] We also extracted data on the use of antenatal corticosteroids, inclusion or exclusion of multiple gestations or ruptured membranes, and whether or not the tocolytic therapy was short term (a predefined length of time such as 48 hours or until contractions stopped) or long term (usually until 36 or 37 weeks estimated gestational age). Primary neonatal outcomes extracted were rates of respiratory distress syndrome and death. Secondary neonatal outcomes extracted were birth weight, chronic lung disease or bronchopulmonary dysplasia, fetal sepsis, intraventricular hemorrhage, necrotizing enterocolitis, hyperbilirubinemia, and premature closure of the ductus arteriosus. These outcomes were defined in the individual trials. Because of the way outcomes were reported in the trials we were unable to report an overall composite score for neonatal morbidity or mortality.

We classified the drugs as placebo (placebo or usual or standard care without a tocolytic drug); beta mimetics (ritodrine, terbutaline, nylidrin, salbutamol, fenoterol, hexoprenaline, isoxsuprine); calcium channel blockers (nifedipine, nicardipine); magnesium sulfate; nitrates (nitroglycerine, nitric oxide); oxytocin receptor blockers (atosiban, barusiban); others (alcohol, human chorionic gonadotropin, combination tocolytic drugs); and prostaglandin inhibitors (indomethacin, celecoxib, sulindac, ketorolac, rofecoxib). Two authors (DMH, PK, or JJMcI) then examined the complete set of trials to assess whether the characteristics of the trials and participants were similar enough to be combined in the network meta-analysis—that is, that the sets of trials did not differ for distribution of potential effect modifiers. This assumption of "consistency"[17] underpins the

validity of a network meta-analysis and is akin to an assumption that the direct and indirect evidence estimate the same underlying treatment effect variable.

Statistical analysis

The primary effectiveness outcome for the network meta-analysis was delivery successfully delayed for 48 hours. We chose this outcome as it was most commonly reported and is a surrogate for the ability to administer a complete course of antenatal corticosteroids or to allow for maternal transport to a tertiary care facility. The secondary outcomes were neonatal mortality, neonatal respiratory distress syndrome, and all cause maternal side effects. All are binary outcomes. From the analysis we excluded studies with zero or 100% events on all arms. A list of excluded trials is available from the authors.

Analyses were done within a Bayesian framework using WinBUGS 1.4.3.[18] We carried out a random effects network meta-analysis[19-21] to simultaneously compare the 18 treatments and eight tocolytic classes for each outcome. Where head to head data were available we also carried out pairwise "direct" meta-analyses using a random effects model. Heterogeneity was assessed using the posterior median between trial variance, τ^2. However, for ease of interpretation we report the χ^2 test for heterogeneity and I^2 statistic for the pairwise meta-analyses (these were calculated using Stata). Owing to the lack of power associated with the χ^2 test we used P=0.10 for our assessment of heterogeneity.[15] In the case of two or fewer trials we carried out a fixed effect meta-analysis. The pairwise meta-analyses were done using the drug classes and not individual treatments as the subject of interest. Posterior median odds ratios and 95% credible intervals were calculated. For the primary outcome we carried out a planned sensitivity analysis based on risk of bias assessment. Additional sensitivity analyses examined were restricted to studies that excluded multiple gestations or excluded participants with ruptured membranes. A metaregression analyzed the impact of planned duration of treatment (acute or short term tocolysis versus prolonged therapy) on the results.

For the network meta-analysis we implemented a class effect model where each treatment effect in the same class is assumed to come from a family of treatment effects with a class specific mean effect and between treatment variability within class (assumed equal across all classes).[22 23] Further details on alternative models evaluated are available from the authors together with the WinBUGS code. Goodness of fit was measured by the posterior mean of the residual deviance. In a well fitting model the residual deviance should be close to the number of data points.[24] Owing to the way in which the residual deviance is calculated, zero cells on the baseline (control) arm can cause computational difficulties. For the purposes of model selection we removed these trials but included them in the final model on which the results are based. A key assumption of network meta-analysis is that of consistency between the direct and indirect evidence.[17] We assessed whether there was inconsistency in each of the three networks by comparing a model assuming consistency with that of an inconsistency model[25] using the deviance information criterion. A difference of 3 or more points is considered meaningful.[26] Convergence was assessed using two chains and was achieved by 25 000 simulations for delivery delayed by 48 hours, 30 000 for neonatal mortality and respiratory distress syndrome, and 35 000 for maternal side effects (based on the Brooks-Gelman-Rubin diagnostic tool in WinBUGS). We run a further 50 000 updates after convergence for delivery delayed by 48 hours, 60 <000 for neonatal mortality and respiratory distress syndrome, and 70 000 for maternal side effects. All reported results are based on these further samples.

Results

Of 3263 titles initially identified, 159 full text articles were retrieved, of which 95 satisfied the study inclusion and exclusion criteria. Fig 1⇓ summarizes the steps of the systematic review. Nine articles were translated (four in Chinese, two in French, and one each in German, Portuguese, and Spanish). One of the French articles was later excluded for not being a true randomized controlled trial, leaving eight reports in a language other than English (8%). Details of the characteristics of the trials and comparison of the quality of tocolytic studies retrieved are reported elsewhere.[16] Twenty five trials contained a placebo arm (26%),[27-51] 60 included beta mimetics (63%),[27-29 31 34-37 39 43 46 48 50 52-98] 29 included magnesium sulfate (31%),[29 30 33 38 47 53 56 61 68 80 89 90 93 95 96 99-112] 29 included calcium channel blockers (31%),[43 52 59 63 64 70-72 74 76 78 79 84-86 91 92 94 101-104 106 112-117] 18 included prostaglandin inhibitors (19%),[40-42 49 57 73 75 81 88 99 107 109-111 116 118-120] 13 included oxytocin receptor blockers (atosiban or barusiban) (14%),[32 44 50 51 54 55 65 66 82 83 87 113 114] four included nitrates (4%),[45 58 100 115] and five included other drugs (5%).[47 50 77 88 108] For the outcome of delivery delayed by 48 hours 64 trials assessing 16 treatments from eight drug classes were eligible for inclusion in the network meta-analysis. For the outcome of respiratory distress syndrome 60 trials assessing 19 treatments from seven drug classes were eligible for inclusion, and for the outcome maternal side effects (all cause) 68 trials assessing 18 treatments from seven drug classes were eligible for inclusion. Trials included a mean of 111.9 (SD 108.8, range 20-708) participants and were published from 1966-2011 (see supplementary table 1 for details of the trials, along with their quality assessments). Fig 2⇓ presents the complete network of the 95 randomised controlled trials on tocolytics. No trials compared atosiban with magnesium sulfate.

Statistical findings

For each outcome no meaningful differences in residual deviance or deviance information criterion values were observed between the inconsistency and consistency models. Furthermore, overlap was substantial between the direct estimates (where available) and network meta-analysis estimates (figs 3-6⇓⇓⇓⇓). This provides support for the assumption of consistency required for network meta-analysis—that is, the estimates for direct treatment effect agree with those generated from the network meta-analysis. Direct pairwise meta-analyses involve the trial results that only directly compare the classes, whereas results from the network meta-analysis utilize the network calculations described previously. The class effect model provided an adequate fit to the data, with a posterior mean residual deviance of 114.4 (112 data points) for the outcome of delivery delayed by 48 hours, 82.6 (81 data points) for neonatal mortality, 78.9 (81 data points) for respiratory distress syndrome, and 123.2 (118 data points) for maternal side effects.

Delivery delayed by 48 hours: efficacy

For the outcome delivery delayed by 48 hours 55 studies were included in the network meta-analysis and 54 in the pairwise meta-analysis (one trial was excluded as it compared two treatments from the same class). Direct evidence was available for 14 class versus class pairwise comparisons. Heterogeneity was evident in two of the random effects pairwise meta-analyses, both of which compared active treatment with placebo (beta mimetic *v* placebo, n=4 trials, P=0.08; and magnesium sulfate *v* placebo, n=3 trials, P=0.001); the I^2 values for these two comparisons were 50% or more (95% for magnesium sulfate

versus placebo). Where available, the supplementary file reports the full results from the pairwise meta-analyses; however, the posterior median odds ratios (95% credible intervals) are also shown in fig 3 alongside the posterior median odds ratios for the full suite of 28 class comparisons available from the network meta-analysis.

The table⇓ reports the effectiveness of tocolytic therapies and adverse events from the network meta-analysis using placebo as the reference class to which all other drug classes were compared. All active classes were superior to placebo in successfully delaying delivery by 48 hours, although the category for others and for nitrates did not achieve conventional significance. The results from the network meta-analysis also suggested that prostaglandin inhibitors had a greater beneficial effect than all the other active classes. However, uncertainty in these estimates was considerable (table, fig 3, and supplementary table 3). Calcium channel blockers and magnesium sulfate had a greater effect than oxytocin receptor blockers, nitrates, beta mimetics, others, and placebo. Fig 7⇓ shows the distribution of probabilities of each class being ranked at each of the possible eight positions. The most efficacious treatment class was prostaglandin inhibitors, which had an 83% probability of being the "best" class. This meant there was still a 17% probability of prostaglandin inhibitors not being the best class. The probability of being ranked in the top three most efficacious classes was 96% for prostaglandin inhibitors, 63% for magnesium sulfate, 57% for calcium channel blockers, 33% for beta mimetics, 24% for nitrates, 14% for oxytocin receptor blockers, 13% for others, and 0% for placebo. The probability of being ranked in the bottom three (least efficacious) was 99% for placebo, 79% for others, 52% for nitrates, 41% for oxytocin receptor blockers, 13% for beta mimetics, 10% for calcium channel blockers, 5% for magnesium sulfate, and 1% for prostaglandin inhibitors. See supplementary table 4 for the predicted odds ratios and intervals.

Neonatal mortality

For the outcome neonatal mortality 40 trials were included in the network meta-analysis and 34 in the pairwise meta-analysis. Two trials were excluded from the pairwise analysis as they compared treatments from the same class (beta mimetics). Direct evidence was available for 13 class versus class pairwise comparisons (see supplementary appendix 2 for the full results). Heterogeneity was evident in one of the seven random effects analyses (beta mimetic *v* placebo, n=6 trials, P<0.001, I^2=78%).

Fig 4 shows the posterior median odds ratios (95% credible intervals) from each pairwise meta-analysis alongside the posterior median odds ratios for the full suite of 21 class comparisons available from the network meta-analysis. The results from the direct, head to head meta-analyses were consistent with those of the network meta-analysis.

No clear evidence was found for the relative effectiveness of any tocolytic versus placebo being beneficial for neonatal mortality (table). The point estimates for the active versus comparisons suggested that calcium channel blockers were more effective than all the other classes of tocolytics (fig 4). However, the credible intervals for the comparisons crossed 1 and uncertainty in all estimates was considerable. This uncertainty is reflected in the rankograms shown in fig 4, which suggest that although calcium channel blockers are the best class for reducing neonatal mortality, the probability was only 41% and so indicative of considerable uncertainty. Indeed, there was a 59% probability that they were not the best class for reducing neonatal mortality. Prostaglandin inhibitors had the next highest

probability of being the best (28%). The probability that calcium channel blockers were ranked in the top three classes for reducing neonatal mortality was 85%, followed by beta mimetics (58%), oxytocin receptor blockers (56%), and prostaglandin inhibitors (54%). The worst performing classes were placebo (1%) and others (2%), followed by magnesium sulfate, with a 3% probability of being the best class for reducing neonatal mortality. These results reflect uncertainty around which class is associated with the fewest neonatal deaths.

Neonatal respiratory distress syndrome

For the outcome neonatal respiratory distress syndrome 42 trials were included in the network meta-analysis and 37 in the pairwise meta-analyses. Five trials were excluded from the pairwise meta-analyses as they compared treatments from within the same class. Direct evidence was available for 13 class versus class pairwise comparisons (see supplementary appendix 2 for the full results). Heterogeneity between trials was not evident in any comparison.

Fig 5 shows the posterior median odds ratios (95% credible intervals) from each pairwise meta-analysis alongside the posterior median odds ratios for the full suite of 21 class comparisons available from the network meta-analysis for respiratory distress syndrome.

The results from the network meta-analysis (table and fig 5) suggested no evidence of a difference between the classes in reducing respiratory distress syndrome and no clear effect compared with placebo. The rankograms (fig 7) suggested that calcium channel blockers were the best class for reducing respiratory distress syndrome; however, this probability was only 47% and so indicative of considerable uncertainty. The probability that calcium channel blockers were ranked in the top three classes for reducing respiratory distress syndrome was 80%. The worst performing class was the others, with only an 11% probability of being in the top three classes for reducing respiratory distress syndrome.

All cause maternal side effects

For the outcome all cause maternal side effects 58 trials were included in the network meta-analysis and 55 in the pairwise meta-analyses. Three trials were excluded from the pairwise meta-analyses as they compared treatments from the same class. Direct evidence was available for 16 class comparisons versus class pairwise comparisons.

Fig 6 shows the posterior median odds ratios (95% credible intervals) from each pairwise meta-analysis (see supplementary appendix 2 for full results). Heterogeneity was statistically significant in four of the pairwise meta-analyses: magnesium sulfate versus prostaglandin inhibitors (n=3 trials), magnesium sulfate versus betamimetic (n=6), magnesium sulfate versus calcium channel blockers (n=5), and calcium channel blockers versus beta mimetics (n=15). The I^2 statistic for these four comparisons was 70% or more.

The network meta-analysis provided treatment effect estimates for all 21 of the pairwise comparisons that are possible from the seven drug classes (fig 6). Although the point estimates indicated that placebo was responsible for fewer maternal side effects than all active classes, some evidence suggested that prostaglandin inhibitors and oxytocin receptor blockers were reasonably well tolerated, with credible intervals consistent with a reduction in all cause maternal side effects (table). Fig 7 reports the probabilities of every drug class being ranked at each of the possible seven positions. For all cause maternal side effects, placebo was ranked first, with a probability of 61%,

suggesting that women given placebo experienced fewer maternal side effects. Placebo had a 98% probability of being ranked in the top three drug classes. The closest active competitor for reducing all cause maternal side effects was prostaglandin inhibitors, with a probability of 79%, followed by oxytocin receptor blockers (70%). Calcium channel blockers had only a 15% probability of being ranked in the top three drug classes for maternal side effects.

Other outcomes

The four outcomes analyzed were the most commonly and consistently reported across the network of trials. It was not possible to fit either pairwise or network meta-analysis models to all the outcomes extracted for the systematic review. This was because of concerns about multiplicity and because of the small number of studies that consistently reported other outcomes such as delivery delayed for seven days, delivery delayed until 37 weeks' gestation, bronchopulmonary dysplasia, fetal sepsis, intraventricular hemorrhage, necrotizing enterocolitis, hyperbilirubinemia, and premature closure of the ductus arteriosus. Only seven trials reported neonatal composite outcomes.[42 50 62 106 108 117 120] The components of the composites in each of these trials were different. Because of this and the way outcomes were reported by individual trials, no overall result for composite outcome could be analyzed.

Sensitivity analyses

Two sensitivity analyses were carried out for the primary outcome; restricted to studies at low risk of bias and restricted to those excluding women at high risk of preterm delivery (women with multiple gestation and ruptured membranes). Neither analysis changed the finding that prostaglandin inhibitors were the best class for delaying preterm delivery by 48 hours. Excluding these trials removed the class other from the analysis. Class treatment effects were not modified by using metaregression to explore the effect of length of treatment delivery (whether it was of short duration or prolonged).

Discussion

Balancing the results relating to benefits and harms, this systematic review and network meta-analysis on trials of tocolytics found that prostaglandin inhibitors and calcium channel blockers have the highest probability of being the best therapy for preterm delivery on the basis of the four outcomes: delivery delayed by 48 hours, neonatal mortality, neonatal respiratory distress syndrome, and maternal side effects (all causes). Of all the classes considered, prostaglandin inhibitors had the highest probability of being the most effective class for delaying preterm delivery and had the most favorable maternal side effect profile. They did not, however, perform as well for the neonatal outcomes. When the probability of being ranked in the top three treatments for delaying delivery was considered, calcium channel blockers also performed reasonably well, with a 57% probability of being the best class. For the two neonatal outcomes, calcium channel blockers also had the highest probability of being the best class. Uncertainty was, however, considerable and the benefit from calcium channel blockers must be considered in light of a somewhat higher probability of being associated with maternal side effects. To fully encapsulate the uncertainty arising from the random effects analysis we also calculated the predicted treatment effect for a new study on a randomly chosen drug from each class reported with a 95% prediction interval (see supplementary table 4). For the outcome of delivery being delayed by 48 hours the credible

intervals were wider; only prostaglandin inhibitors continued to show a statistically significant beneficial treatment effect. For maternal side effects, only beta mimetics continued to show a statistically significant harmful effect, when compared with placebo. For all other outcomes and all comparisons the prediction interval crossed the line of null effect, indicating the true extent of the uncertainty. Weighing the balance of the results seems to indicate that prostaglandin inhibitors would be reasonable first-line agents, followed by calcium channel blockers. This conclusion is true even when the analysis is limited to studies with the least risk of bias and with the most clinically homogenous participant populations. In the evidence base considered here, however, only one, small (n=79) head to head trial compared prostaglandin inhibitors with a calcium channel blocker.[116] The findings in this trial suggest that nifedipine (calcium channel blocker) was more effective than indomethacin (prostaglandin inhibitor) for rapid treatment effect but that the delay in delivery was similar between women in both groups who initially responded to treatment. Respiratory distress syndrome was not reported in the study and two neonatal deaths occurred in the indomethacin group compared with none in the nifedipine group.[116] Given the small amount of direct evidence and considerable uncertainty we identified for the neonatal outcomes, the findings from our network meta-analysis suggest that a head to head trial of these agents is needed to investigate further the effectiveness, adverse effects, and costs of these regimens to women. We therefore plan on carrying out an expected value of information analysis.

The finding that these two classes of medications have the best outcomes was similar to the findings of a recent pooled meta-analysis and decision analysis.[1] In our analysis, however, we used a hierarchical class model, which also retains the individual identity of the within class treatments. Therefore we were not able to include all the available evidence. The consistency of findings despite two different methods for meta-analysis strengthens the argument for these agents being the first-line choice for tocolysis. Similar to other findings, tocolysis has been shown to be beneficial for delaying delivery for at least 48 hours compared with no tocolysis.[121]

Prostaglandin inhibitors have been studied widely but their use is limited in practice. Some data indicate a possible association between neonatal complications and antenatal prostaglandin inhibitors, including reversible premature closure of the ductus arteriosus.[5 122 123] A Cochrane review of trials on prostaglandin inhibitors to prevent preterm delivery found that data, albeit limited, did not show increased adverse neonatal outcomes.[8] One of the reasons why prostaglandin inhibitors may be effective in delaying delivery is because of the large proportion of women at risk of preterm delivery owing to intrauterine inflammation and infection.[124] Clinically, practitioners who utilize indomethacin typically limit its use to pregnancies under 32 weeks' gestation. Because of the limited trial data, in our network meta-analysis we were unable to determine a difference in the rate of premature closure of the ductus arteriosus with prostaglandin inhibitors. Indomethacin is commonly utilized to close a patent ductus in the newborn period.[125 126]

Network meta-analysis

To our knowledge this is the first application of network meta-analysis in obstetrics, particularly in tocolysis for preterm delivery. In obstetrics, where multiple treatment options are usual, often only compared pairwise, network meta-analysis has several benefits. We were able to combine all the available evidence on tocolytic treatments in a single pooled analysis, even if none had been compared in a direct trial. Network

meta-analysis makes the assumption of consistency, which should always be checked.[25] We found no evidence of inconsistency, and empirical studies comparing direct and indirect evidence have found no systematic differences.[127] Where differences have been found, they seem to have been in situations where the doses or treatment combinations in the direct and indirect evidence were not comparable.[128 129] As long as the assumption of consistency is fulfilled, all relevant treatments, even those that are deemed outdated or ineffectual, can be included in the network to utilize the information on relative treatment effects and inform the rankings.[130] Furthermore, a randomized controlled trial with all of these arms designed to answer the same questions presented here would require the randomization of many more participants.[1 131]

Strengths and limitations of the study

A strength of our systematic review was the inclusion of several trials in non-English. These trials are often excluded from meta-analyses but help to inform the model. We therefore believe that we have included all relevant randomized controlled trials on tocolytics up to the search date.

Our analysis was limited by the data in the included studies and the structure of the reported data. For example, neonatal mortality was included as an outcome because of its high clinical importance when evaluating the harms and benefits of tocolytic therapy. However, few neonatal deaths were reported across the network of trials. Indeed, even after excluding trials with zero deaths on both arms, 24/40 trials (60%) reported one or fewer deaths on at least one arm. Meta-analysis of rare events is known to be problematic.[132 133] This is further compounded in network meta-analysis if there are also few trials per comparison, as here for the neonatal mortality outcome (median 3 (range 1-8) trials per comparison). While we understand the importance of this outcome and realize that excluding these trials may lead to a problematic overestimation of rates of neonatal mortality, including these trials in the analysis invalidated the statistical models. Therefore, extra caution should be exercised when interpreting the treatment rankings for this outcome. Furthermore, outcomes were not consistently reported across the ensemble of trials. For example, the most commonly reported efficacy outcome was delivery delayed by 48 hours. However, some studies reported delays at 72 hours or at seven days. Similarly, some studies reported effectiveness as delivery after 34 weeks, 36 weeks, or after 38 weeks' gestation. This is essentially the same underlying variable, and with better reporting these data would allow a more thorough synthesis.[134] A further finding from our systematic review was the need for consistent outcome reporting across trials on tocolytics.

We were also unable to utilize the maternal and neonatal data for several studies that did not include the same outcome measures as the other trials—that is, they were not included in the network meta-analysis although they met the eligibility criteria for the systematic review. This is one of the reasons we were unable to create a composite outcome. Owing to the way data were presented by the individual studies, it was impossible to discern if neonates had more than one outcome, such as respiratory distress syndrome, intraventricular hemorrhage, and necrotizing enterocolitis. Different components of a composite outcome were reported by only seven trials.[42 50 62 106 108 117 120] Attempting to extrapolate a composite score for the individual studies would have essentially duplicated the analysis on respiratory distress syndrome. An individual participant meta-analysis might overcome this limitation. It is also possible that some trials, such as potentially some non-English language trials, were not included. We diligently searched the accessible

literature on tocolytic studies. The trials included represent the major accessible published literature on tocolytic therapy.

The results may also be limited by the modeling assumptions. We preferred a class effect model because of clinical characteristics and mechanisms of the drugs within the class. When compared with a distinct treatment effect model, the class model reported here showed reasonable fit and provided some support for the assumption of a tocolytic class effect. In addition to the class model reported, we also explored a model with the strong assumption of a single class effect, such that all treatments within a class had the same effect. This model gave an adequate fit to the data. However it did not allow estimation of individual treatment effects, which we considered desirable. We also attempted to relax the assumption of equal within class variances, by fitting a model assuming each individual treatment has a distinct effect, but from a common class, with common class effect and a class specific variance. However, we were unable to obtain results owing to computational problems from insufficient data to estimate the variances.

Because of the multitude of doses used for many trials, we were unable to stratify for dose of drug. However, to deal with potential concerns about heterogeneity caused by dose, we carried out a metaregression by treatment duration, which did not alter the findings. During our systematic review, we assessed the risk of bias in the retrieved studies. To aid in doing a sensitivity analysis by risk of bias, we arbitrarily combined the risk of bias assessments into a composite score. Although this is often discouraged,[135] it was necessary for our planned sensitivity analysis. To facilitate a more global view of the risk of bias in the presented studies the full assessments are presented in the supplementary file. In addition, some of these drugs are not licensed in some countries and may not be available to practitioners.

The data contained in the different trials highlight inconsistencies in reporting of outcomes. Delay of delivery was consistently reported and thus the model fits this outcome best. It was more difficult when focusing on maternal safety and on important neonatal outcomes. Long term morbidity and mortality outcomes were inconsistently reported. In future trials, a standard list of both maternal safety and neonatal short term and long term outcomes should be reported to allow researchers to understand the benefits or lack of benefits of tocolytic therapy.

Conclusion

Tocolytic therapy can delay delivery and has an impact on short term neonatal outcomes. In this network meta-analysis, prostaglandin inhibitors and calcium channel blockers had the highest probability of delaying delivery and improving neonatal outcomes.

> ## What is already known on this topic
>
> Tocolytics are used to delay preterm delivery to allow antenatal corticosteroids to be administered to improve neonatal outcomes
>
> Many different drugs have been utilized as tocolytic therapy, but a standard first line drug has not emerged
>
> A multitude of trials have compared a few drugs with each other but no comprehensive trial has compared all commonly used drug classes
>
> ## What this study adds
>
> In this network meta-analysis of tocolytic therapy, prostaglandin inhibitors and calcium channel blockers had the highest probability of delaying delivery and improving neonatal outcomes
>
> Network meta-analysis techniques can be applied to obstetric interventions with heterogeneous treatment options
>
> Reporting of clinically relevant outcome comparisons should be improved and made more consistent

any organisations that might have an interest in the submitted work in the previous three years; and no other relationships or activities that could appear to have influenced the submitted work. DMC and NJW have taught for Pfizer and provided training on network meta-analysis to a research organization that undertakes systematic reviews and network meta-analysis for industry. None of these activities directly conflicts with the network meta-analysis presented.

Ethical approval: This study was approved by the Indiana University-Purdue University Indianapolis-Clarian institutional review board.

Data sharing: No additional data available.

1 Haas DM, Imperiale TF, Kirkpatrick PR, Klein RW, Zollinger TW, Golichowski AM. Tocolytic therapy: a meta-analysis and decision analysis. *Obset Gynecol* 2009;113:585-94.
2 Roberts D, Dalziel SR. Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth. *Cochrane Database Syst Rev* 2006;(3):CD004454.
3 Hamilton BE, Martin JA, Ventura SJ. Births: preliminary data for 2008. *Natl Vital Stat Rep* 2010;58:1-17.
4 ACOG practice bulletin. Management of preterm labor. No 43, May 2003. *Obstet Gynecol* 2003;101:1039-47.
5 Goldenberg RL. The management of preterm labor. *Obstet Gynecol* 2002;100:1020-37.
6 Crowther CA, Hiller JE, Doyle LW. Magnesium sulphate for preventing preterm birth in threatened preterm labour. *Cochrane Database Syst Rev* 2006;(3):CD001060.
7 Anotayanonth S, Subhedar NV, Garner P, Neilson JP, Harigopal S. Betamimetics for inhibiting preterm labour. *Cochrane Database Syst Rev* 2004;(2):CD004352.
8 King J, Flenady V, Cole S, Thornton S. Cyclo-oxygenase (COX) inhibitors for treating preterm labour. *Cochrane Database Syst Rev* 2005;(2):CD001992.
9 King JF, Flenady VJ, Papatsonis DN, Dekker GA, Carbonne B. Calcium channel blockers for inhibiting preterm labour. *Cochrane Database Syst Rev* 2003;(1):CD002255.
10 Papatsonis D, Flenady V, Cole S, Liley H. Oxytocin receptor antagonists for inhibiting preterm labour. *Cochrane Database Syst Rev* 2005;(3):CD004452.
11 Caritis SN. Metaanalysis and labor inhibition therapy. *Am J Obstet Gynecol* 2011;204:95-6.
12 Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;331:897-900.
13 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. *J Clin Epidemiol* 2010;63:875-82.
14 Ioannidis JP. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ* 2009;181:488-93.
15 Higgins JP, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions 5.1.0 [updated March 2011]. Chapter 8. In: The Cochrane Library. Wiley, 2009. www.cochrane-handbook.org.
16 Haas DM, Kirkpatrick PR, McIntosh JJ, Caldwell DM. Assessing the quality of the evidence for preterm labor tocolysis trials. *J Matern Fetal Neonatal Med* 2012;25:1646-52.
17 Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *J Am Stat Assoc* 2006;101:447-59.
18 Spiegelhalter D, Thomas A, Best N, Lunn DJ. WinBUGS user manual v1.4. MRC Biostatistics Unit, 2003.
19 Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009;373:746-58.
20 Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105-24.
21 Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002;21:2313-24.
22 Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. *Stat Med* 2009;28:1861-81.
23 Dakin H, Welton NJ, Ades AE. Mixed treatment comparison of repeated measurements of a continuous endpoint: an example using topical treatments for primary open-angle glaucoma and ocular hypertension. *Stat Med* 2011; published online 5 Jul.
24 Dempster AP. The direct use of likelihood for significance testing. *Stat Comput* 1997;7:247-52.
25 Dias S, Welton NJ, Sutton AJ, Caldwell DM, Guobing L, Ades AE. NICE DSU technical support document 4: inconsistency in networks of evidence based on randomised controlled trials. 2011.
26 Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc: Series B Stat Method* 2002;64:583-639.

27 Treatment of preterm labor with the beta-adrenergic agonist ritodrine. The Canadian Preterm Labor Investigators Group. *N Engl J Med* 1992;327:308-12.
28 Adam GS. Isoxuprine and premature labour. *Aust N Z J Obstet Gynaecol* 1966;6:294-8.
29 Cotton DB, Strassner HT, Hill LM, Schifrin BS, Paul RH. Comparison of magnesium sulfate, terbutaline and a placebo for inhibition of preterm labor. A randomized study. *J Reprod Med* 1984;29:92-7.
30 Cox SM, Sherman ML, Leveno KJ. Randomized investigation of magnesium sulfate for prevention of preterm birth. *Am J Obstet Gynecol* 1990;163:767-72.
31 Csapo AI, Herczeg J. Arrest of premature labor by isoxsuprine. *Am J Obstet Gynecol* 1977;129:482-91.
32 Goodwin TM, Paul R, Silver H, Spellacy W, Parsons M, Chez R, et al. The effect of the oxytocin antagonist atosiban on preterm uterine activity in the human. *Am J Obstet Gynecol* 1994;170:474-8.
33 How HY, Zafaranchi L, Stella CL, Recht K, Maxwell RA, Sibai BM, et al. Tocolysis in women with preterm labor between 32 0/7 and 34 6/7 weeks of gestation: a randomized controlled pilot study. *Am J Obstet Gynecol* 2006;194:976-81.
34 Ingemarsson I. Effect of terbutaline on premature labor. A double-blind placebo-controlled study. *Am J Obstet Gynecol* 1976;125:520-4.
35 Larsen JF, Eldon K, Lange AP, Leegaard M, Osler M, Olsen JS, et al. Ritodrine in the treatment of preterm labor: second Danish Multicenter Study. *Obstet Gynecol* 1986;67:607-13.
36 Larsen JF, Hansen MK, Hesseldahl H, Kristoffersen K, Larsen PK, Osler M, et al. Ritodrine in the treatment of premature labour. A clinical trial to compare a standard treatment with three regimens involving the use of ritodrine. *Br J Obstet Gynaecol* 1980;87:949-57.
37 Leveno KJ, Klein VR, Guzick DS, Young DC, Hankins GD, Williams ML. Single-centre randomised trial of ritodrine hydrochloride for preterm labour. *Lancet* 1986;1:1293-6.
38 Ma L. [Magnesium sulfate in prevention of preterm labor]. *Zhonghua Yi Xue Za Zhi* 1992;72:158-61, 91.
39 Merkatz IR, Peter JB, Barden TP. Ritodrine hydrochloride: a betamimetic agent for use in preterm labor. II. Evidence of efficacy. *Obstet Gynecol* 1980;56:7-12.
40 Mital P, Garg S, Khuteta RP, Khuteta S, Mital P. Mefenamic acid in prevention of premature labour. *J R Soc Health* 1992;112:214-6.
41 Niebyl JR, Blake DA, White RD, Kumor KM, Dubin NH, Robinson JC, et al. The inhibition of premature labor with indomethacin. *Am J Obstet Gynecol* 1980;136:1014-9.
42 Panter KR, Hannah ME, Amankwah KS, Ohlsson A, Jefferies AL, Farine D. The effect of indomethacin tocolysis in preterm labour on perinatal outcome: a randomised placebo-controlled trial. *Br J Obstet Gynaecol* 1999;106:467-73.
43 Read MD, Wellby DE. The use of a calcium antagonist (nifedipine) to suppress preterm labour. *Br J Obstet Gynaecol* 1986;93:933-7.
44 Romero R, Sibai BM, Sanchez-Ramos L, Valenzuela GJ, Veille JC, Tabor B, et al. An oxytocin receptor antagonist (atosiban) in the treatment of preterm labor: a randomized, double-blind, placebo-controlled trial with tocolytic rescue. *Am J Obstet Gynecol* 2000;182:1173-83.
45 Smith GN, Walker MC, Ohlsson A, O'Brien K, Windrim R. Randomized double-blind placebo-controlled trial of transdermal nitroglycerin for preterm labor. *Am J Obstet Gynecol* 2007;196:37 e1-8.
46 Spellacy WN, Cruz AC, Birk SA, Buhi WC. Treatment of premature labor with ritodrine: a randomized controlled study. *Obstet Gynecol* 1979;54:220-3.
47 Steer CM, Petrie RH. A comparison of magnesium sulfate and alcohol for the prevention of premature labor. *Am J Obstet Gynecol* 1977;129:1-4.
48 Weiner CP, Renk K, Klugman M. The therapeutic efficacy and cost-effectiveness of aggressive tocolysis for premature labor associated with premature rupture of the membranes. *Am J Obstet Gynecol* 1988;159:216-22.
49 Zuckerman H, Shalev E, Gilad G, Katzuni E. Further study of the inhibition of premature labor by indomethacin. Part II double-blind study. *J Perinat Med* 1984;12:25-9.
50 Husslein P, Cabero Roura L, Dudenhausen JW, Helmer H, Frydman R, Rizzo N, et al. Atosiban versus usual care for the management of preterm labor. *J Perinat Med* 2007;35:305-13.
51 Thornton S, Goodwin TM, Greisen G, Hedegaard M, Arce J-C. The effect of barusiban, a selective oxytocin antagonist, in threatened preterm labor at late gestational age: a randomized, double-blind, placebo-controlled trial. *Am J Obstet Gynecol* 2009;200:627.e1-10.
52 Al-Qattan F, Omu AE, Labeeb N. A prospective randomized study comparing nifedipine versus ritodrine for the suppression of preterm labour. *Med Princ Pract* 2000;9:164-73.
53 Jimenez Aramayo JF, Jimenez Martinez F, Lopez Rosales C. [Tocolytic therapy with magnesium sulfate and terbutaline for inhibition of premature labor]. *Ginecol Obstet Mex* 1990;58:265-9.
54 Treatment of preterm labor with the oxytocin antagonist atosiban: a double-blind, randomized, controlled comparison with salbutamol. *Eur J Obstet Gynecol Reprod Biol* 2001;98:177-85.
55 The oxytocin antagonist atosiban versus the beta-agonist terbutaline in the treatment of preterm labor. A randomized, double-blind, controlled study. *Acta Obstet Gynecol Scand* 2001;80:413-22.
56 Beall MH, Edgar BW, Paul RH, Smith-Wallace T. A comparison of ritodrine, terbutaline, and magnesium sulfate for the suppression of preterm labor. *Am J Obstet Gynecol* 1985;153:854-9.

57   Besinger RE, Niebyl JR, Keyes WG, Johnson TR. Randomized comparative trial of indomethacin and ritodrine for the long-term treatment of preterm labor. *Am J Obstet Gynecol* 1991;164:981-6; discussion 86-8.

58   Bisits A, Madsen G, Knox M, Gill A, Smith R, Yeo G, et al. The Randomized Nitric Oxide Tocolysis Trial (RNOTT) for the treatment of preterm labor. *Am J Obstet Gynecol* 2004;191:683-90.

59   Cararach V, Palacio M, Martinez S, Deulofeu P, Sanchez M, Cobo T, et al. Nifedipine versus ritodrine for suppression of preterm labor. Comparison of their efficacy and secondary effects. *Eur J Obstet Gynecol Reprod Biol* 2006;127:204-8.

60   Caritis SN, Toig G, Heddinger LA, Ashmead G. A double-blind study comparing ritodrine and terbutaline in the treatment of preterm labor. *Am J Obstet Gynecol* 1984;150:7-14.

61   Chau AC, Gabert HA, Miller JM Jr. A prospective comparison of terbutaline and magnesium for tocolysis. *Obstet Gynecol* 1992;80:847-51.

62   Essed GGM, Eskes TKAB, Jongsma HW. A randomized trial of two beta-mimetic drugs for the treatment of threatening early labor: clinical results in a prospective comparative study with ritodrine and fenoterol. *Eur J Obstet Gynecol Reprod Biol* 1978;8:341-8.

63   Fan L, Wu L, Huang X. [The effect of calcium entry blocker on the management of preterm labor—a randomized controlled study]. *J Pract Gynecol Obstet* 2003;19:87-9.

64   Garcia-Velasco JA, Gonzalez Gonzalez A. A prospective, randomized trial of nifedipine vs. ritodrine in threatened preterm labor. *Int J Gynaecol Obstet* 1998;61:239-44.

65   Goodwin TM, Valenzuela GJ, Silver H, Creasy G. Dose ranging study of the oxytocin antagonist atosiban in the treatment of preterm labor. Atosiban Study Group. *Obstet Gynecol* 1996;88:331-6.

66   Grignaffini A, Soncini E, Ronzoni E, Lo Cane F, Anfuso S, Nardelli GB. Clinical practice evaluation of combination of atosiban, ritodrine and ketoprofen for inhibiting preterm labor. *Minerva Ginecol* 2007;59:481-9.

67   Gummerus M. [Tocolysis with hexoprenalin and salbutamol in a clinical comparison]. *Geburtshilfe Frauenheilkd* 1983;43:151-5.

68   Hollander DI, Nagey DA, Pupkin MJ. Magnesium sulfate and ritodrine hydrochloride: a randomized comparison. *Am J Obstet Gynecol* 1987;156:631-7.

69   Holleboom CA, Merkus JM, van Elferen LW, Keirse MJ. Randomised comparison between a loading and incremental dose model for ritodrine administration in preterm labour. *Br J Obstet Gynaecol* 1996;103:695-701.

70   Janky E, Leng JJ, Cormier PH, Salamon R, Meynard J. [A randomized study of the treatment of threatened premature labor. Nifedipine versus ritodrine]. *J Gynecol Obstet Biol Reprod (Paris)* 1990;19:478-82.

71   Jannet D, Abankwa A, Guyard B, Carbonne B, Marpeau L, Milliez J. Nicardipine versus salbutamol in the treatment of premature labor. A prospective randomized study. *Eur J Obstet Gynecol Reprod Biol* 1997;73:11-6.

72   Koks CA, Brolmann HA, de Kleine MJ, Manger PA. A randomized comparison of nifedipine and ritodrine for suppression of preterm labor. *Eur J Obstet Gynecol Reprod Biol* 1998;77:171-6.

73   Kramer WB, Saade GR, Belfort M, Dorman K, Mayes M, Moise KJ Jr. A randomized double-blind study comparing the fetal effects of sulindac to terbutaline during the management of preterm labor. *Am J Obstet Gynecol* 1999;180:396-401.

74   Kupfermic M, Lessing JB, Yaron Y, Peyser MR. Nifedipine versus ritodrine for suppression of preterm labour. *Br J Obstet Gynaecol* 1993;100:1090-4.

75   Kurki T, Eronen M, Lumme R, Ylikorkala O. A randomized double-dummy comparison between indomethacin and nylidrin in threatened preterm labor. *Obstet Gynecol* 1991;78:1093-7.

76   Laohapojanart N, Soorapan S, Wacharaprechanont T, Ratanajamit C. Safety and efficacy of oral nifedipine versus terbutaline injection in preterm labor. *J Med Assoc Thai* 2007;90:2461-9.

77   Lauersen NH, Merkatz IR, Tejani N, Wilson KH, Roberson A, Mann LI, et al. Inhibition of premature labor: a multicenter comparison of ritodrine and ethanol. *Am J Obstet Gynecol* 1977;127:837-45.

78   Maitra N, Christian V, Kavishvar A. Tocolytic efficacy of nifedipine versus ritodrine in preterm labor. *Int J Gynaecol Obstet* 2007;97:147-8.

79   Mawaldi L, Duminy P, Tamim H. Terbutaline versus nifedipine for prolongation of pregnancy in patients with preterm labor. *Int J Gynaecol Obstet* 2008;100:65-8.

80   Miller JM Jr, Keane MW, Horger EO 3rd. A comparison of magnesium sulfate and terbutaline for the arrest of premature labor. A preliminary report. *J Reprod Med* 1982;27:348-51.

81   Morales WJ, Smith SG, Angel JL, O'Brien WF, Knuppel RA. Efficacy and safety of indomethacin versus ritodrine in the management of preterm labor: a randomized study. *Obstet Gynecol* 1989;74:567-72.

82   Moutquin JM, Sherman D, Cohen H, Mohide PT, Hochner-Celnikier D, Fejgin M, et al. Double-blind, randomized, controlled trial of atosiban and ritodrine in the treatment of preterm labor: a multicenter effectiveness and safety study. *Am J Obstet Gynecol* 2000;182:1191-9.

83   Neri I, Monari F, Valensise H, Vasapollo B, Facchinetti F, Volpe A. Computerized evaluation of fetal heart rate during tocolytic treatment: comparison between atosiban and ritodrine. *Am J Perinatol* 2009;26:259-63.

84   Papatsonis DN, Van Geijn HP, Ader HJ, Lange FM, Bleker OP, Dekker GA. Nifedipine and ritodrine in the management of preterm labor: a randomized multicenter trial. *Obstet Gynecol* 1997;90:230-4.

85   Raymajhi R, Pratap K. A comparative study between nifedipine and isoxsuprine in the suppression of preterm labour. *Kathmandu Univ Med J (KUMJ)* 2003;1:85-90.

86   Roy UK, Pan S. Use of calcium antagonist (nifedipine) in premature labour. *J Indian Med Assoc* 1993;91:8-10.

87   Shim JY, Park YW, Yoon BH, Cho YK, Yang JH, Lee Y, et al. Multicentre, parallel group, randomised, single-blind study of the safety and efficacy of atosiban versus ritodrine in the treatment of acute preterm labour in Korean women. *BJOG* 2006;113:1228-34.

88   Spearing G. Alcohol, indomethacin, and salbutamol. A comparative trial of their use in preterm labor. *Obstet Gynecol* 1979;53:171-4.

89   Surichamorn P. The efficacy of terbutaline and magnesium sulfate in the management of preterm labor. *J Med Assoc Thai* 2001;84:98-104.

90   Tchilinguirian NG, Najem R, Sullivan GB, Craparo FJ. The use of ritodrine and magnesium sulfate in the arrest of premature labor. *Int J Gynaecol Obstet* 1984;22:117-23.

91   Trabelsi K, Hadj Taib H, Amouri H, Abdennadheur W, Ben Amar H, Kallel W, et al. Nicardipine versus salbutamol in the treatment of premature labor: comparison of their efficacy and side effects. *Tunis Med* 2008;86:43-8.

92   Van de Water M, Kessel ET, De Kleine MJ, Oei SG. Tocolytic effectiveness of nifedipine versus ritodrine and follow-up of newborns: a randomised controlled trial. *Acta Obstet Gynecol Scand* 2008;87:340-5.

93   Wang H, Zeng W, Liu H, Ou Y. [A randomized controlled trial on the treatment of preterm labor with ritodrine hydrochloride and magnesium sulfate]. *J West China Univ Med Sci* 2000;31:515-7.

94   Weerakul W, Chittacharoen A, Suthutvoravut S. Nifedipine versus terbutaline in management of preterm labor. *Int J Gynaecol Obstet* 2002;76:311-3.

95   Wilkins IA, Lynch L, Mehalek KE, Berkowitz GS, Berkowitz RL. Efficacy and side effects of magnesium sulfate and ritodrine as tocolytic agents. *Am J Obstet Gynecol* 1988;159:685-9.

96   Zhu B, Fu Y. [Treatment of preterm labor with ritodrine]. *Zhonghua Fu Chan Ke Za Zhi* 1996;31:721-3.

97   Lin CH, Lin SY, Shyu MK, Chen SU, Lee CN. Randomized trial of oxytocin antagonist atosiban versus beta-adrenergic agonists in the treatment of spontaneous preterm labor in Taiwanese women. *J Formos Med Assoc* 2009;108:493-501.

98   Motazedian S, Ghaffarpasand F, Mojtahedi K, Asadi N. Terbutaline versus salbutamol for suppression of preterm labor: a randomized clinical trial. *Ann Saudi Med* 2010;30:370-5.

99   Borna S, Saeidi FM. Celecoxib versus magnesium sulfate to arrest preterm labor: randomized trial. *J Obstet Gynaecol Res* 2007;33:631-4.

100   El-Sayed YY, Riley ET, Holbrook RH Jr, Cohen SE, Chitkara U, Druzin ML. Randomized comparison of intravenous nitroglycerin and magnesium sulfate for treatment of preterm labor. *Obstet Gynecol* 1999;93:79-83.

101   Floyd RC, McLaughlin BN, Perry KG, Martin RW, Sullivan CA, Morrison JC. Magnesium sulfate or nifedipine hydrochloride for acute tocolysis of preterm labor: efficacy and side effects. *J Matern Fetal Invest* 1995;5:25-9.

102   Glock JL, Morales WJ. Efficacy and safety of nifedipine versus magnesium sulfate in the management of preterm labor: a randomized study. *Am J Obstet Gynecol* 1993;169:960-4.

103   Haghighi L. Prevention of preterm delivery: nifedipine or magnesium sulfate. *Int J Gynaecol Obstet* 1999;66:297-8.

104   Larmon JE, Ross BS, May WL, Dickerson GA, Fischer RG, Morrison JC. Oral nicardipine versus intravenous magnesium sulfate for the treatment of preterm labor. *Am J Obstet Gynecol* 1999;181:1432-7.

105   Lorzadeh N, Kazemirad S, Lorzadeh M, Dehnori A. A comparison of human chorionic gonadotropin with magnesium sulphate in inhibition of preterm labor. *J Med Sci* 2007;7:640-4.

106   Lyell DJ, Pullen K, Campbell L, Ching S, Druzin ML, Chitkara U, et al. Magnesium sulfate compared with nifedipine for acute tocolysis of preterm labor: a randomized controlled trial. *Obstet Gynecol* 2007;110:61-7.

107   McWhorter J, Carlan SJ, OLeary TD, Richichi K, OBrien WF. Rofecoxib versus magnesium sulfate to arrest preterm labor: a randomized trial. *Obstet Gynecol* 2004;103:923-30.

108   Mittendorf R, Dambrosia J, Pryde PG, Lee KS, Gianopoulos JG, Besinger RE, et al. Association between the use of antenatal magnesium sulfate in preterm labor and adverse health outcomes in infants. *Am J Obstet Gynecol* 2002;186:1111-8.

109   Morales WJ, Madhav H. Efficacy and safety of indomethacin compared with magnesium sulfate in the management of preterm labor: a randomized study. *Am J Obstet Gynecol* 1993;169:97-102.

110   Parilla BV, Tamura RK, Cohen LS, Clark E. Lack of effect of antenatal indomethacin on fetal cerebral blood flow. *Am J Obstet Gynecol* 1997;176:1166-9; discussion 69-71.

111   Schorr SJ, Ascarelli MH, Rust OA, Ross EL, Calfee EL, Perry KG Jr, et al. A comparative study of ketorolac (Toradol) and magnesium sulfate for arrest of preterm labor. *South Med J* 1998;91:1028-32.

112   Taherian A, Dehdar P. Comparison of efficacy and safety of nifedipine versus magnesium sulfate in treatment of preterm labor. *J Res Med Sci* 2007;12:136-41.

113   Al-Omari WR, Al-Shammaa HB, Al-Tikriti EM, Ahmed KW. Atosiban and nifedipine in acute tocolysis: a comparative study. *Eur J Obstet Gynecol Reprod Biol* 2006;128:129-34.

114   Kashanian M, Akbarian AR, Soltanzadeh M. Atosiban and nifedipin for the treatment of preterm labor. *Int J Gynaecol Obstet* 2005;91:10-4.

115   Amorim MMR, Lippo LAM, Costa AAR, Coutinho IC, Souza ASR. [Transdermal nitroglycerin versus oral nifedipine administration for tocolysis: a randomized clinical trial]. *Rev* 2009;31:552-8.

116   Kashanian M, Bahasadri S, Zolali B. Comparison of the efficacy and adverse effects of nifedipine and indomethacin for the treatment of preterm labor. *Int J Gynaecol Obstet* 2011;113:192-5.

117   Nassar AH, Abu-Musa AA, Awwad J, Khalil A, Tabbara J, Usta IM. Two dose regimens of nifedipine for management of preterm labor: a randomized controlled trial. *Am J Perinatol* 2009;26:575-81.

118   Rasanen J, Jouppila P. Fetal cardiac function and ductus arteriosus during indomethacin and sulindac therapy for threatened preterm labor: a randomized study. *Am J Obstet Gynecol* 1995;173:20-5.

119   Sawdy RJ, Lye S, Fisk NM, Bennett PR. A double-blind randomized study of fetal side effects during and after the short-term maternal administration of indomethacin, sulindac, and nimesulide for the treatment of preterm labor. *Am J Obstet Gynecol* 2003;188:1046-51.

120   Stika CS, Gross GA, Leguizamon G, Gerber S, Levy R, Mathur A, et al. A prospective randomized safety trial of celecoxib for treatment of preterm labor. *Am J Obstet Gynecol* 2002;187:653-60.

121   Macones GA, Bader TJ, Asch DA. Optimising maternal-fetal outcomes in preterm labour: a decision analysis. *Br J Obstet Gynaecol* 1998;105:541-50.

122   Groom KM, Shennan AH, Jones BA, Seed P, Bennett PR. TOCOX—a randomised, double-blind, placebo-controlled trial of rofecoxib (a COX-2-specific prostaglandin inhibitor) for the prevention of preterm delivery in women at high risk. *BJOG* 2005;112:725-30.

123   Norton ME, Merrill J, Cooper BA, Kuller JA, Clyman RI. Neonatal complications after the administration of indomethacin for preterm labor. *N Engl J Med* 1993;329:1602-7.

124   Goldenberg RL, Hauth JC, Andrews WW. Intrauterine infection and preterm delivery. *N Engl J Med* 2000;342:1500-7.

125   Sekar KC, Corff KE. Treatment of patent ductus arteriosus: indomethacin or ibuprofen? *J Perinatol* 2008;28(Suppl 1):S60-2.

126   Peckham GJ, Miettinen OS, Ellison RC, Kraybill EN, Gersony WM, Zierler S, et al. Clinical course to 1 year of age in premature infants with patent ductus arteriosus: results of a multicenter randomized trial of indomethacin. *J Pediatr* 1984;105:285-91.

127   Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003;326:472.

128   Caldwell DM, Gibb DM, Ades AE. Validity of indirect comparisons in meta-analysis. *Lancet* 2007;369:270; author reply 71.

129   Chou R, Fu R, Huffman LH, Korthuis PT. Initial highly-active antiretroviral therapy with a protease inhibitor versus a non-nucleoside reverse transcriptase inhibitor: discrepancies between direct and indirect meta-analyses. *Lancet* 2006;368:1503-15.

130 Dias S, Welton NJ, Sutton A, Ades A. NICE DSU technical support document 2: a generalised linear modelling framework for pair-wise and network meta-analysis. Technical support document series No 2, 2011. www.nicedsu.org.uk.

131 Caldwell DM. Using mixed treatment comparisons in health technology assessment. University of Bristol, 2008.

132 Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med* 2004;23:1351-75.

133 Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med* 2007;26:53-77.

134 Lu G, Ades AE, Sutton AJ, Cooper NJ, Briggs AH, Caldwell DM. Meta-analysis of mixed treatment comparisons at multiple follow-up times. *Stat Med* 2007;26:3681-99.

135 Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999;282:1054-60.

Table

Table 1| Efficacy, neonatal respiratory distress syndrome, and tolerability of tocolytics, using placebo as reference class. Values are posterior median odds ratios (95% credible intervals) unless stated otherwise

Drug class	48 hour delay in delivery*	Probability of being best†	Neonatal mortality‡	Probability of being best	Neonatal respiratory distress syndrome‡	Probability of being best	Maternal side effects§	Probability of being best
Placebo or control	—	<0.01	—	0.01	—	0.02	—	0.61
Beta mimetic	2.41 (1.27 to 4.55)	0.01	0.62 (0.14 to 2.48)	0.12	0.85 (0.50 to 1.45)	0.14	22.68 (7.51 to 73.67)	<0.01
Prostaglandin inhibitor	5.39 (2.14 to 12.34)	0.83	0.62 (0.04 to 4.63)	0.28	0.87 (0.40 to 1.75)	0.20	1.63 (0.40 to 6.85)	0.21
Calcium channel blocker	2.71 (1.17 to 5.91)	0.06	0.39 (0.09 to 1.49)	0.41	0.71 (0.37 to 1.43)	0.47	3.80 (1.02 to 16.92)	0.01
Other	0.93 (0.13 to 6.14)	0.04	2.79 (0.28 to 31.75)	0.02	1.54 (0.55 to 4.71)	0.04	—	—
Magnesium sulfate	2.76 (1.58 to 4.94)	0.02	0.97 (0.29 to 3.29)	0.03	0.99 (0.58 to 1.71)	0.03	8.15 (2.47 to 27.70)	<0.01
Oxytocin receptor blocker (atosiban)	2.02 (1.10 to 3.80)	0.01	0.62 (0.16 to 2.35)	0.13	0.89 (0.55 to 1.37)	0.10	1.99 (0.61 to 6.94)	0.08
Nitrates	1.91 (0.64 to 5.33)	0.04	—	—			3.19 (0.41 to 20.84)	0.10

*Odds ratios >1 favor active class.

†Probability that given drug class is best agent to use for given outcome based on rankings over all eight drug classes.

‡Odds ratios <1 favor active class.

§Odds ratios >1 favor placebo.

Figures
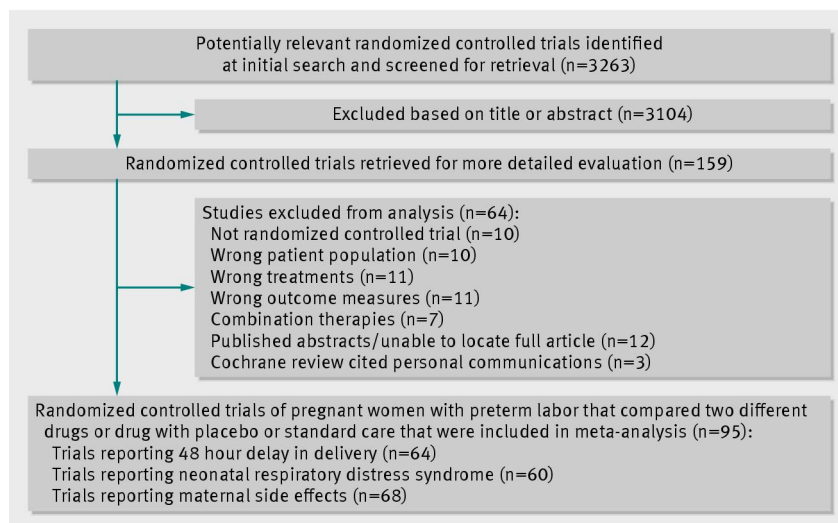
Fig 1 Summary of steps for trial retrieval for network meta-analysis

Fig 2 Graphic representation of tocolytic trials retrieved for network meta-analysis. Lines represent trials comparing two classes of drug for treatment of preterm delivery. Numbers on lines represent number of trials and total number of participants in those trials

Analyses

●— Network meta-analysis   ○— Direct pairwise

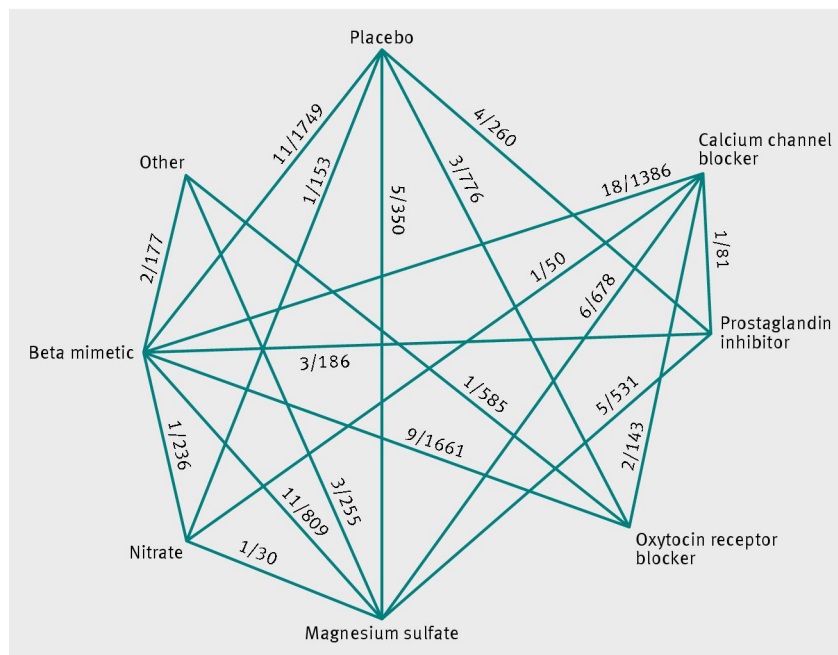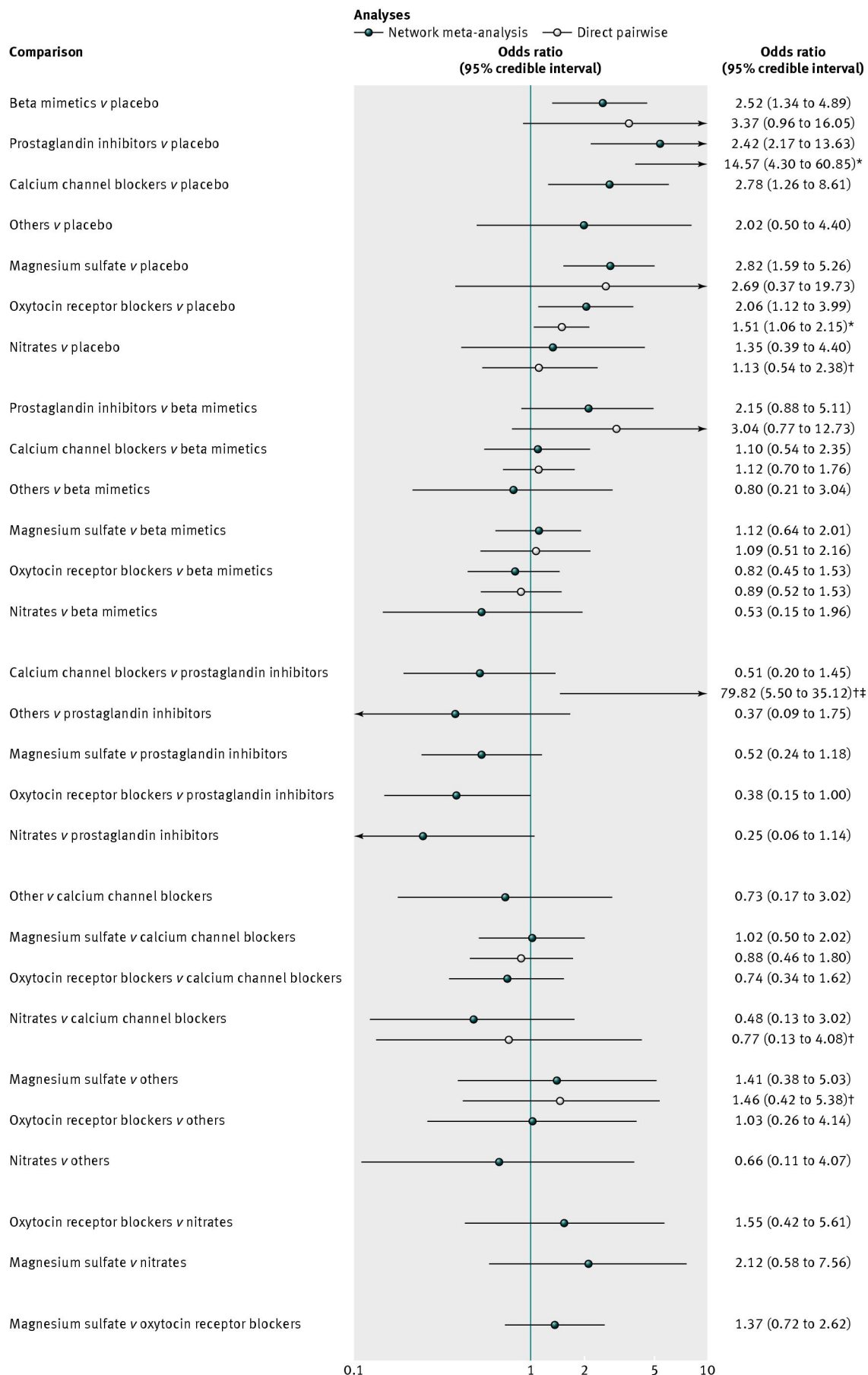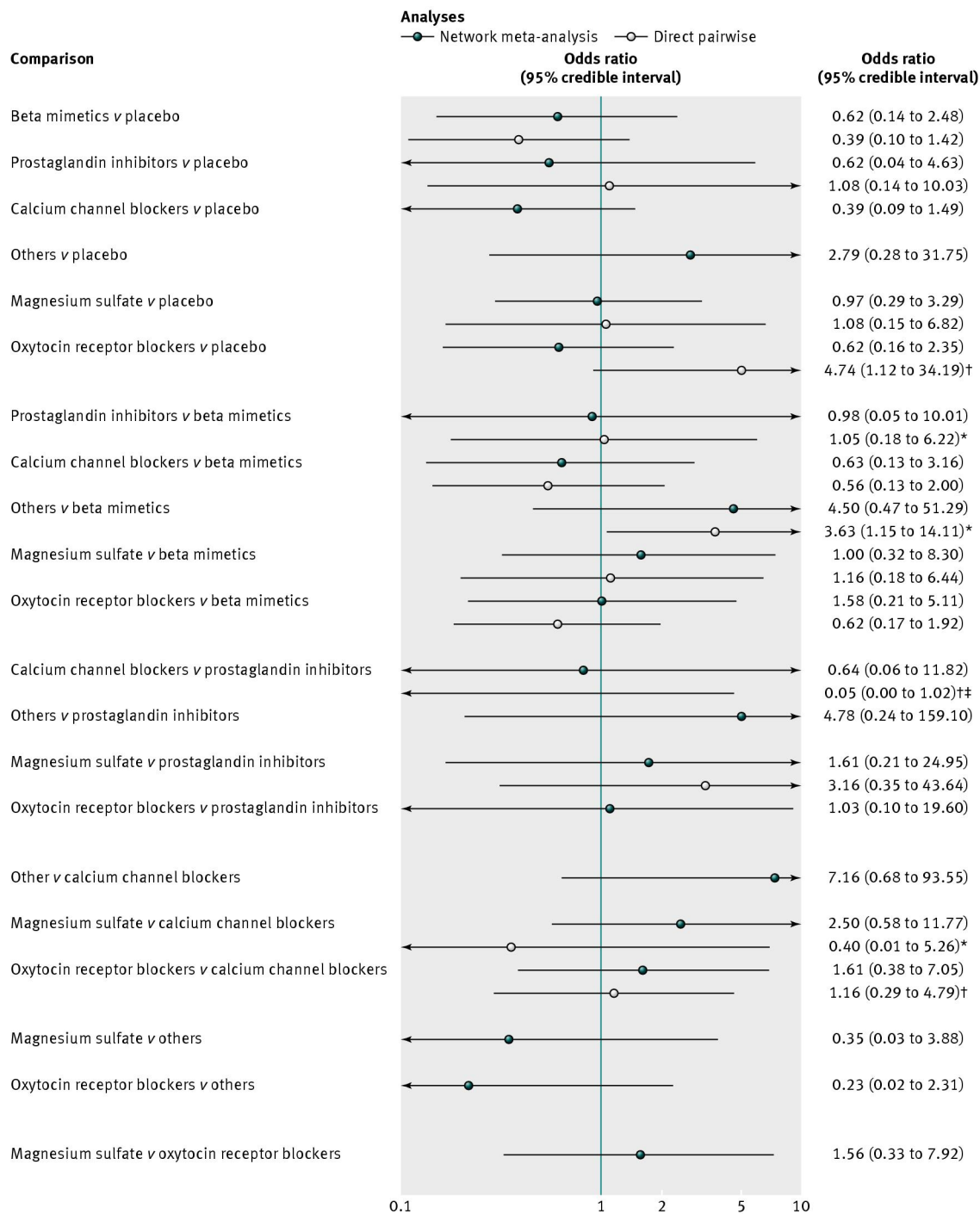| Comparison	Odds ratio (95% credible interval)	Odds ratio (95% credible interval)
Beta mimetics v placebo		2.52 (1.34 to 4.89)
		3.37 (0.96 to 16.05)
Prostaglandin inhibitors v placebo		2.42 (2.17 to 13.63)
		14.57 (4.30 to 60.85)*
Calcium channel blockers v placebo		2.78 (1.26 to 8.61)
Others v placebo		2.02 (0.50 to 4.40)
Magnesium sulfate v placebo		2.82 (1.59 to 5.26)
		2.69 (0.37 to 19.73)
Oxytocin receptor blockers v placebo		2.06 (1.12 to 3.99)
		1.51 (1.06 to 2.15)*
Nitrates v placebo		1.35 (0.39 to 4.40)
		1.13 (0.54 to 2.38)†
Prostaglandin inhibitors v beta mimetics		2.15 (0.88 to 5.11)
		3.04 (0.77 to 12.73)
Calcium channel blockers v beta mimetics		1.10 (0.54 to 2.35)
		1.12 (0.70 to 1.76)
Others v beta mimetics		0.80 (0.21 to 3.04)
Magnesium sulfate v beta mimetics		1.12 (0.64 to 2.01)
		1.09 (0.51 to 2.16)
Oxytocin receptor blockers v beta mimetics		0.82 (0.45 to 1.53)
		0.89 (0.52 to 1.53)
Nitrates v beta mimetics		0.53 (0.15 to 1.96)
Calcium channel blockers v prostaglandin inhibitors		0.51 (0.20 to 1.45)
		79.82 (5.50 to 35.12)†‡
Others v prostaglandin inhibitors		0.37 (0.09 to 1.75)
Magnesium sulfate v prostaglandin inhibitors		0.52 (0.24 to 1.18)
Oxytocin receptor blockers v prostaglandin inhibitors		0.38 (0.15 to 1.00)
Nitrates v prostaglandin inhibitors		0.25 (0.06 to 1.14)
Other v calcium channel blockers		0.73 (0.17 to 3.02)
Magnesium sulfate v calcium channel blockers		1.02 (0.50 to 2.02)
		0.88 (0.46 to 1.80)
Oxytocin receptor blockers v calcium channel blockers		0.74 (0.34 to 1.62)
Nitrates v calcium channel blockers		0.48 (0.13 to 3.02)
		0.77 (0.13 to 4.08)†
Magnesium sulfate v others		1.41 (0.38 to 5.03)
		1.46 (0.42 to 5.38)†
Oxytocin receptor blockers v others		1.03 (0.26 to 4.14)
Nitrates v others		0.66 (0.11 to 4.07)
Oxytocin receptor blockers v nitrates		1.55 (0.42 to 5.61)
Magnesium sulfate v nitrates		2.12 (0.58 to 7.56)
Magnesium sulfate v oxytocin receptor blockers		1.37 (0.72 to 2.62)

0.1   1   2   5   10

Fig 3 Results from network and pairwise meta-analyses for 48 hour delay in delivery. Direct meta-analysis refers to trials that compared two drug classes directly. In most cases analyses were undertaken using a random effects model. *Fixed effect meta-analyses. †Single trial. ‡Continuity correction used (0.5 added to each cell of 2×2 table)
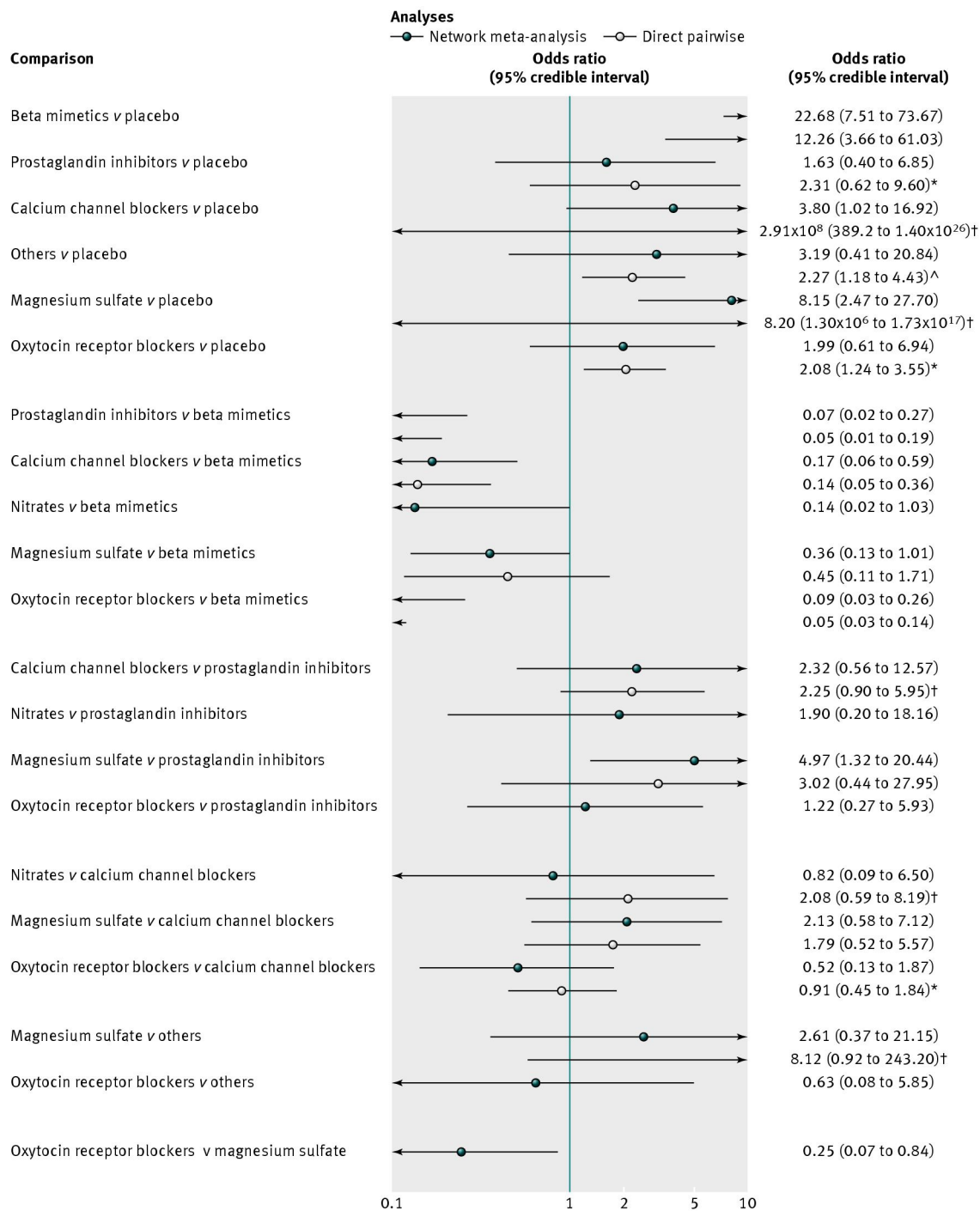
Fig 4 Results from network and pairwise meta-analyses for neonatal mortality. Direct meta-analysis refers to trials that compared two drug classes directly. In most cases analyses were undertaken using a random effects model. *Fixed effect meta-analyses. †Single trial. ‡Continuity correction used (0.5 added to each cell of 2×2 table)
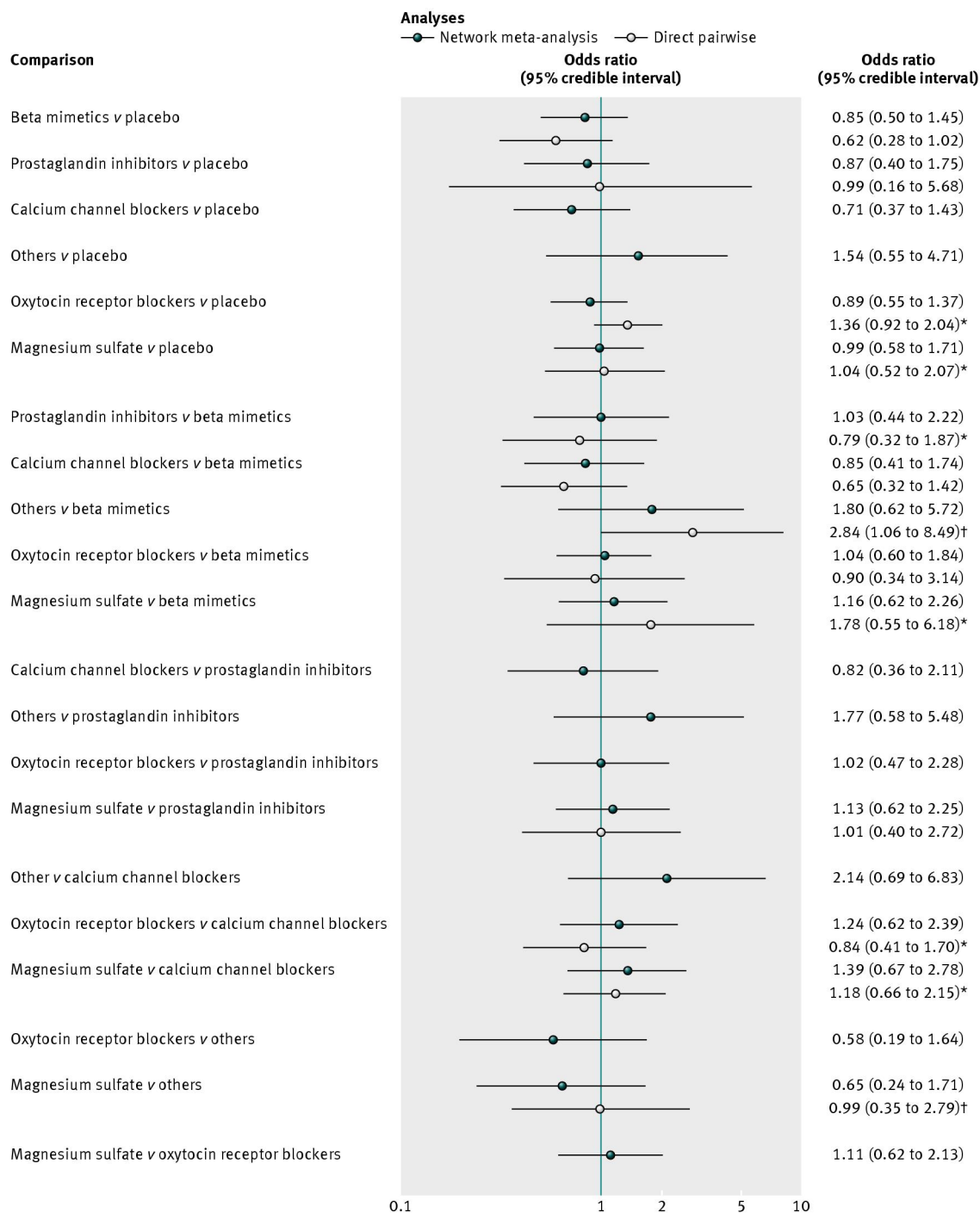
Fig 5 Results from network and pairwise meta-analyses for neonatal respiratory distress syndrome. Direct meta-analysis refers to trials that compared two drug classes directly. In most cases analyses were undertaken using a random effects model. *Fixed effect meta-analyses. †Single trial

Fig 6 Results from network and pairwise meta-analyses for maternal side effects. Direct meta-analysis refers to trials that compared two drug classes directly. In most cases analyses were undertaken using a random effects model. *Fixed effect meta-analyses. †Single trial
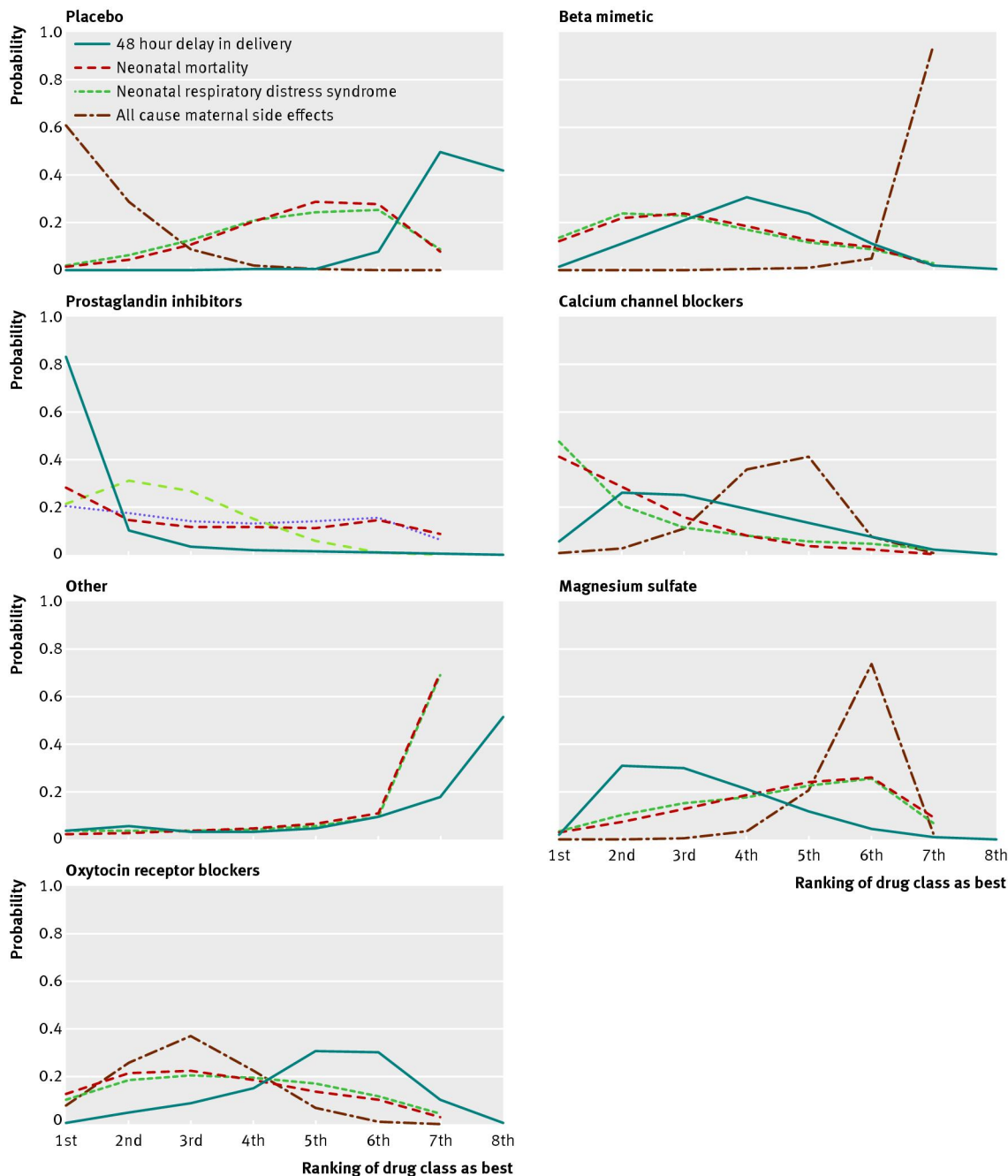
Fig 7 Rankings for efficacy of tocolytics and adverse events. Graph displays distribution of probabilities for each outcome. Ranking indicates probability that drug class is first "best," second "best," etc. Dot-dashed line represents 48 hour delay in delivery. Solid line indicates neonatal mortality. Dashed line indicates respiratory distress syndrome. Dotted line represents all cause maternal side effects