Cochrane Library

Cochrane Database of Systematic Reviews

Caffeine dosing regimens in preterm infants with or at risk for apnea of prematurity (Review)

Bruschettini M, Brattström P, Russo C, Onland W, Davis PG, Soll R

Bruschettini M, Brattström P, Russo C, Onland W, Davis PG, Soll R.
Caffeine dosing regimens in preterm infants with or at risk for apnea of prematurity.
Cochrane Database of Systematic Reviews 2023, Issue 4. Art. No.: CD013873.
DOI: 10.1002/14651858.CD013873.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Caffeine dosing regimens in preterm infants with or at risk for apnea of prematurity

Matteo Bruschettini[1,2], Petter Brattström[3], Chiara Russo[4], Wes Onland[5], Peter G Davis[6,7,8], Roger Soll[9]

[1]Paediatrics, Department of Clinical Sciences Lund, Lund University, Skåne University Hospital, Lund, Sweden. [2]Cochrane Sweden, Department of Research and Education, Lund University, Skåne University Hospital, Lund, Sweden. [3]Blekinge Hospital, Karlskrona, Sweden. [4]University of Genoa, Genoa, Italy. [5]Department of Neonatology, Amsterdam University Medical Centers, VU University Medical Center, Emma Children's Hospital, University of Amsterdam, Amsterdam, Netherlands. [6]Newborn Research Centre and Neonatal Services, The Royal Women's Hospital, Melbourne, Australia. [7]Murdoch Children's Research Institute, Melbourne, Australia. [8]Department of Obstetrics and Gynecology, University of Melbourne, Melbourne, Australia. [9]Division of Neonatal-Perinatal Medicine, Department of Pediatrics, Larner College of Medicine at the University of Vermont, Burlington, Vermont, USA

Contact: Matteo Bruschettini, matteo.bruschettini@med.lu.se.

Editorial group: Cochrane Neonatal Group.
Publication status and date: New, published in Issue 4, 2023.

Citation: Bruschettini M, Brattström P, Russo C, Onland W, Davis PG, Soll R. Caffeine dosing regimens in preterm infants with or at risk for apnea of prematurity. *Cochrane Database of Systematic Reviews* 2023, Issue 4. Art. No.: CD013873. DOI: 10.1002/14651858.CD013873.pub2.

ABSTRACT

Background

Very preterm infants often require respiratory support and are therefore exposed to an increased risk of bronchopulmonary dysplasia (chronic lung disease) and later neurodevelopmental disability. Caffeine is widely used to prevent and treat apnea (temporal cessation of breathing) associated with prematurity and facilitate extubation. Though widely recognized dosage regimes have been used for decades, higher doses have been suggested to further improve neonatal outcomes. However, observational studies suggest that higher doses may be associated with harm.

Objectives

To determine the effects of higher versus standard doses of caffeine on mortality and major neurodevelopmental disability in preterm infants with (or at risk of) apnea, or peri-extubation.

Search methods

We searched CENTRAL, MEDLINE, Embase, CINAHL, the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP), and clinicaltrials.gov in May 2022. The reference lists of relevant articles were also checked to identify additional studies.

Selection criteria

We included randomized (RCTs), quasi-RCTs and cluster-RCTs, comparing high-dose to standard-dose strategies in preterm infants. High-dose strategies were defined as a high-loading dose (more than 20 mg of caffeine citrate/kg) or a high-maintenance dose (more than 10 mg of caffeine citrate/kg/day). Standard-dose strategies were defined as a standard-loading dose (20 mg or less of caffeine citrate/kg) or a standard-maintenance dose (10 mg or less of caffeine citrate/kg/day). We specified three additional comparisons according to the indication for commencing caffeine: 1) prevention trials, i.e. preterm infants born at less than 34 weeks' gestation, who are at risk for apnea; 2) treatment trials, i.e. preterm infants born at less than 37 weeks' gestation, with signs of apnea; 3) extubation trials: preterm infants born at less than 34 weeks' gestation, prior to planned extubation.

Data collection and analysis

We used standard methodological procedures expected by Cochrane. We evaluated treatment effects using a fixed-effect model with risk ratio (RR) for categorical data and mean, standard deviation (SD), and mean difference (MD) for continuous data.

Main results

We included seven trials enrolling 894 very preterm infants (reported in Comparison 1, i.e. any indication).

Two studies included infants for apnea prevention (Comparison 2), four studies for apnea treatment (Comparison 3) and two for extubation management (Comparison 4); in one study, indication for caffeine administration was both apnea treatment and extubation management (reported in Comparison 1, Comparison 3 and Comparison 4).

In the high-dose groups, loading and maintenance caffeine doses ranged from 30 mg/kg to 80 mg/kg, and 12 mg/kg to 30 mg/kg, respectively; in the standard-dose groups, loading and maintenance caffeine doses ranged from 6 mg/kg to 25 mg/kg, and 3 mg/kg to 10 mg/kg, respectively.

Two studies had three study groups: infants were randomized in three different doses (two of them matched our definition of high dose and one matched our definition of standard dose); high-dose caffeine and standard-dose caffeine were compared to theophylline administration (the latter is included in a separate review).

Six of the seven included studies compared high-loading and high-maintenance dose to standard-loading and standard-maintenance dose, whereas in one study standard-loading dose and high-maintenance dose was compared to standard-loading dose and standard-maintenance dose.

High-dose caffeine strategies (administration for any indication) may have little or no effect on mortality prior to hospital discharge (risk ratio (RR) 0.86, 95% confidence of interval (CI) 0.53 to 1.38; risk difference (RD) -0.01, 95% CI -0.05 to 0.03; I^2 for RR and RD = 0%; 5 studies, 723 participants; low-certainty evidence). Only one study enrolling 74 infants reported major neurodevelopmental disability in children aged three to five years (RR 0.79, 95% CI 0.51 to 1.24; RD -0.15, 95% CI -0.42 to 0.13; 46 participants; very low-certainty evidence). No studies reported the outcome mortality or major neurodevelopmental disability in children aged 18 to 24 months and 3 to 5 years. Five studies reported bronchopulmonary dysplasia at 36 weeks' postmenstrual age (RR 0.75, 95% CI 0.60 to 0.94; RD -0.08, 95% CI -0.15 to -0.02; number needed to benefit (NNTB) = 13; I^2 for RR and RD = 0%; 723 participants; moderate-certainty evidence). High-dose caffeine strategies may have little or no effect on side effects (RR 1.66, 95% CI 0.86 to 3.23; RD 0.03, 95% CI -0.01 to 0.07; I^2 for RR and RD = 0%; 5 studies, 593 participants; low-certainty evidence). The evidence is very uncertain for duration of hospital stay (data reported in three studies could not be pooled in meta-analysis because outcomes were expressed as medians and interquartile ranges) and seizures (RR 1.42, 95% CI 0.79 to 2.53; RD 0.14, 95% CI -0.09 to 0.36; 1 study, 74 participants; very low-certainty evidence).

We identified three ongoing trials conducted in China, Egypt, and New Zealand.

Authors' conclusions

High-dose caffeine strategies in preterm infants may have little or no effect on reducing mortality prior to hospital discharge or side effects. We are very uncertain whether high-dose caffeine strategies improves major neurodevelopmental disability, duration of hospital stay or seizures. No studies reported the outcome mortality or major neurodevelopmental disability in children aged 18 to 24 months and 3 to 5 years. High-dose caffeine strategies probably reduce the rate of bronchopulmonary dysplasia.

Recently completed and future trials should report long-term neurodevelopmental outcome of children exposed to different caffeine dosing strategies in the neonatal period. Data from extremely preterm infants are needed, as this population is exposed to the highest risk for mortality and morbidity. However, caution is required when administering high doses in the first hours of life, when the risk for intracranial bleeding is highest. Observational studies might provide useful information regarding potential harms of the highest doses.

PLAIN LANGUAGE SUMMARY

High-dose caffeine compared to standard-dose caffeine in preterm newborns at risk of lung disease

Key messages

• Although caffeine is commonly given to babies born too early, its most effective dose is unclear.

• Higher doses of caffeine might improve breathing and long-term development, but it potentially has unwanted effects.

Why give caffeine to babies born too early?

Newborns born too early (preterm), especially before 28 weeks of pregnancy, have a higher risk of death, lung disease, and brain impairment than those born at or near term. For instance, some of these babies develop intellectual disabilities, blindness, or deafness.

Caffeine is widely used in preterm infants, mainly to improve breathing and reduce apneic spells (that is temporal cessation of breathing) and the need for breathing machines.

What did we want to find out?

We wanted to find out if high-caffeine dose was better than standard-dose caffeine in newborns born too early to improve:

• mortality prior to hospital discharge;

• long-term development at age 18 to 24 months of age and at 3 to 5 years.

We also wanted to find out if high-caffeine dose was associated with any unwanted effects.

What did we do?

We searched for studies that looked at high doses of caffeine compared with standard doses of caffeine in babies born too early. We compared and summarized the results of the studies and rated our confidence in the evidence, based on factors such as study methods and sizes.

What did we find?

We included seven studies in our review, on a total of 894 preterm newborns. Two studies evaluated the use of caffeine for apnea prevention, three studies its use for apnea treatment; one for extubation management (that is removing the tube placed in the windpipe); and one study evaluated caffeine used either to treat apnea or help extubation.

In the high-dose groups, the loading dose (i.e. the very first dose) ranged from 30 mg/kg to 80 mg/kg; the maintenance (i.e. regular daily) caffeine doses ranged from 12 mg/kg to 30 mg/kg.

High-dose caffeine strategies may have little or no effect on death prior to hospital discharge and unwanted side effects.

It is unclear whether the high-dose caffeine strategies reduce death prior to hospital discharge, duration of hospital stay or seizures; only one small study reported on long-term development.

No studies reported the outcome "death or long-term development" in children aged 18 to 24 months and 3 to 5 years.

High-dose caffeine strategies probably reduce rates of chronic lung disease.

There are three ongoing studies.

What are the limitations of the evidence?

We are moderately confident in the evidence on chronic lung disease because the studies were small and used methods likely to introduce errors in their results. We have little confidence in the evidence on mortality and unwanted effects because of how the studies were conducted and there are not enough studies to be certain about the results of our outcomes. We are not confident in the evidence on long-term development because of the reasons mentioned above and because only one small studies reported information. Overall, the results of the studies are unlikely to reflect the results of all the studies that have been conducted in this area, some of which have not made their results public yet.

How up to date is this evidence?

The evidence is up to date to May 2022.

SUMMARY OF FINDINGS

Summary of findings 1. High-dose compared to standard-dose strategies for any indication for preterm infants with or at risk for apnea of prematurity

High-dose compared to standard-dose strategies for any indication for preterm infants with or at risk for apnea of prematurity

Patient or population: preterm infants with or at risk for apnea of prematurity
Setting: neonatal intensive care units
Intervention: high-dose strategies
Comparison: standard-dose strategies

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with standard-dose strategies	Risk with high-dose strategies				
All-cause mortality prior to hospital discharge	Study population		RR 0.86 (0.53 to 1.38)	723 (5 RCTs)	⊕⊕⊖⊖ **Low**[a,b]	High-dose caffeine may have little or no effect in reducing all-cause mortality prior to hospital discharge compared with standard-dose caffeine.
	91 per 1000	78 per 1000 (48 to 125)	RD -0.01 (-0.05 to 0.03)			
Major neurodevelopmental disability in children aged 18 to 24 months CA	See comments	-	-	-	-	This outcome was not reported.
Major neurodevelopmental disability in children aged 3 to 5 years CA	Study population		RR 0.79 (0.51 to 1.24	46 (1 RCT)	⊕⊖⊖⊖ **Very low**[c,d]	We are uncertain whether high-dose caffeine improves major neurodevelopmental disability in children aged three to five years compared with standard-dose caffeine.
	720 per 1000		RD -0.15 (-0.42 to 0.13)			
Mortality or major neurodevelopmental disability in children aged	See comments	-	-	-	-	This outcome was not reported.

18 to 24 months and 3 to 5 years CA						
Bronchopulmonary dysplasia/chronic lung disease (BPD) at 36 weeks' postmenstrual age	Study population		RR 0.75 (0.60 to 0.94) RD -0.08 (-0.15 to -0.02) NNTB = 13	723 (5 RCTs)	⊕⊕⊕⊖ **Moderate**[a]	High-dose caffeine probably reduces bronchopulmonary dysplasia at 36 weeks' postmenstrual age compared with standard-dose caffeine.
	335 per 1000	251 per 1000 (201 to 315)				
Side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine	Study population		RR 1.66 (0.86 to 3.23) RD 0.03 (-0.01 to 0.07)	593 (5 RCTs)	⊕⊕⊖⊖ **Low**[b,e]	High-dose caffeine may have little or no effect in reducing side effects (tachycardia, agitation, or feed intolerance) compared with standard-dose caffeine.
	43 per 1000	72 per 1000 (37 to 139)				
Duration of hospital stay (days)	See comments		-	(3 RCTs)	-	Three studies reported on this outcome (Mohammed 2015; Mohd 2021; Zhao 2016). Data could not be pooled in a meta-analysis because outcomes were expressed as medians and IQRs. Mohammed 2015: high dose (median 30.5 [IQR 20 to 51.5] days), standard dose (median 35 [IQR 25 to 51.5] days). Mohd 2021: high dose (median 52 days), standard dose (median 49.50 days). IQR not reported. Zhao 2016: high dose (median 33 [IQR 25 to 49] days), standard dose (median 39 [IQR 27 to 54] days).
Seizures (clinically diagnosed; diagnosed by electroencephalography)	Study population		RR 1.42 (0.79 to 2.53) RD 0.14 (-0.09 to 0.36)	74 (1 RCT)	⊕⊖⊖⊖ **Very low**[d,f]	We are uncertain whether high-dose caffeine reduces seizures compared with standard-dose caffeine.
	324 per 1000	461 per 1000 (256 to 821)				

*The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CA: corrected age; **CI:** confidence interval; **NNTB**: number needed to treat to benefit; **RD:** risk difference; **RR:** risk ratio

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

[a]Downgraded for study limitations by one level: unclear risk of bias in four domains (selection, detection, reporting, and other bias)
[b]Downgraded for imprecision by one level: few events and wide confidence intervals.
[c]Downgraded for study limitations by one level: unclear risk of selection and reporting bias.
[d]Downgraded for imprecision by two levels: one small trial with few events.
[e]Downgraded for study limitations by one level: unclear risk of bias in four domains (selection, detection, reporting, and other bias) and high risk of performance bias.
[f]Downgraded for study limitations by one level: unclear risk of bias in two domains (selection and other bias).

Summary of findings 2. High-dose compared to standard-dose strategies for prevention of apnea for preterm infants with or at risk for apnea of prematurity

High-dose compared to standard-dose strategies for prevention of apnea for preterm infants at risk for apnea of prematurity

Patient or population: preterm infants at risk for apnea of prematurity
Setting: neonatal intensive care units
Intervention: high-dose strategies
Comparison: standard-dose strategies

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with standard-dose strategies	Risk with high-dose strategies				
All-cause mortality prior to hospital discharge	Study population		RR 1.32 (0.50 to 3.53)	152 (2 RCTs)	⊕⊖⊖⊖ **Very low**[a,b]	We are uncertain whether high-dose caffeine reduces compared with standard-dose caffeine.
	80 per 1000	106 per 1000 (40 to 282)				

			RD 0.03 (-0.06 to 0.11)			
Major neurodevelopmental disability in children aged 18 to 24 months CA	See comments	-	-	-	-	This outcome was not reported.
Major neurodevelopmental disability in children aged 3 to 5 years CA	Study population 720 per 1000		RR 0.79 (0.51 to 1.24) RD -0.15 (-0.42 to 0.13)	46 (1 RCT)	⊕⊝⊝⊝ **Very low**[c,d]	We are uncertain whether high-dose caffeine improves major neurodevelopmental disability in children aged three to five year compared with standard-dose caffeine.
Mortality or major neurodevelopmental disability in children aged 18 to 24 months and 3 to 5 years CA	See comments	-	-	-	-	This outcome was not reported.
Bronchopulmonary dysplasia/chronic lung disease (BPD) at 36 weeks' postmenstrual age	Study population 427 per 1000	380 per 1000 (260 to 555)	RR 0.89 (0.61 to 1.30) RD -0.05 (-0.20 to 0.10)	152 (2 RCTs)	⊕⊝⊝⊝ **Very low**[a,b]	We are uncertain whether high-dose caffeine reduces bronchopulmonary dysplasia compared with standard-dose caffeine.
Side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine	See comments	-	-	-	-	This outcome was not reported.
Duration of hospital stay (days)	See comments	-	-	1 RCT	⊕⊝⊝⊝ **Very low**[e,f]	Mohd 2021 reported a median hospital stay of 52 days in the high-dose group and 49.5 days in the standard-dose group (IQR not reported).
Seizures (clinically diagnosed; diagnosed by electroencephalography)	-	-	-	-	-	This outcome was not reported.

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CA: corrected age; **CI:** confidence interval; **IQR**: interquartile range; **RD:** risk difference; **RR:** risk ratio;

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

[a]Downgraded for study limitations by one level: unclear risk of bias in two domains (selection and other bias).
[b]Downgraded for imprecision by two levels: two small trials with few events and very wide confidence intervals.
[c]Downgraded for study limitations by one level: unclear selection and reporting risk of bias.
[d]Downgraded for imprecision by two levels: one small trial with few events.
[e]Downgraded for study limitations by two levels: unclear risk of bias in most domains.
[f]Downgraded for imprecision by one level: three small trials.

Summary of findings 3. High-dose compared to standard-dose strategies for treatment of apnea for preterm infants with or at risk for apnea of prematurity

High- dose compared to standard-dose strategies for treatment of apnea for preterm infants with or at risk for apnea of prematurity

Patient or population: preterm infants with or at risk for apnea of prematurity
Setting: neonatal intensive care units
Intervention: high-dose strategies
Comparison: standard-dose strategies

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with standard-dose strategies	Risk with high-dose strategies				
All-cause mortality prior to hospital discharge	Study population		RR 0.75 (0.40 to 1.40) RD -0.03 (-0.10 to 0.04)	333 (3 RCTs)	⊕⊖⊖⊖ **Very low**[a,b]	We are uncertain whether high-dose caffeine reduces all-cause mortality prior to hospital discharge compared with standard-dose caffeine.
	120 per 1000	90 per 1000 (48 to 168)				

Major neurodevelopmental disability in children aged 18 to 24 months CA	See comments	-	-	-	-	This outcome was not reported.
Major neurodevelopmental disability in children aged 3 to 5 years CA	See comments	-	-	-	-	This outcome was not reported.
Mortality or major neurodevelopmental disability in children aged 18 to 24 months and 3 to 5 years CA	See comments	-	-	-	-	This outcome was not reported.
Bronchopulmonary dysplasia/chronic lung disease (BPD) at 36 weeks' postmenstrual age	Study population 234 per 1000	168 per 1000 (110 to 259)	RR 0.72 (0.47 to 1.11) RD -0.07 (-0.15 to 0.02)	333 (3 RCTs)	⊕⊖⊖⊖ **Very low**[a,b]	We are uncertain whether high-dose caffeine reduces bronchopulmonary dysplasia compared with standard-dose caffeine.
Side effects (tachycardia, agitation or feed intolerance) leading to a reduction in dose or withholding of caffeine: high-loading and high-maintenance dose standard-loading and standard-maintenance dose	Study population 39 per 1000	76 per 1000 (21 to 272)	RR 1.92 (0.53 to 6.90) RD 0.04 (-0.04 to 0.12)	150 (2 RCTs)	⊕⊖⊖⊖ **Very low**[b,c]	We are uncertain whether high-dose caffeine reduces side effects (tachycardia, agitation, or feed intolerance) compared with standard-dose caffeine.
Duration of hospital stay (days)	See comments	-	-	-	-	Two studies reported this outcome (Mohammed 2015; Zhao 2016). Data could not be pooled in a meta-analysis because outcomes were expressed as medians and IQRs. Mohammed 2015: high dose (median 30.5 [IQR 20 to 51.5] days), standard dose (median 35 [IQR 25 to 51.5] days). Zhao 2016: high dose (median 33 [IQR 25 to 49] days), standard dose (median 39 [IQR 27 to 54] days).
Seizures (clinically diagnosed; diagnosed by electroencephalography)	See comments	-	-	-	-	This outcome was not reported.

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CA: corrected age; **CI:** confidence interval; **RD:** risk difference; **RR:** risk ratio;

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect

[a]Downgraded for study limitations by one level: unclear risk of bias in three domains (selection, detection, and reporting bias).
[b]Downgraded for imprecision by two levels: few events and very wide confidence intervals.
[c]Downgraded for study limitations by one level: unclear risk of bias in three domains (selection, detection, and reporting bias) and high risk of performance bias.

Summary of findings 4. High-dose compared to standard-dose strategies for the prevention of re-intubation for preterm infants with or at risk for apnea of prematurity

High-dose compared to standard-dose strategies for the prevention of re-intubation for preterm infants with or at risk for apnea of prematurity

Patient or population: preterm infants with or at risk for apnea of prematurity
Setting: neonatal intensive care units
Intervention: high-dose strategies
Comparison: standard-dose strategies

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with standard-dose strategies	Risk with high-dose strategies				
All-cause mortality prior to hospital discharge	Study population		RR 0.75 (0.25 to 2.30)	238 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,b]	We are uncertain whether high-dose caffeine reduces all-cause mortality prior to hospital discharge compared with standard-dose caffeine.
	57 per 1000	43 per 1000 (14 to 132)	RD -0.01 (-0.07 to 0.04)			
Major neurodevelopmental disability in children aged 18 to 24 months CA	See comments	-	-	-	-	This outcome was not reported.

Major neurodevelopmental disability in children aged 3 to 5 years CA	See comments	-	-	-	-	This outcome was not reported.
Mortality or major neurodevelopmental disability in children aged 18 to 24 months and 3 to 5 years CA	See comments	-	-	-	-	This outcome was not reported.
Bronchopulmonary dysplasia/chronic lung disease (BPD) at 36 weeks' postmenstrual age	Study population 418 per 1000	284 per 1000 (201 to 405)	RR 0.68 (0.48 to 0.97) RD -0.13 (-0.25 to -0.01) NNTB = 8	238 (1 RCT)	⊕⊕⊖⊖ **Low**[a,c]	High-dose caffeine might reduce slightly bronchopulmonary dysplasia at 36 weeks' postmenstrual age compared with standard-dose caffeine.
Side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine	See comments	-	-	-	-	This outcome was not reported.
Duration of hospital stay (days)	See comments	-	-	-	-	This outcome was not reported.
Seizures (clinically diagnosed; diagnosed by electroencephalography)	See comments	-	-	-	-	This outcome was not reported.

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CA: corrected age; **CI:** confidence interval; **NNTB**: number need to benefit; **RD:** risk difference; **RR:** risk ratio;

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

[a]Downgraded for study limitations by one level: unclear risk of bias in three domains (selection, detection, and reporting bias).
[b]Downgraded for imprecision by two levels: few events and very wide confidence intervals.
[c]Downgraded for imprecision by one level: few events.

BACKGROUND

Description of the condition

Preterm infants have increased risk of respiratory morbidity and long-term neurodevelopmental impairment. They are also susceptible to apnea of prematurity, a developmental disorder in preterm infants that is a consequence of immature respiratory control. Apnea of prematurity occurs at increasing incidence with decreasing gestational age (Henderson-Smart 1981). The pathogenesis of this condition is linked to immaturity of the central nervous system and reduced ventilatory drive. Apnea of prematurity exposes the infant to intermittent hypoxia (low oxygen in body tissues) and, potentially, hypotension (low blood pressure). It is associated with prolonged hospital stay (Eichenwald 1997), and adverse neurodevelopmental outcome (Janvier 2004), although this association has not been seen in other studies (Koons 1993). Weaning from respiratory support may be prolonged in those with poor respiratory drive. Even if extubation is achieved, frequent episodes of apnea may occur in association with respiratory failure (hypercarbia [increased carbon dioxide in the blood], hypoxemia [low oxygen in the blood], and acidosis [excessive acid in the body fluids]), leading to reintubation (Lee 2002). In addition, intermittent hypoxic episodes in the first two months of life cause free radical (unstable atoms) damage, which increases the risk of developing chronic conditions such as retinopathy of prematurity (eye disease that can happen in preterm infants), and is associated with adverse neurodevelopmental outcome (Di Fiore 2010; Poets 2015).

Description of the intervention

Interventions designed to prevent extubation failure and prevent or treat apnea include methylxanthines, various forms of respiratory support, and kinesthetic stimulation. Theophylline, aminophylline, and caffeine are three forms of methylxanthine used to increase rates of successful extubation (Henderson-Smart 2010a), and prevent or treat apnea in preterm infants (Henderson-Smart 2010b; Henderson-Smart 2010c). Of these preparations, only caffeine has a wide enough margin between therapeutic and toxic levels to allow dosage to be substantially increased while minimizing the risk of harm (Blanchard 1992). Plasma concentrations of caffeine as low as three milligrams per liter (mg/L) to four mg/L have been shown to eliminate apnea, but optimal levels range from 8 mg/L to 20 mg/L (Aranda 1979). Typically, to maintain these caffeine plasma concentration levels, a 'standard regime' is used, which comprises a loading dose of 20 milligrams per kilogram of body weight (mg/kg) of caffeine citrate (10 mg/kg of caffeine base) and a maintenance dose of 5 mg/kg/day (2.5 mg/kg/day of caffeine base) (Blanchard 1992). Preterm infants can tolerate higher doses of caffeine very well, even at serum concentrations of 70 mg/L or above (Lee 1997). The Caffeine for Apnea of Prematurity (CAP) trial (Schmidt 2006), which is the largest trial of caffeine in preterm infants, used an intravenous loading dose of 20 mg/kg caffeine citrate followed by a daily maintenance dose of 5 mg/kg. If apneas persisted, the daily maintenance dose could be increased to a maximum of 10 mg/kg caffeine citrate. Serum caffeine levels were not measured and dose adjustment (or titration) was made in response to infants' symptoms. Since this trial, most neonatal units use a similar approach.

Potential adverse effects of caffeine include jitteriness, tachycardia, and tremors (Howell 1981). Although there have been concerns that caffeine could affect the developing nervous system (Millar 2004), in the doses used in the CAP trial, caffeine reduced the incidence of cerebral palsy and cognitive delay without significantly affecting mortality, hearing loss, or blindness at two years of corrected age (Schmidt 2007). Of note, caffeine metabolism in very preterm infants increases with postnatal age, with a clearance of 1 milliliter per minute per kilogram of body weight (mL/min/kg) on the first day of life, and 12 mL/min/kg at 45 days of life (Charles 2008), resulting in a shorter half-life over time (Doyle 2016). Therefore, the use of higher doses of caffeine (i.e. maintenance dose > 10 mg/kg) after the first weeks of life might be beneficial, as suggested by a study investigating the effects on intermittent hypoxic episodes (Dobson 2017). In addition, a retrospective analysis of 198 infants born at less than 29 weeks' gestation, suggested that high caffeine concentrations were associated with lower rates of bronchopulmonary dysplasia (chronic lung disease) and a shorter hospital stay (Alur 2015). However, this observation was not confirmed by Tabacaru 2017. The use of higher doses should be monitored by measuring circulating levels, ideally with micro-sampling methods to minimize blood loss (Bruschettini 2016). The CAP trial allowed clinicians to commence and stop the study drug when they considered it clinically indicated. This led to variation in the time at starting (median 3 [interquartile range [IQR] 1 to 5] days) and stopping (median 34 [IQR 33 to 36] weeks' postmenstrual age) caffeine administration. Similar variability in practice continues to the present day, although some authors suggest continuing therapy to 37 weeks' postmenstrual age and beyond (Rhein 2014). In addition, the CAP trial did not specify which infants should be started on caffeine but asked clinicians to record their reason for starting (i.e. apnea prophylaxis, apnea treatment, or pre-extubation).

How the intervention might work

Methylxanthines have a number of mechanisms of action. They are respiratory center stimulants, they block adenosine receptors, and they improve the function of the respiratory muscles, particularly the diaphragm. Caffeine acts as an antioxidative and anti-inflammatory drug by reducing apoptosis (programmed cell death) and apoptosis-associated factors in models of oxygen-induced lung injury (Endesfelder 2020; Nagatomo 2016). In addition, beneficial effects of caffeine might be mediated through a diuretic effect, as reported in clinical (Gillot 1990), and preclinical studies (Crossley 2012). Due to differences in the maturity of hepatic and renal function amongst preterm infants, the response to varying doses may be considerably different, with regard to both potential benefits and harms (Stevenson 2007). Caffeine has been shown to reduce the duration of ventilation in preterm infants, prevent extubation failure, and reduce the incidence of patent ductus arteriosus (hole in the blood vessel that connects the aorta to the pulmonary artery) and bronchopulmonary dysplasia (Schmidt 2006). It has also been shown to increase the rate of survival without neurodevelopmental disability (Schmidt 2007). A follow-up analysis of the CAP trial showed a reduced risk of motor impairment at 11 years of age (Schmidt 2017), improved visuomotor, visuoperceptual, and visuospatial abilities (Murner-Lavanchy 2018), and no negative effects on general intelligence, attention, and behavior. Shortening duration of endotracheal intubation and overall respiratory support seem to be important to improve respiratory function when the infant grows up, as well as long-term neurodevelopmental outcomes that correlate with these short-term respiratory outcomes. The CAP trial showed that earlier discontinuation of any positive airway pressure in the caffeine

group alone, accounted for half of the beneficial long-term effects of the drug (Schmidt 2007).

Why it is important to do this review

A systematic review comparing methylxanthines with control (no treatment) suggested that methylxanthines increase the likelihood of successful extubation of preterm infants within one week of commencing treatment (Henderson-Smart 2010a). Furthermore, methylxanthines were shown to be effective in reducing the number of apneic episodes in preterm infants (Henderson-Smart 2010c). Caffeine has been shown to increase the rate of survival without neurodevelopmental disability (Schmidt 2007), and improve neurobehavioral outcomes (Murner-Lavanchy 2018), and motor outcomes at 11 years of age (Schmidt 2017). However, the optimal dose of caffeine has not been established. It is important to determine the optimal dosing regimen for caffeine to facilitate extubation in preterm infants and to prevent and treat apnea without increasing the risk of harms. Moreover, it has been suggested that maintaining higher caffeine levels might have economic benefits because of reduced length of hospital stay and need for oxygen at discharge (Alur 2015; Montenegro 2017).

Publication of the CAP trial led to "treatment creep" in the use of caffeine (Jobe 2017a; Jobe 2017b), whereby more infants were treated either earlier, at higher doses, or for longer than previously. The safety profile demonstrated in earlier trials may not persist under these conditions.

OBJECTIVES

To determine the effects of higher versus standard doses of caffeine on mortality and major neurodevelopmental disability in preterm infants with (or at risk of) apnea, or peri-extubation.

METHODS

Criteria for considering studies for this review

Types of studies

We included randomized controlled trials (RCTs) and quasi-RCTs (i.e. with parallel groups). We planned to include cluster-RCTs.

We did not consider cross-over RCTs as eligible for inclusion. We also deemed non-randomized cohort studies not to be eligible for this review, because they are prone to bias due to confounding by indication or by residual confounding – both of which may influence results of the studies (Fewell 2007; Kyriacou 2016).

Types of participants

- For prevention trials: preterm infants born at less than 34 weeks' gestation, who are at risk of apnea
- For treatment trials: preterm infants born at less than 37 weeks' gestation, with signs of apnea
- For extubation trials: preterm infants born at less than 34 weeks' gestation, prior to planned extubation

We included preterm and term infants of postmenstrual age (PMA) up to 46 weeks and 0 days.

Types of interventions

We considered any combination of high-dose and standard-dose strategies, which were defined as follows.

High-dose strategies:

- high-loading dose (> 20 mg of caffeine citrate/kg);
- high-maintenance dose (> 10 mg of caffeine citrate/kg/day).

Standard-dose strategies:

- standard-loading dose (≤ 20 mg of caffeine citrate/kg);
- standard-maintenance dose (≤ 10 mg of caffeine citrate/kg/day).

We included any duration of treatment. We defined loading dose as the total amount of caffeine administered in the first 23 hours after the initial dose. We included both parenteral and enteral administration.

Types of outcome measures

Outcomes were measured after trial entry.

Primary outcomes

- All-cause mortality prior to hospital discharge.
- Major neurodevelopmental disability: cerebral palsy, developmental delay (Bayley Mental Developmental Index (Bayley 1993; Bayley 2006), or Griffiths Mental Development Scale (Griffiths 1954), assessment greater than two standard deviations (SDs) below the mean), intellectual impairment (intelligence quotient (IQ) greater than two SDs below the mean), blindness (vision less than 6 out of 60 in both eyes), or sensorineural deafness requiring amplification (Jacobs 2013). We separately assessed outcomes at age 18 months to 24 months corrected age (CA) and at 3 to 5 years CA.
- Mortality or major neurodevelopmental disability. We separately assessed outcomes at age 18 to 24 months CA and at 3 to 5 years CA.

Secondary outcomes

- Failure to extubate within one week of commencing treatment.
- Reintubation within one week of commencing treatment.
- Failed apnea reduction (less than 50% reduction in apnea) after two to seven days, for infants treated for apnea.
- Apnea: number of episodes (defined as interruption of breathing for more than 20 seconds) after 24 hours from commencing treatment, in a 24-hour period and over one week.
- Apnea: number of infants with at least one episode (defined as interruption of breathing for more than 20 seconds) after 24 hours of commencing treatment, in a 24-hour period and over one week.
- Side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine.
- Bronchopulmonary dysplasia or chronic lung disease: defined as 28 days of oxygen exposure (NIH 1979), at 36 weeks' postmenstrual age (Jobe 2001), and using the 'physiological definition' (Walsh 2004).
- Number of days of respiratory support (mechanical ventilation, continuous positive airway pressure [CPAP], high-flow nasal cannula, non-invasive positive-pressure ventilation [NIPPV]).

- Number of days of mechanical ventilation.
- Number of days of supplemental oxygen.
- Need for mechanical ventilation (yes/no).
- Need for non-invasive respiratory support (CPAP, NIPPV, high-flow nasal cannulae).
- Duration of hospital stay.
- Neonatal mortality (up to 28 days).
- Intraventricular hemorrhage on brain ultrasound (any and severe [Papile grade three to four]) (Papile 1978).
- Cerebellar hemorrhage on brain ultrasound (yes/no).
- Magnetic resonance imaging (MRI) abnormalities at term equivalent age (yes/no), defined as white matter lesions (i.e. cavitations (Rutherford 2010) and punctate lesions (Cornette 2002), germinal matrix-intraventricular hemorrhage (Parodi 2015), or cerebellar hemorrhage (Limperopoulos 2007)).
- Periventricular leukomalacia.
- Necrotizing enterocolitis (proven [Bell stage two or greater]) (Bell 1978).
- Patent ductus arteriosus (PDA) requiring treatment (cyclo-oxygenase inhibitors or surgical ligation).
- Retinopathy of prematurity (ROP) (any and severe [stage three or greater]) (International Committee 2005).
- Seizures (clinically diagnosed or diagnosed by electroencephalography).
- Cost of neonatal care.
- Each component of the composite outcome 'major neurodevelopmental disability' (described in Primary outcomes).

Search methods for identification of studies

Electronic searches

We searched the following databases:

- Cochrane Central Register of Controlled Trials (CENTRAL; 2022, Issue 5) in the Cochrane Library (searched 17 May 2022);
- MEDLINE via Pubmed (1946 to 17 May 2022);
- Embase (1974 to 20 May 2022);
- CINAHL EBSCO (Cumulative Index to Nursing and Allied Health Literature; 1982 to 17 May 2022);
- US National Institutes of Health Ongoing Trial Register ClinicalTrials.gov (www.clinicaltrials.gov; searched 20 May 2022);
- World Health Organization International Clinical Trials Registry Platform (apps.who.int/trialsearch; searched 20 May 2022).

The full search strategy is available in Appendix 1. We did not apply language restrictions.

Searching other resources

We searched for errata or retractions for included studies published in full text on PubMed (www.ncbi.nlm.nih.gov/pubmed).

We also reviewed the reference lists of all identified articles for relevant studies not located in the primary search.

Data collection and analysis

We used standard methods of Cochrane Neonatal, as described below.

Selection of studies

We planned to use Cochrane's Screen4Me workflow to help assess the search results (not performed as we identified fewer than 1000 records). Screen4Me comprises the following three components:

- known assessments (a service that matches records in the search results to records that have already been screened and labeled as 'RCT' or 'not an RCT' in Cochrane Crowd (Cochrane's citizen science platform where the crowd help to identify and describe health evidence; www.crowd.cochrane.org);
- the RCT classifier (a machine-learning model that distinguishes RCTs from non-RCTs);
- Cochrane Crowd (www.crowd.cochrane.org), if appropriate.

For more information about Screen4Me and the evaluations that have been undertaken, please go to the Screen4Me webpage on the Cochrane Information Specialist's portal: https://community.cochrane.org/sites/default/files/uploads/Reporting_Guidance_Screen4Me_FINAL.pdf. In addition, more detailed information regarding evaluations of the Screen4Me components can be found in the following publications: Marshall 2018; McDonald 2017; Noel-Storr 2018; Thomas 2017.

Two review authors (PB, CR) independently searched for and identified eligible studies that met the inclusion criteria. We screened the titles and abstracts to identify potentially relevant citations and retrieved the full texts of all potentially relevant articles. We independently assessed the eligibility of studies by filling out eligibility forms designed in accordance with the specified inclusion criteria. We excluded studies published only in abstract form unless the final results of the study were reported and all necessary information could be ascertained from the abstract or authors, or both. We reviewed studies for relevance by assessing study design, types of participants, interventions provided, and outcome measures reported. We resolved disagreements by discussion and, if necessary, by consultation with a third review author (MB). We provided details of excluded studies in the Characteristics of excluded studies tables, along with reasons for exclusion. We contacted the study authors if details of a primary study were unclear. We recorded the selection process in sufficient detail to complete a PRISMA flow diagram (Moher 2009), and Characteristics of included studies table.

Data extraction and management

Two review authors (PB, CR) independently extracted data using a data extraction form integrated with a modified version of the Cochrane Effective Practice and Organisation of Care Group data collection checklist (Cochrane EPOC Group 2017). We piloted the form within the review team, using a sample of included studies. We extracted the following characteristics from each included study:

- methods: study design, type, duration, and completeness of follow-up (e.g. > 80%), country and location of study, informed consent, ethics approval;
- participants: sex, birth weight, gestational age, number of participants;

- interventions: initiation, dose, and duration of caffeine administration;
- outcomes: as mentioned above under Types of outcome measures;
- administrative details: study author(s), published or unpublished, year of publication, year in which the study was conducted, presence of vested interest, details of other relevant papers cited.

We resolved disagreements by discussion or by consultation with a third review author (MB). We described ongoing studies identified by our search, when available, detailing the primary author, research question(s), methods, and outcome measures, together with an estimate of the reporting date, in the Characteristics of ongoing studies tables.

If any queries arose (e.g. discrepancies in the way outcomes were defined in the studies and the definitions in Types of outcome measures), or if additional data had been required, we would have contacted the study authors for clarification. Two review authors (PB, CR) used Review Manager Web for data entry (RevMan Web 2020).

Assessment of risk of bias in included studies

Two review authors (PB, CR) used the Cochrane risk of bias tool to independently assess the risk of bias (low, high, or unclear) of all included studies for the following domains (Higgins 2011):

- sequence generation (selection bias);
- allocation concealment (selection bias);
- blinding of participants and personnel (performance bias);
- blinding of outcome assessment (detection bias);
- incomplete outcome data (attrition bias);
- selective reporting (reporting bias);
- any other bias.

We resolved any disagreements by discussion or by consultation with a third review author (WO). A more detailed description of risk of bias for each domain is given in Appendix 2.

Measures of treatment effect

We used risk ratios (RR), risk differences (RD), number needed to treat for an additional beneficial outcome (NNTB) or number needed to treat for an additional harmful outcome (NNTH) for categorical variables. We used mean differences (MD) for continuous variables. We reported 95% confidence intervals (CI) for each statistic. We replaced any within-group standard error of the mean (SEM) reported in a study with its corresponding standard deviation (SD).

Unit of analysis issues

We included all RCTs and quasi-RCTs in which the unit of allocation was the individual infant. We planned to include eligible cluster-RCTs (none were identified), and adjust the analysis using the methods stated in Chapter 23 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2022).

Dealing with missing data

We calculated a follow-up rate for each study. If we identified a loss to follow-up over 20%, we contacted the study author(s) to request additional data. If a study reported outcomes only for participants completing the study or only for participants who followed the protocol, we contacted study author(s) to ask them to provide additional information to facilitate an intention-to-treat (ITT) analysis. In instances when this was not possible, we performed a complete-case analysis.

Assessment of heterogeneity

We assessed clinical heterogeneity by comparing the distribution of important participant factors between studies and study factors (allocation concealment, blinding of outcome assessment, loss to follow-up, and treatment type and co-interventions). We assessed statistical heterogeneity by examining the I^2 statistic (Higgins 2019), a quantity that describes the proportion of variation in point estimates that is due to variability across studies rather than to sampling error. We interpreted the I^2 statistic as follows:

- 0% to 40%: might not be important;
- 30% to 60%: may represent moderate heterogeneity;
- 50% to 90%: may represent substantial heterogeneity;
- 75% to 100%: considerable heterogeneity.

We considered statistical heterogeneity to be substantial when the I^2 value was 50% or greater. In addition, we employed the Chi^2 test of homogeneity to determine the strength of evidence that heterogeneity is genuine. We explored clinical variation across studies by comparing the distribution of participant-important factors amongst studies and study factors (randomization concealment, blinding of outcome assessment, loss to follow-up, and treatment types and co-interventions). We considered a threshold of P value less than 0.10 as an indicator of whether heterogeneity (genuine variation in effect sizes) was present.

Assessment of reporting biases

We examined the possibility of within-study selective outcome reporting for each study included in the review. We searched for protocols of included studies on electronic sources such as PubMed, ClinicalTrials.gov, and the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) to assess whether outcome reporting seemed to be sufficiently complete and transparent. We planned to investigate publication bias using funnel plots if 10 or more clinical studies were included in the systematic review (Egger 1997; Higgins 2019). However, we did not investigate publication bias as fewer than 10 studies were included in our review.

Data synthesis

We performed statistical analyses according to the recommendations of Cochrane Neonatal, using RevMan Web 2020. We analyzed all infants randomized on an ITT basis. We analyzed treatment effects in the individual studies. We used a fixed-effect model to combine the data. For all meta-analyses, we synthesized data using RRs, RDs, NNTB, NNTH, MDs, and 95% CIs. We analyzed and interpreted individual studies separately when we judged meta-analysis to be inappropriate.

Subgroup analysis and investigation of heterogeneity

Tests for subgroup differences in effects are to be interpreted with caution given the potential for confounding with other study

characteristics and the observational nature of the comparisons (Deeks 2022). In particular, subgroup analyses with fewer than five studies per category are unlikely to be adequate to ascertain valid difference in effects and are not planned to be shown in our results. When subgroup comparisons were possible, we planned stratified meta-analysis and a formal statistical test for interaction to examine subgroup differences that could account for effect heterogeneity (e.g. Cochran's Q test, meta-regression) (Borenstein 2013; Higgins 2020).

Given the potential differences in the intervention effectiveness related to gestational age, birth weight, need for mechanical ventilation, dose of caffeine, and presence of titration, we planned to conduct subgroup comparisons to see if the intervention is more effective for the primary outcomes of this review.

We planned to carry out the following subgroup analyses of factors that may contribute to heterogeneity in the effects of the intervention (not performed as few studies reported the same outcomes):

- gestational age: < 28 weeks versus 28 weeks to 32 weeks versus ≥ 32 weeks, as infants with lower gestational age are more vulnerable, at high risk of mortality and morbidity, and have different caffeine metabolism;
- birth weight: < 1000 g versus 1000 g to 1500 g versus ≥ 1500 g, as infants with lower birth weight are more vulnerable, at high risk of mortality and morbidity, and have different caffeine metabolism;
- for prevention studies only: intubated newborns versus non-intubated newborns, as intubated infants are likely to be sicker and outcomes such as apneic spells, would be affected;
- loading dose within the high-dose group: very high (≥ 60 mg/kg/day) versus moderately high (between 20 mg/kg/day and 60 mg/kg/day), as higher doses might have additional benefit or harms;
- titration: dose titrated versus dose not titrated, as dose titration might affect the outcomes in either direction.

Sensitivity analysis

We planned to conduct sensitivity analyses to explore the effect of the methodological quality of studies and check to ascertain whether studies with a high risk of bias (in at least two domains) overestimated the effect of treatment (not performed as few studies reported the same outcomes). Differences in study design of included studies might also affect the results of the systematic review. We planned to perform a sensitivity analysis to compare the effects of caffeine in truly randomized trials as opposed to quasi-randomized trials (not performed as no quasi-randomized trials were identified). We also planned to perform a sensitivity analysis to evaluate the overall results with and without inclusion of studies with a high dropout rate.

Summary of findings and assessment of the certainty of the evidence

We used the GRADE approach to assess the certainty of evidence (Schünemann 2013), for the following clinically relevant outcomes:

- all-cause mortality prior to hospital discharge;
- major neurodevelopmental disability: cerebral palsy, developmental delay (Bayley Mental Developmental Index or Griffiths Mental Development Scale assessment greater than two SDs below the mean), intellectual impairment (IQ greater than two SDs below the mean), blindness (vision less than 6/60 in both eyes), or sensorineural deafness requiring amplification, in children aged 18 to 24 months and 3 to 5 years CA;
- mortality or major neurodevelopmental disability in children aged 18 to 24 months and 3 to 5 years CA;
- bronchopulmonary dysplasia (chronic lung disease) at 36 weeks' postmenstrual age;
- side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine;
- duration of hospital stay;
- clinical seizures.

We created four tables that addressed the effect of caffeine dosage in all enrolled infants:

- for caffeine administration for any indication (Summary of findings 1);
- for prevention trials (Summary of findings 2);
- for treatment trials (Summary of findings 3);
- for extubation trials (Summary of findings 4).

Two review authors (MB, CR) independently assessed the certainty of the evidence for each of the seven outcomes above. We considered evidence from RCTs as high-certainty, but downgraded the evidence by one level for serious (or two levels for very serious) limitations based upon the following: design (risk of bias), consistency across studies, directness of the evidence, precision of estimates, and presence of publication bias. We used GRADEpro GDT to create a summary of findings table to report the certainty of the evidence. The GRADE approach resulted in an assessment of the certainty of a body of evidence in one of the following four grades:

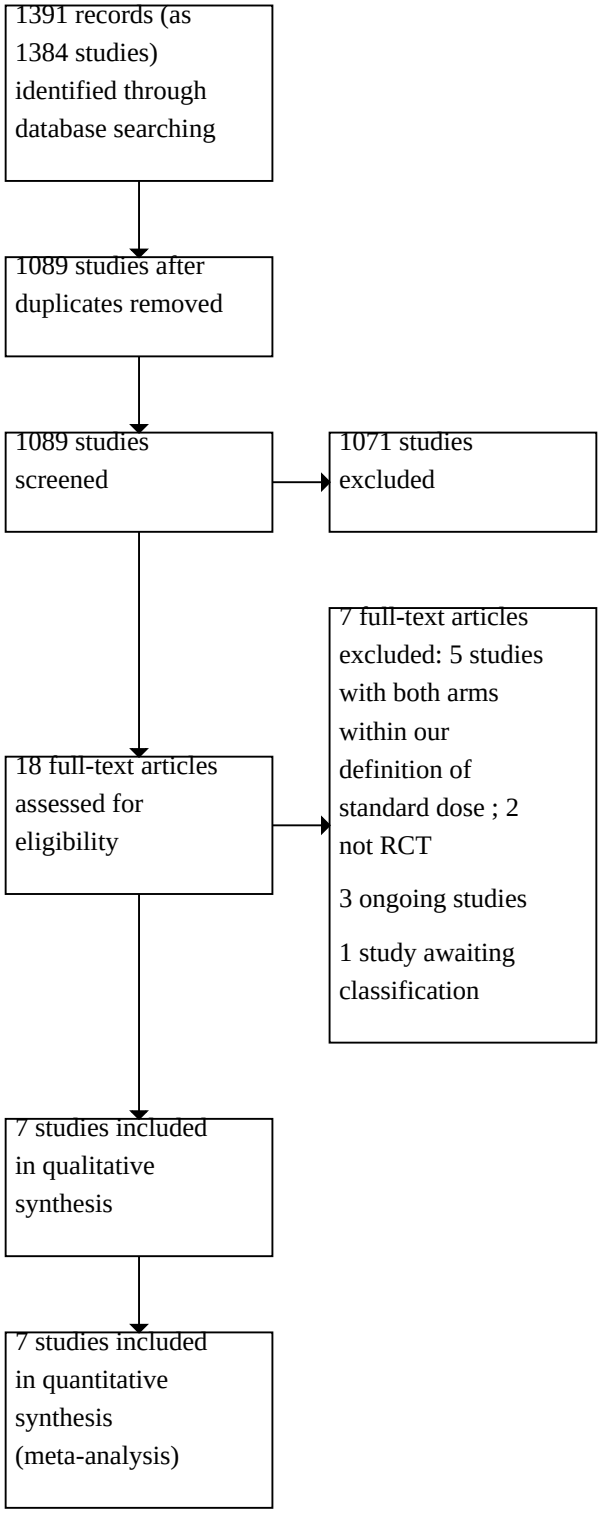
- high: we are very confident that the true effect lies close to that of the estimate of the effect;
- moderate: we are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different;
- low: our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect;
- very low: we have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of effect.

RESULTS

Description of studies

We have provided results of the search for this review update in the study flow diagram (Figure 1).

Figure 1. Study flow diagram.



For the included studies, see Characteristics of included studies and Table 1.

For the other studies, see Characteristics of excluded studies, Characteristics of studies awaiting classification and Characteristics of ongoing studies, respectively.

Results of the search

We searched the databases in May 2022 and identified 1391 references. After screening, we assessed 18 full-text articles for eligibility and included seven studies (McPherson 2015; Mohammed 2015; Mohd 2021; Scanlon 1992; Steer 2003; Steer 2004; Zhao 2016). We excluded seven studies (Autret 1985; Cherif 2003; Gray 2016; Romagnoli 1992; Wan 2020; Yao 2021; Zhang 2019), and classified one as awaiting classification (Gray 2018). We found three relevant ongoing studies by searching clinical trial registries (NCT03298347; NCT04144712; Oliphant 2020).

Included studies

Completed studies

Seven included studies (enrolling 894 preterm infants) were described in 13 separate reports: three reports for two studies (Mohammed 2015; Steer 2004), two reports for two studies (McPherson 2015; Mohd 2021), and one report each for the three remaining studies (Scanlon 1992; Steer 2003; Zhao 2016). Two studies were conducted in Australia; the others were conducted in China, Egypt, Malaysia, UK, and the USA. There were differences in methods, participants, and interventions amongst the seven included studies. Two studies included preterm infants with a gestational age less than 33 weeks (Mohd 2021; Zhao 2016), two less than 32 weeks (Mohammed 2015; Steer 2003), two less than 31 weeks (McPherson 2015; Scanlon 1992), and one less than 30 weeks (Steer 2004). The exact gestational age means values in each study and group are reported in Table 1.

In McPherson 2015, infants were randomized within 24 hours of life; in Mohd 2021 median age at enrollment was two weeks; in the other studies infants were exposed to caffeine within the first week of life.

Indication for caffeine administration was apnea prevention in two studies (McPherson 2015; Mohammed 2015), apnea treatment in three (Mohammed 2015; Scanlon 1992; Zhao 2016), and extubation management in one (Steer 2003). In one study, multiple indications were included, i.e. both apnea treatment and extubation management (Steer 2004).

In the high-dose groups, the loading and maintenance caffeine doses ranged from 30 mg/kg to 80 mg/kg, and 12 mg/kg to 30 mg/kg, respectively. In the standard-dose groups, the loading and maintenance caffeine doses ranged from 6 mg/kg to 25 mg/kg, and 3 mg/kg to 10 mg/kg, respectively.

Two studies had three study groups: in Steer 2003, infants were randomized in three different doses (two of them matched our definition of high dose and one matched out definition of standard dose); in Scanlon 1992, high- and standard-dose caffeine were compared to theophylline administration (which was not included in this review). Six of the seven included studies compared high-loading and high-maintenance dose to standard-loading and standard-maintenance dose, whereas in Zhao 2016 standard-loading and high-maintenance dose was compared to standard-loading and standard-maintenance dose.

Comparison 1 reports outcome data of all seven included trials, whereas the following analyses refer to apnea prevention, apnea treatment, and extubation management, respectively.

Detailed characteristics of the included studies are reported in Characteristics of included studies and Table 1.

Ongoing studies

We identified three ongoing trials enrolling preterm infants in China (NCT03298347), Egypt (NCT04144712), and New Zealand (Oliphant 2020).

NCT03298347 has a planned sample size of 100 preterm infants with a gestational age less than 32 weeks. Indication for administration is treatment of apnea. Infants will be randomized to a loading dose of either 80 mg/kg or 20 mg/kg of caffeine, but it is not clear if multiple or maintenance doses are planned.

NCT04144712 has a planned sample size of 80 preterm infants with two different gestational age thresholds: less than 32 weeks for prophylactic caffeine administration and 32 weeks to 34 weeks for treatment of apnea of prematurity within the first 10 days of life. All infants were either in room air or CPAP at enrollment. The loading and maintenance doses of caffeine have not been specified.

Oliphant 2020 has a planned sample size of 120 late preterm infants (i.e. gestational age 34 weeks to 36 weeks' and 6 days). Infants will be enrolled within 72 hours of birth and administered caffeine or placebo daily until term corrected age. Four different enteral caffeine doses are planned: loading doses of 10 mg/kg, 20 mg/kg, 30 mg/kg, or 40 mg/kg, and maintenance doses of 5 mg/kg, 10 mg/kg, 15 mg/kg, or 20 mg/kg of caffeine citrate.

Excluded studies

Characteristics of excluded studies

In five studies, infants were randomized to two caffeine doses which both fell within the "standard dose" definition used in this review (Autret 1985; Romagnoli 1992; Wan 2020; Yao 2021; Zhang 2019), as shown below:

- Autret 1985: loading and maintenance doses were 20 mg/kg and 5 mg/kg caffeine in group one, respectively and 20 mg/kg and 3 mg/kg in group two, respectively; Romagnoli 1992: loading and maintenance dose were 10 mg/kg and 5 mg/kg caffeine in group one, respectively and 10 mg/kg and 2.5 mg/kg in group two, respectively;
- Zhang 2019: loading and maintenance doses were 20 mg/kg and 10 mg/kg caffeine in group one, respectively and 20 mg/kg and 5 mg/kg in group two, respectively; Wan 2020: loading and maintenance doses were 20 mg/kg and 10 mg/kg caffeine in group one, respectively and 20 mg/kg and 5 mg/kg in group two, respectively;
- Yao 2021: loading and maintenance doses were 20 mg/kg and 5 mg/kg caffeine in group one and two, whereas group one did receive an additional dose caffeine one hour before ventilator weaning.

Finally, two studies were not randomized trials and were therefore excluded (Cherif 2003; Gray 2016).

Risk of bias in included studies

The risk of bias in included studies is presented in Figure 2 and Figure 3. Details of the methodological quality of each study are described in the Characteristics of included studies table.

Figure 2. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

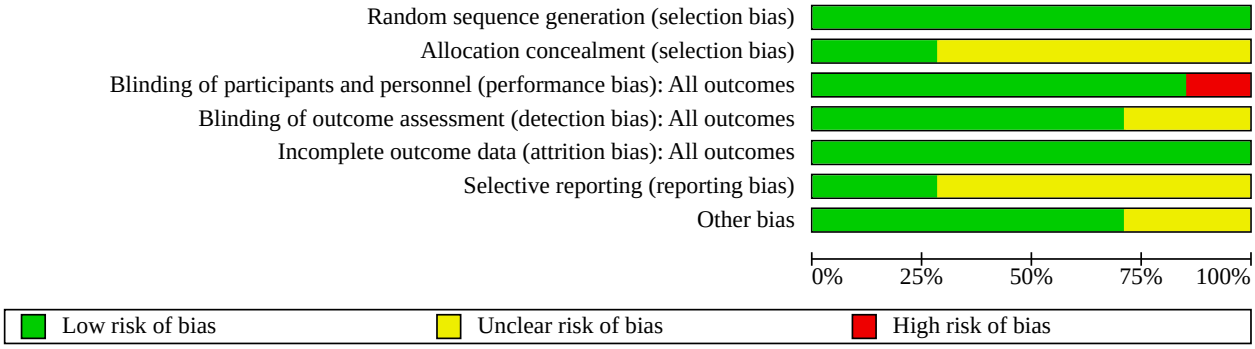


Figure 3. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.

Allocation

We judged random sequence generation to be adequate in all seven included studies. We judged two studies to be low risk of bias for allocation concealment (Mohammed 2015; Mohd 2021). We judged the other five studies to be at unclear risk of selection bias because they provided no information on allocation concealment (McPherson 2015; Scanlon 1992; Steer 2003; Steer 2004; Zhao 2016).

Blinding

We judged six of the included studies at low risk of performance bias because they were double blinded (McPherson 2015; Mohammed 2015; Mohd 2021; Steer 2003; Steer 2004; Zhao 2016). We deemed one study to be at high risk of performance bias as the two concentrations of caffeine citrate used were not identical in appearance (Scanlon 1992). We judged five studies to be at low risk of detection bias because they adequately reported that outcome assessors were blinded to treatment (McPherson 2015; Mohammed 2015; Mohd 2021; Steer 2003; Steer 2004). Scanlon 1992 did not report that outcome assessors were blinded to treatment allocation, while in Zhao 2016, assessors were not blinded for all outcomes. We judged these two studies to be at unclear risk of detection bias.

Incomplete outcome data

We judged all seven included studies to be at low risk of attrition bias because follow-up was almost complete. In Steer 2004, 41 of 287 infants were excluded after randomization and were not analyzed. This corresponds to 14% of the infants, which we judged to be still acceptable for our review.

Selective reporting

We judged two studies to be at low risk of reporting bias (McPherson 2015; Mohd 2021). We judged five studies to be at unclear risk of reporting bias. In four studies, the protocol was not available (Scanlon 1992; Steer 2003; Steer 2004; Zhao 2016). In Mohammed 2015, hydrocephalus was listed as an outcome in the trial protocol but not reported in the publication. Also, duration of CPAP and postnatal steroid therapy for bronchopulmonary dysplasia were reported in the publication but were not pre-specified outcomes in the protocol.

Other potential sources of bias

We judged five studies to be at low risk of other bias (Mohammed 2015; Scanlon 1992; Steer 2003; Steer 2004; Zhao 2016). We judged two studies to be at unclear risk of other bias. In McPherson 2015, the maternal age was higher in the high-dose caffeine group than in the standard-dose group (P = 0.03). In Mohd 2021, more infants in the standard-dose caffeine group were intubated at baseline (97.4% and 92.5 % in the standard-dose caffeine and high-dose caffeine groups, respectively) and needed surfactant (94.7% and 80%, in the standard-dose caffeine and high-dose caffeine groups, respectively).

Effects of interventions

See: **Summary of findings 1** High-dose compared to standard-dose strategies for any indication for preterm infants with or at risk for apnea of prematurity; **Summary of findings 2** High-dose compared to standard-dose strategies for prevention of apnea for preterm infants with or at risk for apnea of prematurity; **Summary of findings 3** High-dose compared to standard-dose strategies for treatment of apnea for preterm infants with or at risk for apnea of prematurity; **Summary of findings 4** High-dose compared to standard-dose strategies for the prevention of re-intubation for preterm infants with or at risk for apnea of prematurity

The seven included studies (894 infants) comparing different doses of caffeine are reported in Comparison 1. These studies are also analyzed by indication for caffeine administration:

- two studies on apnea prevention are reported in Comparison 2 (McPherson 2015; Mohd 2021);
- four studies on apnea treatment are reported in Comparison 3 (Mohammed 2015; Scanlon 1992; Steer 2004 Zhao 2016);
- two studies on extubation management are reported in Comparison 4 (Steer 2003; Steer 2004).

The study with multiple indications, i.e. both apnea treatment and extubation management (Steer 2004), is pooled in the overall Comparison 1, and not in a separate comparison "multiple indications". Steer 2004 provided separate outcome data for infants treated for apnea (see Comparison 3) and extubation management (see Comparison 4).

Comparison 1: high-dose versus standard-dose strategies for any indication

Within Comparisons 1 and 3, outcome data from Zhao 2016, when available, are reported within the second subgroup, i.e. "standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose". The other six studies are pooled in the first subgroup, i.e. "high-loading and high-maintenance dose standard-loading and standard-maintenance dose".

The certainty of the evidence is reported for the seven outcomes specified for the summary of findings table (see Summary of findings 1).

Primary outcome: all-cause mortality prior to hospital discharge

Five studies reported this outcome (McPherson 2015; Mohammed 2015; Mohd 2021; Steer 2004; Zhao 2016). High-dose caffeine may have little or no effect in reducing all-cause mortality prior to hospital discharge compared with standard-dose caffeine (RR 0.86, 95% CI 0.53 to 1.38; I^2 = 0%; RD -0.01, 95% CI -0.05 to 0.03; I^2 = 0%; 5 studies, 723 participants; low-certainty evidence; Analysis 1.1).

Subgroup analysis: high-loading and high-maintenance dose versus standard-loading and standard-maintenance dose

Four studies reported this outcome (McPherson 2015; Mohammed 2015; Mohd 2021; Steer 2004). We are uncertain whether high-loading and high-maintenance dose reduces all-cause mortality prior to hospital discharge compared with standard-loading and standard-maintenance dose (RR 0.92, 95% CI 0.52 to 1.63; I^2 = 0%; RD -0.01, 95% -0.05 to 0.04; I^2 = 0%; 4 studies, 559 participants; low-certainty evidence; subgroup analysis 1.1.1 in Analysis 1.1).

Subgroup analysis: standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose

Zhao 2016 reported this outcome. We are uncertain whether standard-loading and high-maintenance dose reduces all-cause mortality prior to hospital discharge compared with standard-loading and standard-maintenance dose (RR 0.73, 95% CI

0.31 to 1.71; RD -0.04, 95% CI -0.13 to 0.06; 164 participants; very low-certainty evidence; subgroup analysis 1.1.2 in Analysis 1.1).

Primary outcome: major neurodevelopmental disability

No studies reported this outcome in children aged 18 to 24 months CA.

McPherson 2015 reported this outcome in children aged three years to five years CA. We are uncertain whether high-dose caffeine reduces major neurodevelopmental disability compared with standard-dose caffeine (RR 0.79, 95% CI 0.51 to 1.24; RD -0.15, 95% CI -0.42 to 0.13; 46 participants; very low-certainty evidence; Analysis 1.2).

Primary outcome: mortality or major neurodevelopmental disability

No studies reported death or disability in children aged 18 to 24 months and those aged 3 to 5 years CA.

Steer 2004 (outcome data in the latest publication for this study in 2011) reported death or disability at 12 months. The number of events for this outcome was 16 out of 140 and 25 out of 147 infants in the high-dose and standard-dose caffeine group, respectively. We are uncertain whether high-dose caffeine reduces mortality or major neurodevelopmental disability at 12 months compared with standard-dose caffeine.

Secondary outcome: failure to extubate within one week of commencing treatment

Four studies reported this outcome (Mohammed 2015; Mohd 2021; Steer 2003; Steer 2004). High-dose caffeine likely results in a large reduction of failure to extubate within one week of commencing treatment compared with standard-dose caffeine (RR 0.54, 95% CI 0.40 to 0.73; I^2 = 0%; RD -0.15, 95% CI -0.23 to -0.08; I^2 = 10%; number needed to treat for an additional benefit outcome [NNTB] = 7; 4 studies, 521 participants; moderate-certainty evidence; Analysis 1.3).

Subgroup analysis: high-loading and high-maintenance dose versus standard-loading and standard-maintenance dose

Four studies reported this outcome (Mohammed 2015; Mohd 2021; Steer 2003; Steer 2004). High-loading and high-maintenance dose likely results in a large reduction of failure to extubate within one week of commencing treatment compared with standard-loading and standard-maintenance dose (RR 0.54, 95% CI 0.40 to 0.73; I^2 = 0%; RD -0.15, 95% CI -0.23 to -0.08; I^2 = 10%; NNTB = 7; 4 studies, 521 participants; Analysis 1.3).

Standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose

Zhao 2016 reported this outcome. However, without specifying the number of intubated infants, we could not calculate the effect size.

Secondary outcome: reintubation within one week of commencing treatment

Only one study reported this outcome (Steer 2004). High-loading and high-maintenance dosing of caffeine may reduce the risk of reintubation within one week of commencing treatment compared with standard-loading and standard-maintenance dose (RR 0.36, 95% CI 0.19 to 0.71; RD -0.15, 95% CI -0.24 to -0.06; NNTB = 7, 95% CI 4 to 17; 238 participants; Analysis 1.4).

Secondary outcome: failed apnea reduction after two to seven days, for infants treated with apnea

Scanlon 1992 reported this outcome within 48 hours, rather than after two to seven days. We are uncertain whether high-loading and high-maintenance dosing of caffeine reduces failed apnea reduction after two to seven days compared with standard-loading and standard-maintenance dose (RR 0.38, 95% CI 0.02 to 8.59; RD -0.06, 95% CI -0.23 to 0.10; 10 participants). This study compared high-loading and high-maintenance dose to standard-loading and standard-maintenance.

Secondary outcome: apnea: number of episodes after 24 hours from commencing treatment, in a 24-hour period and over one week

Four studies reported this outcome (Mohammed 2015; Steer 2003; Steer 2004; Zhao 2016). Data could not be pooled in a meta-analysis because outcomes were expressed as medians and interquartile ranges (IQRs):

- Mohammed 2015: median 9 (IQR 6 to 16) episodes in the high-dose group and 16 (IQR 14 to 17) episodes in the standard-dose group:
- Steer 2003: median 0.2 (IQR 0 to 13) episodes in the very high-dose group, 0.4 (IQR 0 to 11) episodes in the high-dose group, and 1.3 (IQR 0 to 14) episodes in the standard-dose group:
- Steer 2004: median 4 (IQR 1 to 12) episodes in the high-dose group and 7 (IQR 2 to 22) episodes in the standard-dose group (only data on extubation management were reported);
- Zhao 2016: median 10 (IQR 8 to 15) episodes in the high-dose group and 18 IQR 13 to 22) episodes in the standard-dose group.

Secondary outcome: apnea: number of infants with at least one episode

One study reported this outcome (Mohd 2021). We are uncertain whether high-loading and high-maintenance dosing of caffeine reduces apnea (number of infants with at least one episode) compared with standard-loading and standard-maintenance dose reported this outcome (RR 1.01, 95% CI 0.59 to 1.75; RD 0.01, 95% CI -0.21 to 0.22; 78 participants; Analysis 1.5). This study compared high-loading and high-maintenance dose to standard-loading and standard-maintenance.

Secondary outcome: side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine

Five studies reported this outcome (Mohammed 2015; Mohd 2021; Scanlon 1992; Steer 2003; Steer 2004). High-dose caffeine may have little or no effect in reducing side effects (tachycardia, agitation, or feed intolerance) compared with standard-dose caffeine (RR 1.66, 95% CI 0.86 to 3.23; I^2 = 0%; RD 0.03, 95% CI -0.01 to 0.07; I^2 = 0%; 593 participants; low-certainty evidence; Analysis 1.6). These studies compared high-loading and high-maintenance dose to standard-loading and standard-maintenance dose.

Secondary outcome: BPD: 28 days of oxygen exposure

One study reported this outcome (Steer 2004). We are uncertain whether high-loading and high-maintenance dosing of caffeine reduces BPD defined as 28 days of oxygen exposure compared with standard-loading and standard-maintenance dose reported

this outcome(RR 0.84, 95% CI 0.68 to 1.04; RD -0.10, 95% CI -0.23 to 0.02; 238 participants; Analysis 1.7).

Secondary outcome: BPD: at 36 weeks' postmenstrual age

Five studies reported this outcome (McPherson 2015; Mohammed 2015; Mohd 2021; Steer 2004; Zhao 2016). High-dose caffeine probably reduces bronchopulmonary dysplasia at 36 weeks' postmenstrual age compared with standard-dose caffeine (RR 0.75, 95% CI 0.60 to 0.94; I^2 = 0%; RD -0.08, 95% CI -0.15 to -0.02; I^2 = 0%; NNTB = 13; 723 participants; moderate-certainty evidence; Analysis 1.8).

Subgroup analyses: high-loading and high-maintenance dose versus standard-loading and standard-maintenance dose

Four studies reported this outcome (McPherson 2015; Mohammed 2015; Mohd 2021; Steer 2004). High-loading and high-maintenance dose probably reduces bronchopulmonary dysplasia at 36 weeks' postmenstrual age compared with standard-loading and standard-maintenance dose (RR 0.76, 95% CI 0.60 to 0.97; I^2 = 0%; RD -0.09, 95% CI -0.16 to -0.01; I^2 = 0%; NNTB = 13; 4 studies, 559 participants; subgroup analysis 1.11.1 in Analysis 1.8).

Subgroup analyses: standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose

Zhao 2016 reported this outcome. We are uncertain whether standard-loading and high-maintenance dose reduces bronchopulmonary dysplasia at 36 weeks' postmenstrual age compared with standard-loading and standard-maintenance dose (RR 0.68, 95% CI 0.36 to 1.29; RD -0.07, 95% CI -0.19 to 0.05; 164 participants; subgroup analysis 1.11.2 in Analysis 1.8).

Secondary outcome: number of days using respiratory support

Four studies reported this outcome (McPherson 2015; Mohammed 2015; Mohd 2021; Steer 2004). Data could not be pooled in a meta-analysis because they were expressed as medians and IQRs:

- McPherson 2015: median 4 (IQR 1 to 22) ventilator days in the high-dose group and 3 (IQR 1 to 22) in the standard-dose group, type of respiratory support not specified;
- Mohammed 2015: median 3 (IQR 1 to 10) CPAP days in the high-dose group compared to 5 (IQR 1 to 10) days in those given a standard dose of caffeine;
- Mohd 2021: median 18.50 in the high-dose group and 22 in the standard-dose group (IQRs not reported);
- Steer 2004: median 5.0 (IQR 2.0 to 14.5) days in the high-dose group and 6.9 (IQR 2 to 16.7) days in the standard-dose group; median duration of NCPAP was 7.2 (IQR 2 to 17) days in the high-dose group and 8.4 (IQR 4 to 19) days in the standard-dose group.

Secondary outcome: number of days using mechanical ventilation

Five studies reported this outcome (McPherson 2015; Mohammed 2015; Steer 2003; Steer 2004; Zhao 2016). McPherson 2015 and Steer 2003 were pooled in a meta-analysis (MD 0.01, 95% CI -0.85 to 0.87; I^2 = 0%, 2; Analysis 1.9). Three studies were not pooled in a meta-analysis because outcomes were expressed as medians and IQRs:

- Mohammed 2015: median 3.5 (IQR 1 to 10) days in the high-dose group and 5 (IQR 2 to 13) days in the standard-dose group;
- Steer 2004: median 5.0 (IQR 2.0 to 14.5) in the high-dose group and 6.9 (IQR 2 to 16.7) days in the standard-dose group;
- Zhao 2016: median 5.5 (IQR 1 to 8) days in the high-dose group and 8.5 (IQR 2 to 11) days in the standard-dose group.

Secondary outcome: number of days using supplemental oxygen

Three studies reported this outcome (Mohammed 2015; Mohd 2021; Zhao 2016). Data could not be pooled in a meta-analysis because outcomes were expressed as medians and IQRs:

- Mohammed 2015 median 14.5 (IQR 5 to 28) days in the high-dose group and 20 (IQR 9 to 39) days in the standard-dose group;
- Mohd 2021: high dose 0 (IQR not clearly reported) and standard dose 0 (IQR not clearly reported);
- Zhao 2016: median 15 (IQR 7 to 26) days in the high-dose group and 21 (IQR 11 to 35) days in the standard-dose groups.

Secondary outcome: duration of hospital stay

Three studies reported this outcome (Mohammed 2015; Mohd 2021; Zhao 2016). We are uncertain whether high-dose caffeine reduces the duration of hospital stay compared with standard-dose caffeine. Data could not be pooled in a meta-analysis because outcomes were expressed as medians and IQRs:

- Mohammed 2015: median 30.5 (IQR 20 to 51.5) days in the high-dose group and 35 (IQR 25 to 51.5) days in the standard-dose group;
- Mohd 2021: median 52 days in the high-dose group and 49.5 days in the standard-dose group (IQRs not reported);
- Zhao 2016: median 33 (IQR 25 to 49) days in the high-dose group and 39 (IQR 27 to 54) days in the standard-dose group.

Secondary outcome: intraventricular hemorrhage at brain ultrasound, any grade

Two studies reported this outcome (McPherson 2015; Steer 2004. High dose may result in little to no difference in intraventricular hemorrhage at brain ultrasound any grade compared to standard dose (RR 0.90, 95% CI 0.63 to 1.27; I^2 = 0%; RD -0.03, 95% CI -0.12 to 0.06; I^2 = 0%; 361 participants; Analysis 1.10).

Secondary outcome: intraventricular hemorrhage at brain ultrasound, severe (Papile grade 3 to 4)

Five studies reported this outcome (McPherson 2015; Mohammed 2015; Mohd 2021; Steer 2003; Steer 2004). High dose may result in little to no difference in severe intraventricular hemorrhage at brain ultrasound compared to standard dose (RR 1.26, 95% CI 0.67 to 2.36; I^2 = 0%; RD 0.01, 95% CI -0.02 to 0.05; I^2= 0%; 686 participants; Analysis 1.11).

Secondary outcome: cerebellar hemorrhage at brain ultrasound

One study reported this outcome (McPherson 2015). High dose caffeine may increase cerebellar hemorrhage at brain ultrasound compared to standard dose (RR 3.33, 95% CI 1.00 to 11.15; RD 0.19, 95% CI 0.02 to 0.36; number needed to treat for an additional harmful outcome [NNTH] = 5 (3 to 50); 74 participants; Analysis 1.12).

Secondary outcome: MRI abnormalities at term equivalent age

One study reported this outcome (McPherson 2015) reported this outcome. High dose may result in little to no difference in MRI abnormalities at term equivalent age compared to standard dose (RR 0.50, 95% CI 0.10 to 2.56; RD -0.05, 95% CI -0.18 to 0.07; 74 participants; Analysis 1.13).

Secondary outcome: periventricular leukomalacia

Three studies (McPherson 2015; Mohammed 2015; Mohd 2021), reported this outcome. High dose may result in little to no difference in periventricular leukomalacia compared to standard dose (RR 1.42, 95% CI 0.56 to 3.61; I^2 = 0%; RD 0.02, 95% CI -0.04 to 0.08; I^2 = 0%; 272 participants; Analysis 1.14).

Steer 2004 reported periventricular leukomalacia as part of the composite outcome "Major cerebral abnormalities on ultrasound", defined as one or more of porencephalic cysts, cystic periventricular leukomalacia, or hydrocephalus. Major cerebral abnormalities were found on ultrasound in 5 out of 140 participants in the high-dose group and 11 out of 147 in the standard-dose group, However, data could not be pooled as separate data for periventricular leukomalacia was not reported.

Secondary outcome: necrotizing enterocolitis (proven = Bell stage 2 or greater)

Five studies reported this outcome (McPherson 2015; Mohammed 2015; Mohd 2021; Steer 2003; Steer 2004). High dose may result in little to no difference in necrotizing enterocolitis compared to standard dose (RR 0.80, 95% CI 0.43 to 1.51; I^2 = 0%; RD -0.01, 95% CI -0.05 to 0.02; I^2 = 0%; 637 participants; Analysis 1.15).

Secondary outcome: patent ductus arteriosus (PDA) requiring treatment

Two studies reported this outcome (McPherson 2015; Steer 2003). High dose may result in little to no difference in PDA requiring treatment compared to standard dose (RR 0.87, 95% CI 0.59 to 1.27; I^2 = 0%; RD -0.05, 95% CI -0.17 to 0.08; I^2 = 0%; 201 participants; Analysis 1.16).

Secondary outcome: retinopathy of prematurity (ROP), any ROP

Two studies reported this outcome (Mohd 2021; Steer 2004). High dose may result in little to no difference in any ROP compared to standard dose (RR 0.85, 95% CI 0.56 to 1.29; I^2 = 38%; RD -0.03, 95% CI -0.11 to 0.05; I^2 = 51%; 365 participants; Analysis 1.17).

Secondary outcome: retinopathy of prematurity (ROP), severe ROP (stage 3 or greater)

Three studies (McPherson 2015; Mohammed 2015; Steer 2004) reported this outcome. High dose may result in little to no difference in severe ROP compared to standard dose (RR 0.57, 95% CI 0.27 to 1.20; I^2 = 0%; RD -0.03, 95% CI -0.07 to 0.01; I^2= 0%; 481 participants; Analysis 1.18).

Secondary outcome: seizures (clinically diagnosed; diagnosed by electroencephalography)

McPherson 2015 reported this outcome. We are uncertain whether high-dose caffeine reduces seizures compared with standard-dose caffeine (RR 1.42, 95% CI 0.79 to 2.53; RD 0.14, 95% CI -0.09 to 0.36; 74 participants; very low-certainty evidence; Analysis 1.19).

Secondary outcome: developmental delay (Bayley Mental Developmental Index or Griffiths Mental Development Scale in children aged 18 to 24 months)

One study (McPherson 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces developmental delay compared with standard-dose caffeine (RR 1.45, 95% CI 0.52 to 4.03; RD 0.10, 95% CI -0.17 to 0.37; 44 participants; Analysis 1.20).

Secondary outcome: Bayley-III cognitive score in children at 18 to 24 months CA

One study (McPherson 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces Bayley-III cognitive score compared with standard-dose caffeine (MD -1.90, 95% CI -8.60 to 4.80; 44 participants; Analysis 1.21).

Secondary outcome: cerebral palsy in children aged 18 to 24 months CA

One study (McPherson 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces cerebral palsy compared with standard-dose caffeine (RR 1.82, 95% CI 0.34 to 9.77; RD 0.07, 95% CI -0.13 to 0.27, 44 participants, 1 study; Analysis 1.22).

Secondary outcome: blindness in children aged 18 to 24 months CA

One study (McPherson 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces blindness compared with standard-dose caffeine (no events, RR not estimable; RD 0.00, 95% CI -0.09 to 0.09; 44 participants; Analysis 1.23).

Secondary outcome: deafness in children aged 18 to 24 months CA

One study (McPherson 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces deafness compared with standard-dose caffeine (no events, RR not estimable; RD 0.00, 95% CI -0.09 to 0.09; 44 participants; Analysis 1.24).

Secondary outcomes

The following secondary outcomes were not reported:

- BPD using the 'physiological definition';
- need for mechanical ventilation;
- need for non-invasive respiratory support;
- neonatal mortality; and
- cost of neonatal care.

Comparison 2: high-dose versus standard-dose strategies for prevention of apnea

Two studies enrolled infants administered caffeine for prevention of apnea (McPherson 2015; Mohd 2021). Within this comparison, the certainty of the evidence is very low for all outcomes (see Summary of findings 2).

Within Comparisons 2 and 4, no subgroups by dose strategy are created because all studies are pooled in "high-loading and high-maintenance dose standard-loading and standard-maintenance dose".

Primary outcome: all-cause mortality prior to hospital discharge

Two studies reported this outcome (McPherson 2015; Mohd 2021). We are uncertain whether high-dose caffeine reduces all-cause mortality prior to hospital discharge compared with standard-dose caffeine (RR 1.32, 95% CI 0.50 to 3.53; I^2 = 0%; RD 0.03, 95% CI -0.06 to 0.11; I^2 = 0%; 152 participants; Analysis 2.1).

Primary outcome: major neurodevelopmental disability

McPherson 2015 reported Bayley-III score in children aged 18 to 24 months CA (analyzed 24 out of 37 and 22 out of 37 infants in the high-dose and standard-dose caffeine group, respectively). Cognitive scores were 85.6 and 88.0 in the high-dose and standard-dose caffeine group, respectively. Language scores were 90.5 and 88.9 in the high-dose and standard-dose caffeine group, respectively. Motor scores were 85.3 and 85.9 in the high-dose and standard-dose caffeine group, respectively. Data were reported without SD and are not represented in any forest plot.

McPherson 2015 reported this outcome in children aged three to five years. We are uncertain whether high-dose caffeine reduces major neurodevelopmental disability compared with standard-dose caffeine (RR 0.79, 95% CI 0.51 to 1.24; RD -0.15, 95% CI -0.42 to 0.13; 46 participants; Analysis 2.2).

Primary outcome: mortality or major neurodevelopmental disability

No studies reported death or disability in children aged 18 to 24 months CA and those aged 3 to 5 years.

Secondary outcome: failure to extubate within one week of commencing treatment

Mohd 2021 reported this outcome. We are uncertain whether high-dose caffeine reduces failure to extubate within one week of commencing treatment compared with standard-dose caffeine (RR 0.95 95% CI 0.34 to 2.69; RD -0.01, 95 % CI -0.17 to 0.15; 78 participants; Analysis 2.3).

Secondary outcome: apnea: number of infants with at least one episode

One study reported this outcome (Mohd 2021). We are uncertain whether high-dose caffeine reduces apnea (number of infants with at least one episode) compared with standard-dose caffeine (RR 1.01, 95% CI 0.59 to 1.75; RD 0.01 95% CI -0.21 to 0.22; 78 participants; Analysis 2.4).

Secondary outcome: side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine

Mohd 2021 reported this outcome. We are uncertain whether high-dose caffeine reduces side effects (tachycardia, agitation, or feed intolerance) compared with standard-dose caffeine (RR 1.19, 95% CI 0.34 to 4.09; RD 0.02, 95% CI: -0.12 to 0.16; 78 participants; Analysis 2.5).

Secondary outcome: BPD: at 36 weeks' postmenstrual age

Two studies reported this outcome (McPherson 2015; Mohd 2021). We are uncertain whether high-dose caffeine reduces BPD at 36 weeks' postmenstrual age compared with standard-dose caffeine (RR 0.89, 95% CI 0.61 to 1.30; I^2 = 13%; RD -0.05, 95% CI -0.20 to 0.10; I^2 = 0%; 152 participants; Analysis 2.6).

Secondary outcome: number of days using respiratory support

Two studies reported this outcome (McPherson 2015; Mohd 2021). Data could not be pooled in a meta-analysis because they were expressed as medians and IQRs:

- McPherson 2015: median 4 (IQR 1 to 22) ventilator days in the high-dose group and 3 (IQR 1 to 22) in the standard-dose group, type of respiratory support not specified;
- Mohd 2021: median 18.50 days in the high-dose group and 22 days in the standard-dose group (IQRs not reported).

Secondary outcome: number of days using mechanical ventilation

One study (McPherson 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces number of days using mechanical ventilation compared with standard-dose caffeine (MD 3.50, 95% CI -5.64 to 12.64; 74 participants; Analysis 2.7).

Secondary outcome: number of days using supplemental oxygen

One study (Mohd 2021) reported this outcome. We are uncertain whether high-dose caffeine reduces number of days using supplemental oxygen compared with standard-dose caffeine: high dose 0 (IQR not reported) standard dose 0 (IQR not reported).

Secondary outcome: duration of hospital stay

One study reported this outcome (Mohd 2021). We are uncertain whether high-dose caffeine reduces duration of hospital stay compared with standard-dose caffeine (median duration of hospital stay was 52 days in infants treated with high-dose caffeine compared to 49.5 days for infants treated with standard-dose caffeine; The IQRs were not reported).

Secondary outcome: intraventricular hemorrhage at brain ultrasound, any grade

One study (McPherson 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces intraventricular hemorrhage any grade compared with standard-dose caffeine (RR 0.83, 95% CI 0.41 to 1.69; RD -0.05, 95% CI -0.26 to 0.15; 74 participants; Analysis 2.8).

Secondary outcome: intraventricular hemorrhage at brain ultrasound, severe (Papile grade 3 to 4)

Two studies (McPherson 2015; Mohd 2021) reported this outcome. We are uncertain whether high-dose caffeine reduces severe intraventricular hemorrhage compared with standard-dose caffeine (RR 0.68, 95% CI 0.22 to 2.15; I^2 = 2%; RD -0.03, 95% CI -0.11 to 0.05; I^2 = 0%; 152 participants; Analysis 2.9).

Secondary outcome: cerebellar hemorrhage at brain ultrasound; MRI abnormalities at term equivalent age

One study reported this outcome (McPherson 2015). High dose caffeine may increase cerebellar hemorrhage at brain ultrasound compared to standard dose (RR 3.33, 95% CI 1.00 to 11.15; RD 0.19, 95% CI 0.02 to 0.36; 74 participants; Analysis 2.10).

Secondary outcome: MRI abnormalities at term equivalent age

One study reported this outcome (McPherson 2015). High dose may result in little to no difference in MRI abnormalities at term equivalent age compared to standard dose (RR 0.50, 95% CI 0.10 to 2.56; RD -0.05, 95% CI -0.18 to 0.07; 74 participants; Analysis 2.11).

Secondary outcome: periventricular leukomalacia

Two studies (McPherson 2015; Mohd 2021) reported this outcome. High dose may result in little to no difference in periventricular leukomalacia compared to standard dose (RR 1.64, 95% CI 0.41 to 6.59; RD 0.03, 95% CI -0.05 to 0.10; I^2 = 0%; 152 participants; Analysis 2.12).

Secondary outcome: necrotizing enterocolitis (proven = Bell stage of 2 or greater)

Two studies (McPherson 2015; Mohd 2021) reported this outcome. High dose may result in little to no difference in necrotizing enterocolitis compared to standard dose (RR 1.23, 95% CI 0.51 to 2.93; I^2 = 0%; RD 0.02, 95% CI -0.08 to 0.13; I^2 = 0%; 152 participants; Analysis 2.13).

Secondary outcome: patent ductus arteriosus (PDA) requiring treatment

One study reported this outcome (McPherson 2015). High dose may result in little to no difference in PDA requiring treatment compared to standard dose (RR 1.00, 95% CI 0.66 to 1.52; RD 0.00, 95% CI -0.23 to 0.23; 74 participants; Analysis 2.14).

Secondary outcome: retinopathy of prematurity (ROP), any ROP

One study reported this outcome (McPherson 2015). High dose may result in little to no difference in any ROP compared to standard dose (RR 1.66, 95% CI 0.53 to 5.23; RD 0.07, 95% CI -0.08 to 0.22; 74 participants; Analysis 2.15).

Secondary outcome: retinopathy of prematurity (ROP), severe ROP (stage 3 or greater)

One study reported this outcome (McPherson 2015). High dose may result in little to no difference in severe ROP compared to standard dose (RR 0.50, 95% CI 0.10 to 2.56; RD -0.05, 95% CI -0.18 to 0.07; 74 participants; Analysis 2.16).

Secondary outcome: seizures (clinically diagnosed; diagnosed by electroencephalography)

McPherson 2015 reported this outcome. We are uncertain whether high-dose caffeine reduces seizures compared with standard-dose caffeine (RR 1.42, 95% CI 0.79 to 2.53; RD 0.14, 95% CI -0.09 to 0.36; 74 participants; Analysis 2.17).

Secondary outcome: developmental delay (Bayley Mental Developmental Index or Griffiths Mental Development Scale in children aged 18 to 24 months

One study (McPherson 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces developmental delay compared with standard-dose caffeine (RR 1.45, 95% CI 0.52 to 4.03; RD 0.10, 95% CI -0.17 to 0.37; 44 participants; Analysis 2.18).

Secondary outcome: Bayley-III cognitive score in children at 18 to 24 months CA

One study (McPherson 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces Bayley-III cognitive score compared with standard-dose caffeine (MD -1.90, 95% CI -8.60 to 4.80; 44 participants; Analysis 2.19).

Secondary outcome: cerebral palsy in children aged 18 to 24 months CA

One study (McPherson 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces cerebral palsy compared with standard-dose caffeine (RR 1.82, 95% CI 0.34 to 9.77; RD 0.07, 95% CI -0.13 to 0.27; 44 participants; Analysis 2.20).

Secondary outcome: blindness in children aged 18 to 24 months CA

One study (McPherson 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces blindness compared with standard-dose caffeine (no events, RR not estimable; RD 0.00, 95% CI -0.09 to 0.09; 44 participants; Analysis 2.21).

Secondary outcome: deafness in children aged 18 to 24 months CA

One study (McPherson 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces deafness compared with standard-dose caffeine (no events, RR not estimable; RD 0.00, 95% CI -0.09 to 0.09; 44 participants, 1 study; Analysis 2.22).

Secondary outcomes

The following secondary outcomes were not reported:

- mortality or major neurodevelopmental disability;
- reintubation within one week of commencing treatment;
- failed apnea reduction after two to seven days, for infants treated with apnea;
- apnea: number of episodes;
- BPD: 28 days of oxygen exposure;
- BPD: using the 'physiological definition';
- need for mechanical ventilation;
- need for non-invasive respiratory support;
- neonatal mortality; and
- cost of neonatal care.

Comparison 3: high-dose versus standard-dose strategies for treatment of apnea

Four studies enrolled infants administered caffeine for treatment of apnea (Mohammed 2015; Scanlon 1992; Steer 2004; Zhao 2016). For Steer 2004, within comparison 3, we include outcome data for infants treated for apnea and exclude outcome data for infants treated for prevention of re-intubation, which are reported in comparison 4. Within this comparison, the certainty of the evidence is very low (see Summary of findings 3).

Within comparison 1 and 3, outcome data from Zhao 2016, when available, are reported within the second subgroup, i.e. "standard-loading and high-maintenance dose standard-loading and standard-maintenance dose". The other studies are pooled in

the first subgroup, i.e. "high-loading and high-maintenance dose standard-loading and standard-maintenance dose".

Primary outcome: all-cause mortality prior to hospital discharge

Three studies reported this outcome (Mohammed 2015; Steer 2004; Zhao 2016). We are uncertain whether high-dose caffeine reduces all-cause mortality prior to hospital discharge compared with standard-dose caffeine(RR 0.75, 95% CI 0.40 to 1.40; RD -0.03, 95% CI -0.10 to 0.04; I^2 for RR and RD = 0%; 333 participants; Analysis 3.1).

Subgroup analysis: high-loading and high-maintenance dose versus standard-loading and standard-maintenance dose

Two studies reported this outcome (Mohammed 2015; Steer 2004). We are uncertain whether high-loading and high-maintenance dose reduces all-cause mortality prior to hospital discharge compared with standard-loading and standard-maintenance dose (RR 0.78, 95% CI 0.31 to 1.95; I^2 = 0%; RD -0.02, 95% CI -0.11 to 0.07; I^2 = 0%; 169 participants; subgroup analysis 3.1.1 in Analysis 3.1).

Subgroup analysis: standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose

Zhao 2016 reported this outcome. We are uncertain whether standard-loading and high-maintenance dose reduces all-cause mortality prior to hospital discharge compared with standard-loading and standard-maintenance dose (RR 0.73, 95% CI 0.31 to 1.71; RD -0.04, 95% CI -0.13 to 0.06; 164 participants; subgroup analysis 3.1.1 in Analysis 3.1).

Primary outcome: major neurodevelopmental disability

No studies reported this outcome in children aged 18 to 24 months and those aged 3 to 5 years.

Steer 2004 (outcome data in the latest publication for this study, i.e. 2011) reported major disability at 12 months corrected for prematurity. This time point was not specified in our protocol as it is a difficult age to obtain accurate data on neurodevelopment. We are uncertain whether high-dose caffeine reduces major neurodevelopmental disability compared with standard-dose caffeine. Data could not be pooled with the other studies or represented in a forest plot. Major disability (cerebral palsy, bilateral blindness, and need for hearing aids) at 12 months corrected for prematurity: 3 out of 24 and 1 out of 25 infants in the high and standard dose group, respectively; Deaths up to 12 months of age: 0 out of 24 and 2 out of 25 infants in the high and standard dose group, respectively.

Primary outcome: mortality or major neurodevelopmental disability

No studies reported death or disability in children aged 18 to 24 months CA and those aged 3 to 5 years.

Steer 2004 (outcome data in the latest publication for this study, i.e. 2011) reported death or disability at 12 months: 3 out of 24 and 3 out of 25 infants in the high and standard dose group, respectively. We are uncertain whether high-dose caffeine reduces mortality or major neurodevelopmental disability compared with standard-dose caffeine.

Secondary outcome: failure to extubate within one week of commencing treatment

Mohammed 2015 reported this outcome. High-dose caffeine probably reduces failure to extubate within one week of commencing treatment compared with standard-dose caffeine (RR 0.47, 95% CI 0.24 to 0.92; RD -0.25, 95% CI -0.45 to -0.04; NNTB = 4; 78 participants; Analysis 3.2).

Subgroup analysis: high-loading and high-maintenance dose versus standard-loading and standard-maintenance dose

Mohammed 2015 reported this outcome. High-dose caffeine probably reduces failure to extubate within one week of commencing treatment compared with standard-dose caffeine (RR 0.47, 95% CI 0.24 to 0.92; RD -0.25, 95% CI -0.45 to -0.04; 78 participants; subgroup analysis 3.2.1 in Analysis 3.2).

Subgroup analysis: standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose

One study (Zhao 2016) reported this outcome however without specifying the number of intubated infants: effect size can not be calculated.

Secondary outcome: failed apnea reduction after two to seven days, for infants treated with apnea

One study (Scanlon 1992) reported this outcome however within 48 hours, and not after two to seven days. We are uncertain whether high dose reduces failed apnea reduction compared with standard dose (RR 0.38, 95% CI 0.02 to 8.59; RD -0.06, 95% CI -0.23 to 0.10; 30 participants).

Secondary outcome: apnea: number of episodes after 24 hours from commencing treatment, in a 24-hour period and over one week

Mohammed 2015 and Zhao 2016 reported this outcome. Data could not be pooled in a meta-analysis because outcomes were expressed as medians and IQRs:

- Mohammed 2015: median 9 (IQR 6 to 16) episodes in the high-dose group and 16 (IQR 14 to 17) episodes in the standard-dose group;
- Zhao 2016: median 10 (IQR 8 to 15) episodes in the high-dose group and 18 (IQR 13 to 22) episodes in the standard-dose group.

Secondary outcome: side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine

Two studies reported this outcome (Mohammed 2015; Scanlon 1992). We are uncertain whether high-dose caffeine reduces side effects (tachycardia, agitation, or feed intolerance) compared with standard-dose caffeine (RR 1.92, 95% CI 0.53 to 6.90; I^2 = 26%; RD 0.04, 95% CI -0.04 to 0.12; I^2 = 47%; 150 participants; Analysis 3.3).

Secondary outcome: BPD: at 36 weeks' postmenstrual age

Three studies reported this outcome (Mohammed 2015; Steer 2004; Zhao 2016). We are uncertain whether high-dose caffeine reduces BPD at 36 weeks' postmenstrual age compared with standard-dose caffeine (RR 0.72, 95% CI 0.47 to 1.11; I^2 = 0%; RD -0.07, 95% CI -0.15 to 0.02; I^2 = 26%; 333 participants; Analysis 3.4).

Subgroup analysis: high-loading and high-maintenance dose versus standard-loading and standard-maintenance dose

Two studies (Mohammed 2015; Steer 2004) reported this outcome. We are uncertain whether high-dose caffeine reduces BPD at 36 weeks' postmenstrual age compared with standard-dose caffeine (RR 0.75, 95% CI 0.42 to 1.35; I^2 = 0%; RD -0.06, 95% CI -0.18 to 0.06; 169 participants; I^2 = 59%; subgroup analysis 3.4.1 in Analysis 3.4).

Subgroup analysis: standard-loading and high-maintenance dose standard-loading and standard-maintenance dose

One study (Zhao 2016) reported this outcome. We are uncertain whether high-dose caffeine reduces BPD at 36 weeks' postmenstrual age compared with standard-dose caffeine (RR 0.68, 95% CI 0.36 to 1.29; RD -0.07, 95% CI -0.19 to 0.05; 164 participants; subgroup analysis 3.4.2 in Analysis 3.4).

Secondary outcome: number of days using respiratory support

One study reported this outcome (Mohammed 2015). Infants treated with high-dose-caffeine spent a median of 3 (IQR 1 to 10) days on CPAP compared to 5 (IQR 1 to 10) days for infants treated with standard-dose caffeine.

Secondary outcome: number of days using mechanical ventilation

Two studies reported this outcome (Mohammed 2015; Zhao 2016). Data could not be pooled in a meta-analysis because they were expressed as medians and IQRs:

- Mohammed 2015: median 3.5 (IQR 1 to 10) days in the high-dose group and 5 (IQR 2 to 13) days in the standard-dose group;
- Zhao 2016: median 5.5 (IQR 1 to 8) days in the high-dose group and 8.5 (IQR 2 to 11) days in the standard-dose group.

Secondary outcome: number of days using supplemental oxygen

Two studies reported this outcome (Mohammed 2015; Zhao 2016). Data could not be pooled in a meta-analysis because outcomes were expressed as median and IQR:

- Mohammed 2015: median 14.5 (IQR 5 to 28) days in the high-dose group and 20 (IQR 9 to 39) days in the standard-dose group;
- Zhao 2016: median 15 (IQR 7 to 26) days in the high-dose group and 21 (IQR 11 to 35) days in the standard-dose group.

Secondary outcome: duration of hospital stay

Two studies reported this outcome (Mohammed 2015; Zhao 2016). We are uncertain whether high-dose caffeine reduces duration of hospital stay compared with standard-dose caffeine. Data could not be pooled in a meta-analysis because outcomes were expressed as medians and IQRs:

- Mohammed 2015: median 30.5 (IQR 20 to 51.5) days in the high-dose group and 35 (IQR 25 to 51.5) days in the standard-dose group;
- Zhao 2016: median 33 (IQR 25 to 49) days in the high-dose group and 39 (IQR 27 to 54) days in the standard-dose group.

Secondary outcome: intraventricular hemorrhage at brain ultrasound, any grade

One study (Steer 2004) reported this outcome. We are uncertain whether high-dose caffeine reduces intraventricular hemorrhage any grade compared with standard-dose caffeine (RR 0.21, 95% CI 0.03 to 1.66; RD -0.16, 95% CI -0.33 to 0.02; 49 participants; Analysis 3.5).

Secondary outcome: intraventricular hemorrhage at brain ultrasound, severe (Papile grade 3 to 4)

Two studies (Mohammed 2015; Steer 2004) reported this outcome. We are uncertain whether high-dose caffeine reduces severe intraventricular hemorrhage compared with standard-dose caffeine (RR 1.14, 95% CI 0.46 to 2.82; I^2 = 0%; RD 0.01, 95% CI -0.08 to 0.10; I^2 = 0%; 169 participants; Analysis 3.6).

Secondary outcome: periventricular leukomalacia

One study (Mohammed 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces periventricular leukomalacia compared with standard-dose caffeine (RR 1.25, 95% CI 0.35 to 4.43; RD 0.02, 95% CI -0.08 to 0.11; 120 participants; Analysis 3.7).

Secondary outcome: necrotizing enterocolitis (proven = Bell stage of 2 or greater)

One study (Mohammed 2015) reported this outcome. We are uncertain whether high-dose caffeine reduces necrotizing enterocolitis compared with standard-dose caffeine (RR 0.67, 95% CI 0.20 to 2.24; RD -0.03, 95% CI -0.13 to 0.07; 120 participants; Analysis 3.8).

Secondary outcome: retinopathy of prematurity (ROP), any ROP

One study (Steer 2004) reported this outcome. We are uncertain whether high-dose caffeine reduces any ROP compared with standard-dose caffeine (RR 4.17, 95% CI 0.50 to 34.66; RD 0.13, 95% CI -0.04 to 0.29; 49 participants; Analysis 3.9).

Secondary outcome: retinopathy of prematurity (ROP), severe ROP (stage 3 or greater)

Two studies (Mohammed 2015; Steer 2004) reported this outcome. We are uncertain whether high-dose caffeine reduces severe ROP compared with standard-dose caffeine (RR 0.85, 95% CI 0.29 to 2.54; I^2 = 0%; RD -0.01, 95% CI -0.09 to 0.07; I^2 = 0%; 169 participants; Analysis 3.10).

Secondary outcomes

The following secondary outcomes were not reported:

- major neurodevelopmental disability children aged three to five years;
- mortality or major neurodevelopmental disability in children aged three to five years;
- apnea: number of infants with at least one episode;
- BPD: 28 days of oxygen exposure;
- BPD: using the 'physiological definition';
- need for mechanical ventilation;
- need for non-invasive respiratory support;
- neonatal mortality;
- cerebellar hemorrhage at brain ultrasound;
- MRI abnormalities at term equivalent age;
- patent ductus arteriosus (PDA) requiring treatment (cyclo-oxygenase inhibitors or surgical ligation);

- seizures (clinically diagnosed; diagnosed by electroencephalography); and
- cost of neonatal care.

Comparison 4: high-dose versus standard-dose strategies for prevention of re-intubation

Two studies enrolled infants administered caffeine for prevention of re-intubation (Steer 2003; Steer 2004). For Steer 2004, within Comparison 4, we include outcome data for infants treated for prevention of re-intubation and exclude for infants treated for apnea, which are reported in Comparison 3. Within this comparison, the certainty of the evidence is very low for all outcomes due to imprecision of the estimate, except for bronchopulmonary dysplasia at 36 weeks' (Analysis 4.6) which was low certainty evidence (see Summary of findings 4).

Within comparison 2 and 4, no subgroups by dose strategy were created because all studies were pooled in "high-loading and high-maintenance dose standard-loading and standard-maintenance dose".

Primary outcome: all-cause mortality prior to hospital discharge

Steer 2004 reported this outcome. We are uncertain whether high-dose caffeine reduces all-cause mortality prior to hospital discharge compared with standard-dose caffeine (RR 0.75, 95% CI 0.25 to 2.30; RD: -0.01, 95% CI -0.07 to 0.04; 238 participants; Analysis 4.1).

Primary outcome: major neurodevelopmental disability

No studies reported this outcome in children aged 18 to 24 months CA and those aged 3 to 5 years.

Steer 2004 reported about major disability at 12 months corrected for prematurity. This time point was not specified in our protocol as it is a difficult age to obtain accurate data on neurodevelopment. We are uncertain whether high-dose caffeine reduces major neurodevelopmental disability compared with standard-dose caffeine. Data could not be pooled with the other studies or represented in a forest plot. Major disability (cerebral palsy, bilateral blindness, and need for hearing aids) at 12 months corrected for prematurity: 6 out of 116 and 14 out of 122 infants in the high and standard dose group, respectively; deaths up to 12 months of age: 7 out of 116 and 8 out of 122 infants in the high and standard dose group, respectively; Death or disability at 12 months: 13 out of 116 and 22 out of 122 infants in the high and standard dose group, respectively.

Primary outcome: mortality or major neurodevelopmental disability

No studies reported this outcome in children aged 18 to 24 months CA and those aged 3 to 5 years

Steer 2004 reported death or disability at 12 months: 13 out of 116 and 22 out of 122 infants in the high and standard dose group, respectively. We are uncertain whether high-dose caffeine reduces mortality or major neurodevelopmental disability compared with standard-dose caffeine.

Secondary outcome: failure to extubate within one week of commencing treatment

Two studies reported this outcome (Steer 2003; Steer 2004). We are uncertain whether high-dose caffeine reduces failure to extubate within one week of commencing treatment compared with standard-dose caffeine (RR 0.52 95% CI 0.36 to 0.74; I^2 = 0%; RD -0.17, 95 % CI -0.26 to -0.08; I^2 = 0%; 365 participants; Analysis 4.2).

Secondary outcome: reintubation within one week of commencing treatment

Only one study reported this outcome (Steer 2004). High-loading and high-maintenance dosing of caffeine may reduce the risk of reintubation within one week of commencing treatment compared with standard-loading and standard-maintenance dose (RR 0.36, 95% CI 0.19 to 0.71; RD -0.15, 95% CI -0.24 to -0.06; NNTB = 7, 95% CI 4 to 17; 238 participants; Analysis 4.3).

Secondary outcome: apnea: number of episodes after 24 hours from commencing treatment, in a 24-hour period and over one week

Two studies reported this outcome (Steer 2003; Steer 2004). Data could not be pooled in a meta-analysis because outcomes were expressed as medians and IQRs:

- Steer 2003: median 0.2 (IQR 0 to 13) episodes in the very high-dose group, 0.4 (IQR 0 to 11) episodes in the high-dose group, and 1.3 (IQR 0 to 14) episodes in the standard-dose group;
- Steer 2004: median 4 (IQR 1 to 12) episodes in the high-dose group and 7 (IQR 2 to 22) episodes in the standard-dose group.

Secondary outcome: side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine

Two studies reported this outcome (Steer 2003; Steer 2004). We are uncertain whether high-dose caffeine reduces side effects (tachycardia, agitation, or feed intolerance) compared with standard-dose caffeine (RR 1.86, 95% CI 0.68 to 5.09; I^2 = 0%; RD 0.03, 95% CI -0.02 to 0.07; I^2 = 0%; 365 participants; Analysis 4.4).

Secondary outcome: BPD: 28 days of oxygen exposure

One study reported this outcome (Steer 2004). We are uncertain whether high-loading and high-maintenance dosing of caffeine reduces BPD defined as 28 days of oxygen exposure compared with standard-loading and standard-maintenance dose reported this outcome(RR 0.84, 95% CI 0.68 to 1.04; RD -0.10, 95% CI -0.23 to 0.02; 238 participants; Analysis 4.5).

Secondary outcome: BPD: at 36 weeks' postmenstrual age

Steer 2004 reported this outcome. High-dose caffeine might reduce slightly bronchopulmonary dysplasia at 36 weeks' postmenstrual age compared with standard-dose caffeine (RR 0.68, 95% CI 0.48 to 0.97; RD -0.13, 95% CI: -0.25 to -0.01; NNTB = 8; 238 participants; Analysis 4.6).

Secondary outcome: number of days using respiratory support

Steer 2004 reported the median duration of mechanical ventilation was 7.4 (IQR 3.3 to 16.5) days in the high-dose caffeine group and 9.0 (IQR 0.5 to 77) days in the standard-dose caffeine group. For NCPAP,

median duration was 10.1 (IQR 2.3 to 21.2) and 9.8 (IQR 4.3 to 20.1) in the high-dose and standard-dose groups, respectively.

Secondary outcome: number of days using mechanical ventilation

Two studies reported this outcome (Steer 2003; Steer 2004). Steer 2004 could not be included in the meta-analysis as outcomes were expressed as median and IQR: high dose 7.4 (IQR 3.3 to 16.5) days and standard dose 9.0 (IQR 0.5 to 77) days. For Steer 2003: MD -0.02, 95% CI -0.89 to 0.85; I^2 = 0%; Analysis 4.7

Secondary outcome: intraventricular hemorrhage at brain ultrasound, any grade

One study (Steer 2004) reported this outcome. High dose may result in little to no difference in intraventricular hemorrhage at brain ultrasound any grade compared to standard dose (RR 1.02, 95% CI 0.68 to 1.53; RD 0.01, 95% CI -0.11 to 0.12; 238 participants; Analysis 4.8).

Secondary outcome: intraventricular hemorrhage at brain ultrasound, severe (Papile grade 3 to 4)

Two studies (Steer 2003; Steer 2004) reported this outcome. High dose may result in little to no difference in severe intraventricular hemorrhage at brain ultrasound compared to standard dose (RR 4.08, 95% CI 0.74 to 22.55; I^2 = 0%; RD 0.03, 95% CI -0.00 to 0.06; I^2 = 0%; 365 participants; Analysis 4.9).

Secondary outcome: periventricular leukomalacia

Steer 2004 reported periventricular leukomalacia as part of the composite outcome "major cerebral abnormality at six weeks", defined as one or more of the following: cerebral cystic formation (porencephalic cystic or periventricular leucomalacia or encephaloclastic porencephaly) or hydrocephalus. We are uncertain whether high dose reduces he composite outcome "major cerebral abnormality at six weeks" compared with standard dose (7 out of 116 and 10 out of 122 infants in the high-dose and standard-dose caffeine group, respectively).

Secondary outcome: necrotizing enterocolitis (proven = Bell stage of 2 or greater)

Two studies (Steer 2003; Steer 2004) reported this outcome. High dose may result in little to no difference in necrotizing enterocolitis compared to standard dose (RR 0.38, 95% CI 0.08 to 1.79; I^2 = 60%; RD -0.02, 95% CI -0.05 to 0.01; I^2 = 43%; 365 participants; Analysis 4.10).

Secondary outcome: patent ductus arteriosus (PDA) requiring treatment

One study (Steer 2003) reported this outcome. High dose may result in little to no difference in PDA requiring treatment compared to standard dose (RR 0.65, 95% CI 0.30 to 1.41; RD -0.08, 95% CI -0.22 to 0.07; 127 participants; Analysis 4.11).

Secondary outcome: retinopathy of prematurity (ROP), any ROP

One study (Steer 2004) reported this outcome. High dose may result in little to no difference in any ROP compared to standard dose (RR 0.65, 95% CI 0.40 to 1.05; RD -0.10, 95% CI -0.20 to 0.01; 238 participants; Analysis 4.12).

Secondary outcome: retinopathy of prematurity (ROP), severe ROP (stage 3 or greater)

One study (Steer 2004) reported this outcome. High dose may result in little to no difference in severe ROP compared to standard dose (RR 0.39, 95% CI 0.11 to 1.45; RD -0.04, 95% CI -0.09 to 0.01; 238 participants; Analysis 4.13).

Secondary outcomes

The following outcomes were not reported:

- major neurodevelopmental disability children aged 18 to 24 months and those aged 3 to 5 years;
- mortality or major neurodevelopmental disability in children aged 18 to 24 months and those aged 3 to 5 years;
- failed apnea reduction after two to seven days, for infants treated with apnea;
- apnea: number of infants with at least one episode;
- BPD: using the 'physiological definition';
- number of days using supplemental oxygen;
- need for mechanical ventilation;
- need for non-invasive respiratory support;
- duration of hospital stay;
- neonatal mortality; cerebellar hemorrhage at brain ultrasound;
- MRI abnormalities at term equivalent age;
- periventricular leukomalacia;
- seizures (clinically diagnosed; diagnosed by electroencephalography); and
- cost of neonatal care.

DISCUSSION

Summary of main results

We evaluated the benefits and harms of high compared to low dose strategies for preterm infants. Amongst the seven included studies (894 infants), the indication for caffeine administration was apnea prevention in two studies (McPherson 2015; Mohammed 2015), apnea treatment in three (Mohammed 2015; Scanlon 1992; Zhao 2016), and extubation management in one (Steer 2003). In one study, multiple indications were included, i.e. both apnea treatment and extubation management (Steer 2004). Six of the seven included studies compared high-loading and high-maintenance dose to standard-loading and standard-maintenance dose. The remaining study compared a standard-loading and high-maintenance dose to a standard-loading and standard-maintenance dose (Zhao 2016).

The use of high-dose strategies may have little or no effect on all-cause mortality prior to hospital discharge. Amongst the other primary outcomes of this review, we are uncertain whether high or low dose improves major neurodevelopmental disability because of serious imprecision of the estimates and limitations in study design; none of the studies reported the composite outcome mortality or major neurodevelopmental disability. High dose probably reduces extubation failure (by nearly 50%) and bronchopulmonary dysplasia (approximately 25% lower rate). High-dose caffeine strategies might have little or no effect on side effects; we are uncertain whether they reduce seizure rates and duration of hospital stay.

Different dosage regimens were used across studies, but no clear relationship to effect could be found. We identified three ongoing studies; one of them will enroll late preterm infants, i.e. gestational age 34 weeks to 36 weeks' and six days.

Overall completeness and applicability of evidence

To date, seven studies comparing high-dose versus low-dose caffeine in preterm infants have enrolled 894 newborns. Study authors reported extremely limited data on critical outcomes such as long-term neurodevelopmental assessment. We could not perform an appropriate a priori subgroup analysis to detect differential effects because of the paucity of the included studies. However, we created separate comparisons for caffeine indication, i.e. apnea prevention, apnea treatment, and extubation management. We identified three ongoing studies. Larger trials are required to obtain clear conclusions.

Quality of the evidence

According to the GRADE approach, the overall certainty of evidence for critical outcomes for caffeine administration for any indication ranged from moderate to very low (see Summary of findings 1). All outcomes were downgraded (one level) because of limitations in study design, i.e. unclear high risk of bias in different domains, mainly selection bias, detection bias and reporting bias. Critical outcomes with moderate certainty of the evidence were extubation failure and chronic lung disease. Mortality prior to hospital discharge and side effects were rated low certainty, i.e. were downgraded due to limitations in study design (one level) and imprecision of the estimates (one level). Major neurodevelopmental disability and the composite outcome mortality or major neurodevelopmental disability at 24 months of corrected age were rated very low certainty, i.e. were downgraded due to limitations in study design (one level) and imprecision of the estimates (two levels).

When the studies were analyzed by indication (i.e. apnea prevention, apnea treatment, extubation management), the small number of trials and limited sample size resulted in very low certainty for all outcomes.

We did not explore publication bias using funnel plots because fewer than 10 studies met the inclusion criteria of this Cochrane Review.

Potential biases in the review process

We used the standard methods of Cochrane Neonatal in conducting this systematic review. It is unlikely that the literature search missed relevant studies. We are confident that this systematic review summarizes all the presently available evidence from RCTs comparing different caffeine dosing regimens in preterm infants. We applied no language restrictions. We succeeded in obtaining additional information from study authors. Five studies were excluded because the doses did not match our inclusion criteria, i.e. both doses fell within the "standard dose" definition used in this review (Autret 1985; Romagnoli 1992; Wan 2020; Yao 2021; Zhang 2019). We excluded two studies (Cherif 2003; Gray 2016), because were not randomized trials. The authors of this Cochrane Review are not involved in any of the included studies. However, some of us conducted primary studies (both clinical and preclinical) on caffeine administration to preterm newborns: this might generate an intellectual bias in preparing this review.

The main limitations of this review are the post-hoc changes to our published protocol: to harmonize this review with "Methylxanthine for the prevention and treatment of apnea in preterm infants" (Marques 2021), and to optimally address a request from the World Health Organization, we changed the structure of the comparisons, reporting by indication rather than by dose strategies. As a consequence, we removed the following subgroup analyses: chronological age; timing of caffeine initiation; post-extubation respiratory support; and high- and standard-loading dose. Moreover, we removed intubation (intubated newborns; non-intubated newborns) for indications other than prevention of apnea; changed the definition of high-loading dose (more than 20 mg of caffeine citrate/kg to more than 25 mg of caffeine citrate/kg); and we included studies with different maintenance doses in the comparison high-loading dose versus standard-loading dose (in the protocol we specified "each arm of this comparison had to give identical maintenance doses following the different loading doses"). Overall, we believe that these changes led to improved analyses and reporting of this review without introducing bias to the review process.

Agreements and disagreements with other studies or reviews

A number of non-Cochrane systematic reviews have been published on the same topic (Brattström 2019; Chen 2018; Saroha 2020; Vliegenthart 2018). Two reviews authored by our group (Brattström 2019; Vliegenthart 2018), used lower definitions of high and standard dose and structured the comparisons differently. The main findings of these non-Cochrane reviews are in agreement with this Cochrane Review, despite the addition of several recent trials. Another systematic review, conducted without a pre-registered protocol, identified only three randomized trials (Pakvasa 2018). Major limitations affect other reviews, e.g. Chen 2018, where definition thresholds for high and standard dose were not specified.

The findings of this Cochrane Review contrast with several narrative reviews. In one review "Caffeine for preterm infants: fixed standard dose, adjustments for age or high dose?", reviewers concluded that high-dose caffeine may increase the risk of cerebellar hemorrhage and seizures (Saroha 2020). However, according to the GRADE methodology, the certainty of the evidence for these outcomes is very low. Moreover, these harms have been identified in studies where the caffeine was given in the very first hours of life, when risk for damage is highest: interpretation of these findings is complicated (Davis 2020; Nylander Vujovic 2020), and should probably be restricted to the studies on prevention of apnea.

AUTHORS' CONCLUSIONS

Implications for practice

High-dose caffeine strategies in preterm infants with or at risk for apnea of prematurity may have little or no effect on reducing mortality prior to hospital discharge or side effects (tachycardia, agitation, or feed intolerance). We are uncertain whether high-dose caffeine strategies improves major neurodevelopmental disability or duration of hospital stay. No studies reported the outcome "mortality or major neurodevelopmental disability" in children aged 18 to 24 months and 3 to 5 years. High-dose caffeine strategies probably reduce the rate of bronchopulmonary dysplasia.

Implications for research

Recently completed and future trials should report long-term neurodevelopmental outcome of children exposed to different caffeine dosing strategies in the neonatal period. Data from extremely preterm infants are needed, as this population is exposed to the highest risk for mortality and morbidity. However, caution is required when administering high doses of caffeine in the very first hours, when the risk for intracranial bleeding is highest. Observational studies might provide useful information regarding potential harms of the highest doses.

ACKNOWLEDGEMENTS

The Methods section of this protocol is based on a standard template used by Cochrane Neonatal.

We would like to thank the following people.

- Matthias Bank (Library and ICT services, Lund University) for designing the search strategy.
- Chunli Lu (Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine) and Iris PL Wong (Tian, Xiuying, Department of Neonatology, Tianjin Central Hospital of Gynecology and Obstetrics, Tianjin China) for translating studies published in Mandarin.
- Kerry Dwan (Methods Support Unit, Editorial and Methods department, Cochrane) for methodological support for outcome assessment.
- Cochrane Neonatal: Michelle Fiander and Fiona Russell, Managing Editors, for selecting peer reviewers, collating peer-reviewer comments, providing editorial guidance to authors, and editing support.
- Bill McGuire (Co-coordinating Editor) provided editorial support and was the Sign-off Editor.
- Georg Schmölzer and Gautham Suresh peer reviewed and offered feedback for the protocol of this review.
- Jemma Cleminson (Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, UK); and Sivam Thanigainathan (All India Institute of Medical Sciences, Jodhpur, India) peer reviewed and offered feedback for this review.
- Lindsay Robertson for copy-editing.

REFERENCES

References to studies included in this review

McPherson 2015 *{published data only}*

McPherson C, Lean RE, Cyr PP, Inder TE, Rogers CE, Smyser CD. Five-year outcomes of premature infants randomized to high or standard loading dose caffeine. *Journal of Perinatology* 2022;**42**(5):631-5. [DOI: 10.1038/s41372-022-01333-5] [PMID: 35145209]

* McPherson C, Neil JJ, Tjoeng TH, Pineda R, Inder TE. A pilot randomized trial of high-dose caffeine therapy in preterm infants. *Pediatric Research* 2015;**78**(2):198-204. [DOI: 10.1038/pr.2015.72] [PMID: 25856169]

Vesoulis ZA, McPherson C, Neil JJ, Mathur AM, Inder TE. Early high-dose caffeine increases seizure burden in extremely preterm neonates: a preliminary study. *Journal of Caffeine Research* 2016;**6**(3):101-7. [DOI: 10.1089/jcr.2016.0012] [PMID: 27679737]

Mohammed 2015 *{published data only}*

* Mohammed S, Nour I, Shabaan AE, Shouman B, Abdel-Hady H, Nasef N. High versus low-dose caffeine for apnea of prematurity: a randomized controlled trial. *European Journal of Pediatrics* 2015;**174**(7):949-56. [DOI: 10.1007/s00431-015-2494-8] [PMID: 25644724]

Mohammed S, Nour I, Shabaan AE, Shouman B, Abdel-Hady H, Nasef N. High versus low-dose of caffeine for apnea of prematurity: a randomized controlled trial. In: Pediatric Academic Societies Annual Meeting. 2014.

NCT02103777. High versus low dose of caffeine for apnea of prematurity. www.drugsheet.com/trial/nct02103777-high-versus-low-dose-of-caffeine-for-apnea-of-prematurity-a-double-blind-randomized-control-trial 2014.

Mohd 2021 *{published data only}*

ACTRN. Randomized controlled trial (RCT) comparing between high versus standard dose caffeine citrate for apnoea in prematurity. www.who.int/trialsearch/Trial2.aspx?TrialID=ACTRN12619001708145 2019.

* Mohd Kori AM, Van Rostenberghe H, Ibrahim NR, Yaacob NM, Nasir A. A randomized controlled trial comparing two doses of caffeine for apnoea in prematurity. *International Journal of Environmental Research and Public Health* 2021;**18**(9):4509. [DOI: 10.3390/ijerph18094509] [PMID: 33922783]

Scanlon 1992 *{published data only}*

Scanlon JE, Chin KC, Morgan ME, Durbin GM, Hale KA, Brown SS. Caffeine or theophylline for neonatal apnoea? *Archives of Disease in Childhood* 1992;**67**(4 Spec No):425-8. [DOI: 10.1136/adc.67.4_spec_no.425] [PMID: 1586184]

Steer 2003 *{published data only}*

Steer PA, Flenady VJ, Shearman A, Lee TC, Tudehope DI, Charles BG. Periextubation caffeine in preterm neonates: a randomized dose response trial. *Journal of Paediatrics and Child Health* 2003;**39**(7):511-5. [DOI: 10.1046/j.1440-1754.2003.00207.x] [PMID: 12969204]

Steer 2004 *{published data only}*

Charles BG, Townsend SR, Steer PA, Flenady VJ, Gray PH, Shearman A. Caffeine citrate treatment for extremely premature infants with apnea: population pharmacokinetics, absolute bioavailability, and implications for therapeutic drug monitoring. *Therapeutic Drug Monitoring* 2008;**30**(6):709-16. [DOI: 10.1097/FTD.0b013e3181898b6f] [PMID: 19057373]

Gray PH, Flenady VJ, Charles BG, Steer PA, Caffeine Collaborative Study Group. Caffeine citrate for very preterm infants: Effects on development, temperament and behaviour. *Journal of Paediatrics and Child Health* 2011;**47**(4):167-72. [DOI: 10.1111/j.1440-1754.2010.01943.x] [PMID: 21244548]

* Steer P, Flenady V, Shearman A, Charles B, Gray PH, Henderson-Smart D, et al, Caffeine Collaborative Study Group Steering Group. High dose caffeine citrate for extubation of preterm infants: a randomised controlled trial. *Archives of Disease in Childhood - Fetal and Neonatal Edition* 2004;**89**(6):F499-503. [DOI: 10.1136/adc.2002.023432] [PMID: 15499141]

Zhao 2016 *{published data only}*

Zhao Y, Tian X, Liu G. [Clinical effectiveness of different doses of caffeine for primary apnea in preterm infants]. *Zhonghua Er Ke Za Zhi* 2016;**54**(1):33-6. [DOI: 10.3760/cma.j.issn.0578-1310.2016.01.008] [PMID: 26791921]

References to studies excluded from this review

Autret 1985 *{published data only}*

Autret E, Breteau M, Laugier J, Gold F, Gautier C, Majzoub S. Comparison of 2 different maintenance doses of caffeine in the treatment of apnea in premature infants [Comparaison de deux doses d'entretien différentes de caféine dans le traitement des apnées du prématuré]. *Therapie* 1985;**40**(4):235-9. [PMID: 4024019]

Cherif 2003 *{published data only}*

Cherif A, Klouz A, Zouari B, Belkahia C, Boukef-Larguèche S. Monohydrated caffeine: which dosage is effective in the treatment of apnea of prematurity? [Caféine monohydratée: quelle posologie efficace dans le traitement des apnées du prématuré?]. *Archives de Pédiatrie* 2003;**10**(8):734-5. [DOI: 10.1016/s0929-693x(03)00345-2]

Gray 2016 *{published data only}*

Gray PH, Firman B. Early high dose caffeine citrate: effect on cranial ultrasound findings and neonatal neurological outcomeJournal of paediatrics and child health. In: Conference: 20th Annual Meeting of the Perinatal Society of Australia and New Zealand, PSANZ 2016. Australia. 2016.

Romagnoli 1992 *{published data only}*

Romagnoli C, De Carolis MP, Muzii U, Zecca E, Tortorolo G, Chiarotti M, et al. Effectiveness and side effects of two different doses of caffeine in preventing apnea in premature infants. *Therapeutic Drug Monitoring* 1992;**14**(1):14-9. [DOI: 10.1097/00007691-199202000-00003] [PMID: 1546385]

Wan 2020 *{published data only}*10.1002/ppul.24948

Wan L, Huang L, Chen P. Caffeine citrate maintenance doses effect on extubation and apnea postventilation in preterm infants. *Pediatric Pulmonology* 2020;**55**(10):2635-40. [DOI: 10.1002/ppul.24948] [PMID: 32639634]

Yao 2021 *{published data only}*

Yao LS, Lin XZ, Huang J, Tang LX. Clinical effect of an additional maintenance dose of caffeine before ventilator weaning in preterm infants with respiratory distress syndrome: a prospective randomized controlled trial. *Zhongguo Dang Dai Er Ke Za Zhi* 2021;**23**(1):31-6. [DOI: 10.7499/j.issn.1008-8830.2008044] [PMID: 33476534]

Zhang 2019 *{published data only}*10.7499/j.issn.1008-8830.2019.06.011

A prospective randomized controlled trial for different maintaining doses of caffeine therapy for apnea of very low birthweight infants. www.who.int/trialsearch/Trial2.aspx?TrialID=ChiCTR1800019204 2018.

Zhang X, Zhang HT, Lyu Y, Wang LF, Yang ZY. Clinical effect and safety of different maintenance doses of caffeine citrate in treatment of apnea in very low birth weight preterm infants: a prospective randomized controlled trial. *Zhongguo Dang Dai Er Ke Za Zhi* 2019;**21**(6):558-561. [DOI: 10.7499/j.issn.1008-8830.2019.06.011] [PMID: 31208509]

References to studies awaiting assessment

Gray 2018 *{published data only}*

Gray PH, Molnar A, Firman B. Early high-dose caffeine citrate: neonatal and neurodevelopmental outcomes. *American Journal of Perinatology* 2018;**35**(S01):A025. [DOI: 10.1055/s-0038-1647088]

References to ongoing studies

NCT03298347 *{published data only}*

Yuan S, Juan M. Caffeine for preterm infants with apnea of prematurity (AOP). www.ClinicalTrials.gov. Identifier: NCT03298347. 2 October 2017.

NCT04144712 *{published data only}*

Farid YA, Mazrou EM, Badr El-Deen OG. High versus low dose caffeine as respiratory stimulant in preterm infants. NCT04144712 7 February 2020.

Oliphant 2020 *{published data only}*

ACTRN12618001745235. Caffeine to improve oxygen levels in mildly preterm babies (Latte Dosage Trial). www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=375717 2018.

Oliphant EA, McKinlay CD, McNamara DG, Alsweiler JM. Caffeine prophylaxis to improve intermittent hypoxaemia in infants born late preterm: a randomised controlled dosage trial (Latte Dosage Trial). *British Medical Journal Open* 2020;**10**(10):e038271. [DOI: 10.1136/bmjopen-2020-038271] [PMID: 33082191]

Additional references

Alur 2015

Alur P, Bollampalli V, Bell T, Hussain N, Liss J. Serum caffeine concentrations and short-term outcomes in premature infants of 29 weeks of gestation. *Journal of Perinatology: Official Journal of the California Perinatal Association* 2015;**35**(6):434-8. [DOI: 10.1038/jp.2014.226] [PMID: 25521559]

Aranda 1979

Aranda JV, Turmen T. Methylxanthines in apnea of prematurity. *Clinics in Perinatology* 1979;**6**(1):87-108. [PMID: 383366]

Bayley 1993

Bayley N. Bayley Scales of Infant Development. 2nd edition. San Antonio (TX): The Psychological Corporation, 1993.

Bayley 2006

Bayley N. Bayley Scales of Infant and Toddler Development. San Antonio (TX): Harcourt Assessment, 2006.

Bell 1978

Bell MJ, Ternber J, Feigin RD, Keating JP, Marshall R, Barton L, et al. Neonatal necrotising enterocolitis. Therapeutic decisions based upon clinical staging. *Annals of Surgery* 1978;**187**(1):1-7. [DOI: 10.1097/00000658-197801000-00001] [PMID: PMC1396409]

Blanchard 1992

Blanchard P, Aranda J. Pharmacotherapy of respiratory control disorders. In: Beckerman R, Brouillette R, Hunt C, editors(s). Respiratory control disorders in infants and children. Maryland, USA: Williams and Wilkins, 1992:352-70.

Borenstein 2013

Borenstein M, Higgins JP. Meta-analysis and subgroups. *Prevention Science* 2013;**14**:134-43. [DOI: 10.1007/s11121-013-0377-7] [PMID: 23479191]

Brattström 2019

Brattström P, Russo C, Ley D, Bruschettini M. High-versus low-dose caffeine in preterm infants: a systematic review and meta-analysis. *Acta Paediatrica* 2019;**108**(3):401-10. [DOI: 10.1111/apa.14586] [PMID: 30242903]

Bruschettini 2016

Bruschettini M, Barco S, Romantsik O, Risso F, Gennai I, Chinea B, et al. DBS-LC-MS/MS assay for caffeine: validation and neonatal application. *Bioanalysis* 2016;**8**(18):1893-902. [DOI: 10.4155/bio-2016-0127] [PMID: 27532249]

Charles 2008

Charles BG, Townsend SR, Steer PA, Flenady VJ, Gray PH, Shearman A. Caffeine citrate treatment for extremely premature infants with apnea: population pharmacokinetics, absolute bioavailability, and implications for therapeutic drug monitoring. *Therapeutic Drug Monitoring* 2008;**30**(6):709-16. [DOI: 10.1097/FTD.0b013e3181898b6f] [PMID: 19057373]

Chen 2018

Chen J, Jin L, Chen X. Efficacy and safety of different maintenance doses of caffeine citrate for treatment of apnea

in premature infants: A systematic review and meta-analysis. *BioMed Research International* 2018;**2018**:9061234. [DOI: 10.1155/2018/9061234] [PMID: 30671477]

Cochrane EPOC Group 2017

Cochrane Effective Practice and Organisation of Care (EPOC). Data extraction and management. EPOC resources for review authors, 2017. epoc.cochrane.org/epoc-resources-review-authors (accessed prior to 12 December 2019).

Cornette 2002

Cornette LG, Tanner SF, Ramenghi LA, Miall LS, Childs AM, Arthur RJ, et al. Magnetic resonance imaging of the infant brain: anatomical characteristics and clinical significance of punctate lesions. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2002;**86**(3):F171-7. [DOI: 10.1136/fn.86.3.f171] [PMID: 11978747]

Crossley 2012

Crossley KJ, Allison BJ, Polglase GR, Morley CJ, Harding R, Davis PG, et al. Effects of caffeine on renal and pulmonary function in preterm newborn lambs. *Pediatric Research* 2012;**72**(1):19-25. [DOI: 10.1038/pr.2012.49] [PMID: 22484473]

Davis 2020

Davis PG. When to start and stop caffeine and why respiratory status matters. *Seminars in Fetal & Neonatal Medicine* 2020;**25**(6):101175. [DOI: 10.1016/j.siny.2020.101175] [PMID: 33144042]

Deeks 2022

Deeks JJ, Higgins JP, Altman DG, editor(s). Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.3 (updated February 2022). Cochrane, 2022. Available from www.training.cochrane.org/handbook.

Di Fiore 2010

Di Fiore JM, Bloom JN, Orge F, Schutt A, Schluchter M, Cheruvu VK, et al. A higher incidence of intermittent hypoxemic episodes is associated with severe retinopathy of prematurity. *Journal of Pediatrics* 2010;**157**(1):69-73. [DOI: 10.1016/j.jpeds.2010.01.046] [PMID: 20304417]

Dobson 2017

Dobson NR, Rhein LM, Darnall RA, Corwin MJ, Heeren TC, Eichenwald E, et al. Caffeine decreases intermittent hypoxia in preterm infants nearing term-equivalent age. *Journal of Perinatology: Official Journal of the California Perinatal Association* 2017;**37**(10):1135-40. [DOI: 10.1038/jp.2017.82] [PMID: 28749480]

Doyle 2016

Doyle J, Davidson D, Katz S, Varela M, Demeglio D, DeCristofaro J. Apnea of prematurity and caffeine pharmacokinetics: potential impact on hospital discharge. *Journal of Perinatology: Official Journal of the California Perinatal Association* 2016;**36**(2):141-4. [DOI: 10.1038/jp.2015.167] [PMID: 26562367]

Egger 1997

Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**(7109):629-34. [DOI: 10.1136/bmj.315.7109.629] [PMID: 9310563]

Eichenwald 1997

Eichenwald EC, Aina A, Stark AR. Apnea frequently persists beyond term gestation in infants delivered at 24 to 28 weeks. *Pediatrics* 1997;**100**(3 Pt 1):354-9. [DOI: 10.1542/peds.100.3.354] [PMID: 9282705]

Endesfelder 2020

Endesfelder S, Strauß E, Bendix I, Schmitz T, Bührer C. Prevention of oxygen-induced inflammatory lung injury by caffeine in neonatal rats. *Oxidative Medicine and Cellular Longevity* 2020;**2020**:3840124. [DOI: 10.1155/2020/3840124] [PMID: 32831996]

Fewell 2007

Fewell Z, Davey Smith G, Sterne JA. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. *American Journal of Epidemiology* 2007;**166**(6):646-55. [DOI: 10.1093/aje/kwm165] [PMID: 17615092]

Gillot 1990

Gillot I, Gouyon JB, Guignard JP. Renal effects of caffeine in preterm infants. *Biology of the Neonate* 1990;**58**(3):133-6. [DOI: 10.1159/000243252] [PMID: 2279048]

GRADEpro GDT [Computer program]

GRADEpro GDT. Version accessed 16 March 2018. Hamilton (ON): McMaster University (developed by Evidence Prime), 2015. Available at gradepro.org.

Griffiths 1954

Griffiths R. The Abilities of Babies: a Study in Mental Measurement. New York (NY): McGraw-Hill Book Co. Inc., 1954.

Henderson-Smart 1981

Henderson-Smart DJ. The effect of gestational age on the incidence and duration of recurrent apnoea in newborn babies. *Australian Paediatric Journal* 1981;**17**(4):273-6. [DOI: 10.1111/j.1440-1754.1981.tb01957.x] [PMID: 7347216]

Henderson-Smart 2010a

Henderson-Smart DJ, Davis PG. Prophylactic methylxanthines for extubation in preterm infants. *Cochrane Database of Systematic Reviews* 2010, Issue 12. Art. No: CD000139. [DOI: 10.1002/14651858.CD000139.pub2]

Henderson-Smart 2010b

Henderson-Smart DJ, De Paoli AG. Prophylactic methylxanthine for prevention of apnoea in preterm infants. *Cochrane Database of Systematic Reviews* 2010, Issue 12. Art. No: CD000432. [DOI: 10.1002/14651858.CD000432.pub2]

Henderson-Smart 2010c

Henderson-Smart DJ, De Paoli AG. Methylxanthine treatment for apnoea in preterm infants. *Cochrane Database of*

Systematic Reviews 2010, Issue 12. Art. No: CD000140. [DOI: 10.1002/14651858.CD000140.pub2]

Higgins 2011

Higgins JP, Altman DG, Sterne JA: on behalf of the Cochrane Statistical Methods Group and the Cochrane Bias Methods Group. Chapter 8: Assessing risk of bias in included studies. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2019

Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.

Higgins 2020

Higgins JP, Li T, Deeks JJ. Chapter 6: Choosing effect measures and computing estimates of effect. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (updated September 2020). Cochrane, 2020. Available from www.training.cochrane.org/handbook.

Higgins 2022

Higgins JP, Eldridge S, Li T, editor(s). Chapter 23: Including variants on randomized trials. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.3 (updated February 2022). Cochrane, 2022. . Available from www.training.cochrane.org/handbook.

Howell 1981

Howell J, Clozel M, Aranda JV. Adverse effects of caffeine and theophylline in the newborn infant. *Seminars in Perinatology* 1981;**5**(4):359-69. [PMID: 7302612]

International Committee 2005

International Committee for the Classification of Retinopathy of Prematurity. The international classification of retinopathy of prematurity revisited. *Archives of Ophthalmology* 2005;**123**(7):991-9. [DOI: 10.1001/archopht.123.7.991] [PMID: 16009843]

Jacobs 2013

Jacobs SE, Berg M, Hunt R, Tarnow-Mordi WO, Inder TE, Davis PG. Cooling for newborns with hypoxic ischaemic encephalopathy. *Cochrane Database of Systematic Reviews* 2013, Issue 1. Art. No: CD003311. [DOI: 10.1002/14651858.CD003311.pub3]

Janvier 2004

Janvier A, Khairy M, Kokkotis A, Cormier C, Messmer D, Barrington KJ. Apnea is associated with neurodevelopmental impairment in very low birth weight infants. *Journal of Perinatology* 2004;**24**(12):763-8. [DOI: 10.1038/sj.jp.7211182] [PMID: 15329741]

Jobe 2001

Jobe AH, Bancalari E. Bronchopulmonary dysplasia. *American Journal of Respiratory and Critical Care Medicine* 2001;**163**(7):1723-9. [DOI: 10.1164/ajrccm.163.7.2011060] [PMID: 11401896]

Jobe 2017a

Jobe AH. Caffeine: a lung drug for all very low birth weight preterm infants? *American Journal of Respiratory and Critical Care Medicine* 2017;**196**(10):1241-3. [DOI: 10.1164/rccm.201707-1402ED] [PMID: 28742379]

Jobe 2017b

Jobe AH. Caffeine therapy-when and why? *Journal of Pediatrics* 2017;**190**:2. [DOI: 10.1016/j.jpeds.2017.09.032] [PMID: 29144246]

Koons 1993

Koons AH, Mojica N, Jadeja N, Ostfeld B, Hiatt M, Hegyi T. Neurodevelopmental outcome of infants with apnea of infancy. *American Journal of Perinatology* 1993;**10**(3):208-11. [DOI: 10.1055/s-2007-994719] [PMID: 7686008]

Kyriacou 2016

Kyriacou DN, Lewis RJ. Confounding by Indication in clinical research. *JAMA* 2016;**316**(17):1818-9. [DOI: 10.1001/jama.2016.16435] [PMID: 27802529]

Lee 1997

Lee TC, Charles B, Steer P, Flenady V, Shearman A. Population pharmacokinetics of intravenous caffeine in neonates with apnea of prematurity. *Clinical Pharmacology and Therapeutics* 1997;**61**(6):628-40. [DOI: 10.1016/S0009-9236(97)90097-7] [PMID: 9209245]

Lee 2002

Lee CY, Su BH, Lin TW, Lin HC, Li TC, Wang NP. Risk factors of extubation failure in extremely low birth weight infants: a five year retrospective analysis. *Acta Paediatrica Taiwanica* 2002;**43**(6):319-25. [PMID: 12632784]

Limperopoulos 2007

Limperopoulos C, Bassan H, Gauvreau K, Robertson RL Jr, Sullivan NR, Benson CB, et al. Does cerebellar injury in premature infants contribute to the high prevalence of long-term cognitive, learning, and behavioral disability in survivors? *Pediatrics* 2007;**120**(3):584-93. [DOI: 10.1542/peds.2007-1041] [PMID: 17766532]

Marques 2021

Marques K, Roehr CC, Bruschettini M, Davis PG, Soll R. Methylxanthine for the prevention and treatment of apnea in preterm infants. *Cochrane Database of Systematic Reviews* 2021, Issue 1. Art. No: CD013830. [DOI: 10.1002/14651858.CD013830]

Marshall 2018

Marshall IJ, Noel-Storr AH, Kuiper J, Thomas J, Wallace BC. Machine Learning for Identifying randomized controlled trials: an evaluation and practitioner's guide. *Research Synthesis Methods* 2018;**9**(4):602-14. [DOI: 10.1002/jrsm.1287] [PMID: 29314757]

McDonald 2017

McDonald S, Noel-Storr AH, Thomas J. Harnessing the efficiencies of machine learning and Cochrane Crowd to identify randomised trials for individual Cochrane reviews. GlobalEvidence Summit, Cape Town, South Africa; 13th – 16th September 2017.

Millar 2004

Millar D, Schmidt B. Controversies surrounding xanthine therapy. *Seminars in Neonatology* 2004;**9**(3):239-44. [DOI: 10.1016/j.siny.2003.11.008] [PMID: 15050217]

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of Clinical Epidemiology* 2009;**62**(10):1006-12. [DOI: 10.1016/j.jclinepi.2009.06.005] [PMID: 19631508]

Montenegro 2017

Montenegro BL, Amberson M, Veit L, Freiberger C, Dukhovny D, Rhein LM. Economics of home monitoring for apnea in late preterm infants. *Respiratory Care* 2017;**62**(1):42-8. [DOI: 10.4187/respcare.04954] [PMID: 28003553]

Murner-Lavanchy 2018

Murner-Lavanchy IM, Doyle LW, Schmidt B, Roberts RS, Asztalos EV, Costantini L, et al. Neurobehavioral outcomes 11 years after neonatal caffeine therapy for apnea of prematurity. *Pediatrics* 2018;**141**(5):e20174047. [DOI: 10.1542/peds.2017-4047] [PMID: 29643070]

Nagatomo 2016

Nagatomo T, Jiménez J, Richter J, De Baere S, Vanoirbeek J, Naulaers G, et al. Caffeine prevents hyperoxia-induced functional and structural lung damage in preterm rabbits. *Neonatology* 2016;**109**(4):274-81. [DOI: 10.1159/000442937] [PMID: 26866610]

NIH 1979

National Institutes of Health. Report of workshop on bronchopulmonary dysplasia. In: NIH Publication No. 80-1660; 1978 December 4-6; Bethesda (MD). Washington, DC: National Institutes of Health, 1979.

Noel-Storr 2018

Noel-Storr AH and the Project Transform team. Cochrane Crowd: new ways of working together to produce health evidence. Evidence Live 2018, Oxford, UK; 18–20 June 2018.

Nylander Vujovic 2020

Nylander Vujovic S, Nava C, Johansson M, Bruschettini M. Confounding biases in studies on early- versus late-caffeine in preterm infants: a systematic review. *Pediatric Research* 2020;**88**(3):357-64. [DOI: 10.1038/s41390-020-0757-1] [PMID: 31931506]

Pakvasa 2018

Pakvasa MA, Saroha V, Patel RM. Optimizing caffeine use and risk of bronchopulmonary dysplasia in preterm infants: A systematic review, meta-analysis, and application of grading of recommendations assessment, development, and evaluation methodology. *Clinics in Perinatology* 2018;**45**(2):273-91. [DOI: 10.1016/j.clp.2018.01.012] [PMID: 29747888]

Papile 1978

Papile LA, Burstein J, Burstein R, Koffler H. Incidence and evolution of subependymal and intraventricular haemorrhage: a study of infants with birth weights less than 1,500 gm. *Journal of Pediatrics* 1978;**92**(4):529-34. [DOI: 10.1016/s0022-3476(78)80282-0] [PMID: 305471]

Parodi 2015

Parodi A, Morana G, Severino MS, Malova M, Natalizia AR, Sannia A, et al. Low-grade intraventricular hemorrhage: is ultrasound good enough? *Journal of Maternal-fetal & Neonatal Medicine* 2015;**28 Suppl 1**:2261-4. [DOI: 10.3109/14767058.2013.796162] [PMID: 23968243]

Poets 2015

Poets CF, Roberts RS, Schmidt B, Whyte RK, Asztalos EV, Bader D, et al. Association between intermittent hypoxemia or bradycardia and late death or disability in extremely preterm infants. *JAMA* 2015;**314**(6):595-603. [DOI: 10.1001/jama.2015.8841] [PMID: 26262797]

RevMan Web 2020 [Computer program]

Review Manager Web (RevMan Web). Version 1.22.0. The Cochrane Collaboration, 2020. Available at revman.cochrane.org.

Rhein 2014

Rhein LM, Dobson NR, Darnall RA, Corwin MJ, Heeren TC, Poets CF, et al, Caffeine Pilot Study Group. Effects of caffeine on intermittent hypoxia in infants born prematurely: a randomized clinical trial. *JAMA Pediatrics* 2014;**168**(3):250-7. [DOI: 10.1001/jamapediatrics.2013.4371] [PMID: 24445955]

Rutherford 2010

Rutherford MA, Supramaniam V, Ederies A, Chew A, Bassi L, Groppo M, et al. Magnetic resonance imaging of white matter diseases of prematurity. *Neuroradiology* 2010;**52**(6):505-21. [DOI: 10.1007/s00234-010-0700-y] [PMID: 20422407]

Saroha 2020

Saroha V, Patel RM. Caffeine for preterm infants: Fixed standard dose, adjustments for age or high dose? *Seminars in Fetal & Neonatal Medicine* 2020;**25**(6):101178. [DOI: 10.1016/j.siny.2020.101178] [PMID: 33168464]

Schmidt 2006

Schmidt B, Roberts RS, Davis P, Doyle LW, Barrington KJ, Ohlsson A et al. Caffeine therapy for apnea of prematurity. *New England Journal of Medicine* 2006;**354**(20):2112-21. [DOI: 10.1056/NEJMoa054065]

Schmidt 2007

Schmidt B, Roberts RS, Davis P, Doyle LW, Barrington KJ, Ohlsson A et al. Long-term effects of caffeine therapy for apnea of prematurity. *New England Journal of Medicine* 2007;**357**(19):1893-902. [DOI: 10.1056/NEJMoa073679] [PMID: 17989382]

Schmidt 2017

Schmidt B, Roberts RS, Anderson PJ, Asztalos EV, Costantini L, Davis PG, et al. Academic performance, motor function, and behavior 11 years after neonatal caffeine citrate therapy for apnea of prematurity: an 11-year follow-up of the CAP randomized clinical trial. *JAMA Pediatrics* 2017;**171**(6):564-72. [PMID: 28437520]

Schünemann 2013

Schünemann H, Brożek J, Guyatt G, Oxman A, editor(s). Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach (updated October 2013). GRADE Working Group, 2013. Available from gdt.guidelinedevelopment.org/app/handbook/handbook.html.

Stevenson 2007

Stevenson DK. On the caffeination of prematurity. *New England Journal of Medicine* 2007;**357**(19):1967-8. [DOI: 10.1056/NEJMe078200] [PMID: 17989390]

Tabacaru 2017

Tabacaru CR, Jang SY, Patel M, Davalian F, Zanelli S, Fairchild KD. Impact of caffeine boluses and caffeine discontinuation on apnea and hypoxemia in preterm infants. *Journal of Caffeine Research* 2017;**7**(3):103-10. [PMID: 28875061]

Thomas 2017

Thomas J, Noel-Storr AH, Marshall I, Wallace B, McDonald S, Mavergames C, et al, Living Systematic Review Network. Living systematic reviews: 2. combining human and machine effort. *Journal of Clinical Epidemiology* 2017;**91**:31-7. [DOI: 10.1016/j.jclinepi.2017.08.011] [PMID: 28912003]

Urru 2023

Urru S, Bruschettini M. Early versus late discontinuation of caffeine administration in preterm infants. *Cochrane Database of Systematic Reviews* 2023, Issue [in press].

Vliegenthart 2018

Vliegenthart R, Miedema M, Hutten GJ, van Kaam AH, Onland W. High versus standard dose caffeine for apnoea: a systematic review. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2018;**103**(6):F523-9. [DOI: 10.1136/archdischild-2017-313556] [PMID: 29437799]

Walsh 2004

Walsh MC, Yao Q, Gettner P, Hale E, Collins M, Hensman A, et al, National Institute of Child Health and Human Development Neonatal Research Network. Impact of a physiologic definition on bronchopulmonary dysplasia rates. *Pediatrics* 2004;**114**(5):1305-11. [DOI: 10.1542/peds.2004-0204] [PMID: 15520112]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

McPherson 2015

Study characteristics	
Methods	**Study design:** RCT **Location:** USA **Setting:** NICU (St. Louis Children's Hospital, Missouri) **Duration:** November 2008 to June 2010
Participants	**Inclusion criteria:** 74 infants ≤ 32 weeks' gestational age admitted to the NICU. **Exclusion criteria:** infants who had a known congenital anomaly, were moribund or in respiratory failure (defined as requiring > 80% FiO2 for 6 hours and/or having more than 2 inotropic drugs excluding hydrocortisone), or had severe brain injury (grade 3 to 4 intraventricular hemorrhage (IVH)) in the first 24 hours of life. Infants who were not expected to survive past 72 hours of life were also excluded.
Interventions	**High dose:** loading dose of 40 mg/kg followed by 20 mg/kg 12 hours later, then 10 mg/kg at 24 and 36 hours after the initial dose (80 mg/kg total over 36 hours). **Low dos**e: 20 mg/kg followed by 10 mg/kg 24 hours after the initial dose (30 mg/kg total over 36 hours).
Outcomes	Outcomes reported that are considered for this review: • mortality during first admission; • chronic lung disease at 36 weeks of corrected age;

McPherson 2015 *(Continued)*

- severe IVH ≥ grade 3;
- IVH, any grade;
- cerebellar hemorrhage;
- periventricular leukomalacia (PVL);
- lesions indicative of brain injury (detected by US and MRI): IVH, any grade; PVL; white matter injury, deep grey matter injury; cerebellar hemorrhage;
- duration of MV;
- retinopathy of prematurity ≥ grade 3;
- cognitive delay (at 2 years of age): assessed with Bayley-III scores for cognitive development;
- seizure before discharge;
- need for treatment of PDA;
- necrotizing enterocolitis (NEC); and
- at 5 years of age: standardized neurodevelopmental tests; parent reports of child socioemotional problems.

Notes	**Caffeine initiation time:** within 24 hours of birth **Baseline imbalances:** in the high-dose caffeine group the maternal age was higher than in the low dose group (P = 0.03) This study was reported as two different publications: - McPherson 2015: A pilot randomised trial of high-dose caffeine therapy in preterm infants - Vesoulis 2016: Early high-dose caffeine increases seizure burden in extremely preterm neonates: a preliminary study

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Group assignment was performed by a parallel 1:1 blocked randomization, generated by the dispensing pharmacist who was not involved in clinical care.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	Low risk	The clinical and research team remained blinded to each infant's randomization until completion of developmental assessment at two years of age.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	All testers were blinded to study assignment and past medical history, including imaging findings.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Out of 37 infants allocated to each group: • 28 versus 30 underwent MRI and 28 and 31 underwent neurodevelopmental testing at term-equivalent age, in the high- and low-dose groups respectively. • 24 and 22 underwent two year developmental assessment in the high- and low-dose groups respectively. • 8 infants in the high-dose group and 7 in the low-dose group were excluded due to corrupt data files or recordings equal or less than five hours in length from the EEG recording (for the assessment of seizures) through the 72 first hours of postnatal life

McPherson 2015 *(Continued)*

Selective reporting (reporting bias)	Low risk	The trial protocol is available at www.clinicaltrials.gov/ct2/show/NCT00809055 (25 April 2018). All pre-specified outcomes were reported in the manuscript.
Other bias	Unclear risk	In the high-dose caffeine group, the maternal age was higher than in the standard-dose group (P = 0.03).

Mohammed 2015

Study characteristics	
Methods	**Study design:** double-blind, pilot RCT **Location:** Egypt **Setting:** Neonatal Intensive Care Unit (NICU) of Mansoura University Children's Hospital, Mansoura **Duration:** July 2011 to July 2012
Participants	**Inclusion criteria:** 120 infants < 32 weeks' gestational age who exhibited apnea of prematurity within first 10 days of life. **Exclusion criteria:** major congenital malformations and chromosomal anomalies.
Interventions	**High dose:** loading dose 40 mg/kg/day and maintenance dose 20 mg/kg/day **Low dose:** loading dose 20 mg/kg/day and maintenance dose 10 mg/kg/day
Outcomes	Outcomes reported that are considered for this review: • mortality during first admission; • chronic lung disease at 36 weeks of corrected age; • severe IVH ≥ grade 3; • periventricular leukomalacia (PVL); • lesions indicative of brain injury (detected by US and MRI): IVH, any grade; PVL; • apnea; • retinopathy of prematurity, ≥ grade 3; • extubation failure (defined as need for re-intubation within 72 hours of extubation from mechanical ventilation); • any tachycardia; • need for treatment of PDA; • necrotizing enterocolitis (NEC); and • somatic growth.
Notes	**Caffeine initiation time:** postnatal mean (SD) age of 2.5 (2.6) days in the high-dose group 2.7 (2.8) days in the low-dose group. **Baseline imbalances:** none The trial protocol is available at www.clinicaltrials.gov/ct2/show/NCT02103777 (19 April 2018)

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Mohammed 2015 *(Continued)*

Random sequence generation (selection bias)	Low risk	Enrolled infants were assigned randomly to treatment groups using Internet-based random table technique. A designated pharmacist was responsible for the randomization of selected infants and the preparation of caffeine dose.
Allocation concealment (selection bias)	Low risk	Cards in opaque, sealed envelopes.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Nursing staff and family were blinded to patient's allocation.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Investigators were blinded to patient's allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data available on all of the infants randomized. No exclusions or attrition after randomization.
Selective reporting (reporting bias)	Unclear risk	The protocol is available at www.clinicaltrials.gov/ct2/show/NCT02103777 (24 April 2018). The study did not report on hydrocephalus as pre-specified in the trial protocol. Furthermore, the duration of CPAP and postnatal steroid therapy for BPD were reported as outcomes in the study but not pre-specified in the protocol.
Other bias	Low risk	None known

Mohd 2021

Study characteristics	
Methods	**Study design:** parallel RCT **Location:** Malaysia **Setting:** ICU (Hospital Universiti Sains Malaysia (USM), Kota Bharu, Kelantan) **Duration:** June 2019 to August 2020
Participants	**Inclusion criteria:** preterm infants ≤ 32 weeks' gestational age **Exclusion criteria:** hydrops fetalis
Interventions	**High dose:** loading dose 40 mg/kg/day and maintenance dose 20 mg/kg/day **Low dose:** loading dose 20 mg/kg/day and maintenance dose 10 mg/kg/day
Outcomes	Outcomes reported that are considered for this review: • frequency and number of days with apnea; • extubation failure; • duration of non-invasive ventilation (e.g. Optiflow, Bipap, Duopap and CPAP); • duration of oxygen therapy (nasal prong oxygen); • chronic lung disease defined, as the need for oxygen at 36 weeks' post-menstrual age; • tachycardia; • hypertension;

Mohd 2021 *(Continued)*

- time to reach full enteral feeding;
- weight gain;
- length of hospital stay;
- neonatal mortality;
- necrotizing enterocolitis (NEC);
- intraventricular hemorrhage (IVH);
- periventricular leukomalacia (PVL);
- retinopathy of prematurity (ROP);
- number of infants for whom caffeine was withheld early because of suspected side effects; and
- death before hospital discharge.

Notes

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random sequence was generated by a researcher not involved in the recruitment of patients and data collection, using blocks of variable sizes known only to the randomizer.
Allocation concealment (selection bias)	Low risk	Concealment of allocation was ensured by the use of opaque, sealed, and sequentially numbered envelopes.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Doctors, nursing staff, and family were blinded to the allocation.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Investigators were blinded to the allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	All infants received intended treatment until they exited the study. All included infants were analyzed for primary outcome.
Selective reporting (reporting bias)	Low risk	The trial protocol is available at www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=377580. All prespecified outcomes were reported in the manuscript.
Other bias	Unclear risk	More infants in the standard-dose caffeine were intubated at the baseline (97% and 92% in the standard- and high-dose caffeine groups, respectively) and needed surfactant (95% and 80% in the standard- and high-dose groups, respectively)

Scanlon 1992

Study characteristics

Methods **Study design:** RCT

Scanlon 1992 *(Continued)*

	Location: UK **Setting:** regional Neonatal Intensive Care Unit (Birmingham Maternity Hospital) **Duration:** not reported
Participants	**Inclusion criteria:** 44 infants < 31 weeks' gestational age with either 10 or more apneic attacks in 8 hours or 4 apneas in 1 hour **Exclusion criteria:** not specified
Interventions	**High dose:** loading dose 50 mg/kg/day and maintenance dose 12 mg/kg/day **Low dose:** loading dose 25 mg/kg/day and maintenance dose 6 mg/kg/day
Outcomes	Outcomes reported that are considered for this review: • reduction of number of apneas
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random numbers in sealed envelopes.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	The two strengths of caffeine citrate were not identical in appearance.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	Low risk	Only eight of 44 randomized infants were excluded.
Selective reporting (reporting bias)	Unclear risk	Trial protocol not available.
Other bias	Low risk	None known

Steer 2003

Study characteristics

Methods	**Study design:** double-blind RCT **Location:** Australia

Steer 2003 *(Continued)*

	Setting: one neonatal intensive care nursery of a large maternity teaching hospital (Mater Mothers' Hospital, Brisbane) **Duration:** September 1993 to November 1995
Participants	**Inclusion criteria:** 127 infants < 31 weeks' gestational age who had received or who were anticipated to receive at least 48 hours of mechanical ventilation **Exclusion criteria:** major congenital abnormality, infection, major neurological condition, severe intraventricular hemorrhage, MV for a period greater than 28 days, previous exposure to methylxanthine therapy.
Interventions	This study contained three study arms with different doses. According to our definition outlined in the protocol, two of them were high doses (very high and moderately high) and one was a low dose. **Very high dose:** loading dose 60 mg/kg/day and maintenance dose 30 mg/kg/day **Moderately high dose:** loading dose 30 mg/kg/day and maintenance dose 15 mg/kg/day **Low dose:** loading dose 6 mg/kg/day and maintenance dose 3 mg/kg/day
Outcomes	Outcomes reported that are considered for this review: • severe IVH ≥ grade 3; • lesions indicative of brain injury (detected by US and MRI): severe IVH ≥ grade 3; • duration of MV; • retinopathy of prematurity ≥ grade 3; • extubation failure (defined as either an inability to extubate from mechanical ventilation within 48 hours of caffeine loading for a planned extubation or the use of reintubation or doxapram within seven days of commencing caffeine therapy; • any tachycardia; • tachycardia leading to suspension of study intervention; • necrotizing enterocolitis (NEC); and • somatic growth.
Notes	**Caffeine initiation time:** postnatal mean (SD) age of 4.4 (2.5) days in the very high-dose group, 3.8 (2.1) days in the moderately high-dose group, and 3.5 (1.5) days in the low-dose group **Baseline imbalances:** none The trial was terminated after an interim analysis after enrolling infants corresponding to approximately 50 % of the a priori calculated sample size No trial protocol available

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Enrolled infants were allocated to one of three group using a computer-generated list of random numbers by a hospital pharmacist. The pharmacist was not associated in any other way with the study or with the clinical management of the infants.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	The identity of the treatments was not disclosed to the medical or nursing staff.

Steer 2003 *(Continued)*

All outcomes		
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The identity of the treatments was not disclosed to the investigators.
Incomplete outcome data (attrition bias) All outcomes	Low risk	8 out of 127 randomized infants failed to receive caffeine therapy after randomization; 9 out of 127 did not complete the planned seven day course of caffeine. Outcomes for these 17 infants are included in the ITT analysis.
Selective reporting (reporting bias)	Unclear risk	Trial protocol not available.
Other bias	Low risk	None known.

Steer 2004

Study characteristics	
Methods	**Study design:** multi-centre RCT **Location:** Australia **Setting:** four neonatal intensive care units (Mater Mothers' Hospital, Brisbane; Royal Prince Alfred Hospital, Sydney; Mercy Hospital for Women, Melbourne; and Royal Hobart Hospital, Tasmania) **Duration:** September 1996 to April 1999
Participants	**Inclusion criteria:** 287 infants < 30 weeks' gestational age requiring methylxanthines for treatment of apnea of prematurity or as a part of peri-extubation management. Eligible for the peri-extubation group were infants who received or who were expected to receive at least 48 hours of mechanical ventilation **Exclusion criteria:** infants with major congenital abnormality, infection (sepsis confirmed by blood culture), major neurological condition, grade 3 or 4 IVH, and previous exposure to methylxanthine therapy. Infants who received short-term caffeine (≤ seven days)
Interventions	**High dose:** loading dose 80 mg/kg/day and maintenance dose 20 mg/kg/day **Low dose:** loading dose 20 mg/kg/day and maintenance dose 5 mg/kg/day
Outcomes	Outcomes reported that are considered for this review: • mortality during first admission; • chronic lung disease at 36 weeks of corrected age; • severe IVH, ≥ grade 3; • IVH, any grade; • cerebellar hemorrhage; • periventricular leukomalacia (PVL); • lesions indicative of brain injury (detected by US and MRI): IVH, any grade; major cerebral abnormalities (defined as one or more of porencephalic cysts, cystic periventricular leukomalacia or hydrocephalus); • apnea; • extubation failure (defined as an inability to extubate from mechanical ventilation within 48 hours of caffeine loading for a planned extubation or the use of reintubation or doxapram within 7 days of caffeine loading); • retinopathy of prematurity, ≥ grade 3;

Steer 2004 *(Continued)*

- any tachycardia;
- tachycardia, leading to suspension of study intervention;
- necrotizing enterocolitis (NEC);
- major cerebral abnormalities at 6 weeks; and
- major disability (cerebral palsy, bilateral blindness, need for hearing aids at 12 months corrected for prematurity).

Notes

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Infants were randomized using a computer generated list of random numbers by a hospital pharmacist who was not associated in any other way with the study or with the clinical management of infants.
Allocation concealment (selection bias)	Unclear risk	Not reported.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	The two strengths of caffeine citrate were identical in appearance. Investigators and clinical staff were blind to treatment allocation.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Investigators were blind to treatment allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	41 out 287 infants were excluded after randomization and were not analyzed. This correspond to 14% of the infants, which is still acceptable for our review.
Selective reporting (reporting bias)	Unclear risk	Trial protocol not available.
Other bias	Low risk	None known.

Zhao 2016

Study characteristics

Methods	**Study design:** double-blind RCT **Location:** China **Setting:** one neonatal unit (Tianjin Central Hospital of Gynaecology Obstetrics) **Duration:** October 2013 to December 2014
Participants	**Inclusion criteria:** 164 infants < 32 weeks' gestational age who were diagnosed with apnea **Exclusion criteria:** infants with congenital malformations and chromosomal abnormalities, central nervous system diseases, primary lung diseases, serious infection, metabolic diseases, obstructive apnea etc.
Interventions	**High dose:** loading dose 20 mg/kg/day and maintenance dose 15 mg/kg/day

Zhao 2016 *(Continued)*

	Low dose: loading dose 20 mg/kg/day and maintenance dose 5 mg/kg/day
Outcomes	Outcomes reported that are considered for this review: • mortality during first admission; • chronic lung disease at 36 weeks of corrected age; • apnea; • success in ventilator removal; • any tachycardia.
Notes	**Baseline imbalances:** none No trial protocol available

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Patients were allocated into low-dose or high-dose group using a random number table by a researcher who did not participate in data collection or analysis.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Researchers, nurses, and family members did not know about the grouping arrangement.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Described, but not for all outcomes.
Incomplete outcome data (attrition bias) All outcomes	Low risk	All patients were analyzed for the outcomes.
Selective reporting (reporting bias)	Unclear risk	Trial protocol not available.
Other bias	Low risk	None known

BPD: bronchopulmonary dysplasia; **CPAP**: continuous positive airway pressure; **FiO2**: fraction of inspired oxygen; **ITT**: intention-to-treat; **kg**: kilogram; **mg**: milligram; **MRI**: magnetic resonance imaging; **MV**: mechanical ventilation; **NICU**: neonatal intensive care unit; **PDA**: patent ductus arteriosus; **RCT**: randomized controlled trial; **SD**: standard deviation; **US:** ultrasound

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Autret 1985	Infants were randomized to two caffeine doses which both fall within the "standard dose" definition used in this review. Loading and maintenance dose were 20 mg/kg and 5 mg/kg caffeine in group one, respectively; 20 mg/kg and 3 mg/kg in group two, respectively.
Cherif 2003	Not an RCT

Study	Reason for exclusion
Gray 2016	Not an RCT
Romagnoli 1992	Infants were randomized to two caffeine doses which both fall within the "standard dose" definition used in this review. Loading and maintenance dose were 10 mg/kg and 5 mg/kg caffeine in-group one, respectively; 10 mg/kg and 2.5 mg/kg in group two, respectively.
Wan 2020	Infants were randomized to two caffeine doses which both fall within the "standard dose" definition used in this review. Loading and maintenance dose were 20 mg/kg and 10 mg/kg caffeine in-group one, respectively; 20 mg/kg and 5 mg/kg in group two, respectively.
Yao 2021	Infants were randomized to two caffeine doses which both fall within the "standard dose" definition used in this review. Loading and maintenance dose were 20 mg/kg and 5 mg/kg caffeine in both group one and group two, whereas group one received am additional dose caffeine one hour before ventilator weaning.
Zhang 2019	Infants were randomized to two caffeine doses which both fall within the "standard dose" definition used in this review. Loading and maintenance dose were 20 mg/kg and 10 mg/kg caffeine in-group one, respectively; 20 mg/kg and 5 mg/kg in group two, respectively.

Characteristics of studies awaiting classification *[ordered by study ID]*

Gray 2018

Methods	Conference abstract with no full text; unclear if randomized.
Participants	Infants < 28 weeks' gestation who received a loading dose of caffeine within the first 36 hours of life from 2011 to 2013 were included in the study.
Interventions	High-dose group: median dose of 80 mg/kg Standard-dose group: median dose of 20 mg/kg
Outcomes	Neonatal outcomes including results of cranial ultrasound were compared. The infants were followed up at 2 years of age and had a neurological assessment, the Bayley Scales of Infant and Toddler Development, 3rd Edition and the Neurosensory Motor Developmental Assessment (NSDMA) The incidence of cerebellar hemorrhage was 4 (2.5%) in the high-dose group and 1 (1.7%) in the standard-dose group, with no difference in the incidence of IVH. The mean cognitive score on the Bayley for the high-dose group was 97.4 compared with 91.9 for the standard-dose group (P = 0.06). There were no differences in the language or motor scores, the NSMDA results, or the incidence of cerebral palsy between the groups.
Notes	

IVH: intraventricular hemorrhage; **kg**: kilogram; **mg:** milligram

Characteristics of ongoing studies *[ordered by study ID]*

NCT03298347

Study name	Caffeine for preterm infants with apnea of prematurity (AOP)
Methods	Study type: interventional Estimated enrollment: 100 participants

NCT03298347 *(Continued)*

	Allocation: randomized Intervention model: parallel assignment Intervention model description: 100 preterm infants admitted to Daping Hospital and the Research Institute of Surgery of the Third Military Medical University (Chongqing, China), divided in two groups. Masking: quadruple (participant, care provider, investigator, outcomes assessor) Primary purpose: treatment
Participants	Preterm infants with gestational age less than 32 weeks and AOP
Interventions	80 versus 20 mg/kg of caffeine (not clear if single or multiple administration)
Outcomes	Rate of AOP within 100 days
Starting date	1 October 2017
Contact information	Ma Juan 416767068@qq.com Shi Yuan petshi530@vip.163.com
Notes	

NCT04144712

Study name	High- versus low-dose caffeine as respiratory stimulant in preterm infants
Methods	Study type: interventional (clinical trial) Estimated enrollment: 80 participants Allocation: randomized Intervention model: parallel assignment Intervention model description: 80 preterm infants admitted to the Neonatal Intensive Care Unit (NICU), Ain Shams University (Cairo, Egypt), divided into two groups Masking: double (participant, care provider) Primary purpose: prevention
Participants	Preterm infants with a gestational age < 32 weeks in room air or CPAP (prophylactic). Preterm infants with gestational age 32 to 34 weeks who exhibited apnea of prematurity within the first 10 days of life in room air or CPAP
Interventions	High and low dose of caffeine citrate (doses are not specified)
Outcomes	Rate of occurrence of apnea of prematurity between infants receiving high- and low-dose caffeine
Starting date	1 April 2019
Contact information	Eslam M Mazrou eslammazrou@gmail.com

NCT04144712 *(Continued)*

Notes	Unclear whether this study, once completed, will be included in this review, as doses are not specified in the protocol.

Oliphant 2020

Study name	Caffeine prophylaxis to improve intermittent hypoxemia in infants born late preterm: a randomized controlled dosage trial (Latte Dosage Trial)
Methods	Phase: IIB Study type: controlled trial Allocation: randomized Intervention model: parallel assignment, five-arms Intervention model description: late preterm infants admitted to the neonatal unit and postnatal wards at Auckland City and Middlemore Hospitals (Auckland, New Zealand), randomized into five groups within 72 hours of birth to receive 5 mg/kg, 10 mg/kg, 15 mg/kg, or 20 mg/kg/day caffeine citrate or matching placebo daily until term corrected Masking: double-blind Primary purpose: prevention
Participants	Infants born between 34 weeks and 36 weeks' and 6 days gestation without contradiction to caffeine treatment. Sample size: 120.
Interventions	Enteral loading dose of the study drug (10 mg/kg, 20 mg/kg, 30 mg/kg, or 40 mg/kg of caffeine citrate or water), followed by a daily dose each morning (5 mg/kg, 10 mg/kg, 15 mg/kg, or 20 mg/kg of caffeine citrate or placebo) until term equivalent age (40 weeks' post-menstrual age)
Outcomes	Primary outcome: frequency of intermittent hypoxemia (events/hour) Secondary outcomes: • respiratory: frequency of intermittent hypoxemia on overnight oximetry at term equivalent age; mean overnight oxygen saturation at 2 weeks and term equivalent age; use of respiratory support, including oxygen, until term equivalent age; • growth: growth velocity from birth to term equivalent age for weight gain, length and head circumference; failure to regain birth weight by 2 weeks of age; • side effects: feed intolerance as reported by parents; duration of tube feeding; sleep and arousal as reported by parents (measured by sub scale nine on the Infant Behaviour Questionnaire-Revised, modified for neonates); tachycardia; study drug stopped due to presumed side effects; neonatal seizures requiring anticonvulsant treatment before 44 weeks' postmenstrual age; neonatal or infant death; • maternal and infant salivary caffeine concentration at two weeks after randomization; • readmission to hospital until 44 weeks' postmenstrual age or open-label caffeine use; and • maternal caffeine intake at birth, 2 weeks and term corrected age and mental health (Edinburgh post-natal depression score) at birth and term corrected age.
Starting date	Unknown
Contact information	Elizabeth Anne Oliphant: e.oliphant@auckland.ac.nz

Oliphant 2020 *(Continued)*

Notes

AOP: apnea of prematurity; **CPAP**: continuous positive airway pressure; **kg**: kilogram; **mg**: milligram

DATA AND ANALYSES

Comparison 1. High- versus standard-dose strategies for any indication

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 All-cause mortality prior to hospital discharge	5	723	Risk Ratio (M-H, Fixed, 95% CI)	0.86 [0.53, 1.38]
1.1.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	4	559	Risk Ratio (M-H, Fixed, 95% CI)	0.92 [0.52, 1.63]
1.1.2 Standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	164	Risk Ratio (M-H, Fixed, 95% CI)	0.73 [0.31, 1.71]
1.2 Major neurodevelopmental disability in children aged three to five years	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.2.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.3 Failure to extubate within one week of commencing treatment	4		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.3.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	4	521	Risk Ratio (M-H, Fixed, 95% CI)	0.54 [0.40, 0.73]
1.4 Reintubation within one week of commencing treatment	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.4.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	238	Risk Ratio (M-H, Fixed, 95% CI)	0.36 [0.19, 0.71]
1.5 Apnea: number of infants with at least one episode (defined as interruption of breathing for more than 20 seconds) after 24 hours from commencing treatment, in a 24-hour period and over one week	1	78	Risk Ratio (M-H, Fixed, 95% CI)	1.01 [0.59, 1.75]
1.5.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	78	Risk Ratio (M-H, Fixed, 95% CI)	1.01 [0.59, 1.75]

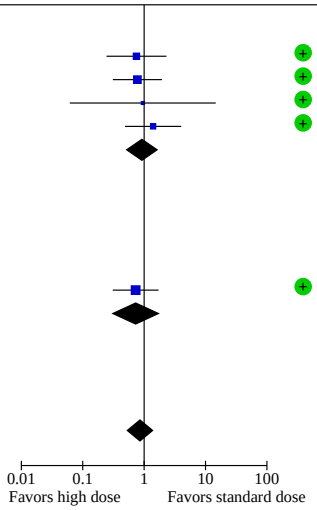
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.6 Side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine	5		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.6.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	5	593	Risk Ratio (M-H, Fixed, 95% CI)	1.66 [0.86, 3.23]
1.7 Bronchopulmonary dysplasia/chronic lung disease: 28 days of oxygen exposure	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.8 Bronchopulmonary dysplasia/chronic lung disease at 36 weeks' postmenstrual age	5	723	Risk Ratio (M-H, Fixed, 95% CI)	0.75 [0.60, 0.94]
1.8.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	4	559	Risk Ratio (M-H, Fixed, 95% CI)	0.76 [0.60, 0.97]
1.8.2 Standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	164	Risk Ratio (M-H, Fixed, 95% CI)	0.68 [0.36, 1.29]
1.9 Number of days using mechanical ventilation	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.9.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	2	201	Mean Difference (IV, Fixed, 95% CI)	0.01 [-0.85, 0.87]
1.10 Intraventricular hemorrhage, any grade	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.10.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	2	361	Risk Ratio (M-H, Fixed, 95% CI)	0.90 [0.63, 1.27]
1.11 Intraventricular hemorrhage, grade 3 to 4	5		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.11.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	5	686	Risk Ratio (M-H, Fixed, 95% CI)	1.26 [0.67, 2.36]
1.12 Cerebellar hemorrhage at brain ultrasound (yes/no)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.12.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	74	Risk Ratio (M-H, Fixed, 95% CI)	3.33 [1.00, 11.15]
1.13 Magnetic resonance imaging (MRI) abnormalities at term equivalent age (yes/no), defined as white matter lesions (i.e. cavitations [Rutherford 2010]) and punctate lesions (Cornette 2002); germinal matrix-intraventricular	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only

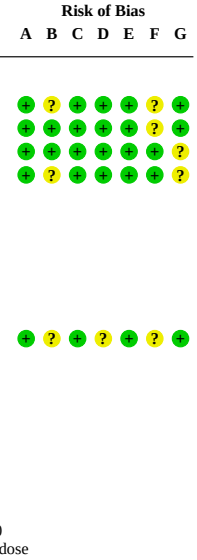
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
hemorrhage (Parodi 2015); or cerebellar hemorrhage (Limperopoulos 2007)				
1.13.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	74	Risk Ratio (M-H, Fixed, 95% CI)	0.50 [0.10, 2.56]
1.14 Periventricular leukomalacia	3	272	Risk Difference (M-H, Fixed, 95% CI)	0.02 [-0.04, 0.08]
1.14.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	3	272	Risk Difference (M-H, Fixed, 95% CI)	0.02 [-0.04, 0.08]
1.15 Necrotizing enterocolitis (proven = Bell stage of 2 or greater) (Bell 1978)	5		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.15.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	5	637	Risk Ratio (M-H, Fixed, 95% CI)	0.80 [0.43, 1.51]
1.16 Patent ductus arteriosus (PDA) requiring treatment (cyclo-oxygenase inhibitors or surgical ligation)	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.16.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	2	201	Risk Ratio (M-H, Fixed, 95% CI)	0.87 [0.59, 1.27]
1.17 Retinopathy of prematurity (ROP) (any ROP) (International Committee 2005)	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.17.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	2	365	Risk Ratio (M-H, Fixed, 95% CI)	0.85 [0.56, 1.29]
1.18 Retinopathy of prematurity (ROP) (severe ROP [stage 3 or greater]) (International Committee 2005)	3		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
1.18.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	3	481	Risk Difference (M-H, Fixed, 95% CI)	-0.03 [-0.07, 0.01]
1.19 Seizures (clinically diagnosed; diagnosed by electroencephalography)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.19.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	74	Risk Ratio (M-H, Fixed, 95% CI)	1.42 [0.79, 2.53]
1.20 Developmental delay (Bayley Mental Developmental Index or Griffiths Mental Development Scale in children aged 18 to 24 months	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.20.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.21 Bayley-III cognitive score in children at 18 to 24 months CA	1	42	Mean Difference (IV, Fixed, 95% CI)	-1.90 [-8.60, 4.80]
1.22 Cerebral palsy in children aged 18 to 24 months	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.22.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.23 Blindness in children aged 18 to 24 months	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.23.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.24 Deafness in children aged 18 to 24 months	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.24.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected

Analysis 1.1. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 1: All-cause mortality prior to hospital discharge

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.1.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Steer 2004	5	140	7	147	20.8%	0.75 [0.24 , 2.31]
Mohammed 2015	7	60	9	60	27.4%	0.78 [0.31 , 1.95]
Mohd 2021	1	40	1	38	3.1%	0.95 [0.06 , 14.65]
McPherson 2015	7	37	5	37	15.2%	1.40 [0.49 , 4.01]
Subtotal (95% CI)		**277**		**282**	**66.5%**	**0.92 [0.52 , 1.63]**
Total events:	20		22			
Heterogeneity: Chi² = 0.87, df = 3 (P = 0.83); I² = 0%						
Test for overall effect: Z = 0.29 (P = 0.77)						
1.1.2 Standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Zhao 2016	8	82	11	82	33.5%	0.73 [0.31 , 1.71]
Subtotal (95% CI)		**82**		**82**	**33.5%**	**0.73 [0.31 , 1.71]**
Total events:	8		11			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.73 (P = 0.47)						
Total (95% CI)		**359**		**364**	**100.0%**	**0.86 [0.53 , 1.38]**
Total events:	28		33			
Heterogeneity: Chi² = 1.08, df = 4 (P = 0.90); I² = 0%						
Test for overall effect: Z = 0.65 (P = 0.52)						
Test for subgroup differences: Chi² = 0.20, df = 1 (P = 0.66), I² = 0%						

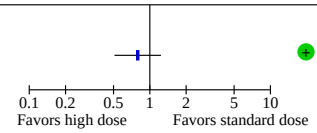


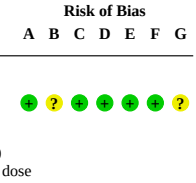


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.2. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 2: Major neurodevelopmental disability in children aged three to five years

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Risk Ratio M-H, Fixed, 95% CI
1.2.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose					
McPherson 2015	12	21	18	25	0.79 [0.51 , 1.24]

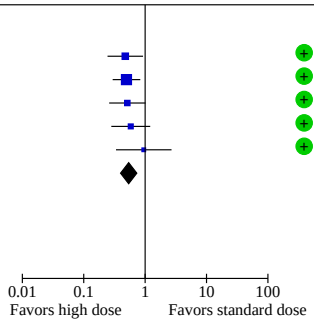


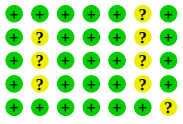


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.3. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 3: Failure to extubate within one week of commencing treatment

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.3.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Mohammed 2015	9	40	18	38	21.7%	0.47 [0.24 , 0.92]
Steer 2004 (1)	17	116	36	122	41.2%	0.50 [0.30 , 0.83]
Steer 2003	11	45	10	21	16.0%	0.51 [0.26 , 1.02]
Steer 2003 (2)	10	40	9	21	13.9%	0.58 [0.28 , 1.21]
Mohd 2021 (3)	6	40	6	38	7.2%	0.95 [0.34 , 2.69]
Subtotal (95% CI)		**281**		**240**	**100.0%**	**0.54 [0.40 , 0.73]**
Total events:	53		79			
Heterogeneity: Chi² = 1.44, df = 4 (P = 0.84); I² = 0%						
Test for overall effect: Z = 3.97 (P < 0.0001)						
Test for subgroup differences: Not applicable						





Footnotes

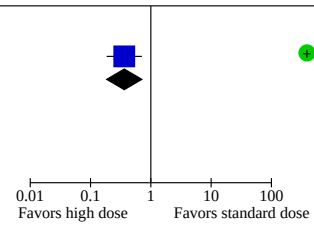
(1) In this forest plot, outcome data refer to the subgroup of infants treated for extubation management (and not to the 140 + 147 infants treated for apnea treatment)
(2) infants with the highest and second highest dose # "Steer 2003" and "(1)Steer 2003", resp.; infants with the lowest dose # "Steer 2003" and "(1)Steer 2003" (21 each)
(3) Time frame not specified

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.4. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 4: Reintubation within one week of commencing treatment

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.4.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Steer 2004 (1)	10	116	29	122	100.0%	0.36 [0.19 , 0.71]
Subtotal (95% CI)		**116**		**122**	**100.0%**	**0.36 [0.19 , 0.71]**
Total events:	10		29			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.96 (P = 0.003)						



Footnotes

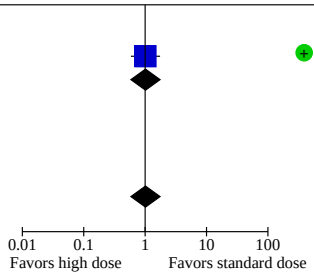
(1) In this forest plot, outcome data refer to the subgroup of infants treated for extubation management (and not to the 140 + 147 infants treated for apnea treatment)

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.5. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 5: Apnea: number of infants with at least one episode (defined as interruption of breathing for more than 20 seconds) after 24 hours from commencing treatment, in a 24-hour period and over one week

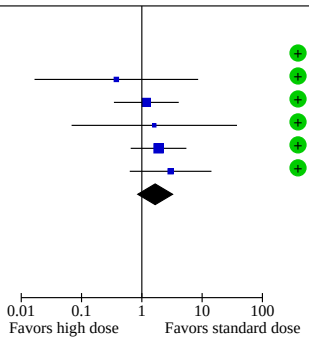
Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.5.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Mohd 2021	16	40	15	38	100.0%	1.01 [0.59 , 1.75]
Subtotal (95% CI)		**40**		**38**	**100.0%**	**1.01 [0.59 , 1.75]**
Total events:	16		15			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.05 (P = 0.96)						
Total (95% CI)		**40**		**38**	**100.0%**	**1.01 [0.59 , 1.75]**
Total events:	16		15			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.05 (P = 0.96)						
Test for subgroup differences: Not applicable						

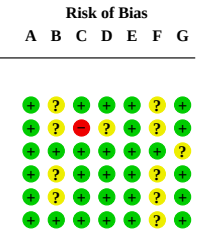


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.6. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 6: Side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.6.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Steer 2003	0	45	0	21		Not estimable
Scanlon 1992	0	14	1	16	10.8%	0.38 [0.02 , 8.59]
Mohd 2021	5	40	4	38	31.5%	1.19 [0.34 , 4.09]
Steer 2003 (1)	1	40	0	21	5.0%	1.61 [0.07 , 37.88]
Steer 2004 (2)	9	116	5	122	37.4%	1.89 [0.65 , 5.48]
Mohammed 2015	6	60	2	60	15.3%	3.00 [0.63 , 14.27]
Subtotal (95% CI)		**315**		**278**	**100.0%**	**1.66 [0.86 , 3.23]**
Total events:	21		12			
Heterogeneity: Chi² = 1.76, df = 4 (P = 0.78); I² = 0%						
Test for overall effect: Z = 1.50 (P = 0.13)						



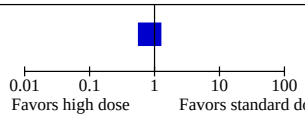


Footnotes
(1) infants with the highest and second highest dose # "Steer 2003" and "(1)Steer 2003", resp.; infants with the lowest dose # "Steer 2003" and "(1)Steer 2003" (21 each)
(2) In this forest plot, outcome data refer to the subgroup of infants treated for extubation management (and not to the 140 + 147 infants treated for apnea treatment)

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.7. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 7: Bronchopulmonary dysplasia/chronic lung disease: 28 days of oxygen exposure

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Risk Ratio M-H, Fixed, 95% CI	Risk Ratio M-H, Fixed, 95% CI	Risk of Bias A B C D E F G
Steer 2004 (1)	64	116	80	122	0.84 [0.68 , 1.04]		+ ? + + + ? +



Footnotes

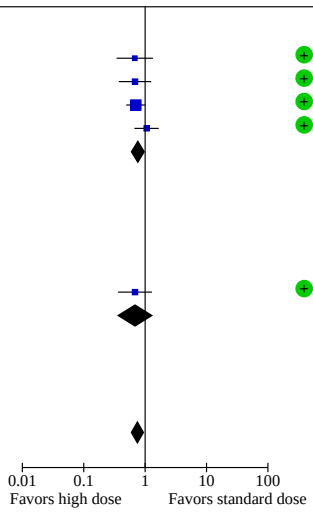
(1) In this forest plot, outcome data refer to the subgroup of infants treated for extubation management (and not to the 140 + 147 infants treated for apnea treatment)

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.8. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 8: Bronchopulmonary dysplasia/chronic lung disease at 36 weeks' postmenstrual age

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI	Risk Ratio M-H, Fixed, 95% CI	Risk of Bias A B C D E F G
1.8.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose								
Mohd 2021	10	40	14	38	11.9%	0.68 [0.34 , 1.34]		+ + + + + + ?
Mohammed 2015	13	60	19	60	15.7%	0.68 [0.37 , 1.26]		+ + + + + ? +
Steer 2004	35	140	52	147	41.9%	0.71 [0.49 , 1.01]		+ ? + + + ? +
McPherson 2015	19	37	18	37	14.9%	1.06 [0.67 , 1.67]		+ ? + + + + ?
Subtotal (95% CI)		**277**		**282**	**84.3%**	**0.76 [0.60 , 0.97]**		
Total events:	77		103					
Heterogeneity: Chi² = 2.37, df = 3 (P = 0.50); I² = 0%								
Test for overall effect: Z = 2.22 (P = 0.03)								
1.8.2 Standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose								
Zhao 2016	13	82	19	82	15.7%	0.68 [0.36 , 1.29]		+ ? + ? + ? +
Subtotal (95% CI)		**82**		**82**	**15.7%**	**0.68 [0.36 , 1.29]**		
Total events:	13		19					
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.17 (P = 0.24)								
Total (95% CI)		**359**		**364**	**100.0%**	**0.75 [0.60 , 0.94]**		
Total events:	90		122					
Heterogeneity: Chi² = 2.52, df = 4 (P = 0.64); I² = 0%								
Test for overall effect: Z = 2.51 (P = 0.01)								
Test for subgroup differences: Chi² = 0.09, df = 1 (P = 0.76), I² = 0%								

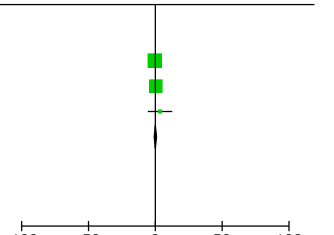


Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.9. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 9: Number of days using mechanical ventilation

Study or Subgroup	High-dose strategies Mean	SD	Total	Standard-dose strategies Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
1.9.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose								
Steer 2003 (1)	3.1	2.2	40	3.5	2.3	21	52.1%	-0.40 [-1.60 , 0.80]
Steer 2003	3.9	2.7	45	3.5	2.3	21	47.0%	0.40 [-0.86 , 1.66]
McPherson 2015	15.2	22.4	37	11.7	17.4	37	0.9%	3.50 [-5.64 , 12.64]
Subtotal (95% CI)			**122**			**79**	**100.0%**	**0.01 [-0.85 , 0.87]**
Heterogeneity: Chi² = 1.38, df = 2 (P = 0.50); I² = 0%								
Test for overall effect: Z = 0.02 (P = 0.98)								

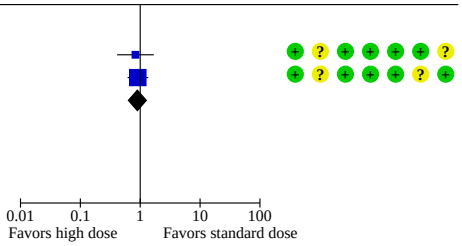


Footnotes

(1) infants with the highest and second highest dose # "Steer 2003" and "(1)Steer 2003", resp.; infants with the lowest dose # "Steer 2003" and "(1)Steer 2003" (21 each

Analysis 1.10. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 10: Intraventricular hemorrhage, any grade

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.10.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
McPherson 2015	10	37	12	37	24.0%	0.83 [0.41 , 1.69]
Steer 2004	34	140	39	147	76.0%	0.92 [0.62 , 1.36]
Subtotal (95% CI)		**177**		**184**	**100.0%**	**0.90 [0.63 , 1.27]**
Total events:	44		51			
Heterogeneity: Chi² = 0.05, df = 1 (P = 0.82); I² = 0%						
Test for overall effect: Z = 0.62 (P = 0.53)						

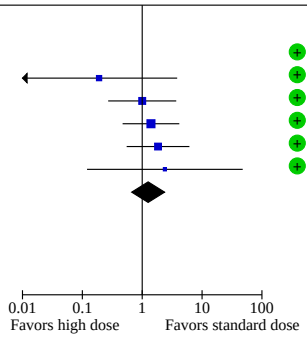


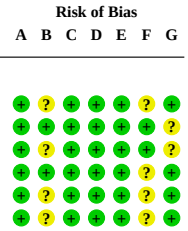
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.11. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 11: Intraventricular hemorrhage, grade 3 to 4

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.11.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Steer 2003 (1)	0	40	0	21		Not estimable
Mohd 2021	0	40	2	38	15.9%	0.19 [0.01 , 3.84]
McPherson 2015	4	37	4	37	24.8%	1.00 [0.27 , 3.70]
Mohammed 2015	7	60	5	60	31.0%	1.40 [0.47 , 4.17]
Steer 2004	7	140	4	147	24.2%	1.84 [0.55 , 6.14]
Steer 2003	2	45	0	21	4.2%	2.39 [0.12 , 47.72]
Subtotal (95% CI)		**362**		**324**	**100.0%**	**1.26 [0.67 , 2.36]**
Total events:	20		15			
Heterogeneity: Chi² = 2.23, df = 4 (P = 0.69); I² = 0%						
Test for overall effect: Z = 0.71 (P = 0.48)						





Footnotes

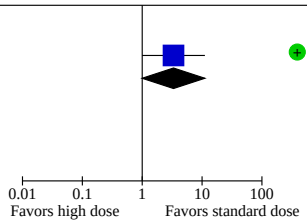
(1) infants with the highest and second highest dose # "Steer 2003" and "(1)Steer 2003", resp.; infants with the lowest dose # "Steer 2003" and "(1)Steer 2003" (21 each)

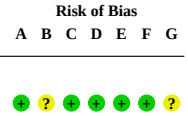
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.12. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 12: Cerebellar hemorrhage at brain ultrasound (yes/no)

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.12.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
McPherson 2015	10	37	3	37	100.0%	3.33 [1.00 , 11.15]
Subtotal (95% CI)		**37**		**37**	**100.0%**	**3.33 [1.00 , 11.15]**
Total events:	10		3			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.95 (P = 0.05)						



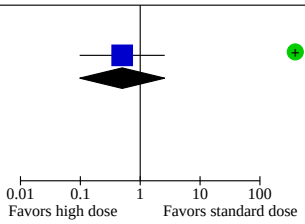


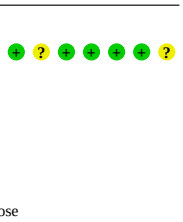
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.13. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 13: Magnetic resonance imaging (MRI) abnormalities at term equivalent age (yes/no), defined as white matter lesions (i.e. cavitations [Rutherford 2010]) and punctate lesions (Cornette 2002); germinal matrix-intraventricular hemorrhage (Parodi 2015); or cerebellar hemorrhage (Limperopoulos 2007)

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.13.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
McPherson 2015	2	37	4	37	100.0%	0.50 [0.10 , 2.56]
Subtotal (95% CI)		**37**		**37**	**100.0%**	**0.50 [0.10 , 2.56]**
Total events:	2		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.83 (P = 0.41)						

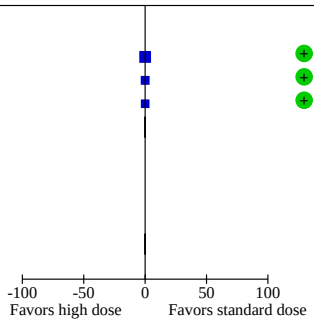


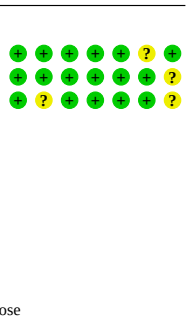


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.14. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 14: Periventricular leukomalacia

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
1.14.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Mohammed 2015	5	60	4	60	44.1%	0.02 [-0.08 , 0.11]
Mohd 2021	2	40	1	38	28.7%	0.02 [-0.06 , 0.11]
McPherson 2015	3	37	2	37	27.2%	0.03 [-0.09 , 0.14]
Subtotal (95% CI)		**137**		**135**	**100.0%**	**0.02 [-0.04 , 0.08]**
Total events:	10		7			
Heterogeneity: Chi² = 0.02, df = 2 (P = 0.99); I² = 0%						
Test for overall effect: Z = 0.74 (P = 0.46)						
Total (95% CI)		**137**		**135**	**100.0%**	**0.02 [-0.04 , 0.08]**
Total events:	10		7			
Heterogeneity: Chi² = 0.02, df = 2 (P = 0.99); I² = 0%						
Test for overall effect: Z = 0.74 (P = 0.46)						
Test for subgroup differences: Not applicable						

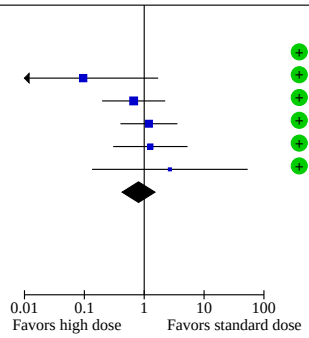


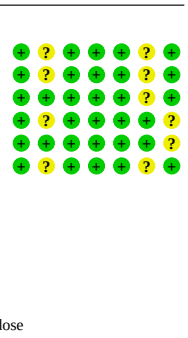


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.15. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 15: Necrotizing enterocolitis (proven = Bell stage of 2 or greater) (Bell 1978)

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.15.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Steer 2003	0	45	0	21		Not estimable
Steer 2004 (1)	0	116	5	122	26.7%	0.10 [0.01 , 1.71]
Mohammed 2015	4	60	6	60	29.9%	0.67 [0.20 , 2.24]
McPherson 2015	6	37	5	37	24.9%	1.20 [0.40 , 3.59]
Mohd 2021	4	40	3	38	15.3%	1.27 [0.30 , 5.29]
Steer 2003 (2)	2	40	0	21	3.2%	2.68 [0.13 , 53.45]
Subtotal (95% CI)		**338**		**299**	**100.0%**	**0.80 [0.43 , 1.51]**
Total events:	16		19			
Heterogeneity: Chi² = 3.71, df = 4 (P = 0.45); I² = 0%						
Test for overall effect: Z = 0.68 (P = 0.50)						





Footnotes

(1) In this forest plot, outcome data refer to the subgroup of infants treated for extubation management (and not to the 140 + 147 infants treated for apnea treatment)

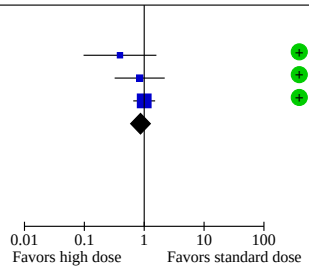
(2) infants with the highest and second highest dose # "Steer 2003" and "(1)Steer 2003", resp.; infants with the lowest dose # "Steer 2003" and "(1)Steer 2003" (21 each)

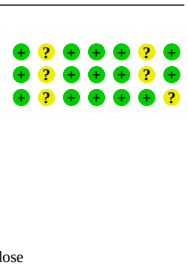
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.16. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 16: Patent ductus arteriosus (PDA) requiring treatment (cyclo-oxygenase inhibitors or surgical ligation)

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.16.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Steer 2003 (1)	3	40	4	21	16.4%	0.39 [0.10 , 1.60]
Steer 2003	9	45	5	21	21.3%	0.84 [0.32 , 2.20]
McPherson 2015	20	37	20	37	62.4%	1.00 [0.66 , 1.52]
Subtotal (95% CI)		**122**		**79**	**100.0%**	**0.87 [0.59 , 1.27]**
Total events:	32		29			
Heterogeneity: Chi² = 1.67, df = 2 (P = 0.43); I² = 0%						
Test for overall effect: Z = 0.74 (P = 0.46)						





Footnotes

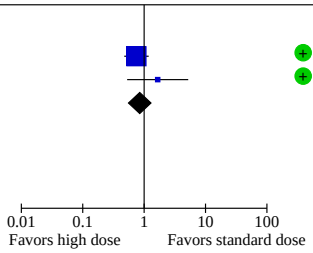
(1) infants with the highest and second highest dose # "Steer 2003" and "(1)Steer 2003", resp.; infants with the lowest dose # "Steer 2003" and "(1)Steer 2003" (21 each)

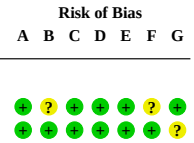
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.17. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 17: Retinopathy of prematurity (ROP) (any ROP) (International Committee 2005)

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.17.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Steer 2004	25	140	35	147	89.3%	0.75 [0.47 , 1.19]
Mohd 2021	7	40	4	38	10.7%	1.66 [0.53 , 5.23]
Subtotal (95% CI)		**180**		**185**	**100.0%**	**0.85 [0.56 , 1.29]**
Total events:	32		39			
Heterogeneity: Chi² = 1.60, df = 1 (P = 0.21); I² = 38%						
Test for overall effect: Z = 0.77 (P = 0.44)						
Test for subgroup differences: Not applicable						

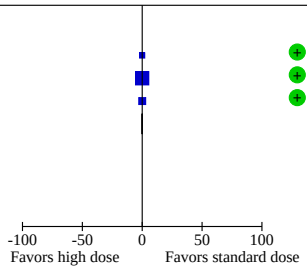


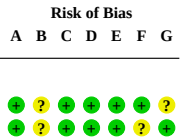


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.18. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 18: Retinopathy of prematurity (ROP) (severe ROP [stage 3 or greater]) (International Committee 2005)

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
1.18.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
McPherson 2015	2	37	4	37	15.4%	-0.05 [-0.18 , 0.07]
Steer 2004	3	140	9	147	59.7%	-0.04 [-0.09 , 0.01]
Mohammed 2015	5	60	5	60	25.0%	0.00 [-0.10 , 0.10]
Subtotal (95% CI)		**237**		**244**	**100.0%**	**-0.03 [-0.07 , 0.01]**
Total events:	10		18			
Heterogeneity: Chi² = 0.64, df = 2 (P = 0.73); I² = 0%						
Test for overall effect: Z = 1.52 (P = 0.13)						

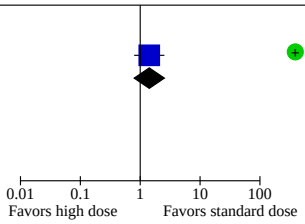




Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.19. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 19: Seizures (clinically diagnosed; diagnosed by electroencephalography)

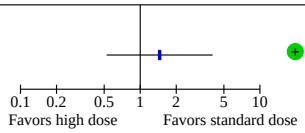
Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.19.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
McPherson 2015	17	37	12	37	100.0%	1.42 [0.79 , 2.53]
Subtotal (95% CI)		**37**		**37**	**100.0%**	**1.42 [0.79 , 2.53]**
Total events:	17		12			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.17 (P = 0.24)						

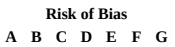


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.20. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 20: Developmental delay (Bayley Mental Developmental Index or Griffiths Mental Development Scale in children aged 18 to 24 months

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Risk Ratio M-H, Fixed, 95% CI
1.20.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose					
McPherson 2015	6	19	5	23	1.45 [0.52 , 4.03]

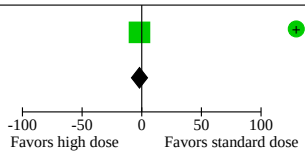


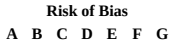


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.21. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 21: Bayley-III cognitive score in children at 18 to 24 months CA

Study or Subgroup	High-dose strategies Mean	SD	Total	Standard-dose strategies Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
McPherson 2015	86.1	12.8	19	88	8.4	23	100.0%	-1.90 [-8.60 , 4.80]
Total (95% CI)			**19**			**23**	**100.0%**	**-1.90 [-8.60 , 4.80]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.56 (P = 0.58)								
Test for subgroup differences: Not applicable								

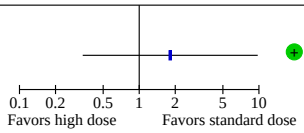




Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.22. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 22: Cerebral palsy in children aged 18 to 24 months

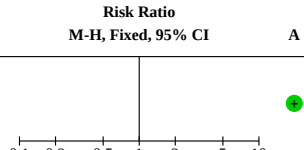
Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Risk Ratio M-H, Fixed, 95% CI
1.22.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose					
McPherson 2015	3	19	2	23	1.82 [0.34 , 9.77]

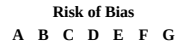


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.23. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 23: Blindness in children aged 18 to 24 months

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Risk Ratio M-H, Fixed, 95% CI
1.23.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose					
McPherson 2015	0	19	0	23	Not estimable

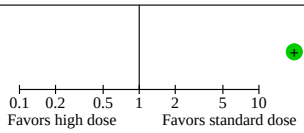




Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.24. Comparison 1: High- versus standard-dose strategies for any indication, Outcome 24: Deafness in children aged 18 to 24 months

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Risk Ratio M-H, Fixed, 95% CI
1.24.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose					
McPherson 2015	0	19	0	23	Not estimable



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2. High- versus standard-dose strategies for prevention of apnea

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1 All-cause mortality prior to hospital discharge	2	152	Risk Ratio (M-H, Fixed, 95% CI)	1.32 [0.50, 3.53]

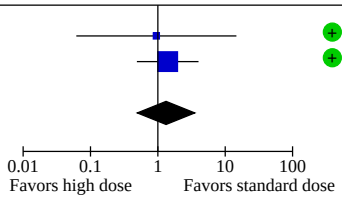
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.2 Major neurodevelopmental disability in children aged three to five years	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
2.2.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
2.3 Failure to extubate within one week of commencing treatment	1	78	Risk Ratio (M-H, Fixed, 95% CI)	0.95 [0.34, 2.69]
2.4 Apnea: number of infants with at least one episode (defined as interruption of breathing for more than 20 seconds) after 24 hours from commencing treatment, in a 24-hour period and over one week	1	78	Risk Ratio (M-H, Fixed, 95% CI)	1.01 [0.59, 1.75]
2.5 Side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine	1	78	Risk Ratio (M-H, Fixed, 95% CI)	1.19 [0.34, 4.09]
2.6 Bronchopulmonary dysplasia/chronic lung disease at 36 weeks' postmenstrual age	2	152	Risk Ratio (M-H, Fixed, 95% CI)	0.89 [0.61, 1.30]
2.7 Number of days using mechanical ventilation	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.8 Intraventricular hemorrhage, any grade	1	74	Risk Ratio (M-H, Fixed, 95% CI)	0.83 [0.41, 1.69]
2.9 Intraventricular hemorrhage, grade 3 to 4	2	152	Risk Ratio (M-H, Fixed, 95% CI)	0.68 [0.22, 2.15]
2.10 Cerebellar hemorrhage at brain ultrasound (yes/no)	1	74	Risk Ratio (M-H, Fixed, 95% CI)	3.33 [1.00, 11.15]
2.11 Magnetic resonance imaging (MRI) abnormalities at term equivalent age (yes/no), defined as white matter lesions (i.e. cavitations [Rutherford 2010]) and punctate lesions (Cornette 2002); germinal matrix-intraventricular hemorrhage (Parodi 2015); or cerebellar hemorrhage (Limperopoulos 2007)	1	74	Risk Ratio (M-H, Fixed, 95% CI)	0.50 [0.10, 2.56]
2.12 Periventricular leukomalacia	2	152	Risk Ratio (M-H, Fixed, 95% CI)	1.64 [0.41, 6.59]
2.13 Necrotizing enterocolitis (proven = Bell stage of 2 or greater) (Bell 1978)	2	152	Risk Ratio (M-H, Fixed, 95% CI)	1.23 [0.51, 2.93]
2.14 Patent ductus arteriosus (PDA) requiring treatment (cyclo-oxygenase inhibitors or surgical ligation)	1	74	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.66, 1.52]
2.15 Retinopathy of prematurity (ROP) (any ROP) (International Committee 2005)	1	78	Risk Ratio (M-H, Fixed, 95% CI)	1.66 [0.53, 5.23]

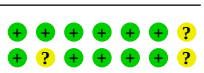
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.16 Retinopathy of prematurity (ROP) (severe ROP [stage 3 or greater]) (International Committee 2005)	1	74	Risk Ratio (M-H, Fixed, 95% CI)	0.50 [0.10, 2.56]
2.17 Seizures (clinically diagnosed; diagnosed by electroencephalography)	1	74	Risk Ratio (M-H, Fixed, 95% CI)	1.42 [0.79, 2.53]
2.18 Developmental delay (Bayley Mental Developmental Index or Griffiths Mental Development Scale in children aged 18 to 24 months	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
2.18.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
2.19 Bayley-III cognitive score in children at 18 to 24 montsh CA	1	42	Mean Difference (IV, Fixed, 95% CI)	-1.90 [-8.60, 4.80]
2.20 Cerebral palsy in children aged 18 to 24 months	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
2.20.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
2.21 Blindness in children aged 18 to 24 months	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
2.21.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
2.22 Deafness in children aged 18 to 24 months	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
2.22.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected

Analysis 2.1. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 1: All-cause mortality prior to hospital discharge

Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Mohd 2021	1	40	1	38	17.0%	0.95 [0.06 , 14.65]
McPherson 2015	7	37	5	37	83.0%	1.40 [0.49 , 4.01]
Total (95% CI)		**77**		**75**	**100.0%**	**1.32 [0.50 , 3.53]**
Total events:	8		6			

Heterogeneity: Chi² = 0.07, df = 1 (P = 0.80); I² = 0%
Test for overall effect: Z = 0.56 (P = 0.58)
Test for subgroup differences: Not applicable

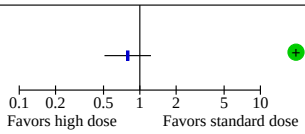


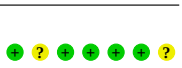


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.2. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 2: Major neurodevelopmental disability in children aged three to five years

Study or Subgroup	High dose strategies Events	Total	Standard dose strategies Events	Total	Risk Ratio M-H, Fixed, 95% CI
2.2.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose					
McPherson 2015	12	21	18	25	0.79 [0.51 , 1.24]



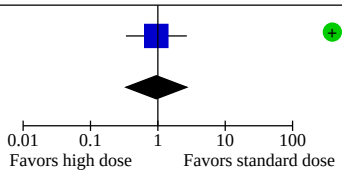


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.3. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 3: Failure to extubate within one week of commencing treatment

Study or Subgroup	High dose Events	High dose Total	Standard dose Events	Standard dose Total	Weight	Risk Ratio M-H, Fixed, 95% CI	Risk Ratio M-H, Fixed, 95% CI	Risk of Bias A B C D E F G
Mohd 2021 (1)	6	40	6	38	100.0%	0.95 [0.34 , 2.69]		+ + + + + + ?
Total (95% CI)		**40**		**38**	**100.0%**	**0.95 [0.34 , 2.69]**		
Total events:	6		6					

Heterogeneity: Not applicable
Test for overall effect: Z = 0.10 (P = 0.92)
Test for subgroup differences: Not applicable



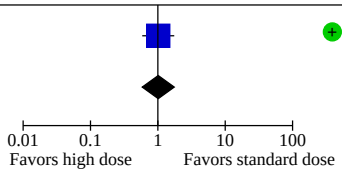
Footnotes
(1) Timeframe not specified

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.4. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 4: Apnea: number of infants with at least one episode (defined as interruption of breathing for more than 20 seconds) after 24 hours from commencing treatment, in a 24-hour period and over one week

Study or Subgroup	High dose Events	High dose Total	Standard dose Events	Standard dose Total	Weight	Risk Ratio M-H, Fixed, 95% CI	Risk Ratio M-H, Fixed, 95% CI	Risk of Bias A B C D E F G
Mohd 2021	16	40	15	38	100.0%	1.01 [0.59 , 1.75]		+ + + + + + ?
Total (95% CI)		**40**		**38**	**100.0%**	**1.01 [0.59 , 1.75]**		
Total events:	16		15					

Heterogeneity: Not applicable
Test for overall effect: Z = 0.05 (P = 0.96)
Test for subgroup differences: Not applicable

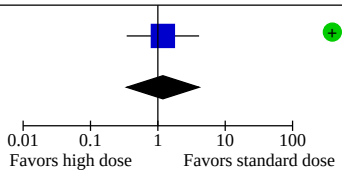


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.5. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 5: Side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine

Study or Subgroup	High dose Events	High dose Total	Standard dose Events	Standard dose Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Mohd 2021	5	40	4	38	100.0%	1.19 [0.34 , 4.09]
Total (95% CI)		**40**		**38**	**100.0%**	**1.19 [0.34 , 4.09]**
Total events:	5		4			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.27 (P = 0.79)
Test for subgroup differences: Not applicable

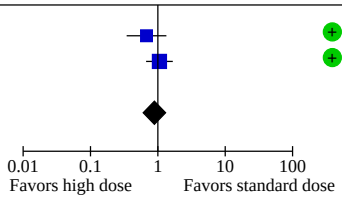


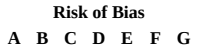
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.6. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 6: Bronchopulmonary dysplasia/chronic lung disease at 36 weeks' postmenstrual age

Study or Subgroup	High dose Events	High dose Total	Standard dose Events	Standard dose Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Mohd 2021	10	40	14	38	44.4%	0.68 [0.34 , 1.34]
McPherson 2015	19	37	18	37	55.6%	1.06 [0.67 , 1.67]
Total (95% CI)		**77**		**75**	**100.0%**	**0.89 [0.61 , 1.30]**
Total events:	29		32			

Heterogeneity: Chi² = 1.15, df = 1 (P = 0.28); I² = 13%
Test for overall effect: Z = 0.61 (P = 0.54)
Test for subgroup differences: Not applicable

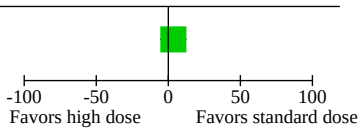




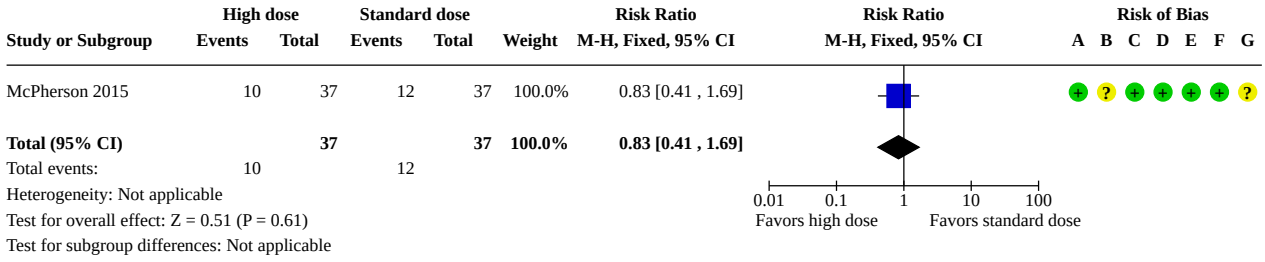
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.7. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 7: Number of days using mechanical ventilation

Study or Subgroup	High dose Mean	High dose SD	High dose Total	Standard dose Mean	Standard dose SD	Standard dose Total	Mean Difference IV, Fixed, 95% CI
McPherson 2015	15.2	22.4	37	11.7	17.4	37	3.50 [-5.64 , 12.64]

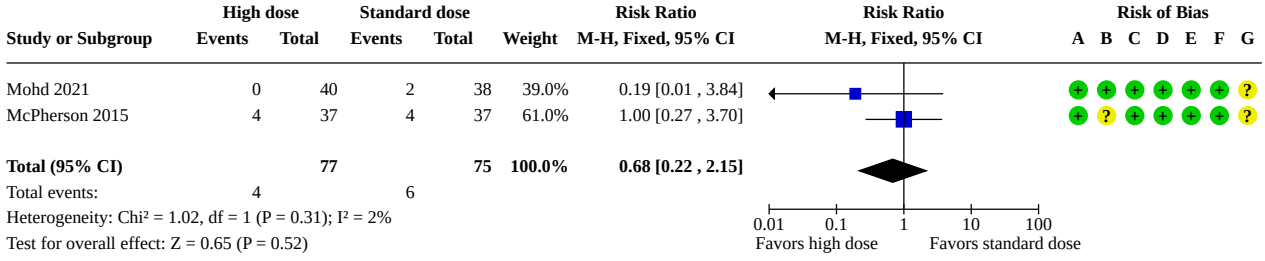


Analysis 2.8. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 8: Intraventricular hemorrhage, any grade



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.9. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 9: Intraventricular hemorrhage, grade 3 to 4

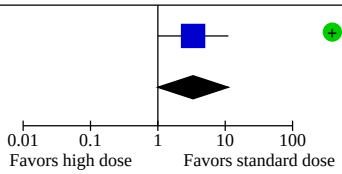


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.10. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 10: Cerebellar hemorrhage at brain ultrasound (yes/no)

Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI	Risk of Bias A B C D E F G
McPherson 2015	10	37	3	37	100.0%	3.33 [1.00 , 11.15]	+ ? + + + + ?
Total (95% CI)		**37**		**37**	**100.0%**	**3.33 [1.00 , 11.15]**	
Total events:	10		3				

Heterogeneity: Not applicable
Test for overall effect: Z = 1.95 (P = 0.05)
Test for subgroup differences: Not applicable

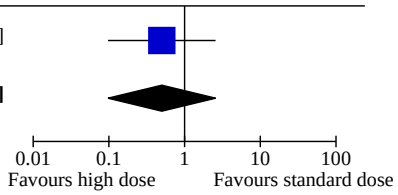


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.11. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 11: Magnetic resonance imaging (MRI) abnormalities at term equivalent age (yes/no), defined as white matter lesions (i.e. cavitations [Rutherford 2010]) and punctate lesions (Cornette 2002); germinal matrix-intraventricular hemorrhage (Parodi 2015); or cerebellar hemorrhage (Limperopoulos 2007)

Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
McPherson 2015	2	37	4	37	100.0%	0.50 [0.10 , 2.56]
Total (95% CI)		**37**		**37**	**100.0%**	**0.50 [0.10 , 2.56]**
Total events:	2		4			

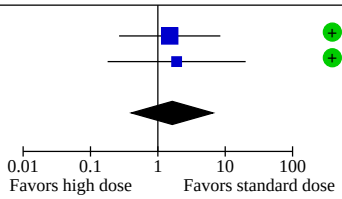
Heterogeneity: Not applicable
Test for overall effect: Z = 0.83 (P = 0.41)
Test for subgroup differences: Not applicable

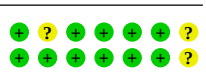


Analysis 2.12. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 12: Periventricular leukomalacia

Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
McPherson 2015	3	37	2	37	66.1%	1.50 [0.27 , 8.46]
Mohd 2021	2	40	1	38	33.9%	1.90 [0.18 , 20.10]
Total (95% CI)		**77**		**75**	**100.0%**	**1.64 [0.41 , 6.59]**
Total events:	5		3			

Heterogeneity: Chi² = 0.03, df = 1 (P = 0.87); I² = 0%
Test for overall effect: Z = 0.69 (P = 0.49)
Test for subgroup differences: Not applicable



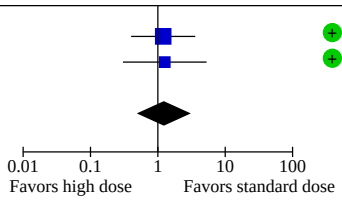


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.13. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 13: Necrotizing enterocolitis (proven = Bell stage of 2 or greater) (Bell 1978)

Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
McPherson 2015	6	37	5	37	61.9%	1.20 [0.40 , 3.59]
Mohd 2021	4	40	3	38	38.1%	1.27 [0.30 , 5.29]
Total (95% CI)		**77**		**75**	**100.0%**	**1.23 [0.51 , 2.93]**
Total events:	10		8			

Heterogeneity: Chi² = 0.00, df = 1 (P = 0.95); I² = 0%
Test for overall effect: Z = 0.46 (P = 0.65)
Test for subgroup differences: Not applicable

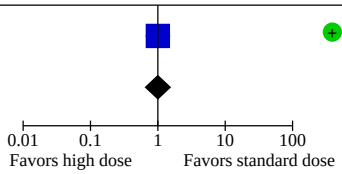


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.14. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 14: Patent ductus arteriosus (PDA) requiring treatment (cyclo-oxygenase inhibitors or surgical ligation)

Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI	Risk Ratio M-H, Fixed, 95% CI	Risk of Bias A B C D E F G
McPherson 2015	20	37	20	37	100.0%	1.00 [0.66 , 1.52]		+ ? + + + + ?
Total (95% CI)		**37**		**37**	**100.0%**	**1.00 [0.66 , 1.52]**		
Total events:	20		20					

Heterogeneity: Not applicable
Test for overall effect: Z = 0.00 (P = 1.00)
Test for subgroup differences: Not applicable

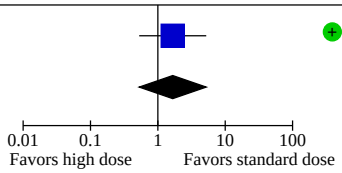


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.15. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 15: Retinopathy of prematurity (ROP) (any ROP) (International Committee 2005)

Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI	Risk Ratio M-H, Fixed, 95% CI	Risk of Bias A B C D E F G
Mohd 2021	7	40	4	38	100.0%	1.66 [0.53 , 5.23]		+ + + + + + ?
Total (95% CI)		**40**		**38**	**100.0%**	**1.66 [0.53 , 5.23]**		
Total events:	7		4					

Heterogeneity: Not applicable
Test for overall effect: Z = 0.87 (P = 0.38)
Test for subgroup differences: Not applicable

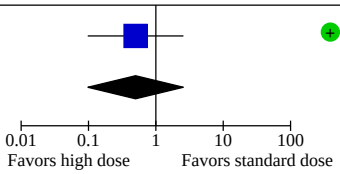


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.16. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 16: Retinopathy of prematurity (ROP) (severe ROP [stage 3 or greater]) (International Committee 2005)

Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
McPherson 2015	2	37	4	37	100.0%	0.50 [0.10 , 2.56]
Total (95% CI)		**37**		**37**	**100.0%**	**0.50 [0.10 , 2.56]**
Total events:	2		4			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.83 (P = 0.41)
Test for subgroup differences: Not applicable

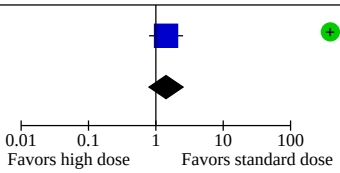


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.17. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 17: Seizures (clinically diagnosed; diagnosed by electroencephalography)

Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
McPherson 2015	17	37	12	37	100.0%	1.42 [0.79 , 2.53]
Total (95% CI)		**37**		**37**	**100.0%**	**1.42 [0.79 , 2.53]**
Total events:	17		12			

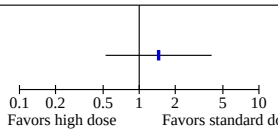
Heterogeneity: Not applicable
Test for overall effect: Z = 1.17 (P = 0.24)
Test for subgroup differences: Not applicable



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.18. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 18: Developmental delay (Bayley Mental Developmental Index or Griffiths Mental Development Scale in children aged 18 to 24 months

Study or Subgroup	High dose strategies Events	Total	Standard dose strategies Events	Total	Risk Ratio M-H, Fixed, 95% CI	Risk Ratio M-H, Fixed, 95% CI	Risk of Bias A B C D E F G
2.18.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose							
McPherson 2015	6	19	5	23	1.45 [0.52 , 4.03]		+ ? + + + + ?

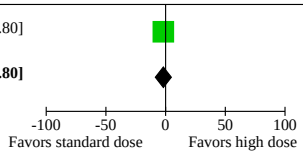


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.19. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 19: Bayley-III cognitive score in children at 18 to 24 montsh CA

Study or Subgroup	High dose strategies Mean	SD	Total	Standard dose strategies Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI	Risk of Bias A B C D E F G
McPherson 2015	86.1	12.8	19	88	8.4	23	100.0%	-1.90 [-8.60 , 4.80]		+ ? + + + + ?
Total (95% CI)			**19**			**23**	**100.0%**	**-1.90 [-8.60 , 4.80]**		

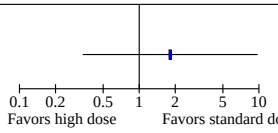
Heterogeneity: Not applicable
Test for overall effect: Z = 0.56 (P = 0.58)
Test for subgroup differences: Not applicable



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

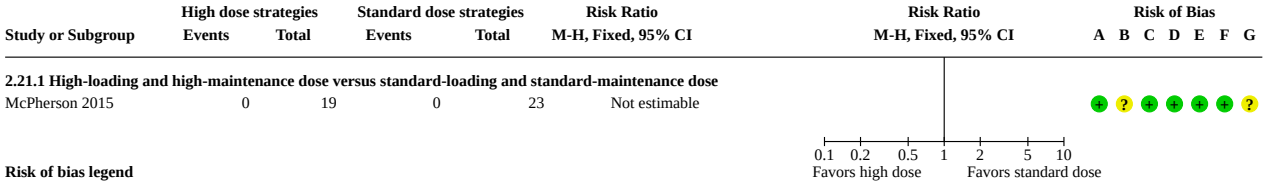
Analysis 2.20. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 20: Cerebral palsy in children aged 18 to 24 months

Study or Subgroup	High dose strategies Events	Total	Standard dose strategies Events	Total	Risk Ratio M-H, Fixed, 95% CI	Risk Ratio M-H, Fixed, 95% CI	Risk of Bias A B C D E F G
2.20.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose							
McPherson 2015	3	19	2	23	1.82 [0.34 , 9.77]		+ ? + + + + ?



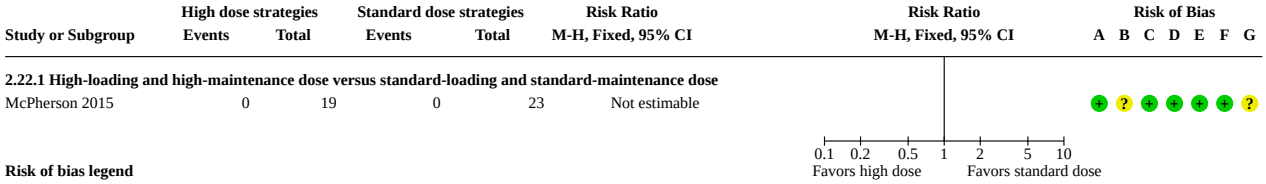
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.21. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 21: Blindness in children aged 18 to 24 months



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 2.22. Comparison 2: High- versus standard-dose strategies for prevention of apnea, Outcome 22: Deafness in children aged 18 to 24 months



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

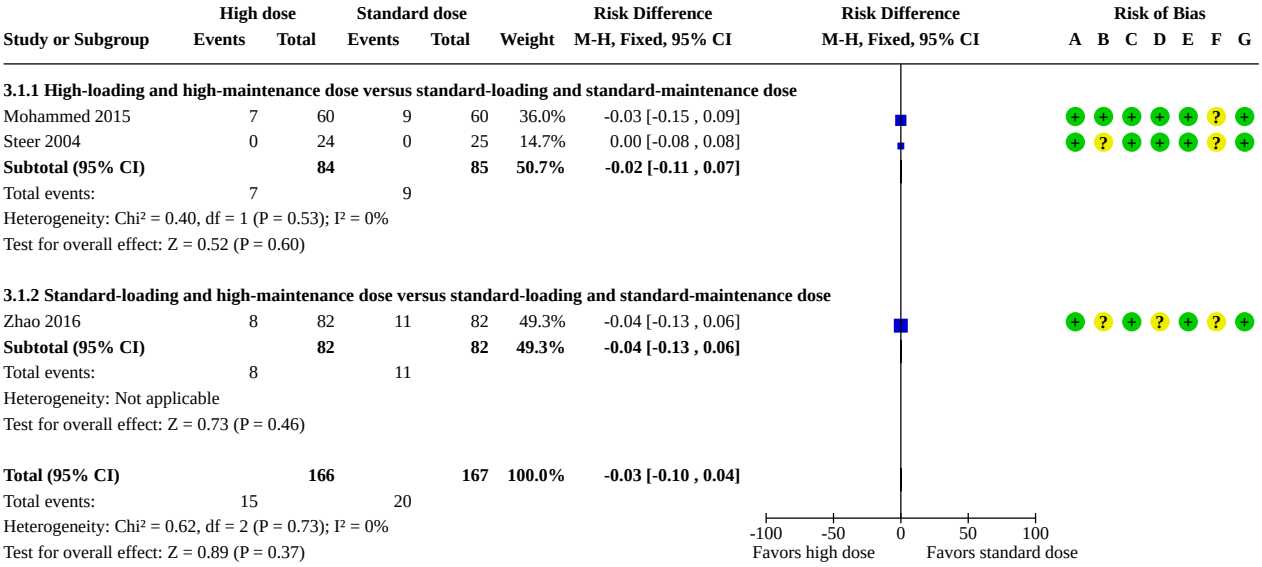
Comparison 3. High- versus standard-dose strategies for treatment of apnea

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.1 All-cause mortality prior to hospital discharge	3	333	Risk Difference (M-H, Fixed, 95% CI)	-0.03 [-0.10, 0.04]
3.1.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	2	169	Risk Difference (M-H, Fixed, 95% CI)	-0.02 [-0.11, 0.07]
3.1.2 Standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	164	Risk Difference (M-H, Fixed, 95% CI)	-0.04 [-0.13, 0.06]
3.2 Failure to extubate within one week of commencing treatment	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.2.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	78	Risk Ratio (M-H, Fixed, 95% CI)	0.48 [0.24, 0.92]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.3 Side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.3.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	2	150	Risk Ratio (M-H, Fixed, 95% CI)	1.92 [0.53, 6.90]
3.4 Bronchopulmonary dysplasia/chronic lung disease at 36 weeks' postmenstrual age	3	333	Risk Ratio (M-H, Fixed, 95% CI)	0.72 [0.47, 1.11]
3.4.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	2	169	Risk Ratio (M-H, Fixed, 95% CI)	0.75 [0.42, 1.35]
3.4.2 Standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	164	Risk Ratio (M-H, Fixed, 95% CI)	0.68 [0.36, 1.29]
3.5 Intraventricular hemorrhage, any grade	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.5.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	49	Risk Ratio (M-H, Fixed, 95% CI)	0.21 [0.03, 1.66]
3.6 Intraventricular hemorrhage, grade 3 to 4	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.6.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	2	169	Risk Ratio (M-H, Fixed, 95% CI)	1.14 [0.46, 2.82]
3.7 Periventricular leukomalacia	1	120	Risk Ratio (M-H, Fixed, 95% CI)	1.25 [0.35, 4.43]
3.7.1 High-loading and high-maintenance dose versus standar-loading and standard-maintenance dose	1	120	Risk Ratio (M-H, Fixed, 95% CI)	1.25 [0.35, 4.43]
3.8 Necrotizing enterocolitis (proven = Bell stage of 2 or greater) (Bell 1978)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.8.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	120	Risk Ratio (M-H, Fixed, 95% CI)	0.67 [0.20, 2.24]
3.9 Retinopathy of prematurity (ROP) (any ROP) (International Committee 2005)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.9.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	1	49	Risk Ratio (M-H, Fixed, 95% CI)	4.17 [0.50, 34.66]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.10 Retinopathy of prematurity (ROP) (severe ROP [stage 3 or greater]) (International Committee 2005)	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.10.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose	2	169	Risk Ratio (M-H, Fixed, 95% CI)	0.85 [0.29, 2.54]

Analysis 3.1. Comparison 3: High- versus standard-dose strategies for treatment of apnea, Outcome 1: All-cause mortality prior to hospital discharge

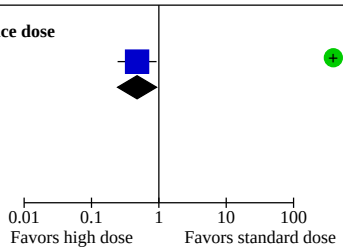


Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.2. Comparison 3: High- versus standard-dose strategies for treatment of apnea, Outcome 2: Failure to extubate within one week of commencing treatment

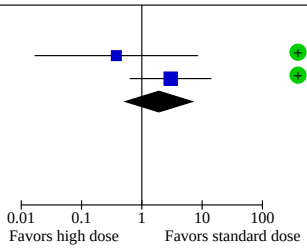
Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.2.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Mohammed 2015	9	40	18	38	100.0%	0.47 [0.24 , 0.92]
Subtotal (95% CI)		**40**		**38**	**100.0%**	**0.47 [0.24 , 0.92]**
Total events:	9		18			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.19 (P = 0.03)						
Test for subgroup differences: Not applicable						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.3. Comparison 3: High- versus standard-dose strategies for treatment of apnea, Outcome 3: Side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine

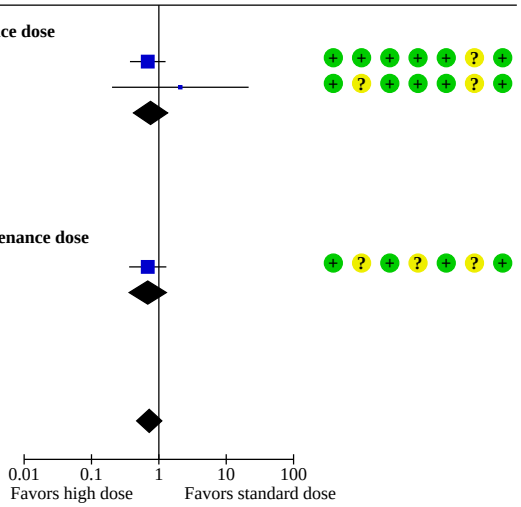
Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.3.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Scanlon 1992	0	14	1	16	41.3%	0.38 [0.02 , 8.59]
Mohammed 2015	6	60	2	60	58.7%	3.00 [0.63 , 14.27]
Subtotal (95% CI)		**74**		**76**	**100.0%**	**1.92 [0.53 , 6.90]**
Total events:	6		3			
Heterogeneity: Chi² = 1.35, df = 1 (P = 0.24); I² = 26%						
Test for overall effect: Z = 1.00 (P = 0.32)						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.4. Comparison 3: High- versus standard-dose strategies for treatment of apnea, Outcome 4: Bronchopulmonary dysplasia/chronic lung disease at 36 weeks' postmenstrual age

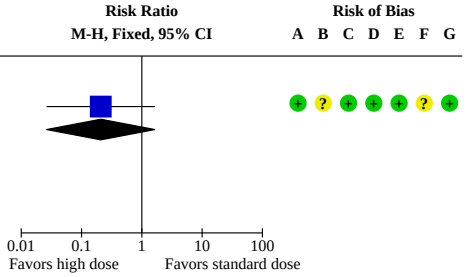
Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.4.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Mohammed 2015	13	60	19	60	48.7%	0.68 [0.37 , 1.26]
Steer 2004	2	24	1	25	2.5%	2.08 [0.20 , 21.50]
Subtotal (95% CI)		**84**		**85**	**51.3%**	**0.75 [0.42 , 1.35]**
Total events:	15		20			
Heterogeneity: Chi² = 0.83, df = 1 (P = 0.36); I² = 0%						
Test for overall effect: Z = 0.95 (P = 0.34)						
3.4.2 Standard-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Zhao 2016	13	82	19	82	48.7%	0.68 [0.36 , 1.29]
Subtotal (95% CI)		**82**		**82**	**48.7%**	**0.68 [0.36 , 1.29]**
Total events:	13		19			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.17 (P = 0.24)						
Total (95% CI)		**166**		**167**	**100.0%**	**0.72 [0.47 , 1.11]**
Total events:	28		39			
Heterogeneity: Chi² = 0.85, df = 2 (P = 0.65); I² = 0%						
Test for overall effect: Z = 1.50 (P = 0.13)						
Test for subgroup differences: Chi² = 0.05, df = 1 (P = 0.83), I² = 0%						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.5. Comparison 3: High- versus standard-dose strategies for treatment of apnea, Outcome 5: Intraventricular hemorrhage, any grade

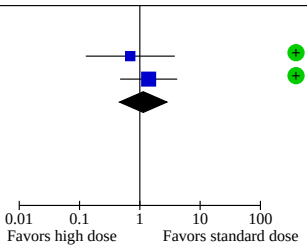
Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.5.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Steer 2004	1	24	5	25	100.0%	0.21 [0.03 , 1.66]
Subtotal (95% CI)		**24**		**25**	**100.0%**	**0.21 [0.03 , 1.66]**
Total events:	1		5			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.48 (P = 0.14)						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.6. Comparison 3: High- versus standard-dose strategies for treatment of apnea, Outcome 6: Intraventricular hemorrhage, grade 3 to 4

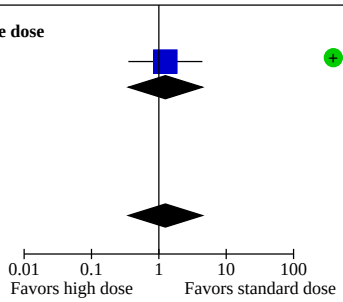
Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.6.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Steer 2004	2	24	3	25	37.0%	0.69 [0.13 , 3.80]
Mohammed 2015	7	60	5	60	63.0%	1.40 [0.47 , 4.17]
Subtotal (95% CI)		**84**		**85**	**100.0%**	**1.14 [0.46 , 2.82]**
Total events:	9		8			
Heterogeneity: Chi² = 0.46, df = 1 (P = 0.50); I² = 0%						
Test for overall effect: Z = 0.28 (P = 0.78)						

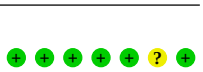


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.7. Comparison 3: High- versus standard-dose strategies for treatment of apnea, Outcome 7: Periventricular leukomalacia

Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.7.1 High-loading and high-maintenance dose versus standar-loading and standard-maintenance dose						
Mohammed 2015	5	60	4	60	100.0%	1.25 [0.35 , 4.43]
Subtotal (95% CI)		**60**		**60**	**100.0%**	**1.25 [0.35 , 4.43]**
Total events:	5		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.35 (P = 0.73)						
Total (95% CI)		**60**		**60**	**100.0%**	**1.25 [0.35 , 4.43]**
Total events:	5		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.35 (P = 0.73)						
Test for subgroup differences: Not applicable						

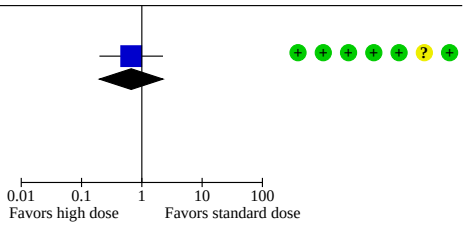




Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.8. Comparison 3: High- versus standard-dose strategies for treatment of apnea, Outcome 8: Necrotizing enterocolitis (proven = Bell stage of 2 or greater) (Bell 1978)

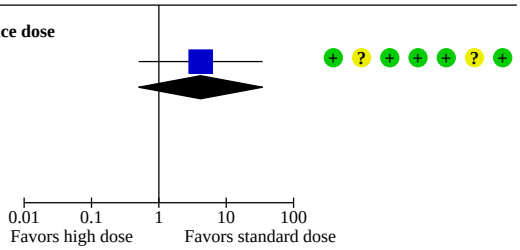
Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.8.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Mohammed 2015	4	60	6	60	100.0%	0.67 [0.20 , 2.24]
Subtotal (95% CI)		**60**		**60**	**100.0%**	**0.67 [0.20 , 2.24]**
Total events:	4		6			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.65 (P = 0.51)						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

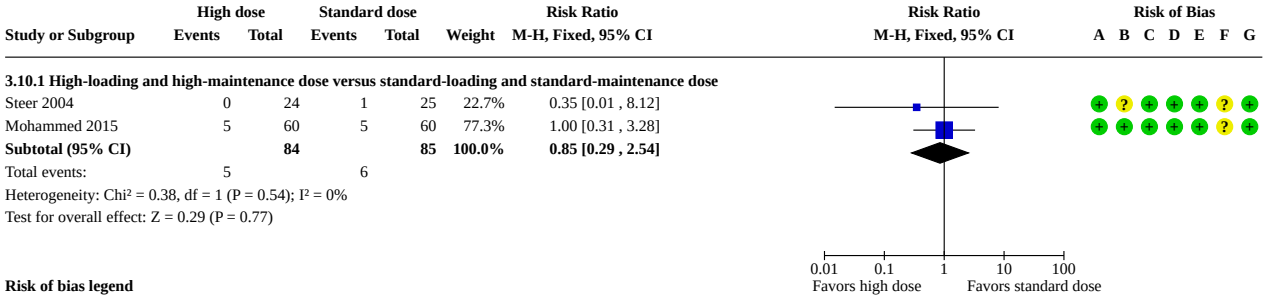
Analysis 3.9. Comparison 3: High- versus standard-dose strategies for treatment of apnea, Outcome 9: Retinopathy of prematurity (ROP) (any ROP) (International Committee 2005)

Study or Subgroup	High dose Events	Total	Standard dose Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.9.1 High-loading and high-maintenance dose versus standard-loading and standard-maintenance dose						
Steer 2004	4	24	1	25	100.0%	4.17 [0.50 , 34.66]
Subtotal (95% CI)		**24**		**25**	**100.0%**	**4.17 [0.50 , 34.66]**
Total events:	4		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.32 (P = 0.19)						
Test for subgroup differences: Not applicable						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 3.10. Comparison 3: High- versus standard-dose strategies for treatment of apnea, Outcome 10: Retinopathy of prematurity (ROP) (severe ROP [stage 3 or greater]) (International Committee 2005)



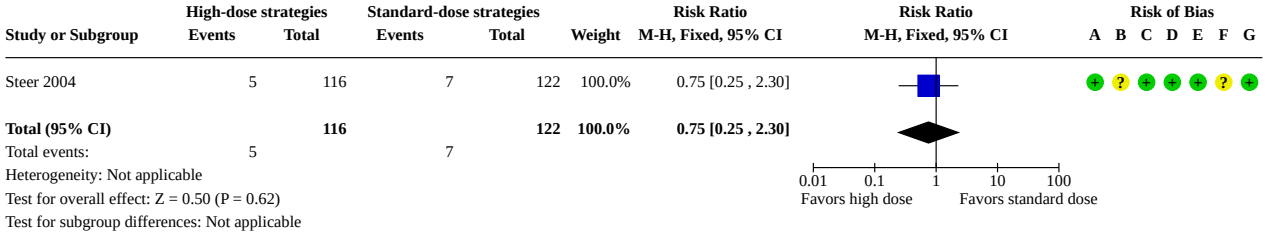
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 4. High- versus standard-dose strategies for the prevention of re-intubation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.1 All-cause mortality prior to hospital discharge	1	238	Risk Ratio (M-H, Fixed, 95% CI)	0.75 [0.25, 2.30]
4.2 Failure to extubate within one week of commencing treatment	2	365	Risk Ratio (M-H, Fixed, 95% CI)	0.52 [0.36, 0.74]
4.3 Reintubation within one week of commencing treatment	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
4.4 Side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine	2	365	Risk Ratio (M-H, Fixed, 95% CI)	1.86 [0.68, 5.09]
4.5 Bronchopulmonary dysplasia/chronic lung disease: 28 days of oxygen exposure	1	238	Risk Ratio (M-H, Fixed, 95% CI)	0.84 [0.68, 1.04]
4.6 Bronchopulmonary dysplasia/chronic lung disease at 36 weeks' postmenstrual age	1	238	Risk Difference (M-H, Fixed, 95% CI)	-0.13 [-0.25, -0.01]
4.7 Number of days using mechanical ventilation	1	127	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.89, 0.85]
4.8 Intraventricular hemorrhage, any grade	1	238	Risk Ratio (M-H, Fixed, 95% CI)	1.02 [0.68, 1.53]
4.9 Intraventricular hemorrhage, grade 3 to 4	2	365	Risk Ratio (M-H, Fixed, 95% CI)	4.08 [0.74, 22.55]
4.10 Necrotizing enterocolitis (proven = Bell stage of 2 or greater) (Bell 1978)	2	365	Risk Ratio (M-H, Fixed, 95% CI)	0.38 [0.08, 1.79]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.11 Patent ductus arteriosus (PDA) requiring treatment (cyclo-oxygenase inhibitors or surgical ligation)	1	127	Risk Ratio (M-H, Fixed, 95% CI)	0.65 [0.30, 1.41]
4.12 Retinopathy of prematurity (ROP) (any ROP) (International Committee 2005)	1	238	Risk Ratio (M-H, Fixed, 95% CI)	0.65 [0.40, 1.05]
4.13 Retinopathy of prematurity (ROP) (severe ROP [stage 3 or greater]) (International Committee 2005)	1	238	Risk Ratio (M-H, Fixed, 95% CI)	0.39 [0.11, 1.45]

Analysis 4.1. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 1: All-cause mortality prior to hospital discharge

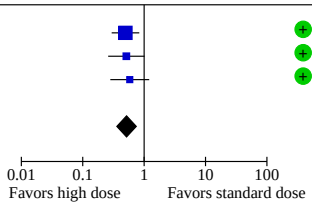


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.2. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 2: Failure to extubate within one week of commencing treatment

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Steer 2004	17	116	36	122	58.0%	0.50 [0.30 , 0.83]
Steer 2003	11	45	10	21	22.5%	0.51 [0.26 , 1.02]
Steer 2003 (1)	10	40	9	21	19.5%	0.58 [0.28 , 1.21]
Total (95% CI)		**201**		**164**	**100.0%**	**0.52 [0.36 , 0.74]**
Total events:	38		55			

Heterogeneity: Chi² = 0.13, df = 2 (P = 0.94); I² = 0%
Test for overall effect: Z = 3.55 (P = 0.0004)
Test for subgroup differences: Not applicable



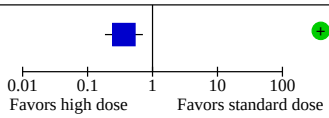
Footnotes
(1) infants with the highest and second highest dose # "Steer 2003" and "(1)Steer 2003", resp.; infants with the lowest dose # "Steer 2003" and "(1)Steer 2003" (21 each)

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.3. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 3: Reintubation within one week of commencing treatment

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Risk Ratio M-H, Fixed, 95% CI
Steer 2004	10	116	29	122	0.36 [0.19 , 0.71]

Test for subgroup differences: Not applicable

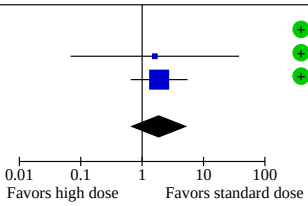


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.4. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 4: Side effects (tachycardia, agitation, or feed intolerance) leading to a reduction in dose or withholding of caffeine

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Steer 2003	0	45	0	21		Not estimable
Steer 2003 (1)	1	40	0	21	11.8%	1.61 [0.07 , 37.88]
Steer 2004	9	116	5	122	88.2%	1.89 [0.65 , 5.48]
Total (95% CI)		**201**		**164**	**100.0%**	**1.86 [0.68 , 5.09]**
Total events:	10		5			

Risk Ratio M-H, Fixed, 95% CI



Risk of Bias: A B C D E F G

Heterogeneity: Chi² = 0.01, df = 1 (P = 0.92); I² = 0%
Test for overall effect: Z = 1.21 (P = 0.23)
Test for subgroup differences: Not applicable

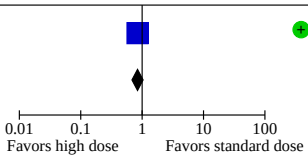
Footnotes
(1) infants with the highest and second highest dose # "Steer 2003" and "(1)Steer 2003", resp.; infants with the lowest dose # "Steer 2003" and "(1)Steer 2003" (21 each)

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.5. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 5: Bronchopulmonary dysplasia/chronic lung disease: 28 days of oxygen exposure

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Steer 2004	64	116	80	122	100.0%	0.84 [0.68 , 1.04]
Total (95% CI)		**116**		**122**	**100.0%**	**0.84 [0.68 , 1.04]**
Total events:	64		80			

Risk Ratio M-H, Fixed, 95% CI



Risk of Bias: A B C D E F G

Heterogeneity: Not applicable
Test for overall effect: Z = 1.62 (P = 0.10)
Test for subgroup differences: Not applicable

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.6. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 6: Bronchopulmonary dysplasia/chronic lung disease at 36 weeks' postmenstrual age

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
Steer 2004	33	116	51	122	100.0%	-0.13 [-0.25 , -0.01]
Total (95% CI)		**116**		**122**	**100.0%**	**-0.13 [-0.25 , -0.01]**
Total events:	33		51			

Risk Difference M-H, Fixed, 95% CI: -100, -50, 0, 50, 100; Favors high dose; Favors standard dose

Risk of Bias: A B C D E F G

Heterogeneity: Not applicable
Test for overall effect: Z = 2.18 (P = 0.03)
Test for subgroup differences: Not applicable

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.7. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 7: Number of days using mechanical ventilation

Study or Subgroup	High-dose strategies Mean	SD	Total	Standard-dose strategies Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Steer 2003 (1)	3.1	2.2	40	3.5	2.3	21	52.6%	-0.40 [-1.60 , 0.80]
Steer 2003	3.9	2.7	45	3.5	2.3	21	47.4%	0.40 [-0.86 , 1.66]
Total (95% CI)			**85**			**42**	**100.0%**	**-0.02 [-0.89 , 0.85]**

Mean Difference IV, Fixed, 95% CI: -100, -50, 0, 50, 100; Favors high dose; Favors standard dose

Heterogeneity: Chi² = 0.81, df = 1 (P = 0.37); I² = 0%
Test for overall effect: Z = 0.05 (P = 0.96)
Test for subgroup differences: Not applicable

Footnotes
(1) infants with the highest and second highest dose # "Steer 2003" and "(1)Steer 2003", resp.; infants with the lowest dose # "Steer 2003" and "(1)Steer 2003" (21 each

Analysis 4.8. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 8: Intraventricular hemorrhage, any grade

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Steer 2004	33	116	34	122	100.0%	1.02 [0.68 , 1.53]
Total (95% CI)		**116**		**122**	**100.0%**	**1.02 [0.68 , 1.53]**
Total events:	33		34			

Risk Ratio M-H, Fixed, 95% CI: 0.01, 0.1, 1, 10, 100; Favors high dose; Favors standard dose

Risk of Bias: A B C D E F G

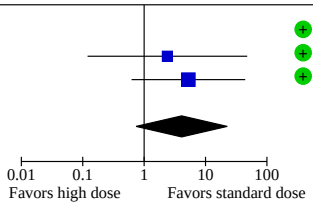
Heterogeneity: Not applicable
Test for overall effect: Z = 0.10 (P = 0.92)
Test for subgroup differences: Not applicable

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.9. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 9: Intraventricular hemorrhage, grade 3 to 4

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Steer 2003 (1)	0	40	0	21		Not estimable
Steer 2003	2	45	0	21	41.0%	2.39 [0.12 , 47.72]
Steer 2004	5	116	1	122	59.0%	5.26 [0.62 , 44.34]
Total (95% CI)		**201**		**164**	**100.0%**	**4.08 [0.74 , 22.55]**
Total events:	7		1			

Heterogeneity: Chi² = 0.18, df = 1 (P = 0.67); I² = 0%
Test for overall effect: Z = 1.61 (P = 0.11)
Test for subgroup differences: Not applicable



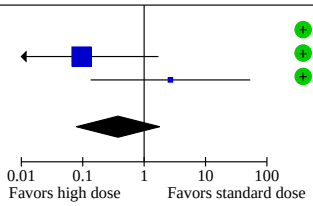
Footnotes
(1) infants with the highest and second highest dose # "Steer 2003" and "(1)Steer 2003", resp.; infants with the lowest dose # "Steer 2003" and "(1)Steer 2003" (21 each)

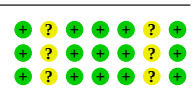
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.10. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 10: Necrotizing enterocolitis (proven = Bell stage of 2 or greater) (Bell 1978)

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Steer 2003	0	45	0	21		Not estimable
Steer 2004	0	116	5	122	89.2%	0.10 [0.01 , 1.71]
Steer 2003 (1)	2	40	0	21	10.8%	2.68 [0.13 , 53.45]
Total (95% CI)		**201**		**164**	**100.0%**	**0.38 [0.08 , 1.79]**
Total events:	2		5			

Heterogeneity: Chi² = 2.52, df = 1 (P = 0.11); I² = 60%
Test for overall effect: Z = 1.23 (P = 0.22)
Test for subgroup differences: Not applicable





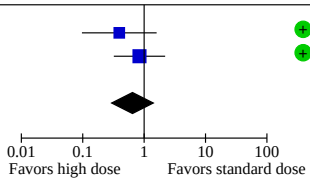
Footnotes
(1) infants with the highest and second highest dose # "Steer 2003" and "(1)Steer 2003", resp.; infants with the lowest dose # "Steer 2003" and "(1)Steer 2003" (21 each)

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.11. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 11: Patent ductus arteriosus (PDA) requiring treatment (cyclo-oxygenase inhibitors or surgical ligation)

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Steer 2003 (1)	3	40	4	21	43.5%	0.39 [0.10 , 1.60]
Steer 2003	9	45	5	21	56.5%	0.84 [0.32 , 2.20]
Total (95% CI)		**85**		**42**	**100.0%**	**0.65 [0.30 , 1.41]**
Total events:	12		9			

Heterogeneity: Chi² = 0.77, df = 1 (P = 0.38); I² = 0%
Test for overall effect: Z = 1.10 (P = 0.27)
Test for subgroup differences: Not applicable



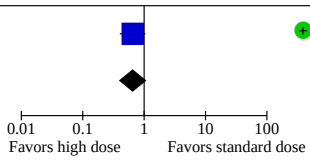
Footnotes
(1) infants with the highest and second highest dose # "Steer 2003" and "(1)Steer 2003", resp.; infants with the lowest dose # "Steer 2003" and "(1)Steer 2003" (21 each)

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.12. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 12: Retinopathy of prematurity (ROP) (any ROP) (International Committee 2005)

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Steer 2004	21	116	34	122	100.0%	0.65 [0.40 , 1.05]
Total (95% CI)		**116**		**122**	**100.0%**	**0.65 [0.40 , 1.05]**
Total events:	21		34			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.76 (P = 0.08)
Test for subgroup differences: Not applicable

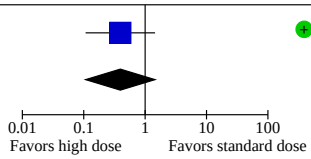


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 4.13. Comparison 4: High- versus standard-dose strategies for the prevention of re-intubation, Outcome 13: Retinopathy of prematurity (ROP) (severe ROP [stage 3 or greater]) (International Committee 2005)

Study or Subgroup	High-dose strategies Events	Total	Standard-dose strategies Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Steer 2004	3	116	8	122	100.0%	0.39 [0.11 , 1.45]
Total (95% CI)		**116**		**122**	**100.0%**	**0.39 [0.11 , 1.45]**
Total events:	3		8			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.40 (P = 0.16)
Test for subgroup differences: Not applicable



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

ADDITIONAL TABLES

Table 1. Overview of the included studies, ordered by indication

High-dose versus standard-dose strategies

Study ID	Enrollment period	Country	GA (weeks)	Sample size	Age at study entry	High dose	Standard dose	Indication	Subgroup by dose
McPherson 2015	2008 to 2010	US	Mean (SD): 25.8 (2.0) high dose Mean (SD): 26.5 (1.9) standard dose.	74	Within 24 h	LD 40 mg/kg + 20 mg/kg after 12 h; MD 10 mg/kg at 24 and 36 h; then every 24 hs	LD 20 mg/kg MD 10 mg/kg	Apnea prevention	High LD versus standard-LD
Mohd 2021	2019 to 2020	Malaysia	Median (IQR): 30 (2.7) high dose Median (IQR): 29.8 (3.3) standard dose	78	Median 2 weeks	LD 40 mg/kg MD 20 mg/kg	LD 20 mg/kg MD 10 mg/kg	Apnea prevention	High LD versus standard-LD
Scanlon 1992	Not reported	UK	Mean (SD): 28.2 (1.1) high dose Mean (SD): 28.7 (1.2) standard dose	44	Mean 6 days	LD 50 mg/kg MD 12 mg/kg	LD 25mg/kg MD 6 mg/kg	Apnea treatment	High LD versus standard LD
Zhao 2016[a]	2013 to 2014	China	Mean (SD): 29.8 (3.4) high dose Mean (SD): 29.9 (2.7) standard dose	164	Mean 4 days	LD 20 mg/kg MD 15 mg/kg	LD 20 mg/kg MD 5 mg/kg	Apnea treatment	Standard LD in both groups
Mohammed 2015	2011 to 2012	Egypt	Mean (SD): 29.4 (2.0) high dose Mean (SD): 29.8 (1.9) standard dose	120	Mean 3 days	LD 40 mg/kg MD 20 mg/kg	LD 20 mg/kg	Apnea treatment	High LD versus standard LD

Table 1. Overview of the included studies, ordered by indication *(Continued)*

							MD 10 mg/kg		
Steer 2004	1996 to 1999	Australia	Mean (SD): 27.3 (1.4) high dose Mean (SD): 27.5 (1.4) standard dose	287	Median 4 days, within 12 days	LD 80 mg/kg MD 20 mg/kg	LD 20 mg/kg MD 5 mg/kg	Apnea treatment and extubation management	High LD versus standard LD
Steer 2003[b]	1993 to 1995	Australia	Mean (SD): 27.8 (1.9) high dose 1 Mean (SD): 28.4 (1.7) high dose 2 Mean (SD): 28 (1.8) standard dose	127	Mean 4 days	LD 60 mg/kg or 30 mg/kg; MD 30 mg/kg or 15 mg/kg	LD 6 mg/kg MD 3 mg/kg	Extubation management	High LD versus standard LD

GA: gestational age; **IQR:** interquartile range; **LD:** loading dose; **MD**: maintenance dose; **SD**: standard deviation

Footnotes

[a]In all studies infants were randomized to high-loading and high-maintenance or standard-loading and standard-maintenance dose except Zhao 2016, where infants were randomized to standard-loading and high-maintenance or standard-loading and standard-maintenance dose.

[b]In Steer 2003, infants were randomized in three groups: two of them match our definition of high dose and one standard dose. Outcome data are reported according to the footnotes in comparisons 1 and 3 (e.g. Analysis 1.1).

Definition of the doses

- High-loading dose: more than 25 mg of caffeine citrate/kg
- Standard-loading dose: 25 mg or less of caffeine citrate/kg
- High-maintenance dose: more than 10 mg of caffeine citrate/kg/day
- Standard-maintenance dose: 10 mg or less of caffeine citrate/kg/day

APPENDICES

Appendix 1. Search strategies

Update searches – May 2022

MEDLINE via Pubmed, 17 May 2022

#1 "Caffeine"[Mesh] OR caffeine OR caffedrine OR coffein OR cafeine OR methylxanthine OR trimethylxanthine 41,316

#2 (infant, newborn[MeSH] OR newborn*[TIAB] OR "new born*"[TIAB] OR "newly born"[TIAB] OR baby*[TIAB] OR babies[TIAB] OR premature[TIAB] OR prematurity[TIAB] OR preterm[TIAB] OR "pre term"[TIAB] OR “low birth weight”[TIAB] OR "low birthweight"[TIAB] OR VLBW[TIAB] OR LBW[TIAB] OR infan*[TIAB] OR neonat*[TIAB]) 1,261,232

#3 (randomized controlled trial [pt] OR controlled clinical trial [pt] OR randomized [tiab] OR randomised [tiab] OR placebo [tiab] OR drug therapy [sh] OR randomly [tiab] OR trial [tiab] OR groups [tiab] OR randomly [tiab]) OR ((single[tiab] OR doubl*[tiab] OR tripl*[tiab] OR treb*[tiab]) AND (blind*[tiab] OR mask*[tiab])) 5,468,966

#4 #2 AND #3 236,831

#5 ("Animals"[Mesh]) NOT "Humans"[Mesh] 5,004,967

#6 #4 NOT #5 215,803
#7 #1 AND #6 636

#8 (("2021/02/15"[Date - Create]: "3000"[Date - Create])) AND #7 42

Cinahl via EBSCOhost, 17 May 2022

S1	infant OR infants OR infant’s OR infantile OR infancy OR newborn* OR "new born" OR "new borns" OR "newly born" OR neonat* OR baby* OR babies OR premature OR prematures OR prematurity OR preterm OR preterms OR "pre term" OR premies OR "low birth weight" OR "low birthweight" OR VLBW OR LBW	520,047
S2	randomized controlled trial OR controlled clinical trial OR randomized OR randomised OR placebo OR clinical trials as topic OR randomly OR trial OR PT clinical trial	587,317
S3	(MH "Caffeine") OR caffeine OR caffedrine OR coffein OR cafeine OR methylxanthine OR trimethylxanthine	6725
S4	S1 AND S2 AND S3	117
S5	Published Date: 20210101-20221231	11

Cochrane Library, 17 May 2022
#1 MeSH descriptor: [Infant, Newborn] explode all trees 17,416

#2 (infant* OR infantile OR infancy OR newborn* or "new born" or "new borns" or "newly born" or neonat* or baby* or babies or premature or prematures or prematurity or preterm or preterms or "pre term" or premies or "low birth weight" or "low birthweight" or VLBW or LBW or ELBW or NICU):ti,ab,kw (Word variations have been searched) 96,587

#3 #1 OR #2 96,587

#4 MeSH descriptor: [Caffeine] explode all trees 2260

#5 (caffeine OR caffedrine OR coffein OR cafeine OR methylxanthine OR trimethylxanthine):ti,ab,kw (Word variations have been searched) 5113

#6 #4 OR #5 5113

#7 #3 AND #6 384

#8 Limit to Cochrane Library publication date from Feb 2021 to May 2022 38

Embase.com, 20 May 2022

#01. 'caffeine'/exp OR 'methylxanthine'/exp OR methylxanthine OR caffeine OR caffedrine OR coffein OR cafeine OR 'trimethylxanthine'/exp OR trimethylxanthine 67,020

#02. 'newborn'/de OR 'prematurity'/de OR 'newborn intensive care'/de OR 'newborn care'/de OR 'gestational age'/de 825,925

#03. (babe or babes or baby* or babies or 'gestational age$' or infant$ or infantile or infancy or 'low birth weight' or 'low birthweight' or neonat* or 'neo-nat*' or newborn* or 'new born$' or 'newly born' or premature or pre-mature or pre-matures or prematures or prematurity or pre-maturity or preterm or preterms or 'pre term$' or preemie or preemies or premies or premie or VLBW or VLBWI or VLBW-I or VLBWs or LBW or LBWI or LBWs or ELBW or ELBWI or ELBWs or NICU or NICUs):ti,ab,kw 1,242,974

#04. #2 OR #3 1,551,172

#05. 'randomized controlled trial'/de OR 'controlled clinical trial'/de 887,470

#06. random*:ti,ab,kw 1,794,554

#07. 'randomization'/de 93,824

#08. placebo:ti,ab,kw 342,172

#09. ((double OR single OR doubly OR singly) NEAR/2 (blind OR blinded OR blindly)):ti,ab,kw 260,000

#10. 'double blind procedure'/de 195,491

#11. (controlled NEAR/7 (study OR design OR trial)):ti,ab,kw 416,999

#12. 'parallel group$':ti,ab 29,309

#13. crossover:ti,ab OR 'cross over':ti,ab 116,629

#14. ((assign* OR match OR matched OR allocation) NEAR/5 (alternate OR group$ OR intervention$ OR patient$ OR subject$ OR participant $)):ti,ab 378,723

#15. (open NEAR/2 label):ti,ab 96,646

#16. quasirandom*:ti,ab,kw OR 'quasi random*':ti,ab,kw OR randomi*:ti,ab,kw OR randomly:ti,ab,kw 1,463,204

#17. (control* NEAR/2 (group$ OR random*)):ti,ab,kw 1,190,410

#18. #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 3,071,763

#19. ('animal'/exp OR 'invertebrate'/exp OR 'animal experiment'/de OR 'animal model'/de OR 'animal tissue'/de OR 'animal cell'/de OR 'nonhuman'/de) AND ('human'/de OR 'normal human'/de OR 'human cell'/de) 24,781,602

#20. 'animal'/exp OR 'invertebrate'/exp OR 'animal experiment'/de OR 'animal model'/de OR 'animal tissue'/de OR 'animal cell'/de OR 'nonhuman'/de 32,278,921

#21. #20 NOT #19 7,497,319

#22. #18 NOT #21 2,635,458

#23. #1 AND #4 AND #22 497

Number of hits from literature databases May 2022: 588

Clinical trial registries

WHO International Clinical Trials Registry Platform (ICTRP), 20 May 2022

Advanced search

Title: infant* OR infantile OR infancy OR newborn* or "new born" or "new borns" or "newly born" or neonat* or baby* or babies or premature or prematures or prematurity or preterm or preterms or "pre term" or premies or "low birth weight" or "low birthweight" or VLBW or LBW or ELBW or NICU

Intervention: caffeine

Limit: search for clinical trials in children **21**

Clinicaltrials.gov, 20 May 2022

Advanced search
Intervention/treatment: caffeine
Other terms: premature OR prematurity OR preterms OR preterm OR “very low birth” OR “low birth weight” OR newborn OR newborns OR neonate OR neonates OR infant OR infants
Age group: Child
First posted from 16/02/2021 to 20/05/2022
No further limits applied **1**

ISRCTN: no new studies

Number of hits from trial registries May 2022: 22

Total number of hits May 2022: 610

Update searches - February 2021

Pubmed, 16 February 2021

#1 "Caffeine"[Mesh] OR caffeine OR caffedrine OR coffein OR cafeine OR methylxanthine OR trimethylxanthine 39,560

#2 (infant, newborn[MeSH] OR newborn*[TIAB] OR "new born*"[TIAB] OR "newly born"[TIAB] OR baby*[TIAB] OR babies[TIAB] OR premature[TIAB] OR prematurity[TIAB] OR preterm[TIAB] OR "pre term"[TIAB] OR “low birth weight”[TIAB] OR "low birthweight"[TIAB] OR VLBW[TIAB] OR LBW[TIAB] OR infan*[TIAB] OR neonat*[TIAB]) 1,198,267

#3 (randomized controlled trial [pt] OR controlled clinical trial [pt] OR randomized [tiab] OR randomised [tiab] OR placebo [tiab] OR drug therapy [sh] OR randomly [tiab] OR trial [tiab] OR groups [tiab] OR randomly [tiab]) OR ((single[tiab] OR doubl*[tiab] OR tripl*[tiab] OR treb*[tiab]) AND (blind*[tiab] OR mask*[tiab])) 5,023,016

#4 #2 AND #3 219,930

#5 ("Animals"[Mesh]) NOT "Humans"[Mesh] 4,789,041

#6 #4 NOT #5 200,262

#7 #1 AND #6 590

Cinahl via EBSCOhost, 16 February 2021

S4	S1 AND S2 AND S3	117
S3	(MH "Caffeine") OR caffeine OR caffedrine OR coffein OR cafeine OR methylxanthine OR trimethylxanthine	6725
S2	randomized controlled trial OR controlled clinical trial OR randomized OR randomised OR placebo OR clinical trials as topic OR randomly OR trial OR PT clinical trial	587,317
S1	infant or infants or infant's or infantile or infancy or newborn* or "new born" or "new borns" or "newly born" or neonat* or baby* or babies or premature or prematures or prematurity or preterm or preterms or "pre term" or premies or "low birth weight" or "low birthweight" or VLBW or LBW	520,047

Cochrane Library, 16 February 2021

#1 MeSH descriptor: [Infant, Newborn] explode all trees 16,060

#2 (infant* OR infantile OR infancy OR newborn* or "new born" or "new borns" or "newly born" or neonat* or baby* or babies or premature or prematures or prematurity or preterm or preterms or "pre term" or premies or "low birth weight" or "low birthweight" or VLBW or LBW or ELBW or NICU):ti,ab,kw (Word variations have been searched) 90,071

#3 #1 OR #2 90,071

#4 MeSH descriptor: [Caffeine] explode all trees 2100

#5 (caffeine OR caffedrine OR coffein OR cafeine OR methylxanthine OR trimethylxanthine):ti,ab,kw (Word variations have been searched) 4756

#6 #4 OR #5 4756

#7 #3 AND #6 352

352 records from Cochrane Library
- 16 Cochrane reviews
- 2 Cochrane protocols
- 332 Trials
- 2 Clinical answers

Number of hits from all databases:1059

Number after deduplication in EndNote: 810

ClinicalTrials.gov

Advanced search
Intervention/treatment: caffeine

Other terms: premature OR prematurity OR preterms OR preterm OR “very low birth” OR “low birth weight” OR newborn OR newborns OR neonate OR neonates OR infant OR infants No further limits applied 44 trials

ISRCTN

Advanced search
Interventions: caffeine
Participant age range: neonate 2 trials

Appendix 2. Risk of bias tool

We will use the standard methods of Cochrane and Cochrane Neonatal to assess the methodological quality of the trials. For each trial, we will seek information regarding the method of randomization, blinding, and reporting of all outcomes of all the infants enrolled in the trial. We will assess each criterion as being at a low, high, or unclear risk of bias. Two review authors will separately assess each study. We will resolve any disagreements by discussion. We will add this information to the Characteristics of included studies table. We will evaluate the following issues and enter the findings into the Risk of bias table.

Sequence generation (checking for possible selection bias). Was the allocation sequence adequately generated?

For each included study, we will categorize the method used to generate the allocation sequence as being at:

- low risk of bias (any truly random process, e.g. random number table; computer random number generator);
- high risk of bias (any non-random process, e.g. odd or even date of birth; hospital or clinic record number); or
- unclear risk of bias.

Allocation concealment (checking for possible selection bias). Was allocation adequately concealed?

For each included study, we will categorize the method used to conceal the allocation sequence as being at:

- low risk of bias (e.g. telephone or central randomization; consecutively numbered sealed opaque envelopes);
- high risk of bias (open random allocation; unsealed or non-opaque envelopes, alternation; date of birth); or
- unclear risk.

Blinding of participants and personnel (checking for possible performance bias). Was knowledge of the allocated intervention adequately prevented during the study?

For each included study, we will categorize the methods used to blind study participants and personnel from knowledge of which intervention a participant received. Blinding will be assessed separately for different outcomes or class of outcomes. We will categorize the methods as being at:

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

Blinding of outcome assessment (checking for possible detection bias). Was knowledge of the allocated intervention adequately prevented at the time of outcome assessment?

For each included study, we will categorize the methods used to blind outcome assessment. Blinding will be assessed separately for different outcomes or class of outcomes. We will categorize the methods as being at:

- low risk of bias for outcome assessors;
- high risk of bias for outcome assessors; or
- unclear risk of bias for outcome assessors.

Incomplete outcome data (checking for possible attrition bias through withdrawals, dropouts, protocol deviations). Were incomplete outcome data adequately addressed?

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

- low risk of bias (less than 20% missing data);
- high risk of bias (20% or greater missing data); or

- unclear risk of bias.

Selective reporting bias. Are reports of the study free of suggestion of selective outcome reporting?

For each included study, we will describe how we investigated the possibility of selective outcome reporting bias and what we found. We will search study protocols of the included trials in ClinicalTrials.gov; the World Health Organization's International Trials Registry and Platform, and the ISRCTN Registry. For studies in which study protocols were published in advance, we will compare prespecified outcomes versus outcomes eventually reported in the published results. If the study protocols were not published in advance, we will contact study authors to gain access to the study protocol. We will assess the likelihood of selective reporting bias as:

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

Other sources of bias. Was the study apparently free of other problems that could put it at a high risk of bias?

For each included study, we will describe any important concerns we had about other possible sources of bias (e.g. whether there was a potential source of bias related to the specific study design or whether the trial was stopped early due to some data-dependent process). We will assess whether each study is at:

- low risk of other bias;
- high risk of other bias; or
- unclear risk of other bias.

If needed, we will explore the impact of the level of bias through undertaking sensitivity analyses.

HISTORY

Protocol first published: Issue 2, 2021

CONTRIBUTIONS OF AUTHORS

MB identified the topic of the review, drafted the protocol, selected eligible studies, resolved disagreements, assessed risk of bias, helped GRADE assessment, and wrote and revised the final review.

RS identified the topic of the review, drafted the protocol, and wrote and revised the final review.

CR screened search outputs, assessed study eligibility, extracted and synthesized data, assessed risk of bias, undertook GRADE assessment, and wrote and revised the final review.

PB screened search outputs, assessed study eligibility, extracted and synthesized data, assessed risk of bias, undertook GRADE assessment, and wrote and revised the final review.

WO wrote and revised the final review.

DECLARATIONS OF INTEREST

MB: no known conflicts of interest. MB is an Associate Editor with the Cochrane Neonatal Group but took no part in editorial acceptance or review of this manuscript.

PB: no known conflicts of interest.

CR: no known conflicts of interest.

WO: no known conflicts of interest.

PGD is co-author of publications relevant to the interventions in the work; works as a neonatologist at The Royal Women's Hospital, Melbourne.

RFS is the Co-ordinating Editor of the Cochrane Neonatal Review Group, President and Director of Clinical Trials of the Vermont Oxford Network, and a professor at the University of Vermont. RFS took no part in the editorial processes for this review.

SOURCES OF SUPPORT

Internal sources

- Institute for Clinical Sciences, Lund University, Lund, Sweden

  MB is employed by this organization
- World Health Organization, Switzerland

  Funding for this review was provided by the World Health Organization Geneva
- Australian National Health and Medical Research Council (NHMRC), Australia

  Research performed by Peter Davis is supported by the Australian NHMRC

External sources

- Region Skåne, Skåne University Hospital, Lund University and Region Västra Götaland, Sweden, Sweden

  Cochrane Sweden is supported from Region Skåne, Skåne University Hospital Lund University and Region Västra Götaland
- Australian Satellite of the Cochrane Neonatal Review Group, Australia

  The Australasian Satellite of Cochrane Neonatal aims to increase the number of Australasian authors with a published review and a published protocol. The Satellite provides high-level support to reviewers to increase capacity within the Australasian neonatal community for new reviews, protocols and review updates.
- Vermont Oxford Network, USA

  Cochrane Neonatal Reviews are produced with support from Vermont Oxford Network, a worldwide collaboration of health professionals dedicated to providing evidence-based care of the highest quality for newborn infants and their families

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We made a number of post-hoc additions and minor changes to the protocol. We changed the Criteria for considering studies for this review to better align it with the structure of a related review, "Methylxanthine for the prevention and treatment of apnea in preterm infants" (Marques 2021). Both this review and Marques 2021 are topics of interest to WHO, and so the parallel structure better addresses questions posed by WHO.

Objectives

- We removed the secondary objective early versus late discontinuation of caffeine administration because this population is distinct from the population of infants with or at risk for apnea of prematurity. We plan a separate Cochrane Review (Urru 2023) to address this issue.

Types of interventions

- The definition of high-loading dose has been changed from more than 20 mg of caffeine citrate/kg to more than 25 mg of caffeine citrate/kg.

Types of outcomes

- We added the outcome apnea: number of infants with at least one episode.
- We added the outcome mortality or major neurodevelopmental disability in children aged 18 to 24 months and 3 to 5 years CA to the summary of findings tables as it is a primary outcome.

Search

- We searched Embase in addition to the databases listed in the protocol.

Comparisons

- We changed the structure of the comparisons, reporting by indication rather than by dose strategies; the latter are analyzed in subgroup analyses.
- In the comparison high-loading dose versus standard-loading dose, we pre-specified in the protocol (Types of interventions) "each arm of this comparison had to give identical maintenance doses following the different loading doses". In the full review studies with different maintenance doses have been included

Subgroup analyses

- The text in the Methods section for subgroup analyses has been updated following the latest version of the Cochrane Neonatal template.
- We removed the following subgroup analyses:
  - chronological age;
  - timing of caffeine initiation;
  - post-extubation respiratory support: high- and standard-loading dose;
  - intubation (intubated newborns versus non-intubated newborns) for indications other than prevention of apnea.

INDEX TERMS

Medical Subject Headings (MeSH)

Apnea; *Bronchopulmonary Dysplasia [prevention & control]; Caffeine; Infant, Extremely Premature; *Infant, Premature, Diseases

MeSH check words

Child; Humans; Infant; Infant, Newborn