Inconsistency checks for review question: What is the effectiveness of uterotonics for the prevention of postpartum haemorrhage?

Guidelines Technical Support Unit (TSU), University of Bristol (Beatrice Downing, Nicky J. Welton)

Introduction

The purpose of this analysis was to assess the consistency assumption in the NMA model used to estimate the comparative effectiveness uterotonics for the prevention of postpartum haemorrhage. The outcomes included in this analysis were 1) PPH ≥ 1000ml, 2) additional uterotonics, 3) blood transfusion, 4) ICU admission (morbidity), and 5) mean blood loss (ml).

Methods

Inconsistency checks

NMA assumes that the included studies are similar in terms of factors that might interact with the intervention effects (effect modifiers). So, the relative effect of intervention B vs intervention A would be expected to be similar in all of the studies (if they had included A and B interventions). This assumption is the same as that made in conventional pairwise meta-analysis, but we have to be particularly careful that the studies making different comparisons do not differ in effect modifiers (the data are consistent). We can assess this assumption by measuring statistical heterogeneity, and also by checking if the direct and indirect estimates are in agreement when there are loops of evidence in the network.

To conduct inconsistency checks, an appropriate base-case model (fixed or random effects) must be determined beforehand. We assessed and compared the fit of a fixed effect model and a random effects model with a standard, uninformative prior distribution for all outcomes on the between-study standard deviation. The vague prior used on the between-study standard deviation was either Uniform (0,5) (for PPH ≥ 1000ml, additional uterotonics, blood transfusion, ICU admission (morbidity)), or Uniform (0, 10,000) (for mean blood loss). To determine if there is evidence of inconsistency, the selected consistency model (fixed or random effects) was compared to an “inconsistency”, or unrelated mean effects (UME), model (Dias 2013, Dias 2014). The latter is equivalent to having separate, unrelated, meta-analyses for every pairwise contrast, with a common variance parameter assumed in the case of random effects models. Note that the consistency assumption can only be assessed when there are closed loops of direct evidence on 3 treatments that are informed by at least 3 independent sources of evidence (Van Valkenhoef 2016).

The posterior mean of the residual deviance, which measures the magnitude of the differences between the observed data and the model predictions of the data, was used to assess and compare the goodness of fit of each model (Spiegelhalter 2002). Smaller values are preferred and in a well-fitting model the posterior mean residual deviance should be close to the number of data points in the network (each study arm contributes 1 data point on average) (Spiegelhalter 2002).

Where the base-case model assumes random effects, if the inconsistency model has smaller heterogeneity (measured by the posterior median between-study standard deviation) compared to the consistency model, then this may also indicate potential inconsistency in the data.

We performed further checks for evidence of inconsistency through node-splitting. The node-splitting method permits the direct and indirect evidence contributing to an estimate of a relative effect to be split and compared (Dias 2014, Dias 2010).

There are some small differences between the NMA estimates produced by the NMA models (presented in the main results) and those produced by the node-splitting models for exploring inconsistency (presented in forest plot below), due to small differences in the software used (WinBUGS or the GeMTC package in R). The NMA estimates presented in the main results were used to compare the safety and effectiveness of the interventions. In a separate exercise, the direct, indirect, and NMA estimates produced by the node-splitting modelling were used to assess how potential inconsistency between the direct and indirect estimates impacted the NMA estimates.

Results

Outcome: PPH >1000ml

Summary

We identified moderate heterogeneity in both full and mode-of-delivery subgroup datasets for this outcome, but little evidence of inconsistency. There was some indication of inconsistency between the studies comparing treatments on the oxytocin (>5 IU and ≤ 10 IU) vs carboprost v ergometrine loop, however these findings were driven by very small numbers of events in the Modi 2014 study.

Full data set

Global inconsistency check

Analysis of the full dataset for the outcome post-partum haemorrhage (>1000ml) included 98 studies (212 arms) of 13 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Comparing fixed (FE) and random-effect (RE) network meta-analysis (NMA) models indicated support for the random-effect model on the basis of a small decrease in DIC and a sizeable decrease in residual deviance (**Error! Reference source not found.**).

Total residual deviance was lower in the inconsistency UME model than in the NMA model; however, DIC was lower for the NMA model and the estimate of between-study SD was similar in the NMA and UME models (Table 49). This suggests that there is little evidence of inconsistency but moderate heterogeneity between study estimates.

Table 49: Model fit statistics for fixed- and random-effect NMA and UME models of the outcome PPH >1000ml, full dataset.

Outcome	Pop.	Model	Posterior total residual deviance[1]	Between-study SD Mean, 95% credible interval	pD	DIC[2]

[1] Posterior mean residual deviance compared to 212 total data points
[2] Deviance information criteria (DIC) – lower values preferred

PPH	Full	FE NMA	270.0	-	106.3	1078.7
PPH	Full	RE NMA	247.2	0.22 (0.03, 0.41)	123.7	1073.0
PPH	Full	FE UME	257.1	-	122.8	1082.3
PPH	Full	RE UME	239.0	0.22 (0.04, 0.44)	136.3	1077.8

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the random effects UME and NMA models (Figure 47), shows that two studies showed inconsistency with the rest of the dataset:

- Begley, 1990
    - Compares Ergometrine (coded 12) and placebo (coded 1)
- Modi, 2014
    - Compares Oxytocin [>5 IU and ≤ 10 IU] (coded 5), Misoprostol ≤600mcg (coded 8), Ergometrine (coded 12) and Carboprost (coded 13)
    - In this four-armed trial events were rare with only two events observed, both on the carboprost arm, which likely explains the high deviance contribution for this study under the NMA model.

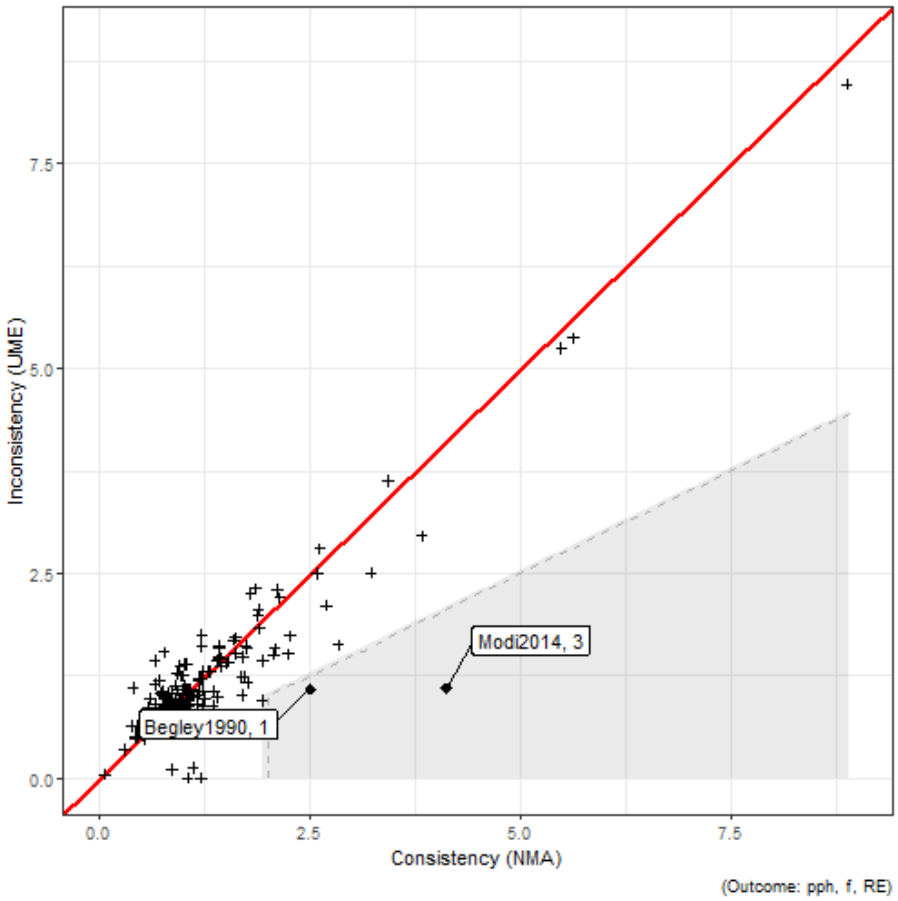

Figure 47. Dev-dev plot of each study arm's residual deviance under the standard and inconsistency models (RE model structure) for PPH >1000ml, full dataset. Labels indicate study arms with high deviance in the NMA model, relative to their deviance in the UME model. The dotted line and grey shaded area denotes where study arms fit poorly in the NMA model but well in the UME model, suggesting that they are predicted poorly when the model enforces consistency in treatment differences.

Checking inconsistency within individual treatment comparisons (Node-splitting)

Node-splitting procedures indicated no evidence of a difference between direct and indirect evidence on most treatment comparisons. Evidence conflicted on five comparisons (Table 50), including treatment comparisons where study arms were identified as inconsistent:

- Placebo and Ergometrine
- Carbetocin and Oxytocin >5 IU and ≤ 10 IU
- Ergometrine and Carboprost
- Carboprost and Oxytocin >5 IU and ≤ 10 IU
- Carboprost and Misoprostol ≤ 600 mcg

Given multiple testing of 33 contrasts, we would expect p-values below a 5% threshold in at least 1 case. Applying a Bonferroni correction suggests that only comparisons between carboprost and misoprostol (≤ 600 mcg) and oxytocin (>5 IU and ≤ 10 IU).

Forest plots for the comparisons where direct evidence conflicts with indirect evidence are presented (Figure 48 and Figure 49). We note that carboprost is only linked to treatments Oxytocin >5 IU and ≤ 10 IU, ergometrine and misoprostol ≤ 600 mcg by a single study (Modi 2014) in which no events were observed on 3 arms, leaving very little evidence with which to reach an estimate of the treatment effect and leading to extremely large treatment differences.

Table 50. Model fit statistics for node-split model (PPH >1000ml, full dataset). Comparisons where there is an indication of inconsistency between direct and indirect estimates (p-values <0.001 following application of a Bonferroni correction) are highlighted in orange.

Comparison	Total Residual Deviance	p-value
Carboprost vs Oxytocin >5 IU and ≤ 10 IU	240.9	<0.001
Carboprost vs Misoprostol ≤ 600 mcg	240.0	<0.001
Placebo vs Ergometrine	246.6	0.029
Carbetocin vs Oxytocin >5 IU and ≤ 10 IU	245.6	0.038
Ergometrine vs Carboprost	240.2	0.015
Placebo vs Oxytocin >1 IU and ≤ 5 IU	244.1	0.741
Placebo vs Oxytocin >5 IU and ≤ 10 IU	246.2	0.922
Placebo vs Ergometrine plus oxytocin	246.8	0.843
Placebo vs Misoprostol ≤ 600 mcg	242.7	0.526
Placebo vs Carboprost	247.8	0.653
Carbetocin vs Oxytocin >1 IU and ≤ 5 IU	246.9	0.863

Carbetocin vs Oxytocin > 10 IU	246.1	0.471
Carbetocin vs Ergometrine plus oxytocin	242.1	0.157
Carbetocin vs Misoprostol ≤ 600 mcg	247.3	0.471
Oxytocin >1 IU and ≤ 5 IU vs Ergometrine plus oxytocin	247.1	0.434
Oxytocin >1 IU and ≤ 5 IU vs Misoprostol ≤ 600 mcg	244.1	0.393
Oxytocin >5 IU and ≤ 10 IU vs Ergometrine plus oxytocin	238.9	0.543
Oxytocin >5 IU and ≤ 10 IU vs Misoprostol ≤ 600 mcg	236.8	0.578
Oxytocin >5 IU and ≤ 10 IU vs Misoprostol >600 mcg and ≤ 800 mcg	248.5	0.779
Oxytocin > 10 IU vs Ergometrine plus oxytocin	247.0	0.256
Oxytocin > 10 IU vs Misoprostol ≤ 600 mcg	247.1	0.870
Oxytocin > 10 IU vs Misoprostol >600 mcg and ≤ 800 mcg	247.6	0.390
Ergometrine plus oxytocin vs Misoprostol ≤ 600 mcg	239.9	0.417
Misoprostol ≤ 600 mcg vs Misoprostol >600 mcg and ≤ 800 mcg	245.7	0.037
Misoprostol plus oxytocin vs Oxytocin >5 IU and ≤ 10 IU	251.4	0.061
Misoprostol plus oxytocin vs Oxytocin > 10 IU	250.3	0.074
Misoprostol plus oxytocin vs Ergometrine plus oxytocin	246.6	0.991
Misoprostol plus oxytocin vs Misoprostol ≤ 600 mcg	244.8	0.067
Ergometrine vs Oxytocin >5 IU and ≤ 10 IU	239.9	0.082
Ergometrine vs Misoprostol ≤ 600 mcg	242.8	0.054
Ergometrine vs Misoprostol >600 mcg and ≤ 800 mcg	245.7	0.508
Carboprost vs Oxytocin >1 IU and ≤ 5 IU	246.8	0.715
Carboprost vs Ergometrine plus oxytocin	245.7	0.250
NMA (no nodes split)	**246.6**	**-**

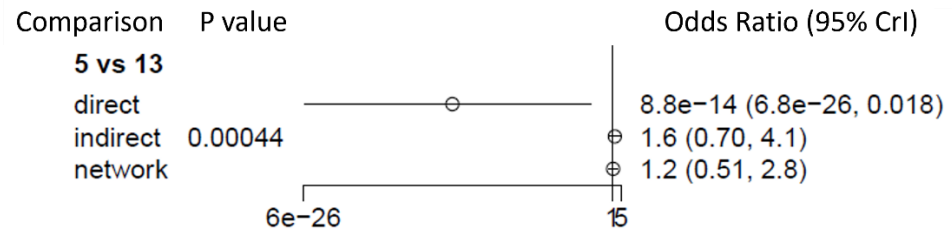

Figure 48. Treatment effect estimates separated by direct and indirect evidence: Oxytocin >5 IU and ≤ 10 IU (5) vs carboprost (13).

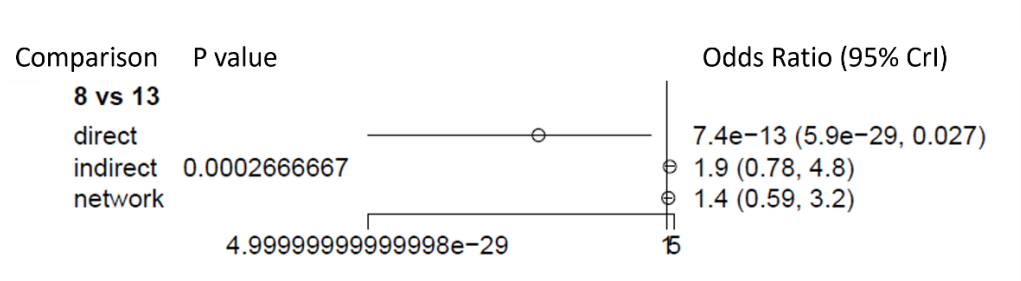

Figure 49. Treatment effect estimates separated by direct and indirect evidence: Misoprostol ≤ 600 mcg (8) vs carboprost (13).

Vaginal birth subgroup

Global inconsistency check

Analysis of the dataset for the vaginal birth subgroup for the outcome post-partum haemorrhage (>1000ml) included 71 studies (157 arms) of 13 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Fitting of both fixed and random-effect network meta-analysis (NMA) models indicated support for the random-effect model on the basis of a small decrease in DIC and a sizeable decrease in residual deviance (Table 51).

Total residual deviance and DIC were lower in the inconsistency UME model than in the NMA model and the estimate of between-study SD was the similar in the NMA and UME models (Table 51). This suggests no evidence of inconsistency, but that there is moderate heterogeneity between study estimates.

Table 51. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome PPH >1000ml, vaginal birth subgroup.

Outcome	Pop.	Model	Posterior total residual deviance[3]	Between-study SD Mean, 95% credible interval	pD	DIC[4]
PPH	VD	FE NMA	205.9	-	79.5	798.4
PPH	VD	RE NMA	190.5	0.20 (0.02, 0.45)	91.6	795.1
PPH	VD	FE UME	190.8	-	90.5	794.2
PPH	VD	RE UME	178.9	0.21 (0.02, 0.47)	100.1	791.9

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the random effects UME and NMA models (Figure 50), shows that two studies showed inconsistency with the rest of the dataset:

- Begley, 1990
  - Compares Ergometrine (coded 12) and placebo (coded 1)
- Modi, 2014
  - Compares Oxytocin [>5 IU and ≤ 10 IU] (coded 5), Misoprostol ≤600mcg (coded 8), Ergometrine (coded 12) and Carboprost (coded 13)
  - In this four-armed trial events were rare with only two events observed, both on the carboprost arm, which likely explains the high deviance contribution for this study under the NMA model.

[3] Posterior mean residual deviance compared to 157 data points
[4] Deviance information criteria (DIC) – lower values preferred

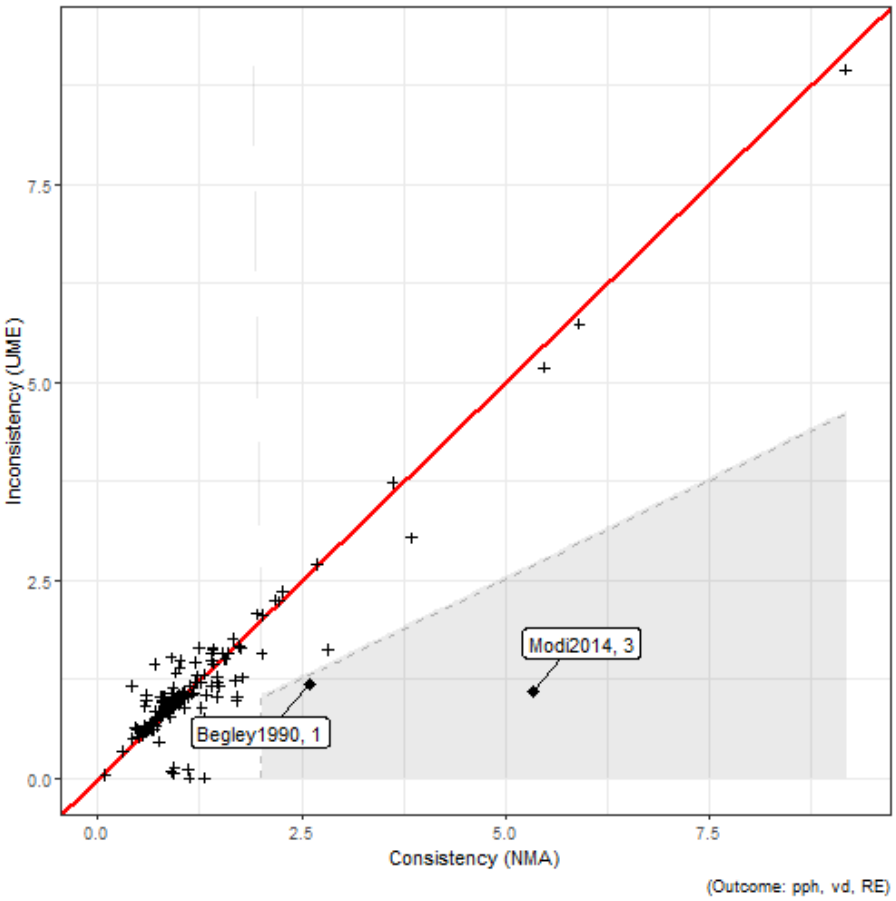

Figure 50. Dev-dev plot of each study arm's residual deviance under the consistency (NMA) and inconsistency models with RE structure for PPH >1000ml, vaginal birth subgroup. Labels indicate study arms with high deviance in the NMA model, relative to their deviance in the UME model. The dotted line and grey shaded area outline a region where study arms fit poorly in the NMA model but well in the UME model, suggesting that they are predicted poorly when the model enforces consistency in treatment differences.

Checking inconsistency within individual treatment comparisons (Node-splitting)

Node-splitting procedures indicate that direct and indirect evidence on most treatment comparisons agree. Evidence conflicted on 1 comparison following Bonferroni correction: Carboprost vs Misoprostol (<600mcg) (Table 52). The direct evidence for this comparison conflicts with indirect evidence (Figure 51), though the direct evidence is weak, being drawn from a single study (Modi 2014) in which no events were observed on 3 arms.

Table 52. Model fit statistics for node-split model (PPH >1000ml, vaginal birth subgroup). Comparisons where there is an indication of inconsistency between direct and indirect estimates (p-values <0.002 [p<0.05 following Bonferroni correction for 29 comparisons]) are highlighted in orange.

Comparison	Total Residual Deviance	p-value
CarbProst vs Mis_b600	186.4	<0.001

Plac vs Erg	189.0	0.019
Plac vs CarbProst	191.5	0.880
Erg vs CarbProst	187.4	0.004
Plac vs Oxy_a1b5	189.8	0.790
CarbProst vs Oxy_a1b5	190.2	0.997
Carb vs Oxy_a1b5	185.7	0.088
Plac vs Oxy_a5b10	191.2	0.928
Erg vs Oxy_a5b10	190.4	0.093
CarbProst vs Oxy_a5b10	189.2	0.002
Carb vs Oxy_a5b10	188.0	0.263
Mis_Oxy vs Oxy_a5b10	182.7	0.988
Mis_Oxy vs Oxy_a10	182.0	0.939
Plac vs Erg_Oxy	186.2	0.801
Oxy_a1b5 vs Erg_Oxy	188.7	0.035
Mis_Oxy vs Erg_Oxy	190.6	0.607
CarbProst vs Erg_Oxy	182.0	0.070
Carb vs Erg_Oxy	187.2	0.409
Oxy_a5b10 vs Erg_Oxy	182.3	0.539
Plac vs Mis_b600	189.0	0.528
Mis_Oxy vs Mis_b600	188.9	0.224
Erg vs Mis_b600	188.7	0.055
Oxy_a1b5 vs Mis_b600	185.9	0.410
Erg_Oxy vs Mis_b600	184.6	0.385
Oxy_a5b10 vs Mis_b600	179.1	0.872
Oxy_a10 vs Mis_b600	190.7	0.903
Erg vs Mis_a600b800	190.5	0.592
Oxy_a5b10 vs Mis_a600b800	184.0	0.280
Mis_b600 vs Mis_a600b800	190.0	0.041
NMA (no nodes split)	**190.5**	**-**

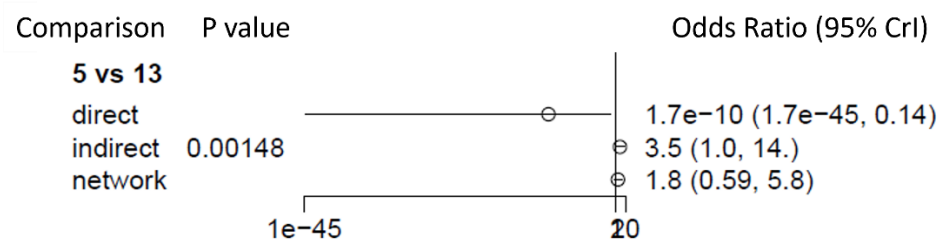

Figure 51. Treatment effect estimates separated by direct and indirect evidence: Misoprostol ≤600mcg (5) vs Carboprost (13) for the vaginal birth subgroup.

Caesarean Section birth subgroup

Global inconsistency check

Analysis of the dataset for the CS birth subgroup for the outcome post-partum haemorrhage (>1000ml) included 26 studies (53 arms) of 8 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Fitting both fixed (FE) and random-effect (RE) network meta-analysis (NMA) models gave similar values for DIC. However, residual deviance was lower by 5.3 for the RE NMA (Table 53) suggesting support for the RE model for these data.

Total residual deviance and DIC were lower in the inconsistency UME model than in the NMA model and the estimate of between-study SD was the similar in the NMA and UME models (Table 53). This suggests that there is no evidence of inconsistency, but moderate heterogeneity between study estimates.

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the random effects UME and NMA models (Figure 52), showed no study arms to have high deviance in the NMA model, relative to their deviance in the UME model. Taken together, the model fit and dev-dev plots suggest there was little evidence of inconsistency in these data. Therefore, node-splitting models were not required.

Table 53. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome PPH >1000ml, CS birth subgroup.

Outcome	Pop.	Model	Posterior total residual deviance[5]	Between-study SD Mean, 95% credible interval	pD	DIC[6]
PPH	CS	FE NMA	61.6	-	32.6	276.6
PPH	CS	RE NMA	56.3	0.34 (0.03, 0.81)	38.1	276.9
PPH	CS	FE UME	56.7	-	36.6	275.7
PPH	CS	RE UME	55.4	0.26 (0.01, 0.79)	39.5	277.3

[5] Posterior mean residual deviance compared to 53 data points
[6] Deviance information criteria (DIC) – lower values preferred

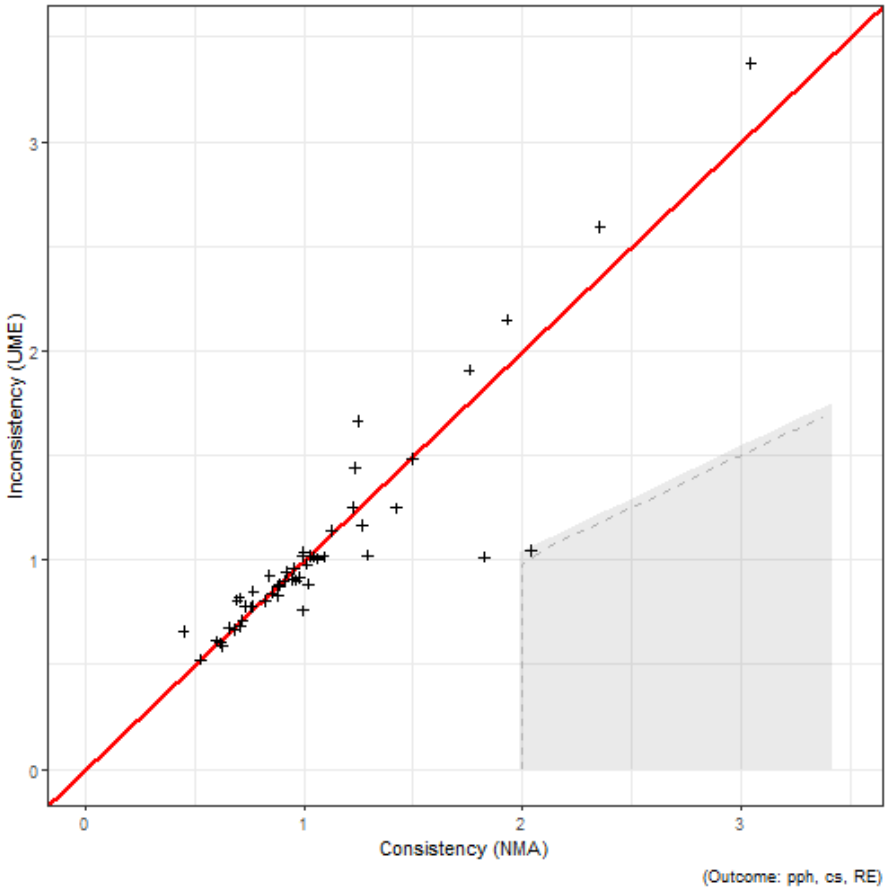

Figure 52. Dev-dev plot of each study arm's residual deviance under the consistency (NMA) and inconsistency models with RE structure for PPH >1000ml, CS birth subgroup. No study arms were identified as inconsistent (i.e. points in the shaded region).

Outcome: Additional uterotonics

Summary

We identified strong heterogeneity in study estimates for this outcome, but little evidence of inconsistency in both the full population and either of the mode-of-delivery subgroups. Estimates from two study arms, Supe 2016 (carboprost arm) and Maged 2020 (carbetocin arm), which were present in the full dataset and the vaginal birth subgroup datasets, were found to have a poor fit to the NMA model. Further investigation using node-splitting showed that although there was some evidence of inconsistency on the carboprost-carbetocin-misoprostol (>600mcg and <800mcg)-placebo loop, this was likely driven by a lack of data and hence very imprecise estimates. Global inconsistency tests detected no inconsistency in the CS birth subgroup.

Full dataset

Global inconsistency check

Analysis of the full dataset for the outcome additional uterotonics included 161 studies (345 arms) of 14 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Fitting of both fixed and random-effect network meta-analysis (NMA) models indicated strong support for the random-effect model on the basis of large reductions in DIC and residual deviance (Table 54**Error! Reference source not found.**). Between-study SD was estimated to be 0.83 (95% CrI 0.71, 0.98), which is large on the odds ratio scale. Modelling treatment differences with a random effects structure results in good model fit, with the total residual deviance equivalent to the number of study arms. Total residual deviance was slightly lower in the inconsistency UME model than in the NMA model, however DIC was lower for the NMA model and the estimate of between-study SD was similarly large in both NMA and UME models. This suggests that there is no evidence of inconsistency but there is evidence of substantial heterogeneity between study estimates.

Table 54. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome additional uterotonics, full dataset.

Outcome	Pop.	Model	Posterior total residual deviance[7]	Between-study SD Mean, 95% credible interval	pD	DIC[8]
Uterotonics	Full	FE NMA	1162.0	-	173.1	2715.7
Uterotonics	Full	RE NMA	366.5	0.83 (0.71, 0.98)	288.3	2035.7
Uterotonics	Full	FE UME	1035.0	-	198.0	2614.0
Uterotonics	Full	RE UME	360.4	0.91 (0.76, 1.08)	298.4	2039.6

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the random effects UME and NMA models (Figure 53), shows that two studies showed inconsistency with the rest of the dataset:

- Supe 2016
    - Compares misoprostol (>600mcg and <800mcg) (coded 10), ergometrine (coded 13), carboprost (coded 14) and placebo (coded 1)

[7] Posterior mean residual deviance compared to 345 total data points
[8] Deviance information criteria (DIC) – lower values preferred

- This trial reports relatively small numbers of events on each arm (with a total of 12 events in 200 participants across all four arms)

- Maged 2020
  - Compares carbetocin (coded 2) and ergometrine plus oxytocin (coded 8)

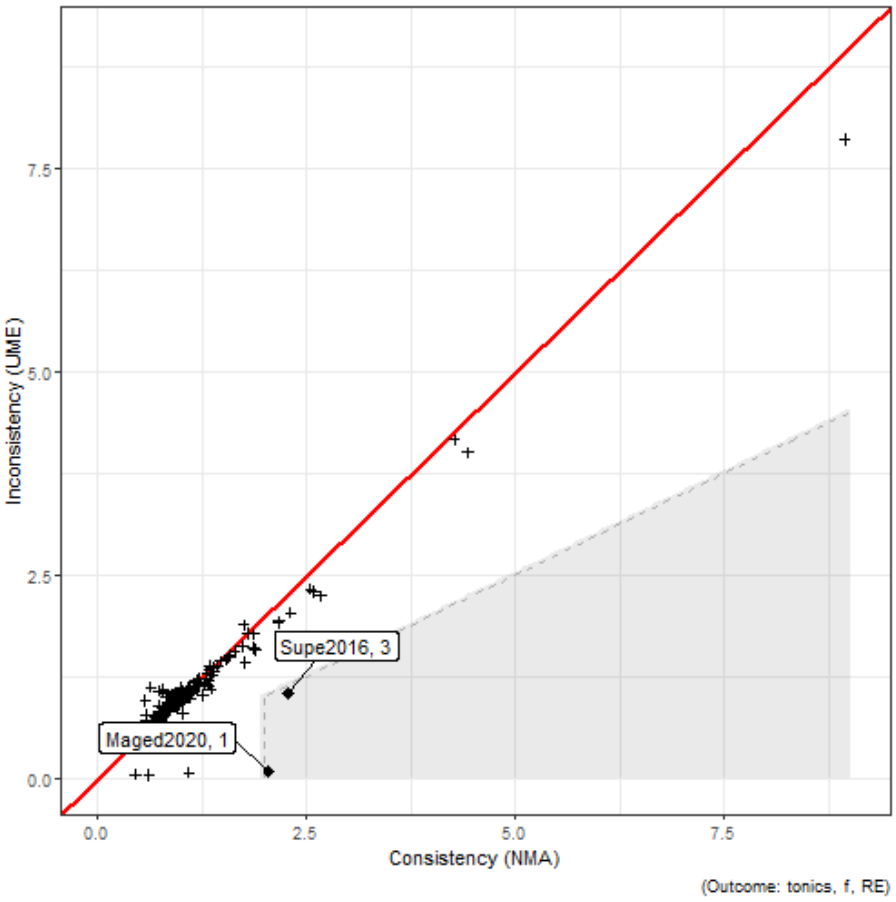

Figure 53. Dev-dev plot of each study arm's residual deviance under the standard and inconsistency models (RE model structure) for additional uterotonics, full dataset. Labels indicate study arms with high deviance in the NMA model, relative to their deviance in the UME model. The dotted line and grey shaded area denotes where study arms fit poorly in the NMA model but well in the UME model, suggesting that they are predicted poorly when the model enforces consistency in treatment differences.

Checking inconsistency within individual treatment comparisons (Node-splitting)

Node-splitting procedures indicate that direct and indirect evidence on most treatment comparisons agree. Applying a Bonferroni correction suggests that there are no comparisons where direct and indirect evidence is inconsistent, given 42 comparisons (Table 55). However, we note a p-value of 0.003 when comparing estimates from direct and indirect evidence for the comparison between misoprostol (>6000mcg and <8000mcg) and carbetocin and present the forest plot for this comparison (Figure 54), as well as for the comparison between misoprostol (>6000mcg and <8000mcg) and carboprost.

Table 55. Model fit statistics for node-split model (additional uterotonics, full dataset). No comparisons had p<0.0012 (Bonferroni correction of p <0.05 given 42 tests) when testing consistency between estimates from direct and indirect evidence.

Comparison	Total Residual Deviance	p-value
Mis_a600b800 vs Carb	365.3	0.003
Plac vs Mis_a600b800	365.8	0.831
Plac vs Erg	367.5	0.340
Plac vs CarbProst	366.6	0.271
Plac vs Carb	366.3	0.184
Plac vs Oxy_a1b5	363.1	0.758
Plac vs Oxy_a5b10	366.7	0.704
Plac vs Erg_Oxy	367.0	0.458
Plac vs Mis_b600	365.2	0.270
Mis_a600b800 vs Erg	366.7	0.955
Mis_a600b800 vs CarbProst	366.9	0.055
Mis_a600b800 vs Oxy_a1b5	366.8	0.988
Mis_a600b800 vs Oxy_a5b10	367.0	0.509
Mis_a600b800 vs Oxy_a10	366.9	0.981
Mis_a600b800 vs Mis_b600	364.1	0.720
Mis_Oxy vs Carb	366.2	0.526
Mis_Oxy vs Oxy_a1b5	366.3	0.866
Mis_Oxy vs Oxy_a5b10	366.3	0.465

Mis_Oxy vs Oxy_a10	365.1	0.323
Mis_Oxy vs Erg_Oxy	366.8	0.501
Mis_Oxy vs Mis_b600	365.2	0.793
Erg vs CarbProst	362.5	0.081
Erg vs Oxy_a1b5	366.4	0.162
Erg vs Oxy_a5b10	362.3	0.963
Erg vs Mis_b600	362.0	0.519
CarbProst vs Oxy_a1b5	364.5	0.567
CarbProst vs Oxy_a5b10	365.1	0.346
CarbProst vs Erg_Oxy	367.3	0.863
CarbProst vs Mis_b600	362.7	0.274
Carb vs Oxy_a1b5	364.7	0.422
Carb vs Oxy_a5b10	365.3	0.886
Carb vs Oxy_a10	364.1	0.643
Carb vs Erg_Oxy	365.3	0.607
Carb vs Mis_b600	364.9	0.644
Oxy_a1b5 vs Oxy_a10	365.0	0.318
Oxy_a1b5 vs Erg_Oxy	366.9	0.944
Oxy_a1b5 vs Mis_b600	364.7	0.489
Oxy_a5b10 vs Erg_Oxy	364.3	0.185
Oxy_a5b10 vs Mis_b600	360.2	0.459
Oxy_a10 vs Erg_Oxy	366.7	0.776
Oxy_a10 vs Mis_b600	364.1	0.854
Erg_Oxy vs Mis_b600	365.6	0.606
NMA (no nodes split)	366.7	**-**

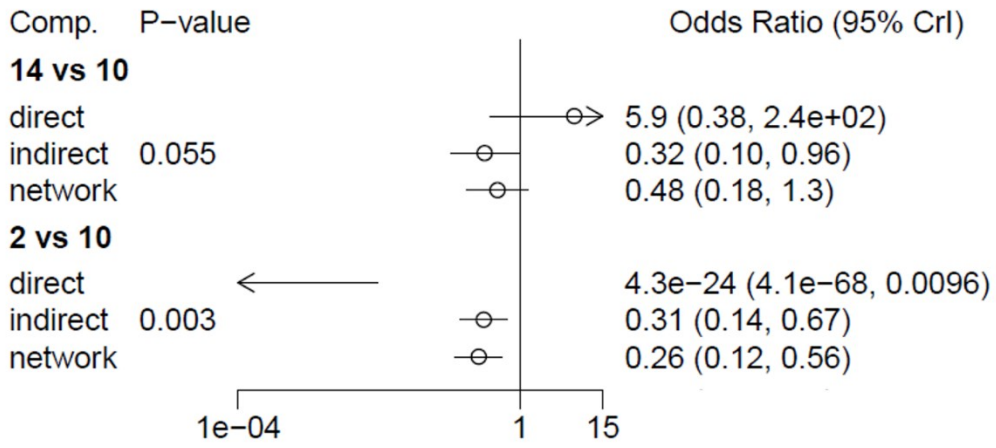

Figure 54. Treatment effect estimates separated by direct and indirect evidence for two comparisons: misoprostol (>600mcg and <800mcg) (10) vs carboprost (14) and misoprostol (>600mcg and <800mcg) (10) vs carbetocin (2).

Vaginal birth subgroup

Global inconsistency check

Analysis of the dataset for the vaginal birth subgroup for the outcome additional uterotonics included 109 studies (236 arms) of 12 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Comparing fixed and random-effect network meta-analysis (NMA) models indicated strong support for the random-effect model on the basis of large decreases in DIC and residual deviance (Table 56). Between-study SD was estimated to be large at 0.73 (95% CrI 0.58, 0.90), suggesting that there is heterogeneity in study estimates of treatment effect for the same treatment comparison. There was no reduction in DIC or between studies SD for the RE UME relative to the RE NMA, however there was an improvement in overall fit (residual deviance) for the RE UME suggesting there may be evidence of inconsistency, which we explore further below.

Table 56. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome additional uterotonics, vaginal birth subgroup.

Outcome	Pop.	Model	Posterior total residual deviance[9]	Between-study SD Mean, 95% credible interval	pD	DIC[10]
Uterotonics	VD	FE NMA	682.1	-	119.1	1768.3
Uterotonics	VD	RE NMA	254.0	0.73 (0.58, 0.90)	193.3	1414.4
Uterotonics	VD	FE UME	579.2	-	137.4	1683.8
Uterotonics	VD	RE UME	249.3	0.74 (0.57, 0.94)	198.8	1415.3

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the random effects UME and NMA models (Figure 55), shows that two studies showed inconsistency with the rest of the dataset:

- Supe 2016
    - Compares misoprostol a600b800 (coded 10), ergometrine (coded 13), carboprost (coded 14) and placebo (coded 1), with the carboprost arm specifically labelled as being inconsistent by residual deviance

[9] Posterior mean residual deviance compared to 236 data points
[10] Deviance information criteria (DIC) – lower values preferred

- This trial reports small numbers of events on each arm (12 events in 200 participants)

- Maged 2020
  - Compares carbetocin (coded 2) and ergometrine plus oxytocin (coded 8), with the carbetocin arm specifically labelled as being inconsistent by residual deviance

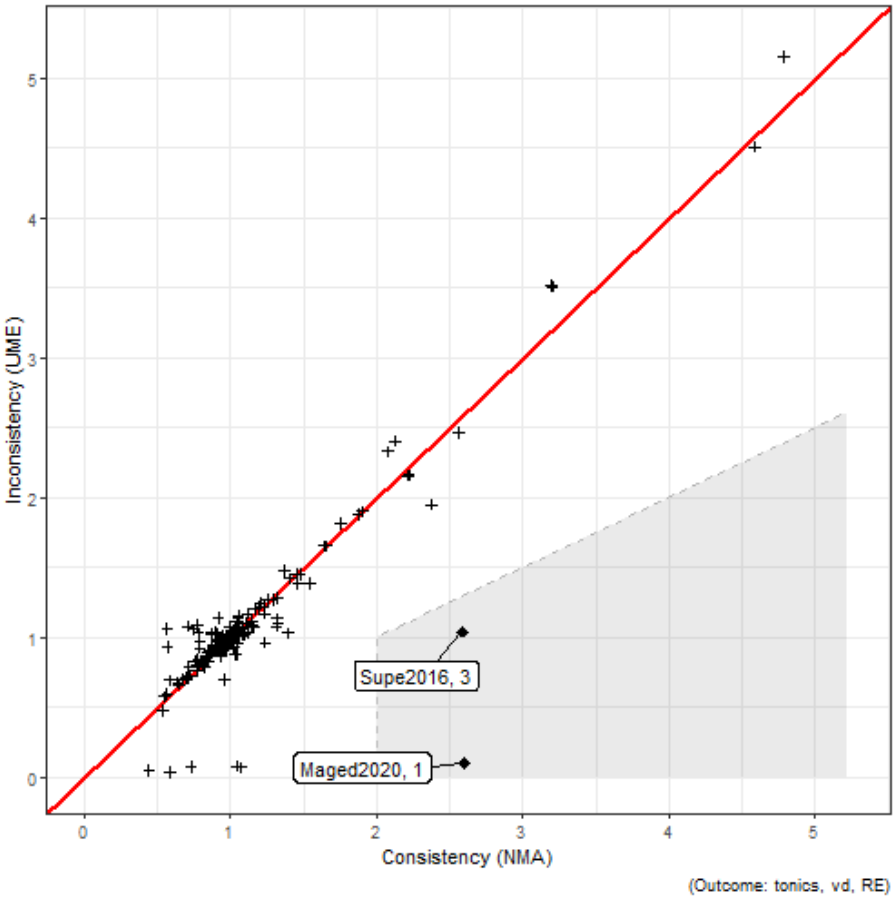

Figure 55. Dev-dev plot of each study arm's residual deviance under the consistency (NMA) and inconsistency models with RE structure for additional uterotonics, vaginal birth subgroup. Labels indicate study arms with high deviance in the NMA model, relative to their deviance in the UME model. The dotted line and grey shaded area outline a region where study arms fit poorly in the NMA model but well in the UME model, suggesting that they are predicted poorly when the model enforces consistency in treatment differences.

Checking inconsistency within individual treatment comparisons (Node-splitting)

Node-splitting procedures indicate that direct and indirect evidence on most treatment comparisons agree. Applying a Bonferroni correction suggests that there are no comparisons where direct and indirect evidence is inconsistent, given 36 comparisons (Table 57). We note some small p-values, e.g. for the comparison between misoprostol (>6000mcg and <8000mcg) and carbetocin and between oxytocin (<10) and ergometrine with oxytocin. We present the forest plot for these comparisons (Figure 56) but the inconsistency is likely to be the result of weak evidence meaning that effect estimates are imprecisely estimated for these comparisons (as can be seen in the credible intervals).

Table 57. Model fit statistics for node-split model (additional uterotonics, vaginal birth subgroup). No comparisons had p<0.0014 (Bonferroni correction of p <0.05 given 36 tests) when testing consistency between direct and indirect evidence.

Comparison	Total Residual Deviance	p-value
Carb vs Mis_a600b800	252.1	0.003
Plac vs Erg	254.8	0.157
Plac vs CarbProst	253.1	0.263
Plac vs Oxy_a1b5	252.4	0.667
Plac vs Oxy_a5b10	253.5	0.745
Plac vs Erg_Oxy	254.4	0.235
Plac vs Mis_b600	252.9	0.353
Plac vs Mis_a600b800	253.0	0.988
Mis_Oxy vs Oxy_a5b10	252.8	0.110
Mis_Oxy vs Oxy_a10	253.4	0.140
Mis_Oxy vs Erg_Oxy	253.6	0.707
Mis_Oxy vs Mis_b600	253.0	0.644
Erg vs CarbProst	249.2	0.046
Erg vs Oxy_a1b5	253.1	0.069
Erg vs Oxy_a5b10	248.6	0.673
Erg vs Mis_b600	249.0	0.553
Erg vs Mis_a600b800	253.2	0.996
CarbProst vs Oxy_a1b5	251.0	0.472
CarbProst vs Oxy_a5b10	252.0	0.211

CarbProst vs Erg_Oxy	254.5	0.819
CarbProst vs Mis_b600	249.5	0.219
CarbProst vs Mis_a600b800	253.6	0.041
Carb vs Oxy_a1b5	254.0	0.714
Carb vs Oxy_a5b10	252.9	0.270
Carb vs Oxy_a10	253.5	0.036
Carb vs Erg_Oxy	251.9	0.774
Carb vs Mis_b600	253.5	0.588
Oxy_a1b5 vs Mis_b600	252.6	0.698
Oxy_a1b5 vs Mis_a600b800	254.2	0.738
Oxy_a5b10 vs Erg_Oxy	250.5	0.217
Oxy_a5b10 vs Mis_b600	247.3	0.788
Oxy_a5b10 vs Mis_a600b800	254.2	0.306
Oxy_a10 vs Erg_Oxy	253.1	0.018
Oxy_a10 vs Mis_b600	254.2	0.483
Erg_Oxy vs Mis_b600	252.7	0.588
Mis_b600 vs Mis_a600b800	251.2	0.648
NMA (no nodes split)	**253.6**	**-**

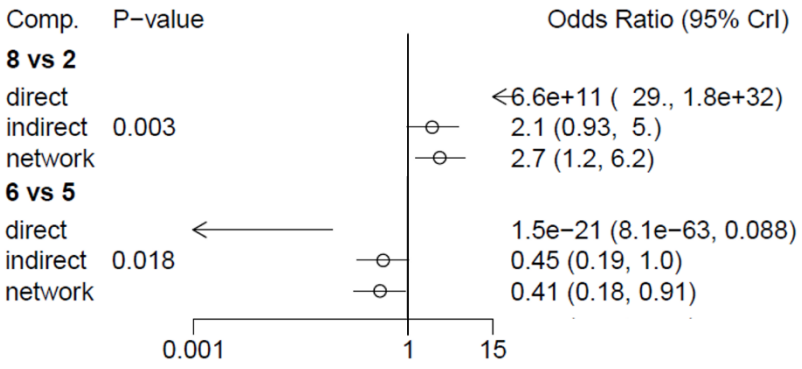

Figure 56. Treatment effect estimates separated by direct and indirect evidence for two comparisons in the vaginal birth subgroup: Misoprostol (>600mcg and < 800mcg) (8) vs Carbetocin (2) and Oxytocin (<10) (5) vs Ergometrine plus Oxytocin (6).

Caesarean Section birth subgroup

Global inconsistency check

Analysis of the dataset for the CS birth subgroup for the outcome additional uterotonics included 51 studies (107 arms) of 12 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Fitting of both fixed (FE) and random-effect (RE) network meta-analysis (NMA) models indicated strong support for the random-effect model on the basis of large reductions in DIC and residual deviance (Table 57). There was no reduction in DIC when the RE UME model was fitted, relative to the RE NMA, suggesting that direct and indirect evidence is consistent.

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the random effects UME and NMA models (Figure 57), showed no study arms to have high deviance in the NMA model, relative to their deviance in the UME model. Taken together, the model fit and dev-dev plots suggest there is no evidence of inconsistency, and so node-splitting models were not required.

Table 58. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome additional uterotonics, CS birth subgroup.

Outcome	Pop.	Model	Posterior total residual deviance[11]	Between-study SD Mean, 95% credible interval	pD	DIC[12]
Uterotonics	CS	FE NMA	315.7	-	60.9	788.3
Uterotonics	CS	RE NMA	111.5	1.03 (0.76, 1.39)	94.7	617.9
Uterotonics	CS	FE UME	305.8	-	70.9	788.3
Uterotonics	CS	RE UME	110.8	1.20 (0.85, 1.68)	98.0	620.6

[11] Posterior mean residual deviance compared to 107 data points
[12] Deviance information criteria (DIC) – lower values preferred

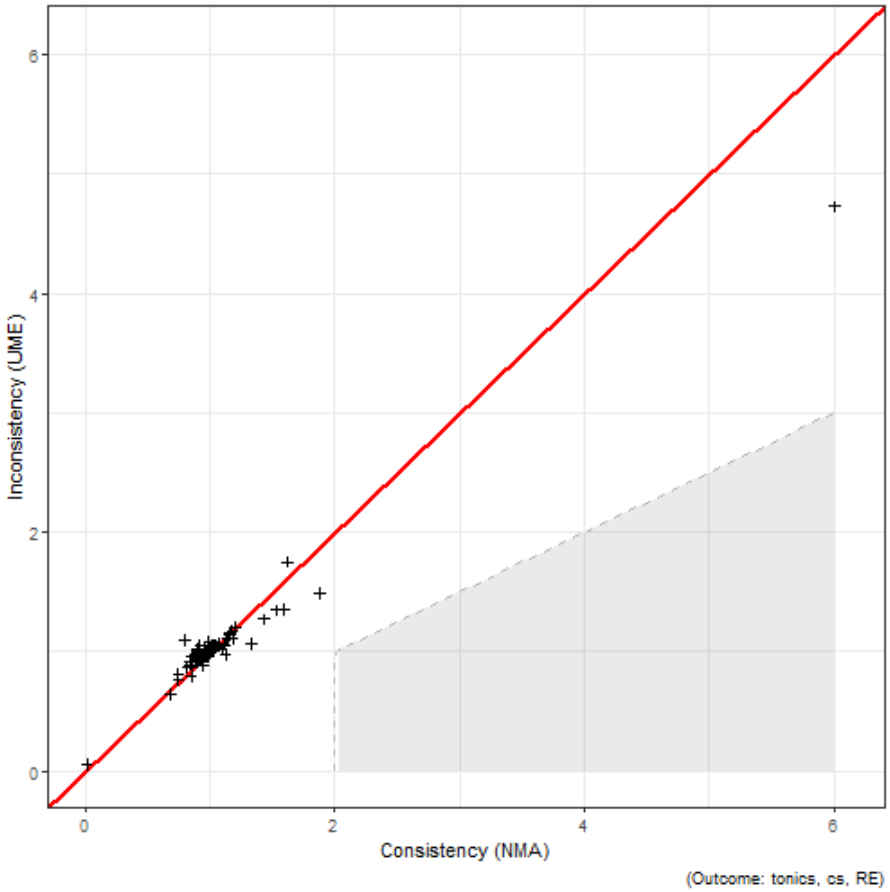

Figure 57. Dev-dev plot of each study arm's residual deviance under the consistency (NMA) and inconsistency models with RE structure for additional uterotonics, CS birth subgroup. No study arms were identified as inconsistent (i.e. there were no points in the shaded region).

Outcome: Blood transfusion

Summary

We identified strong heterogeneity in studies of this outcome that was adequately captured by the random-effects network meta-analysis. There was little evidence of inconsistency between direct and indirect evidence in the full population or either of the mode-of-delivery subgroups. One study in the full dataset (Modi et al. 2014) showed moderate inconsistency on the carboprost arm with indirect evidence from the network. Node-splitting of the full dataset suggests that there may be the potential for inconsistency on the oxytocin (10IU)-misoprostol (>600mcg and <800mcg)-placebo loop. However, this is likely to be the result of weak direct evidence: transfusion being a rare event in this population. Global inconsistency checks support that node-splitting models were not required for the vaginal birth and CS birth subgroups.

Full dataset

Global inconsistency check

Analysis of the full dataset for the outcome Transfusion included 113 studies (242 arms) of 13 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Fitting of both fixed and random-effect network meta-analysis (NMA) models indicated strong support for the random-effect model on the basis of large reductions in DIC and residual deviance (Table 59**Error! Reference source not found.**). Between-study SD was estimated to be 0.74 (95% credible interval [CrI] 0.51, 1.02), which is large on the odds ratio scale. Modelling treatment differences with a random effects structure results in good model fit, with the total residual deviance equivalent to the number of study arms. Total residual deviance was slightly lower in the inconsistency UME model than in the NMA model, however DIC was lower for the NMA model and the estimate of between-study SD was similarly large in both NMA and UME models. This suggests that there is evidence of substantial heterogeneity between study estimates, but no evidence of inconsistency.

Table 59. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome Transfusion, full dataset.

Outcome	Pop.	Model	Posterior total residual deviance[13]	Between-study SD Mean, 95% credible interval	pD	DIC[14]
Transfusion	Full	FE NMA	381.5	-	120.0	1142.9
Transfusion	Full	RE NMA	270.1	0.74 (0.51, 1.02)	163.1	1074.6
Transfusion	Full	FE UME	344.7	-	138.8	1124.9
Transfusion	Full	RE UME	268.1	0.75 (0.45, 1.11)	172.5	1082.0

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the random effects UME and NMA models (Figure 58), shows that one study showed inconsistency with the rest of the dataset: Modi 2014. Modi 2014 compares oxytocin [>5 IU and ≤ 10 IU] (coded 4), misoprostol ≤600mcg (coded 8), ergometrine (coded 12) and carboprost (coded 13). In this four-armed trial events were rare with only two events observed, both on the carboprost arm, which likely explains the high deviance contribution for this study under the NMA model.

[13] Posterior mean residual deviance compared to 242 total data points
[14] Deviance information criteria (DIC) – lower values preferred

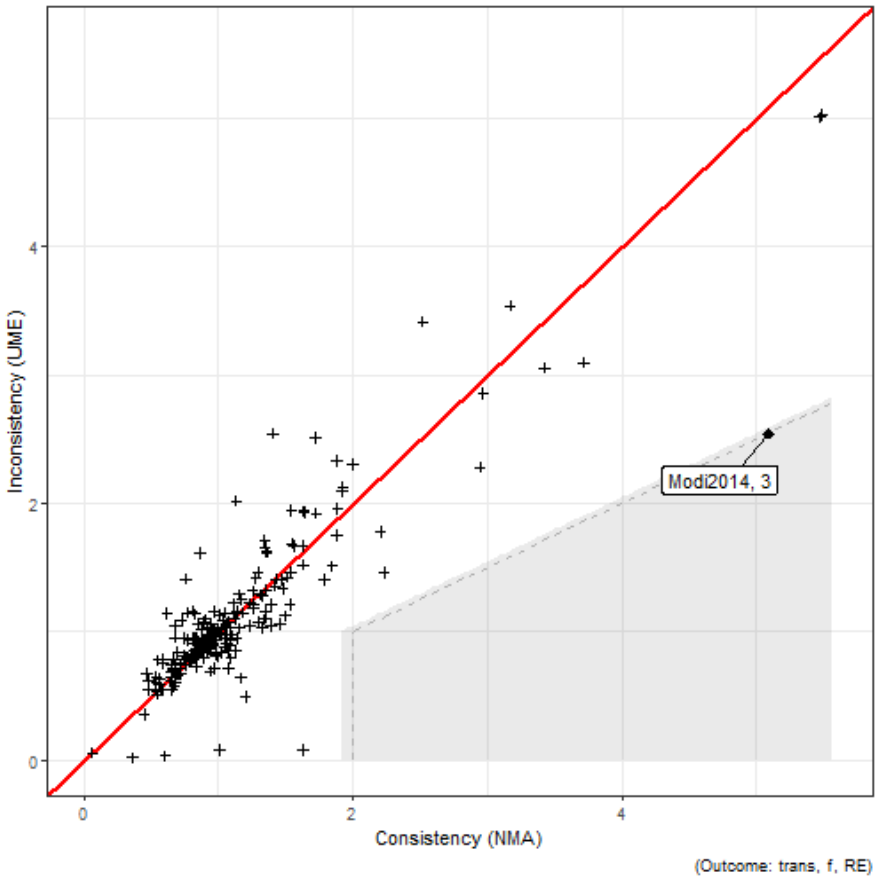

Figure 58. Dev-dev plot of each study arm's residual deviance under the standard and inconsistency models (RE model structure) for Transfusion, full dataset. Labels indicate study arms with high deviance in the NMA model, relative to their deviance in the UME model. The dotted line and grey shaded area denotes where study arms fit poorly in the NMA model but well in the UME model, suggesting that they are predicted poorly when the model enforces consistency in treatment differences.

Checking inconsistency within individual treatment comparisons (Node-splitting)

Node-splitting procedures indicate that direct and indirect evidence on most treatment comparisons agree. Applying a Bonferroni correction suggests that there are no comparisons where direct and indirect evidence is inconsistent, given 42 comparisons (Table 60). However, we note small p-values arising from comparisons of direct and indirect evidence for the treatment effect estimates between misoprostol (>6000mcg and <8000mcg) and placebo, misoprostol (>6000mcg and <8000mcg) and oxytocin (<10IU) and ergometrine and oxytocin (>1IU and <5IU). These are the result of very imprecise estimates from the direct evidence: for example, although misoprostol (>6000mcg and <8000mcg) was used in seven studies, there were either zero or only one person transfused in six of these.

Table 60. Model fit statistics for node-split model (Transfusion, full dataset). No comparisons where were indicated to show inconsistency between direct and indirect estimates (p-values <0.0014 [p<0.05 following Bonferroni correction for 35 comparisons]).

Comparison	Total Residual Deviance	p-value
Erg vs Oxy_a1b5	266.5	0.003
Oxy_a10 vs Mis_a600b800	267.8	0.008
Plac vs Mis_a600b800	267.7	0.018
Plac vs Erg	270.3	0.430
Plac vs Oxy_a1b5	269.0	0.390
Plac vs Oxy_a5b10	269.9	0.671
Plac vs Erg_Oxy	270.1	0.620
Plac vs Mis_b600	268.3	0.520
Mis_Oxy vs Carb	270.6	0.245
Mis_Oxy vs Oxy_a1b5	269.3	0.716
Mis_Oxy vs Oxy_a5b10	273.2	0.042
Mis_Oxy vs Oxy_a10	275.3	0.062
Mis_Oxy vs Erg_Oxy	270.2	0.648
Mis_Oxy vs Mis_b600	268.7	0.552
Erg vs CarbProst	263.1	0.245
Erg vs Oxy_a5b10	266.4	0.468
Erg vs Mis_b600	265.5	0.283
Erg vs Mis_a600b800	268.7	0.743
CarbProst vs Oxy_a5b10	266.0	0.304
CarbProst vs Erg_Oxy	270.2	0.986
CarbProst vs Mis_b600	264.8	0.302
Carb vs Oxy_a1b5	269.2	0.559
Carb vs Oxy_a5b10	265.9	0.627
Carb vs Oxy_a10	271.3	0.508
Carb vs Erg_Oxy	266.7	0.109
Carb vs Mis_b600	269.1	0.838

Oxy_a1b5 vs Mis_b600	262.0	0.511
Oxy_a1b5 vs Mis_a600b800	270.8	0.252
Oxy_a5b10 vs Erg_Oxy	265.4	0.282
Oxy_a5b10 vs Mis_b600	264.3	0.110
Oxy_a5b10 vs Mis_a600b800	271.0	0.235
Oxy_a10 vs Erg_Oxy	270.4	0.546
Oxy_a10 vs Mis_b600	269.7	0.700
Erg_Oxy vs Mis_b600	266.7	0.787
Mis_b600 vs Mis_a600b800	267.5	0.246
NMA (no nodes split)	**270.4**	**-**

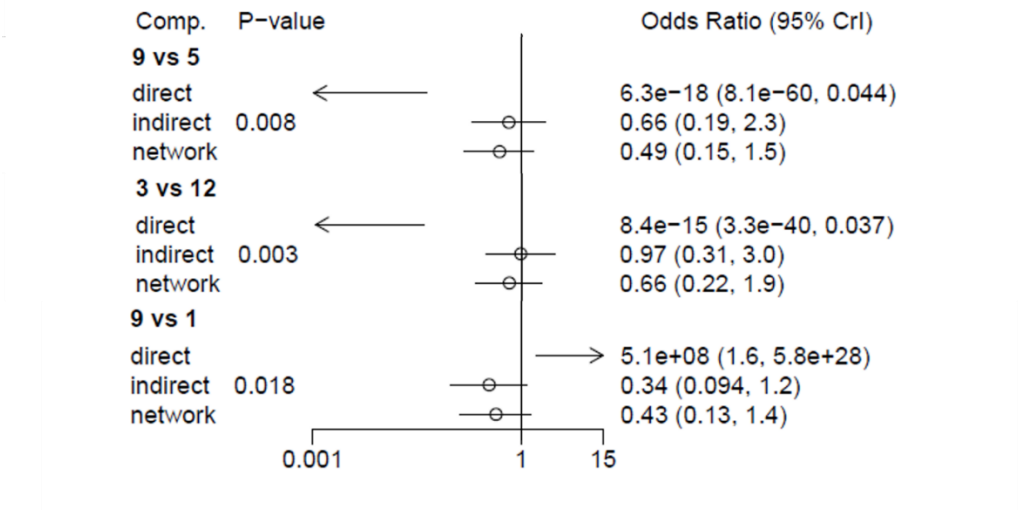

Figure 59. Treatment effect estimates separated by direct and indirect evidence for three comparisons: misoprostol (>600mcg and <800mcg) (9) vs oxytocin (5); oxytocin (>1IU and <5IU) (3) vs ergometrine (12) and misoprostol (>600mcg and <800mcg) (9) vs placebo (1)

Vaginal birth subgroup

Global inconsistency check

Analysis of the dataset for the vaginal birth subgroup for the outcome transfusion included 80 studies (175 arms) of 12 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Fitting of both fixed and random-effect network meta-analysis (NMA) models indicated strong support for the random-effect model on the basis of large decreases in DIC and residual deviance (Table 61). Between-study SD was estimated to be large at 0.53 (95% CrI 0.25, 0.84), suggesting that there is heterogeneity in studies' estimates of treatment effect for the same treatment comparison. There was no reduction in DIC when the RE UME model was fitted, relative to the RE NMA, suggesting no evidence of inconsistency.

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the random effects UME and NMA models (Figure 60), showed no study arms to have high deviance in the NMA model, relative to their deviance in the UME model. Taken together, the model fit and dev-dev plots suggest there was little evidence of inconsistency so node-splitting models were not required.

Table 61. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome Transfusion, vaginal birth subgroup.

Outcome	Pop.	Model	Posterior total residual deviance[15]	Between-study SD Mean, 95% credible interval	pD	DIC[16]
Transfusion	VD	FE NMA	243.9	-	86.9	803.0
Transfusion	VD	RE NMA	200.4	0.53 (0.25, 0.84)	110.7	782.7
Transfusion	VD	FE UME	228	-	102.8	803.0
Transfusion	VD	RE UME	196.5	0.54 (0.21, 0.93)	121.1	789.7

[15] Posterior mean residual deviance compared to 175 data points
[16] Deviance information criteria (DIC) – lower values preferred

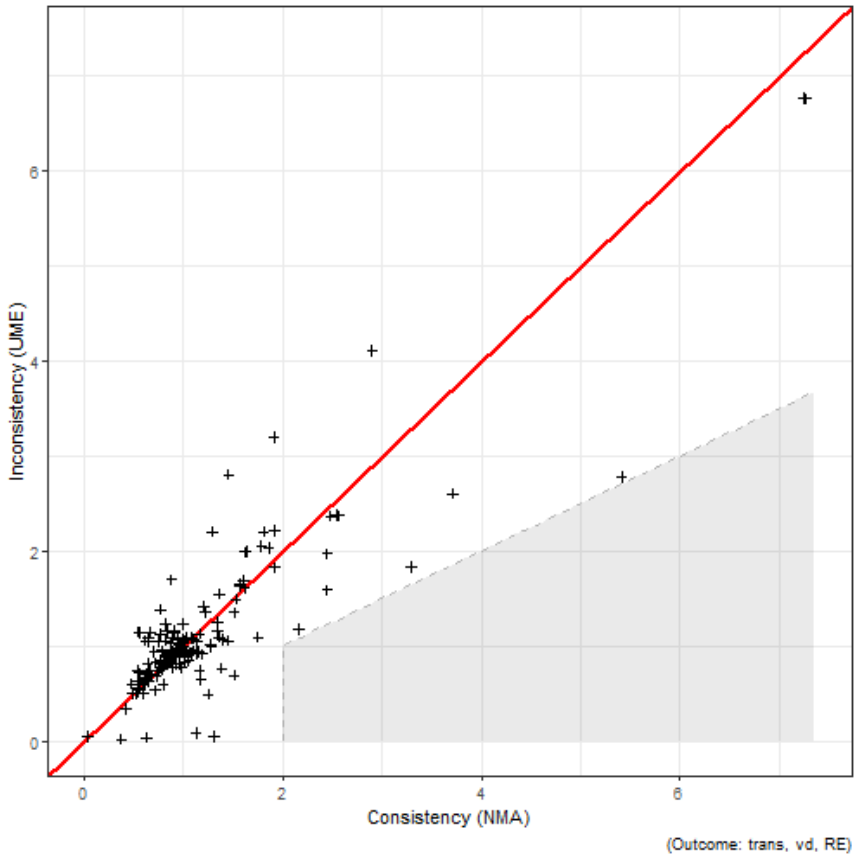

Figure 60. Dev-dev plot of each study arm's residual deviance under the consistency (NMA) and inconsistency models with RE structure for Transfusion, vaginal birth subgroup. No study arms had high deviance in the NMA model, relative to their deviance in the UME model.

Caesarean Section birth subgroup

Global inconsistency check

Analysis of the dataset for the CS birth subgroup for the outcome transfusion included 32 studies (65 arms) of 9 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Fitting of both fixed (FE) and random-effect (RE) network meta-analysis (NMA) models indicated strong support for the random-effects model on the basis of a large reduction in residual deviance and a moderate reduction in DIC (Table 62). There was no reduction in DIC when the RE UME model was fitted, relative to the RE NMA, suggesting that direct and indirect evidence is consistent.

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the random effects UME and NMA models (Figure 61), showed no study arms to have high deviance in the NMA model, relative to their deviance in the UME model. Taken together, the model fit and dev-dev plots suggest there was little evidence of inconsistency so node-splitting models were not required.

Table 62. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome Transfusion, CS birth subgroup.

Outcome	Pop.	Model	Posterior total residual deviance[17]	Between-study SD Mean, 95% credible interval	pD	DIC[18]
Transfusion	CS	FE NMA	94.73	-	37.9	299.0
Transfusion	CS	RE NMA	67.95	1.11 (0.45, 1.99)	49.6	283.9
Transfusion	CS	FE UME	90.01	-	42.4	298.8
Transfusion	CS	RE UME	67.59	1.50 (0.48, 3.09)	52.8	286.8

[17] Posterior mean residual deviance compared to 65 data points
[18] Deviance information criteria (DIC) – lower values preferred

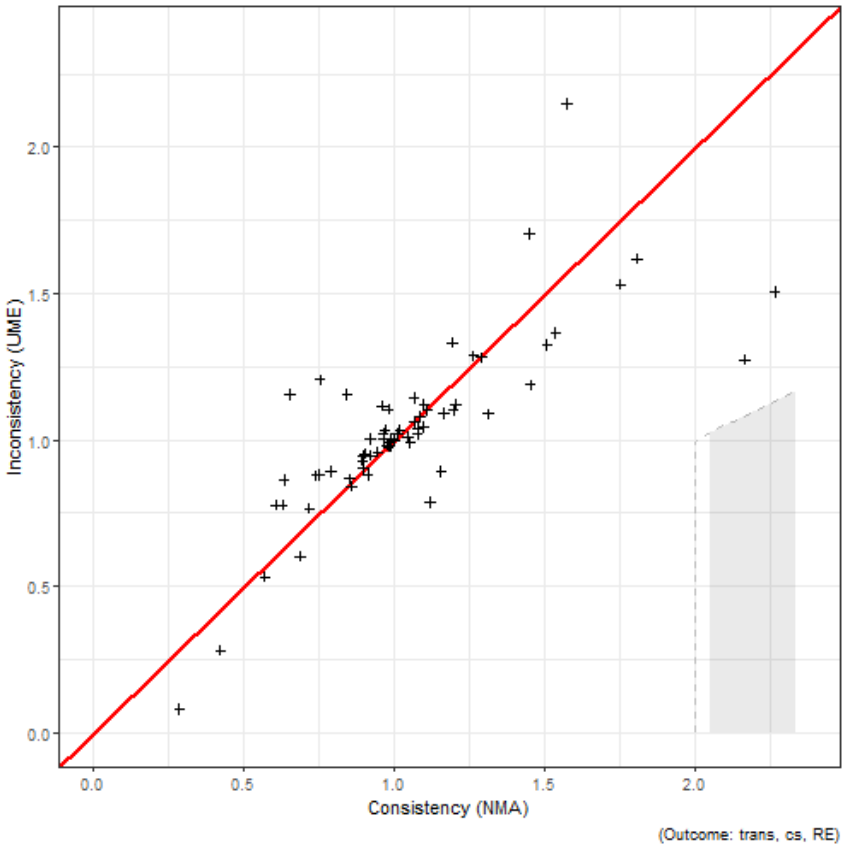

Figure 61. Dev-dev plot of each study arm's residual deviance under the consistency (NMA) and inconsistency models with RE structure for Transfusion, CS birth subgroup. No study arms were identified as inconsistent (i.e. there were no points in the shaded region).

Outcome: ICU admission

Summary

ICU admission was a rare event in these datasets and network meta-analysis (NMA) was possible for only the full and vaginal birth subgroups. Only one treatment comparison was informed by more than one study, providing insufficient evidence to inform a random-effects model structure.

There was only one loop of evidence within the full dataset and vaginal birth subgroup networks, and no evidence of inconsistency was identified within these datasets.

Full dataset

Global inconsistency check

Analysis of the full dataset for the outcome ICU admission included 9 studies (18 arms) of 8 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Fitting both fixed (FE) and random-effect (RE) NMA models indicated that the FE model structure was sufficient for the full dataset, with residual deviance of 17.2 (Table 63), approximately equivalent to the number of study arms (18).

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the fixed effects UME and NMA models (Figure 62), showed no study arms to have high deviance in the NMA model, relative to their deviance in the UME model. Taken together, the model fit, small number of studies on the loop of evidence, and dev-dev plots suggest there was little evidence of inconsistency; therefore, node-splitting models were not required.

Table 63. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome ICU admission, full dataset.

Outcome	Pop.	Model	Posterior total residual deviance[19]	Between-study SD Mean, 95% credible interval	pD	DIC[20]
ICU admission	Full	FE NMA	17.2	-	12.3	72.9
ICU admission	Full	RE NMA	16.3	1.87 (0.06, 4.75)	13.7	73.4
ICU admission	Full	FE UME	14.8	-	12.5	70.6
ICU admission	Full	RE UME	15.1	1.62 (0.04, 4.66)	13.2	71.6

[19] Posterior mean residual deviance compared to 18 total data points
[20] Deviance information criteria (DIC) – lower values preferred

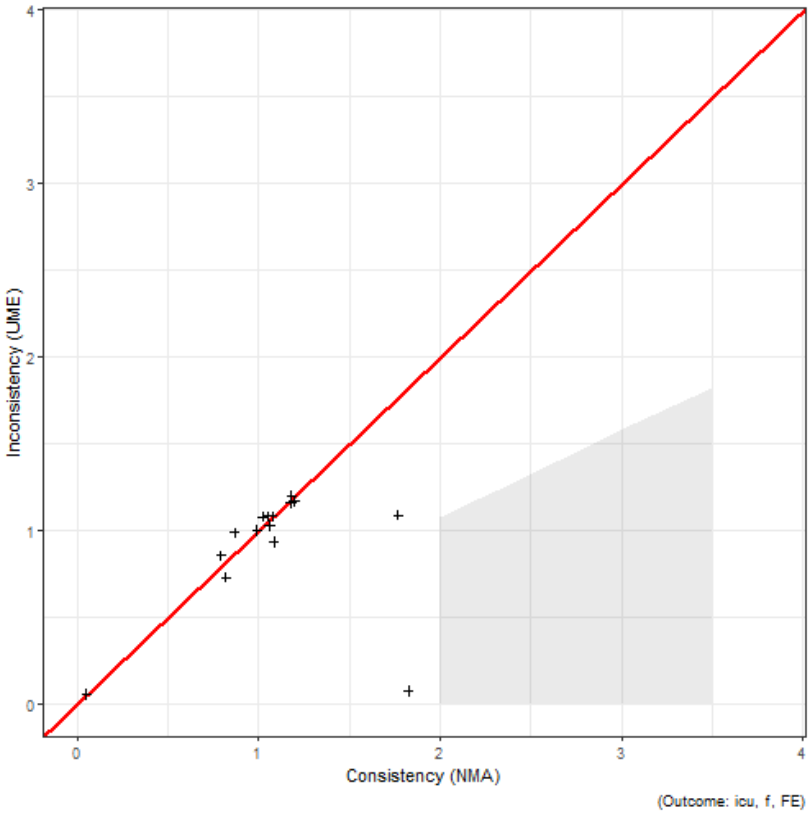

Figure 62. Dev-dev plot of each study arm's residual deviance under the standard and inconsistency models (FE model structure) for ICU admission, full dataset. No study arms had high deviance in the NMA model, relative to their deviance in the UME model.

Vaginal birth subgroup

Global inconsistency check

Analysis of the dataset for the vaginal birth subgroup for the outcome ICU admission included 8 studies (16 arms) of 7 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Fitting both fixed (FE) and random-effect (RE) NMA models indicated that the FE model structure was sufficient for the full dataset, with residual deviance of 16.0 (Table 64), equivalent to the 16 study arms.

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the fixed effects UME and NMA models (Figure 63), showed no study arms to have high deviance in the NMA model, relative to their deviance in the UME model. Taken together, the model fit, small number of studies on the loop of evidence, and dev-dev plots suggest there was little evidence of inconsistency; therefore, node-splitting models were not required.

Table 64. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome ICU admission, vaginal birth subgroup.

Outcome	Pop.	Model	Posterior total residual deviance[21]	Between-study SD Mean, 95% credible interval	pD	DIC[22]
ICU admission	VD	FE NMA	16.0	-	11.4	68.7
ICU admission	VD	RE NMA	15.0	1.83 (0.06, 4.72)	12.7	69.0
ICU admission	VD	FE UME	13.5	-	11.6	66.4
ICU admission	VD	RE UME	13.9	1.64 (0.04, 4.67)	12.3	67.5

[21] Posterior mean residual deviance compared to 16 data points
[22] Deviance information criteria (DIC) – lower values preferred

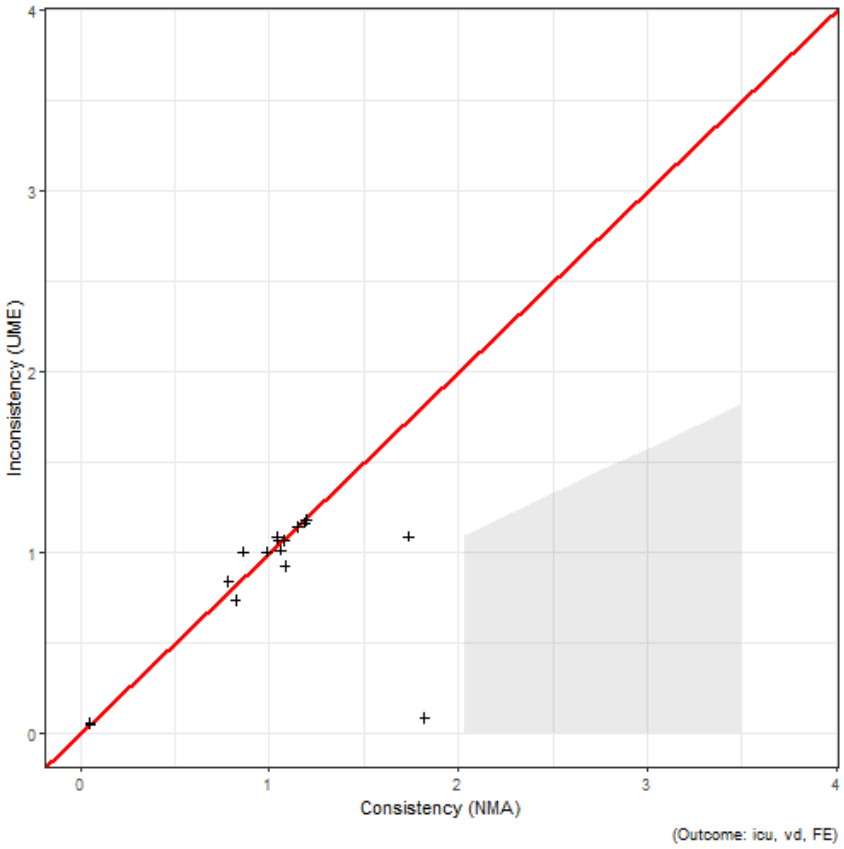

Figure 63. Dev-dev plot of each study arm's residual deviance under the consistency (NMA) and inconsistency models with FE structure for ICU admission, vaginal birth subgroup. No study arms had high deviance in the NMA model, relative to their deviance in the UME model.

Outcome: Mean blood loss

Summary

We identified moderate heterogeneity in both full and mode-of-delivery subgroup datasets for this outcome, but little evidence of inconsistency. Mean blood loss was analysed with treatment effects estimated on the log scale, with the effect of treatment assumed to be proportional rather than additive.

The only indicated inconsistency was connected to the estimate for the treatment effect of ergometrine (coded 13) relative to Oxytocin [>5 IU and ≤ 10 IU] (coded 5) from the study Modi 2014. Arm 3 from Modi 2014, which included carboprost, also showed relatively high residual deviance within the dataset. However, residual deviance for this arm was more similar between NMA and UME models, suggesting heterogeneity between study estimates for this treatment comparison rather than inconsistency.

Node-splitting models indicated potential inconsistency between direct and indirect evidence in the estimation of treatment differences between misoprostol plus oxytocin and oxytocin (>10IU) in the full dataset, and misoprostol (≤600mcg) and oxytocin (>10IU) in the vaginal delivery subgroup. While seven studies in the full dataset inform the comparison between oxytocin (>10IU) and misoprostol (≤600mcg) and nine inform the comparison between oxytocin (>10IU) and misoprostol plus oxytocin, only one study (Caliskan et al. 2003) informs

the comparison between misoprostol plus oxytocin and misoprostol (≤600mcg). This suggests that this loop of evidence (misoprostol (≤600mcg) – oxytocin (>10IU) – misoprostol plus oxytocin) contains inconsistency and we would suggest that Caliskan et al. 2003 is examined to ensure that there are no sources of potential inconsistency.

Full dataset

Global inconsistency check
Analysis of the full dataset for the outcome mean blood loss included 156 studies (332 arms) of 14 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Comparing fixed (FE) and random-effect (RE) network meta-analysis (NMA) models indicated support for the random-effect model on the basis of large decreases in DIC and residual deviance (Table 65).

Total residual deviance was lower in the inconsistency UME model than in the NMA model; however, DIC was lower for the NMA model and the estimate of between-study SD was similar in the NMA and UME models (Table 65). This suggests that there is little evidence of inconsistency but moderate heterogeneity between study estimates.

Table 65. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome Blood loss, full dataset.

Outcome	Pop.	Model	Posterior total residual deviance[23]	Between-study SD Mean, 95% credible interval	pD	DIC[24]
PPH	Full	FE NMA	5125.0	-	168.8	7526.2
PPH	Full	RE NMA	336.1	0.24 (0.23, 0.27)	314.4	2883.0
PPH	Full	FE UME	3533.0	-	199.4	5965.4
PPH	Full	RE UME	334.2	0.23 (0.20, 0.26)	317.4	2884.0

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the random effects UME and NMA models (Figure 64), shows that one study arm showed inconsistency with the rest of the dataset. The point on the border of the shaded area is the third arm from the same study:

- Modi, 2014

[23] Posterior mean residual deviance compared to 332 total data points
[24] Deviance information criteria (DIC) – lower values preferred

- Compares Oxytocin [>5 IU and ≤ 10 IU] (coded 5), Misoprostol ≤600mcg (coded 9), Ergometrine (coded 13) and Carboprost (coded 14)

However, residual deviance is relatively low even for this study arm. Taken together, the model fit and dev-dev plots suggest there was little evidence of inconsistency in these data. Node-splitting models for the loop of evidence informed by Modi et al. 2014 were run in WinBUGS.

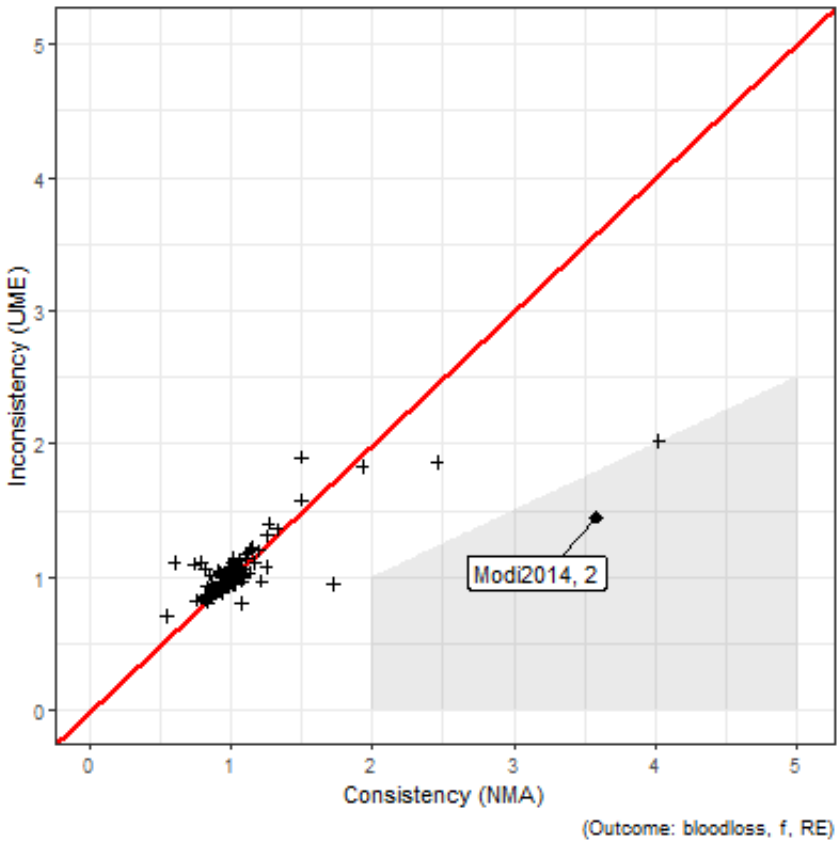

Figure 64. Dev-dev plot of each study arm's residual deviance under the standard and inconsistency models (RE model structure) for Blood loss, full dataset. Labels indicate study arms with high deviance in the NMA model, relative to their deviance in the UME model. The dotted line and grey shaded area denotes where study arms fit poorly in the NMA model but well in the UME model, suggesting that they are predicted poorly when the model enforces consistency in treatment differences.

Checking inconsistency within individual treatment comparisons (Node-splitting)

Node-splitting procedures indicated no evidence of a difference between direct and indirect evidence on most treatment comparisons. Evidence conflicted on a single comparison, for oxytocin (> 10IU) vs misoprostol with oxytocin (Table 66). A forest plot is presented for this comparison (Figure 65).

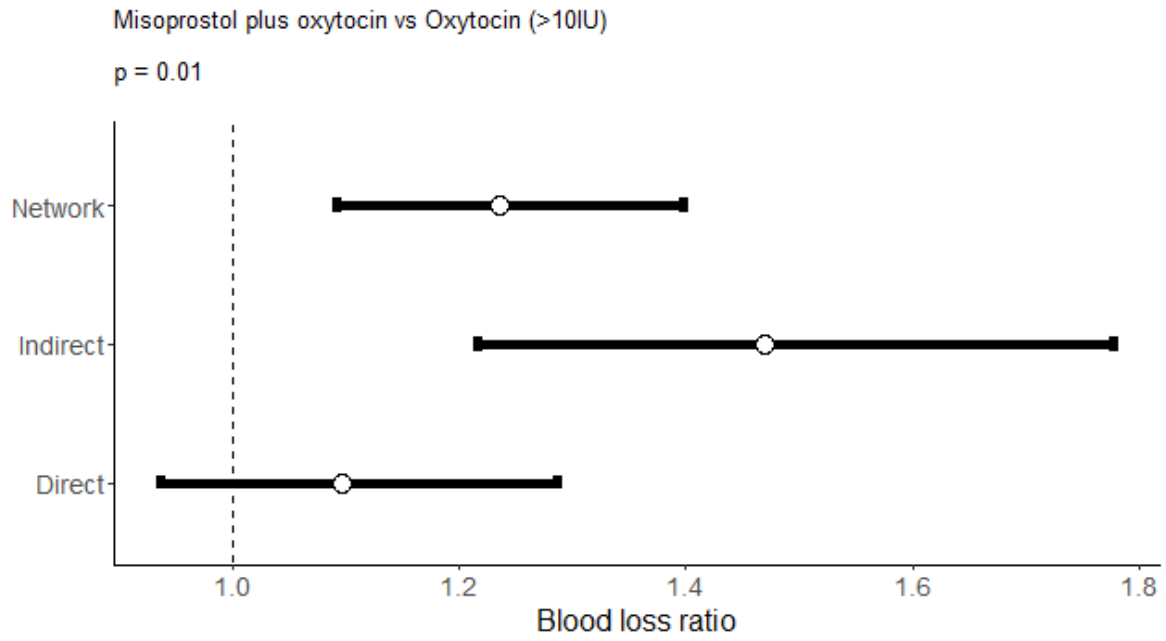

Figure 65. Treatment difference of misoprostol plus oxytocin vs oxytocin on outcome mean blood loss, full dataset, as estimated from the full network, indirect evidence only and direct evidence only. Ratios greater than 1 indicate that blood loss was greater in the misoprostol plus oxytocin group.

Table 66. Estimates of treatment effect from direct and indirect evidence from node-splitting models, full dataset. 'Comparison' indicates the pair of nodes for which evidence was split into direct and indirect evidence. Direct and indirect estimates are log blood loss ratios, where zero indicates no difference in blood loss between treatment groups.

Comparison	Total residual deviance	pD	DIC	Direct	Indirect	p value	Between-study SD
Oxy_a10 vs Mis_Oxy	**343.1**	**317.0**	**2893**	**0.09 (-0.06, 0.25)**	**0.38 (0.20, 0.58)**	**0.01**	**0.23 (0.20, 0.26)**
Carb vs Placebo	486.5	295.8	3015	-0.48 (-1.27, 0.20)	-0.39 (-0.54, -0.23)	0.42	0.26 (0.21, 0.27)
Oxy_b1 vs Placebo	341.8	317.1	2891	0.01 (-0.46, 0.47)	-0.21 (-0.86, 0.39)	0.70	0.24 (0.21, 0.27)
Oxy_a1b5 vs Placebo	377.8	345.8	2956	-0.19 (-0.41, 0.01)	-0.17 (-0.35, 0.02)	0.43	0.24 (0.21, 0.27)
Oxy_a5b10 vs Placebo	341.1	316.1	2890	-0.25 (-0.58, 0.09)	-0.18 (-0.32, -0.05)	0.37	0.24 (0.21, 0.27)
Erg_Oxy vs Placebo	342.3	315.5	2890	-0.11 (-0.45, 0.23)	-0.29 (-0.44, -0.14)	0.82	0.24 (0.21, 0.27)
Mis_b600 vs Placebo	347.0	318.7	2898	-0.19 (-0.35, -0.02)	-0.21 (-0.37, -0.06)	0.58	0.24 (0.21, 0.27)
Mis_a600b800 vs Placebo	531.0	430.0	3193	-0.31 (-0.79, 0.18)	-0.19 (-0.40, 0.02)	0.33	0.25 (0.21, 0.27)
Erg vs Placebo	1074	605.5	3912	-0.28 (-0.63, 0.07)	-0.12 (-0.28, 0.03)	0.21	0.24 (0.22, 0.28)
CarbProst vs Placebo	432.5	348.0	3013	-0.21 (-0.56, 0.14)	-0.40 (-0.59, -0.21)	0.83	0.24 (0.21, 0.27)
Oxy_a1b5 vs Carb	337.1	314.5	2884	0.13 (-0.07, 0.34)	0.23 (0.06, 0.40)	0.23	0.24 (0.21, 0.27)
Oxy_a5b10 vs Carb	339.8	317.7	2890	0.16 (0.01, 0.32)	0.20 (0.06, 0.34)	0.35	0.24 (0.21, 0.27)
Oxy_a10 vs Carb	537.2	416.8	3186	0.25 (0.05, 0.45)	0.21 (0.05, 0.37)	0.62	0.24 (0.21, 0.27)
Erg_Oxy vs Carb	358.9	321.1	2912	0.12 (-0.07, 0.32)	0.12 (-0.03, 0.26)	0.52	0.24 (0.21, 0.27)
Mis_b600 vs Carb	343.5	318.8	2895	0.38 (-0.18,	0.18 (0.07, 0.30)	0.75	0.24 (0.21, 0.27)

				0.96)			
Oxy_a1b5 vs Oxy_b1	537.9	140.0	2910	-0.13 (-0.71, 0.43)	-0.12 (-0.58, 0.34)	0.49	0.30 (0.21, 0.45)
Mis_b600 vs Oxy_b1	644.7	488.5	3366	-0.24 (-0.87, 0.40)	-0.07 (-0.43, 0.28)	0.31	0.24 (0.21, 0.27)
Erg_Oxy vs Oxy_a1b5	341.1	317.3	2891	-0.18 (-0.53, 0.18)	-0.05 (-0.20, 0.10)	0.25	0.24 (0.21, 0.27)
Mis_b600 vs Oxy_a1b5	2615.0	2371.0	7219	0.02 (-0.22, 0.25)	0.01 (-0.14, 0.16)	0.52	0.26 (0.21, 0.27)
Oxy_a10 vs Oxy_a5b10	557.2	430.5	3220	0.04 (-0.43, 0.51)	0.04 (-0.08, 0.16)	0.51	0.24 (0.21, 0.27)
Erg_Oxy vs Oxy_a5b10	866.3	-7.8	3091	0.02 (-0.31, 0.35)	-0.04 (-0.23, 0.22)	0.66	0.39 (0.22, 0.53)
Mis_b600 vs Oxy_a5b10	358.6	323.6	2915	0.03 (-0.06, 0.13)	-0.02 (-0.14, 0.09)	0.77	0.24 (0.21, 0.27)
Erg_Oxy vs Oxy_a10	374.6	330.7	2938	-0.05 (-0.35, 0.15)	-0.10 (-0.25, 0.05)	0.52	0.24 (0.21, 0.27)
Mis_b600 vs Oxy_a10	344.0	316.9	2893	-0.15 (-0.33, 0.03)	0.03 (-0.11, 0.17)	0.06	0.23 (0.21, 0.26)
Mis_b600 vs Erg_Oxy	343.7	320.7	2897	0.12 (-0.05, 0.28)	0.02 (-0.10, 0.14)	0.81	0.23 (0.20, 0.26)
Carb vs Mis_a600b800	336.5	314.5	2883	-0.34 (-0.80, 0.12)	-0.11 (-0.30, 0.07)	0.19	0.23 (0.21, 0.27)
Oxy_a1b5 vs Mis_a600b800	339.6	317.3	2889	-0.19 (-0.72, 0.35)	0.08 (-0.12, 0.28)	0.18	0.24 (0.21, 0.27)
Oxy_a5b10 vs Mis_a600b800	343.9	317.9	2894	0.05 (-0.18, 0.29)	0.04 (-0.17, 0.25)	0.53	0.24 (0.21, 0.27)
Oxy_a10 vs Mis_a600b800	344.1	317.8	2894	0.16 (-0.32, 0.64)	0.07 (-0.12, 0.27)	0.64	0.24 (0.21, 0.27)
Mis_b600 vs Mis_a600b800	387.5	337.3	2957	0.09 (-0.38, 0.56)	0.05 (-0.12, 0.21)	0.56	0.24 (0.21, 0.27)
Mis_a800b1000 vs Mis_a600b800	343.7	319.1	2895	-0.02 (-0.49, 0.44)	0.23 (-0.30, 0.75)	0.24	0.24 (0.21, 0.27)
Erg vs Mis_a600b800	410.0	339.4	2982	0.30 (-0.13,	0.17 (-0.07,	0.68	0.26 (0.21, 0.27)

				0.47)	0.27)		
CarbProst vs Mis_a600b800	362.3	336.3	2931	0.21 (-0.21, 0.62)	-0.16 (-0.37, 0.04)	0.94	0.23 (0.21, 0.27)
Oxy_a5b10 vs Mis_a800b1000	342.8	319.0	2894	-0.17 (-0.67, 0.34)	0.10 (-0.28, 0.49)	0.20	0.24 (0.21, 0.27)
Mis_b600 vs Mis_a800b1000	749.6	450.1	3432	0.11 (-0.37, 0.59)	-0.09 (-0.54, 0.37)	0.72	0.26 (0.21, 0.27)
Erg vs Mis_a800b1000	343.7	316.4	2893	0.15 (-0.25, 0.56)	0.02 (-0.33, 0.37)	0.69	0.24 (0.21, 0.27)
Oxy_a1b5 vs Mis_Oxy	342.2	317.8	2892	0.45 (-0.01, 0.92)	0.14 (-0.03, 0.31)	0.89	0.23 (0.21, 0.27)
Oxy_a5b10 vs Mis_Oxy	342.0	317.3	2892	0.26 (0.10, 0.41)	0.05 (-0.14, 0.25)	0.95	0.23 (0.21, 0.26)
Erg_Oxy vs Mis_Oxy	392.3	340.6	2965	0.06 (-0.42, 0.53)	0.12 (-0.04, 0.26)	0.42	0.24 (0.21, 0.27)
Mis_b600 vs Mis_Oxy	403.3	326.9	2963	0.16 (-0.25, 0.57)	0.18 (0.04, 0.31)	0.46	0.24 (0.21, 0.27)
Oxy_b1 vs Erg	341.9	316.4	2891	-0.17 (-0.81, 0.44)	0.14 (-0.22, 0.49)	0.18	0.23 (0.21, 0.27)
Oxy_a1b5 vs Erg	410.5	348.3	2991	0.00 (-0.29, 0.30)	-0.05 (-0.19, 0.11)	0.61	0.24 (0.21, 0.26)
Oxy_a5b10 vs Erg	350.1	318.0	2900	0.02 (-0.15, 0.19)	-0.07 (-0.20, 0.05)	0.82	0.24 (0.21, 0.27)
Erg_Oxy vs Erg	343.4	315.8	2892	-0.24 (-0.73, 0.26)	-0.09 (-0.22, 0.04)	0.29	0.24 (0.21, 0.27)
Mis_b600 vs Erg	347.6	317.6	2898	-0.04 (-0.16, 0.09)	0.03 (-0.11, 0.17)	0.25	0.24 (0.21, 0.27)
CarbProst vs Erg	361.9	332.6	2927	-0.11 (-0.28, 0.06)	-0.36 (-0.6, -0.11)	0.95	0.24 (0.21, 0.27)
Oxy_a1b5 vs CarbProst	1187.0	-426.7	2993	0.11 (-0.82, 1.14)	0.18 (-0.14, 0.51)	0.42	0.39 (0.21, 0.70)
Oxy_a5b10 vs CarbProst	348.9	322.7	2904	0.12 (-0.17, 0.41)	0.23 (0.06, 0.39)	0.26	0.24 (0.20, 0.26)
Erg_Oxy vs CarbProst	345.0	318.0	2895	0.11 (-0.29,	0.10 (-0.08,	0.53	0.24 (0.21, 0.27)

				0.51)	0.27)		
Mis_b600 vs CarbProst	372.1	335.5	2940	-0.16 (-0.54, 0.22)	0.17 (0.02, 0.33)	0.06	0.24 (0.21, 0.27)

Vaginal birth subgroup

Global inconsistency check

Analysis of the dataset for the vaginal delivery subgroup for the outcome blood loss included 109 studies (235 arms) of 13 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Fitting of both fixed and random-effect network meta-analysis (NMA) models indicated support for the random-effect model on the basis of a small decrease in DIC and a sizeable decrease in residual deviance (Table 67).

Total residual deviance and DIC were lower in the inconsistency UME model than in the NMA model and the estimate of between-study SD was the similar in the NMA and UME models (Table 67). This suggests no evidence of inconsistency, but that there is moderate heterogeneity between study estimates.

Table 67. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome Blood loss, vaginal delivery subgroup.

Outcome	Pop.	Model	Posterior total residual deviance[25]	Between-study SD Mean, 95% credible interval	pD	DIC[26]
PPH	VD	FE NMA	2961.0	-	120.9	4553.8
PPH	VD	RE NMA	239.6	0.25 (0.21, 0.29)	224.6	1936.4
PPH	VD	FE UME	1870.0	-	146.9	3489.5
PPH	VD	RE UME	238.1	0.24 (0.20, 0.28)	227.2	1937.6

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the random effects UME and NMA models (Figure 66), shows that one study arm showed inconsistency with the rest of the dataset:

- Modi, 2014
    - Compares Oxytocin [>5 IU and ≤ 10 IU] (coded 5), Misoprostol ≤600mcg (coded 8), Ergometrine (coded 12) and Carboprost (coded 13)

[25] Posterior mean residual deviance compared to 235 data points
[26] Deviance information criteria (DIC) – lower values preferred

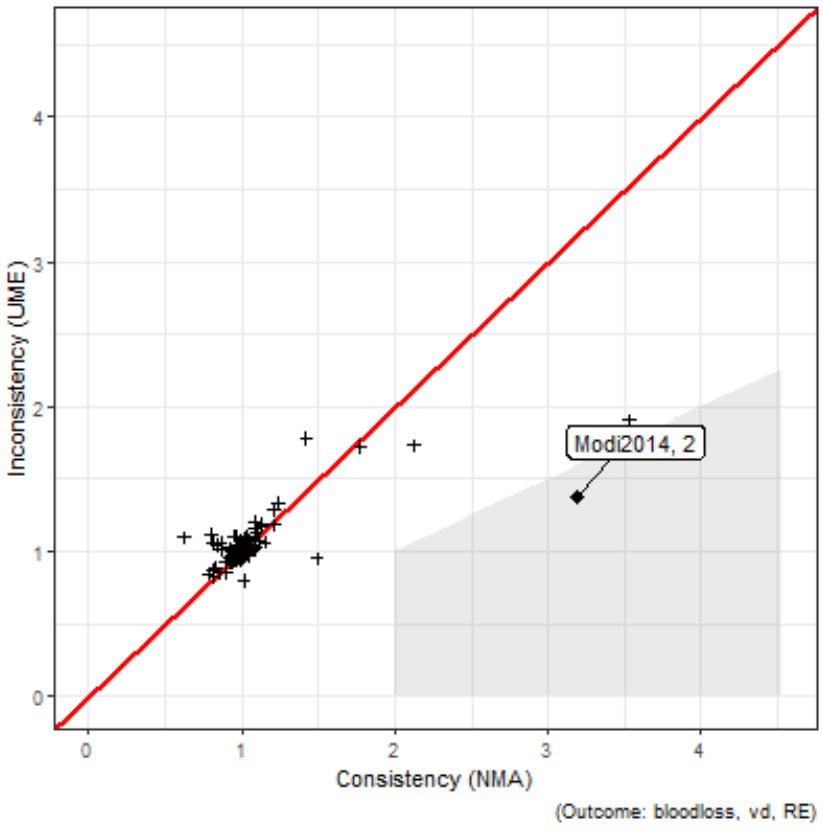

Figure 66. Dev-dev plot of each study arm's residual deviance under the consistency (NMA) and inconsistency models with RE structure for Blood loss, vaginal delivery subgroup. Labels indicate study arms with high deviance in the NMA model, relative to their deviance in the UME model. The dotted line and grey shaded area outline a region where study arms fit poorly in the NMA model but well in the UME model, suggesting that they are predicted poorly when the model enforces consistency in treatment differences.

Checking inconsistency within individual treatment comparisons (Node-splitting)

Node-splitting procedures indicated no evidence of a difference between direct and indirect evidence on most treatment comparisons. Evidence conflicted on a single comparison, for oxytocin (> 10IU) vs misoprostol (<600mcg) (Table 68). A forest plot is presented for this comparison (Figure 67).

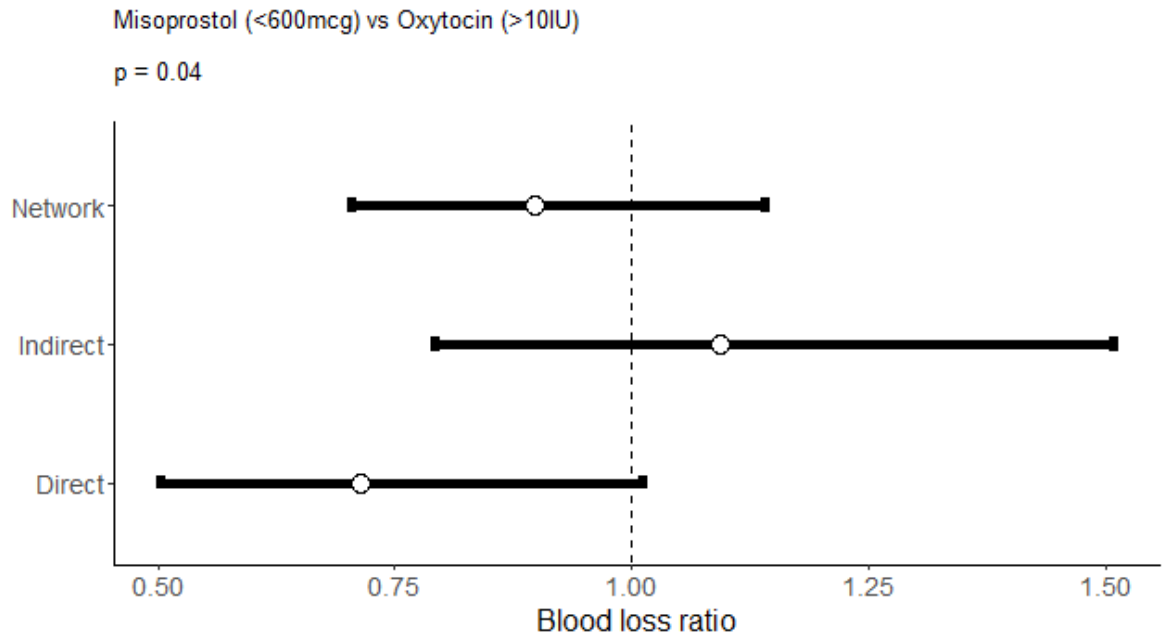

Figure 67. Treatment difference of misoprostol vs oxytocin on outcome mean blood loss, vaginal delivery subgroup, as estimated from the full network, indirect evidence only and direct evidence only. Ratios greater than 1 indicate that blood loss was greater in the misoprostol group.

Table 68. Estimates of treatment effect from direct and indirect evidence from node-splitting models, vaginal delivery subgroup. 'Comparison' indicates the pair of nodes for which evidence was split into direct and indirect evidence. Direct and indirect estimates are log blood loss ratios, where zero indicates no difference in blood loss between treatment groups.

Comparison	Total residual Deviance	pD	DIC	Direct	Indirect	P value	Between-study SD
Mis_b600 vs Oxy_a10	**249.5**	**229.4**	**1951**	**-0.34 (-0.69, 0.01)**	**0.09 (-0.23, 0.41)**	**0.04**	**0.25 (0.21, 0.29)**
Oxy_a1b5 vs Placebo	892.7	197.1	2562	-0.26 (-0.59, 0.07)	-0.02 (-0.34, 0.42)	0.17	0.32 (0.22, 0.49)
Oxy_a5b10 vs Placebo	1044.0	-455.7	2060	-0.25 (-1.12, 0.63)	-0.34 (-0.92, -0.03)	0.52	0.51 (0.22, 0.95)
Erg_Oxy vs Placebo	417.2	235.6	2125	-0.12 (-0.49, 0.25)	-0.27 (-0.45, -0.07)	0.74	0.29 (0.22, 0.29)
Mis_b600 vs Placebo	284.4	259.4	2016	-0.14 (-0.33, 0.06)	-0.23 (-0.41, -0.06)	0.77	0.25 (0.22, 0.29)
Mis_a600b800 vs Placebo	244.6	226.0	1943	-0.31 (-0.81, 0.20)	-0.19 (-0.42, 0.04)	0.34	0.25 (0.21, 0.29)
Erg vs Placebo	256.1	231.1	1959	-0.28 (-0.64, 0.08)	-0.12 (-0.28, 0.05)	0.21	0.25 (0.21, 0.29)
CarbProst vs Placebo	243.1	226.5	1942	-0.21 (-0.58, 0.16)	-0.40 (-0.61, -0.20)	0.82	0.25 (0.21, 0.29)
Oxy_a1b5 vs Carb	246.4	225.9	1944	0.19 (-0.17, 0.55)	0.20 (-0.03, 0.43)	0.49	0.25 (0.22, 0.29)
Oxy_a5b10 vs Carb	254.3	231.6	1958	0.10 (-0.13, 0.34)	0.22 (0.02, 0.42)	0.24	0.25 (0.22, 0.29)
Erg_Oxy vs Carb	278.3	255.1	2006	0.14 (-0.09, 0.37)	0.16 (-0.06, 0.38)	0.46	0.25 (0.22, 0.29)

Mis_b600 vs Carb	246.9	226.8	1946	0.39 (-0.20, 0.98)	0.19 (0.02, 0.36)	0.73	0.25 (0.22, 0.29)
Erg_Oxy vs Oxy_a1b5	244.7	227.4	1944	-0.30 (-0.81, 0.21)	-0.01 (-0.20, 0.18)	0.14	0.25 (0.21, 0.29)
Mis_b600 vs Oxy_a1b5	282.9	259.4	2015	0.01 (-0.22, 0.25)	0.01 (-0.17, 0.19)	0.50	0.25 (0.21, 0.29)
Erg_Oxy vs Oxy_a5b10	247.1	228.1	1947	0.08 (-0.13, 0.29)	-0.04 (-0.2, 0.12)	0.82	0.26 (0.22, 0.30)
Mis_b600 vs Oxy_a5b10	247.2	231.1	1950	0.04 (-0.07, 0.15)	0.02 (-0.12, 0.17)	0.57	0.26 (0.22, 0.30)
Erg_Oxy vs Oxy_a10	246.9	228.7	1948	-0.04 (-0.54, 0.47)	-0.20 (-0.49, 0.10)	0.71	0.25 (0.22, 0.29)
Mis_b600 vs Erg_Oxy	246.0	228.2	1946	0.11 (-0.06, 0.29)	-0.02 (-0.17, 0.12)	0.88	0.24 (0.21, 0.28)
Carb vs Mis_a600b800	239.0	224.0	1935	-0.34 (-0.83, 0.15)	-0.11 (-0.35, 0.13)	0.21	0.25 (0.22, 0.29)
Oxy_a1b5 vs Mis_a600b800	248.3	229.1	1950	-0.19 (-0.75, 0.36)	0.09 (-0.15, 0.33)	0.18	0.25 (0.21, 0.29)
Oxy_a5b10 vs Mis_a600b800	243.4	227.1	1943	0.05 (-0.20, 0.31)	0.01 (-0.24, 0.25)	0.60	0.25 (0.22, 0.29)
Mis_b600 vs Mis_a600b800	368.4	246.9	2087	0.08 (-0.42, 0.59)	0.05 (-0.15, 0.24)	0.56	0.27 (0.22, 0.30)
Mis_a800b1000 vs Mis_a600b800	56420.0	-25440	32450	1.32 (-0.52, 4.45)	0.50 (-0.60, 2.18)	0.58	0.52 (0.22, 0.91)
Erg vs Mis_a600b800	304.5	267.6	2044	0.15 (-0.18, 0.48)	0.09 (-0.11, 0.28)	0.64	0.25 (0.22, 0.29)
CarbProst vs Mis_a600b800	344.4	196.1	2013	0.21 (-0.26, 0.65)	-0.16 (-0.41, 0.06)	0.92	0.27 (0.21, 0.38)
Oxy_a5b10 vs Mis_a800b1000	251.4	226.4	1950	-0.17 (-0.70, 0.37)	0.09 (-0.32, 0.50)	0.23	0.25 (0.21, 0.29)
Mis_b600 vs Mis_a800b1000	285.8	234.2	1992	0.11 (-0.39, 0.62)	-0.13 (-0.52, 0.41)	0.70	0.27 (0.22, 0.29)
Erg vs Mis_a800b1000	961.0	-332.2	2101	0.15 (-0.67, 0.98)	0.02 (-0.67, 0.71)	0.63	0.42 (0.22, 0.65)

Oxy_a5b10 vs Mis_Oxy	249.8	228.9	1951	0.10 (-0.10, 0.30)	-0.24 (-0.71, 0.22)	0.91	0.25 (0.21, 0.29)
Oxy_a10 vs Mis_Oxy	244.9	226.4	1943	0.01 (-0.35, 0.36)	0.35 (-0.01, 0.70)	0.09	0.25 (0.21, 0.29)
Erg_Oxy vs Mis_Oxy	248.7	228.6	1949	0.05 (-0.45, 0.56)	0.03 (-0.20, 0.25)	0.54	0.25 (0.22, 0.29)
Mis_b600 vs Mis_Oxy	2469.0	-1981.0	1960	0.15 (-1.06, 1.35)	0.11 (-0.43, 0.70)	0.58	0.56 (0.22, 1.00)
Oxy_a1b5 vs Erg	244.0	227.5	1944	0.01 (-0.30, 0.31)	-0.03 (-0.22, 0.15)	0.59	0.25 (0.21, 0.29)
Oxy_a5b10 vs Erg	247.8	228.7	1949	0.03 (-0.16, 0.22)	-0.09 (-0.23, 0.04)	0.85	0.25 (0.21, 0.29)
Erg_Oxy vs Erg	250.2	229.9	1952	-0.24 (-0.76, 0.28)	-0.07 (-0.22, 0.08)	0.27	0.25 (0.22, 0.29)
Mis_b600 vs Erg	250.1	227.6	1950	-0.04 (-0.17, 0.10)	0.06 (-0.09, 0.22)	0.17	0.26 (0.22, 0.30)
CarbProst vs Erg	817.5	-190.3	2099	-0.09 (-0.42, 0.25)	-0.49 (-1.07, -0.06)	0.94	0.42 (0.22, 0.66)
Oxy_a1b5 vs CarbProst	753.1	-153.8	2071	0.11 (-0.94, 1.19)	0.14 (-0.31, 0.49)	0.46	0.43 (0.22, 0.69)
Oxy_a5b10 vs CarbProst	250.4	229.9	1952	0.12 (-0.18, 0.42)	0.22 (0.04, 0.40)	0.28	0.24 (0.21, 0.28)
Erg_Oxy vs CarbProst	252.5	237.2	1962	0.22 (-0.45, 0.88)	0.12 (-0.07, 0.31)	0.62	0.25 (0.22, 0.29)
Mis_b600 vs CarbProst	1219.0	-558.2	2133	-0.17 (-1.14, 0.79)	0.07 (-0.49, 0.40)	0.24	0.54 (0.22, 0.93)

Caesarean birth subgroup

Global inconsistency check

Analysis of the dataset for the CS delivery subgroup for the outcome blood loss included 46 studies (95 arms) of 12 treatments.

Results were based on 80,000 iterations following a burn-in of 60,000 iterations, which was sufficient to achieve convergence using a standard, uninformative prior for between-study standard deviation (SD).

Fitting both fixed (FE) and random-effect (RE) network meta-analysis (NMA) models gave similar values for DIC. However, residual deviance was lower by 5.3 for the RE NMA (Table 69) suggesting support for the RE model for these data.

Total residual deviance and DIC were lower in the inconsistency UME model than in the NMA model and the estimate of between-study SD was the similar in the NMA and UME models (Table 69). This suggests that there is no evidence of inconsistency, but moderate heterogeneity between study estimates.

The dev-dev plot, which shows the contribution of each study datapoint to the residual deviance under the random effects UME and NMA models (Figure 68), showed no study arms to have high deviance in the NMA model, relative to their deviance in the UME model. Taken together, the model fit and dev-dev plots suggest there was little evidence of inconsistency in these data. Therefore, node-splitting models were not required.

Table 69. Model fit statistics for fixed- and random-effect NMA and UME models of the outcome Blood loss, CS delivery subgroup.

Outcome	Pop.	Model	Posterior total residual deviance[a]	Between-study SD Mean, 95% credible interval	pD	DIC[b]
PPH	CS	FE NMA	1706.0	-	56.9	2506.8
PPH	CS	RE NMA	93.9	0.21 (0.16, 0.27)	88.9	926.5
PPH	CS	FE UME	717.0	-	65.0	1525.7
PPH	CS	RE UME	94.4	0.19 (0.14, 0.25)	89.2	927.3

[a] Posterior mean residual deviance compared to 95 data points
[b] Deviance information criteria (DIC) – lower values preferred

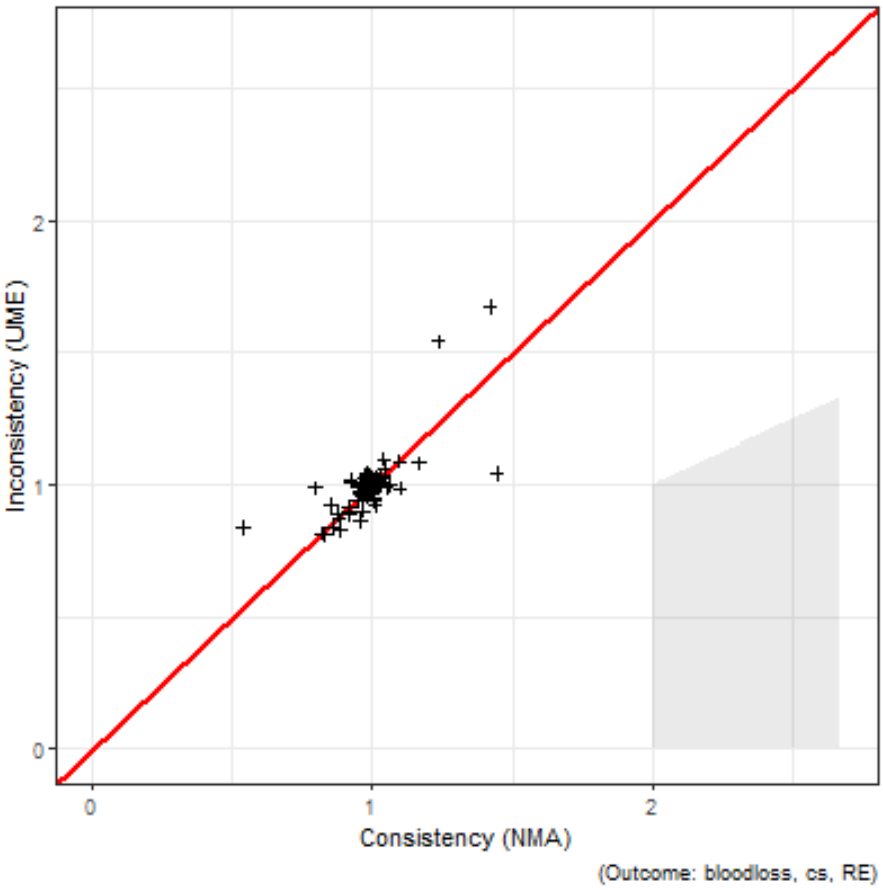

Figure 68. Dev-dev plot of each study arm's residual deviance under the consistency (NMA) and inconsistency models with RE structure for Blood loss, CS delivery subgroup. No study arms were identified as inconsistent (i.e. points in the shaded region).

NMA code

The code below was originally based on information within the TSU evidence synthesis technical support documents (Dias 2011, Dias 2014).

WinBUGS code for fixed effect model – binary outcomes

```
# Binomial likelihood, logit link
# Fixed effects model
model{                          # *** PROGRAM STARTS
for(i in 1:ns){                 # LOOP THROUGH STUDIES
    mu[i] ~ dnorm(0,.0001)      # vague priors for all trial baselines
    for (k in 1:na[i])  {       # LOOP THROUGH ARMS
        r[i,k] ~ dbin(p[i,k],n[i,k])    # binomial likelihood
# model for linear predictor
        logit(p[i,k]) <- mu[i] + d[t[i,k]] - d[t[i,1]]
# expected value of the numerators
        rhat[i,k] <- p[i,k] * n[i,k]
#Deviance contribution
```

```
        dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))
             +  (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k]-
rhat[i,k])))
      }
# summed residual deviance contribution for this trial
    resdev[i] <- sum(dev[i,1:na[i]])
     }
totresdev <- sum(resdev[])      # Total Residual Deviance
d[1]<-0    # treatment effect is zero for reference treatment
# vague priors for treatment effects
for (k in 2:nt){  d[k] ~ dnorm(0,.001) }

# ranking on relative scale
for (k in 1:nt) {
# rk[k] <- nt+1-rank(d[],k) # assumes events are "good"
 rk[k] <- rank(d[],k) # assumes events are "bad"
 best[k] <- equals(rk[k],1) #calculate probability that treat k is best
 for (h in 1:nt){ prob[h,k] <- equals(rk[k],h) } # calculates probability
that treat k is h-th best
 }

}                                                    # *** PROGRAM ENDS
```

WinBUGS code for random effect model – binary outcomes

```
# Binomial likelihood, logit link
# Random effects model for multi-arm trials
model{                                  # *** PROGRAM STARTS
for(i in 1:ns){                         # LOOP THROUGH STUDIES
    w[i,1] <- 0    # adjustment for multi-arm trials is zero for control
arm
    delta[i,1] <- 0             # treatment effect is zero for control arm
    mu[i] ~ dnorm(0,.0001)           # vague priors for all trial baselines
    for (k in 1:na[i]) {             # LOOP THROUGH ARMS
        r[i,k] ~ dbin(p[i,k],n[i,k]) # binomial likelihood
        logit(p[i,k]) <- mu[i] + delta[i,k]  # model for linear predictor
        rhat[i,k] <- p[i,k] * n[i,k] # expected value of the numerators
#Deviance contribution
        dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))
            +  (n[i,k]-r[i,k]) * (log(n[i,k]-r[i,k]) - log(n[i,k]-
rhat[i,k])))         }
#  summed residual deviance contribution for this trial
    resdev[i] <- sum(dev[i,1:na[i]])
    for (k in 2:na[i]) {             # LOOP THROUGH ARMS
# trial-specific LOR distributions
        delta[i,k] ~ dnorm(md[i,k],taud[i,k])
# mean of LOR distributions (with multi-arm trial correction)
        md[i,k] <-  d[t[i,k]] - d[t[i,1]] + sw[i,k]
# precision of LOR distributions (with multi-arm trial correction)
        taud[i,k] <- tau *2*(k-1)/k
# adjustment for multi-arm RCTs
        w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])
# cumulative adjustment for multi-arm trials
        sw[i,k] <- sum(w[i,1:k-1])/(k-1)
      }
  }
totresdev <- sum(resdev[])           # Total Residual Deviance
d[1]<-0       # treatment effect is zero for reference treatment
# vague priors for treatment effects
for (k in 2:nt){
      d[k] ~ dnorm(0,.001)
      }
```

```
sd ~ dunif(0,5)      # vague prior for between-trial SD
tau <- pow(sd,-2)    # between-trial precision = (1/between-trial variance)

# pairwise ORs and LORs for all possible pair-wise comparisons, if nt>2
for (c in 1:(nt-1)) {
 for (k in (c+1):nt) {
      or[c,k] <- exp(d[k] - d[c])
      lor[c,k] <- (d[k]-d[c])
      }
 }

# ranking on relative scale
for (k in 1:nt) {
# rk[k] <- nt+1-rank(d[],k) # assumes events are "good"
 rk[k] <- rank(d[],k) # assumes events are "bad"
 best[k] <- equals(rk[k],1) #calculate probability that treat k is best
 for (h in 1:nt){ prob[h,k] <- equals(rk[k],h) } # calculates probability
that treat k is h-th best
 }


}                                    # *** PROGRAM ENDS
```

WinBUGS code for fixed effect model – continuous outcomes

```
# Normal likelihood, log link
# Fixed effects model
model{                                  # *** PROGRAM STARTS
for(i in 1:ns){                         #   LOOP THROUGH STUDIES
    mu[i] ~ dnorm(0,.0001)              # vague priors for all trial baselines
    for (k in 1:na[i]) {                #  LOOP THROUGH ARMS
        var[i,k] <- pow(se[i,k],2)      # calculate variances
        prec[i,k] <- 1/var[i,k]         # set precisions
        y[i,k] ~ dnorm(theta[i,k],prec[i,k]) # normal likelihood
# model for linear predictor
        log(theta[i,k]) <- mu[i] + d[t[i,k]] - d[t[i,1]]
#Deviance contribution
        dev[i,k] <- (y[i,k]-theta[i,k])*(y[i,k]-theta[i,k])*prec[i,k]
      }
#  summed residual deviance contribution for this trial
    resdev[i] <- sum(dev[i,1:na[i]])
  }
totresdev <- sum(resdev[])              #Total Residual Deviance
d[1]<-0       # treatment effect is zero for control arm

# vague priors for treatment effects
for (k in 2:nt){  d[k] ~ dnorm(0,.0001) }

for (k in 1:nt) {
            blossRatio[k] <- exp(d[k])
      }

}                                       # *** PROGRAM ENDS
```

WinBUGS code for random effect model – continuous outcomes

```
# Normal likelihood, log link
# Random effects model for multi-arm trials
model{                                  # *** PROGRAM STARTS
for(i in 1:ns){                         #   LOOP THROUGH STUDIES
```

```
    w[i,1] <- 0    # adjustment for multi-arm trials is zero for control
arm
    delta[i,1] <- 0             # treatment effect is zero for control arm
    mu[i] ~ dnorm(0,.0001)           # vague priors for all trial baselines
    for (k in 1:na[i]) {             #  LOOP THROUGH ARMS
        var[i,k] <- pow(se[i,k],2)   # calculate variances
        prec[i,k] <- 1/var[i,k]      # set precisions
        y[i,k] ~ dnorm(theta[i,k],prec[i,k]) # binomial likelihood
        log(theta[i,k]) <- mu[i] + delta[i,k]  # model for linear predictor
#Deviance contribution
        dev[i,k] <- (y[i,k]-theta[i,k])*(y[i,k]-theta[i,k])*prec[i,k]
      }
#  summed residual deviance contribution for this trial
    resdev[i] <- sum(dev[i,1:na[i]])
    for (k in 2:na[i]) {             # LOOP THROUGH ARMS
# trial-specific LOR distributions
        delta[i,k] ~ dnorm(md[i,k],taud[i,k])
# mean of LOR distributions, with multi-arm trial correction
        md[i,k] <-  d[t[i,k]] - d[t[i,1]] + sw[i,k]
# precision of LOR distributions (with multi-arm trial correction)
        taud[i,k] <- tau *2*(k-1)/k
# adjustment, multi-arm RCTs
        w[i,k] <- (delta[i,k] - d[t[i,k]] + d[t[i,1]])
# cumulative adjustment for multi-arm trials
        sw[i,k] <- sum(w[i,1:k-1])/(k-1)
      }
  }
totresdev <- sum(resdev[])            #Total Residual Deviance
d[1]<-0       # treatment effect is zero for control arm
# vague priors for treatment effects
for (k in 2:nt){  d[k] ~ dnorm(0,.0001) }
sd ~ dunif(0,5)     # vague prior for between-trial SD
tau <- pow(sd,-2)   # between-trial precision = (1/between-trial variance)
# Provide estimates of treatment effects T[k] on the natural scale
# Given a Mean Effect, meanA, for 'standard' treatment A,
# with precision (1/variance) precA

for (k in 1:nt) {
            blossRatio[k] <- exp(d[k])
      }

}                                    # *** PROGRAM ENDS
```

Acknowledgments

We would like to acknowledge Beatrice Downing and Nicky Welton from the Guidelines Technical Support Unit, at University of Bristol, for providing advice, models, inconsistency checking and quality assurance for the network meta-analyses included in this review.